Available online at www.ijpab.com

DOI: http://dx.doi.org/10.18782/2582-2845.9196

ISSN: 2582 – 2845

Ind. J. Pure App. Biosci. (2026) *14*(1), 10-18

Indian Journal of Pure & Applied Biosciences

Peer-Reviewed, Refereed, Open Access Journal

Research Article

In Silico Genome-Wide Identification and Analysis of Mirror Repeats in Nipah Virus

Kavita[1], Kavita Saini[2] and Vikash Bhardwaj[3*]

[1]Ph.D. Scholar, Department of Zoology,

[*3]Associate Professor, Department of Zoology,

Baba Mastnath University, Asthal Bohar, Rohtak, Haryana, India

[2]Assistant Professor, Govt. College for Women, Gohana, Sonipat, Haryana, India

*Corresponding Author E-mail: vikashbhardwaj@gmail.com

Received: 4.11.2025 | Revised: 28.12.2025 | Accepted: 22.01.2026

ABSTRACT

The Nipah virus (NiV), a highly pathogenic RNA virus recognized by the World Health Organization as an epidemic threat, poses serious public health concerns due to its high mortality rate and human-to-human transmission capability. This study aims to investigate the presence and distribution of mirror repeats (MRs)—a class of repetitive DNA elements characterized by symmetrical sequences—within the complete Nipah virus (NiV) genome using a manual bioinformatics approach. The 18,246-base genome was segmented into 500 bp fragments, and BLAST alignment was conducted between each fragment and its reverse complement to detect MRs. A total of 401 mirror repeats were identified, comprising 375 perfect and 26 imperfect repeats. Filtering for perfect mirror repeats longer than 10 bp yielded 29 significant motifs, of which 24 contained spacer elements. Comparative analysis with other RNA viruses, including SARS-CoV-2, Zika virus (ZIKV), Marburg virus (MARV), and Ebola virus (EBOV), revealed both conserved and unique mirror repeats, suggesting potential evolutionary or functional roles in genome organization, replication, or regulation. Notably, the selected mirror repeats identified in Nipah virus (NiV) were also detected in Zika virus (ZIKV) but were absent in Marburg (MARV), Ebola (EBOV), and SARS-CoV-2, indicating a shared genomic feature between Nipah (NiV) and Zika viruses (ZIKV) that requires further investigation for biological significance and therapeutic relevance.

Keywords: *Nipahvirus, Mirror repeats, BLAST, Bioinformatics, Genome Analysis*

INTRODUCTION

The World Health Organization has identified the Nipah virus (NiV) as an epidemic threat since 2015. It is considered one of the most dangerous emerging viruses becauseit can be transmitted from person to personand exhibits high mortality rates (Skowron et al., 2022; Gurley et al., 2020).

Cite this article: Kavita, Saini, K., & Bhardwaj, V. (2026). In Silico Genome-Wide Identification and Analysis of Mirror Repeats in Nipah Virus, *Ind. J. Pure App. Biosci. 14*(1), 10-18. doi: http://dx.doi.org/10.18782/2582-2845.9196

NiV is classified as a biological safety level 4 (BSL 4) disease because of the lack of effective therapy or vaccines (Rathish et al., 2021; Epstein et al., 2006). The Nipah virus belongs to the family Paramyxoviridae and the genus Henipavirus (Yoneda et al., 2006; Joshi et al., 2023; Bowden et al., 2008; Lou et al., 2006). It is a type of RNA virus that has a single strand of genetic material, a non-segmented genome, negative-sense, has a protective envelope, and helical in shape (Singh et al., 2019; Garbuglia et al., 2023; & Sharma et al., 2019). The initial symptoms are usually a fever, headaches, and respiratory issues. Sleepiness, confusion, a gradual coma, and even death are possible outcomes of subsequent encephalitis (Hafeez et al., 2025; & Banerjee et al., 2019a).

The genome of the Nipah virus (NiV) is about 18.2 kilobases long and produces six structural proteins and three nonstructural proteins. The structural proteins include RNA polymerase, glycoprotein (G), matrix protein (M), fusion protein (F), phosphoprotein (P), and nucleocapsid (N). The P gene also encodes three nonstructural proteins, known as V and W proteins, and an alternate open reading frame (ORF), referred to as the C protein, through a process called RNA editing (Sun et al., 2018; Soman et al., 2020). Genetic studies have classified the Nipah virus (NiV) into two strains: NiV-Bangladesh (NiV-B), found in Bangladesh and India, and NiV-Malaysia (NiV-M), found in Malaysia and Cambodia, respectively. Although the genetic sequences of these two strains are nearly 92% similar, their ability to cause disease and spread between individuals appears to differ significantly. NiV-B seems to be more pathogenic than NiV-M, as it is associated with a higher mortality rate and a greater propensity for human-to-human transmissions (Rahman et al., 2021; Banerjee et al., 2019b; & Mire et al., 2016).

Understanding the Nipah virus's (NiV) adaptive evolution, tracking transmission, and doing epidemiological analysis all depend on whole-genome sequencing (WGS). In order to better understand the biology of pathogens, their pathogenesis, and the development of therapies, WGS is crucial for genome analysis (Rahman et al., 2025). RNA virus genomes have a variety of roles during viral replication, and the complex structural basis of these tasks is becoming clearer. Fundamentally, the genetic information required for virus reproduction is stored in its RNA genome (Nicholson et al., 2015).

Every organism's genome contains repeated sequences (Lupski et al., 1992). The presence of numerous repetitive DNA sequences in various genomes highlights their structural variety, and there are several instances where the functional significance of these repetitive elements has been examined extensively at the molecular level (Shapiro et al., 2005). The family of repetitive DNA can be found extensively within a taxonomic family or genus, or it may be unique to a particular species or chromosome. Repetitions can be located in specific areas of a genome, such as in telomeric regions or distributed unevenly across the genome. Over the course of evolutionary time, these repeats may undergo significant changes in their sequence and the number of copies (Rao et al., 2010). Among the many different kinds of repeat sequences, mirror repeats (MR) are essential to every species' genetic makeup (Yadav et al., 2023). On the same strand of the genome sequence, it is characterized as a repeat with bilateral symmetry or a centre of symmetry. A DNA mirror repeat is a portion of a sequence that is distinguished by the presence of a single strand's centre of symmetry (Langet al., 2007; Yadav et al., 2022).

Mirror repeats (MRs) are widely distributed across various phyla, including microorganisms, plants, animal viruses, and even within the human insulin gene. These sequences are associated with several functional roles, such as participation in H-DNA formation, regulation of transcription and replication, and involvement in the development of disorders affecting the nervous system, etc (Sehrawat et al., 2024). Mirror Repeats (MRs) play a vital function in a

variety of neurological conditions. These MRs have also been used to determine genome structure, create triplex DNA, and perform a variety of other genome-related operations. These triplex DNA or non-B DNA structures, which develop due to the presence of triplet repetitions, have been linked to a variety of human disorders. These structures were discovered to be mutagenic in mammalian cells, causing genomic instability and replication arrest by introducing double-strand breaks (DSBs) (Yadav et al., 2022). A perfect mirror repeat is defined as a sequence that exhibits perfect symmetry around a central axis, with the second half of the sequence being the reverse of the first.

Here, we aim to investigate the presence of mirror repeats across the whole genome of Nipah virus (NiV), with the objective of identifying them through a manual bioinformatics approach.

MATERIALS AND METHODS

The complete genome sequence of the Nipah virus (Accession: NC_002728.1 GI: 13559808), consisting of 18,246 base pairs of single-stranded RNA, was downloaded from the NCBI nucleotide database in FASTA format (https://www.ncbi.nlm.nih.gov/nucleotide/). The genome was fragmented into subsequences of 500 base pairs each. These fragmented sequences were treated as query sequences. Using the reverse complement tool (https://www.bioinformatics.org/sms/rev_comp.html), parallel subject sequences were generated. Both query and subject sequences were then aligned using the BLAST tool (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to identify local similarities.

Mirror repeats were identified by detecting alignments in which the nucleotide sequence was reversed between the query and subject sequences. Special attention was given to breakpoint regions—locations where the genome was segmented into 500 bp fragments—to detect mirror repeats near these junctions. The BLAST algorithm was run using a word size of 7, and alignments were observed across varying E-value thresholds. An E-value of 20 yielded the highest number of hits and was considered optimal for identifying mirror repeats. A reversal in the position numbers between the query and subject sequences characterized alignments indicative of mirror repeats. These mirror repeats were further classified based on the presence and type of spacer elements between them (Figure 1).

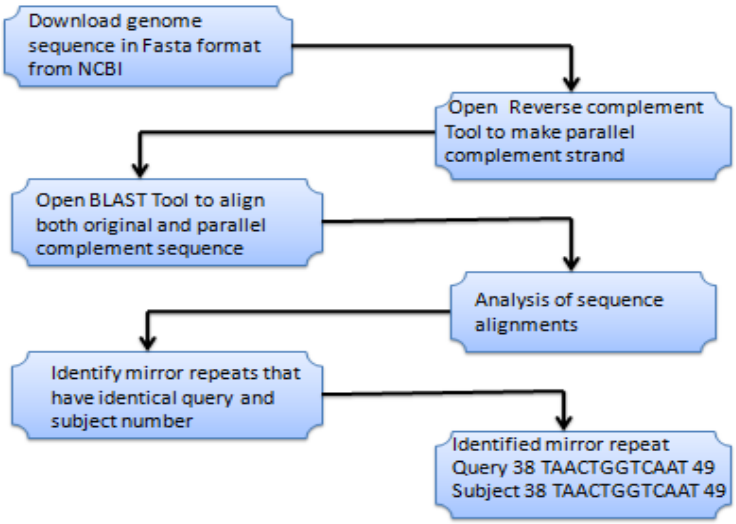

Figure 1. Workflow of the methodology used to identify mirror repeats

For comparative analysis, mirror repeat searches were also conducted in the genomes of other RNA viruses, including Corona virus (SARS-CoV-2), Zika virus (ZIKV), Marburg virus (MARV), and Ebola virus (EBOV), using the MegaBLAST tool. Algorithmic parameters such as word size, expected threshold, and maximum target sequences were adjusted as needed to optimize detection.

RESULTS AND DISCUSSION

In this study, mirror repeats were identified within the Nipah virus (NiV) genome using a simple, manual bioinformatics approach. The proportion of perfect and imperfect mirror repeats is shown in Figure 2. The full genome, consisting of 18,246 nucleotides, was segmented into 500 base pair (bp) fragments to facilitate localized analysis. BLAST alignment between each fragment and its reverse complement enabled the identification of regions containing mirror repeats. A total of 401 mirror repeats were identified across 37 genomic regions of the Nipah virus. The representation of mirror repeats across different genomic regions of the Nipah virus is shown in Figure 3. Among these, 375 repeats were classified as perfect mirror repeats, exhibiting exact sequence symmetry, whereas 26 were imperfect mirror repeats due to the presence of nucleotide mismatches or minor structural deviations. The identified mirror repeats varied in length and genomic distribution, suggesting that they may contribute differently to viral genome organization and potentially influence processes such as replication, transcription, or regulatory element formation.

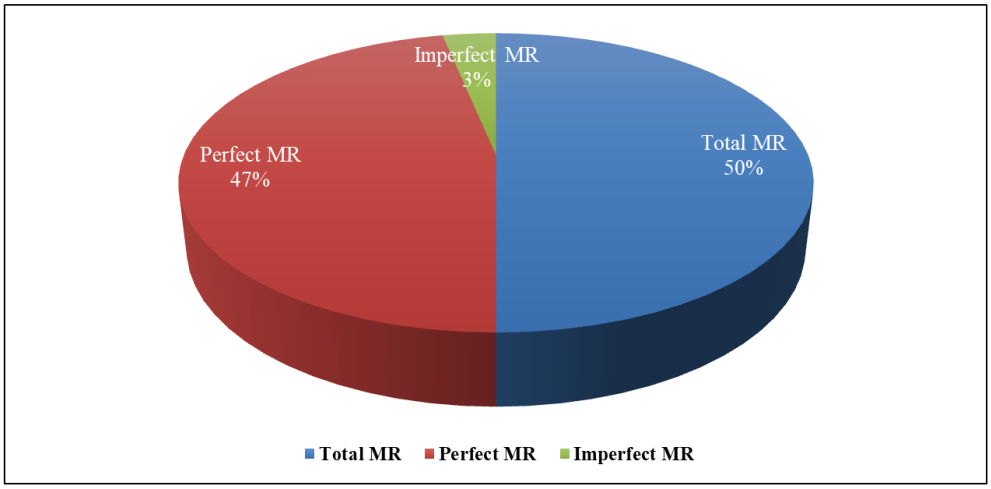

Figure 2: Proportion of perfect and imperfect mirror repeats

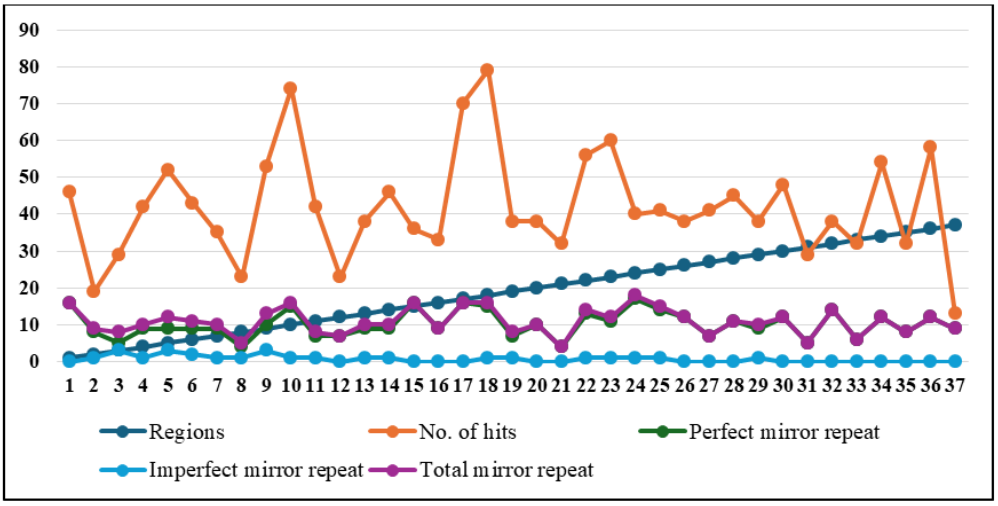

Figure 3: Represent distribution of mirror repeats across different genomic regions of the Nipah virus

To refine our analysis, only perfect mirror repeats longer than 10 base pairs were considered. This filtering step yielded 29 perfect mirror repeats distributed across the 37 regions (Table 1). Among these 29 longer mirror repeats, 24 contained a spacer element, while 5 were contiguous (without spacers).

Table 1: Distribution of MR in each region of Nipah virus genome

Regions	No. of hits	Perfect mirror repeat	Imperfect mirror repeat	Total mirror repeat
1	46	16	0	16
2	19	8	1	9
3	29	5	3	8
4	42	9	1	10
5	52	9	3	12
6	43	9	2	11
7	35	9	1	10
8	23	4	1	5
9	53	10	3	13
10	74	15	1	16
11	42	7	1	8
12	23	7	0	7
13	38	9	1	10
14	46	9	1	10
15	36	16	0	16
16	33	9	0	9
17	70	16	0	16
18	79	15	1	16
19	38	7	1	8
20	38	10	0	10
21	32	4	0	4
22	56	13	1	14
23	60	11	1	12
24	40	17	1	18
25	41	14	1	15
26	38	12	0	12
27	41	7	0	7
28	45	11	0	11
29	38	9	1	10
30	48	12	0	12
31	29	5	0	5
32	38	14	0	14
33	32	6	0	6
34	54	12	0	12
35	32	8	0	8
36	58	12	0	12
37	13	9	0	9

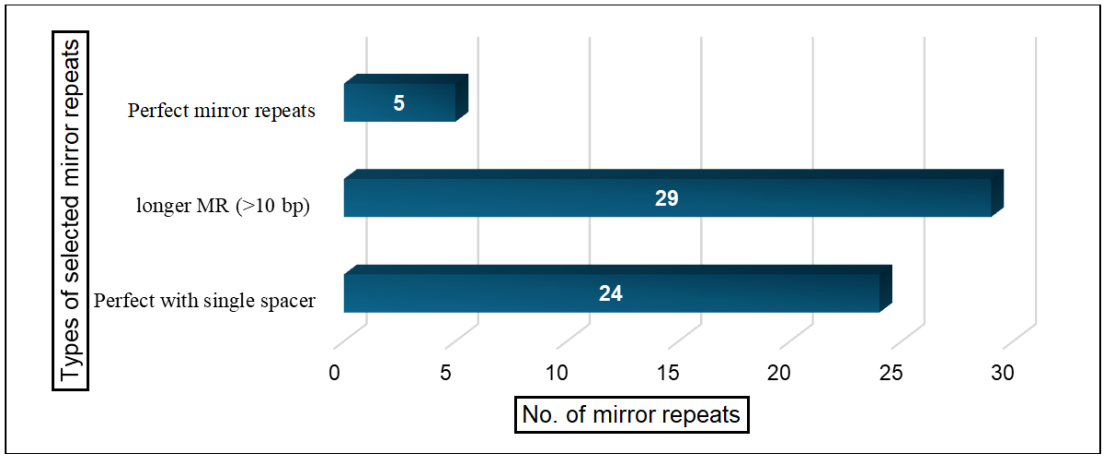

Figure 4: Graph showing the selected mirror repeats

Figure 4 shows the selected mirror repeats. Number of hits show repeat abundance, mirror repeats show structural symmetry To explore the evolutionary or functional conservation of these motifs, the 29 mirror repeats (>10 bp) identified in the Nipah virus (NiV) were compared to the genomes of Zika (ZIKV), Marburg (MARV), Ebola (EBOV), and Coronavirus (SARS-CoV-2) using the MegaBLAST tool. Due to the limitation of MegaBLAST in detecting small sequences (minimum word size of 16), shorter mirror repeats could not be fully analyzed (Table 2). Therefore, only selected mirror repeats exceeding 10 bp were evaluated for presence or absence in the other viral genomes.

Table 2: Classification of selected MR in Nipah virus

S. No.	CDS	Sequence	Length	Location	Type of mirror repeat
1	1-500	AGAGGAGAGGAGA	13	456-468	Perfect with single spacer
2	1-500	GGAGAGGAGAGG	12	454-465	Perfect mirror
3	1500-2000	TAATTATATTAAT	13	312-324	Perfect with single spacer
4	2000-2500	TTTCCTATCCTTT	13	49-61	Perfect with single spacer
5	2000-2500	CTCTAGGATCTC	12	318-329	Perfect mirror
6	4000-4500	GGAAAGGGAAAGG	13	144-156	Perfect with single spacer
7	4000-4500	AGGGAAAGGGA	11	142-152	Perfect with single spacer
8	4500-5000	TCTAACAACAATCT	14	65-78	Perfect mirror
9	4500-5000	ATAACACAATA	11	167-177	Perfect with single spacer
10	4500-5000	TTCTCACTCTT	11	235-245	Perfect with single spacer
11	5000-5500	TAACTGGTCAAT	12	38-49	Perfect mirror
12	6000-6500	GTTTCTCTTTG	11	3 -13	Perfect with single spacer
13	6500-7000	AAAATAAGAATAAAA	15	53-67	Perfect with single spacer
14	6500-7000	ATGTTATTGTA	11	177-187	Perfect with single spacer
15	7000-7500	CATTGGGTTAC	11	387-397	Perfect with single spacer
16	8000-8500	AGAGGATAGGAGA	13	237-249	Perfect with single spacer
17	8000-8500	GAGAAAAAGAG	11	208-218	Perfect with single spacer
18	8500-9000	ATTTCCCTTTA	11	180-190	Perfect with single spacer
19	1050011000	TAATAAATAAT	11	431-441	Perfect with single spacer
20	11000-11500	AATTCCCTTAA	11	378-388	Perfect with single spacer
21	11500-12000	GATTACATTAG	11	122-133	Perfect with single spacer
22	11500-12000	GGGATTTAGGG	11	375-385	Perfect with single spacer
23	12000-12500	TCCCTTGTTCCCT	13	230-242	Perfect with single spacer
24	12000-12500	TTTCTTTCTTT	11	481-491	Perfect with single spacer
25	12500-13000	AAGTTCTTGAA	11	65-75	Perfect with single spacer

26	12500-13000	AACCGAGCCAA	11	408-418	Perfect with single spacer
27	13500-14000	CTAGAAAGATC	11	186-196	Perfect with single spacer
28	14500-15000	AAACAGAAGACAAA	14	445-458	Perfect mirror
29	17500-18000	ATGAACAAGTA	11	165-175	Perfect with single spacer

The comparison revealed variable distribution of these mirror repeats across the selected viruses. The presence or absence of each repeat was recorded using a "+" or "–" sign, respectively. Some repeats were conserved across multiple genomes, while others were unique to Nipah or found only in one or two of the other viruses (Table 3). This comparative analysis highlights both conserved and virus-specific mirror repeat patterns, suggesting possible functional or structural roles of these motifs in viral genome organization or replication.

Table 3: Comparative analysis of MR within Nipah virus and other viruses (+ sign shows the presence of mirror repeats, while – sign shows the absence of mirror repeats)

S. No.	Mirror repeats	Nipah virus (NiV)	Zika virus (ZIKV)	Marburg virus (MARV)	Ebola virus (EBOV)	SARS-CoV-2
1	AGAGGAGAGGAGA	+	+	-	-	-
2	GGAGAGGAGAGG	+	+	-	-	-
3	TAATTATATTAAT	+	+	-	-	-
4	TTTCCTATCCTTT	+	+	-	-	-
5	CTCTAGGATCTC	+	+	-	-	-
6	GGAAAGGGAAAGG	+	+	-	-	-
7	AGGGAAAGGGA	+	+	-	-	-
8	TCTAACAACAATCT	+	+	-	-	-
9	ATAACACAATA	+	+	-	-	-
10	TTCTCACTCTT	+	+	-	-	-
11	TAACTGGTCAAT	+	+	-	-	-
12	GTTTCTCTTTG	+	+	-	-	-
13	AAAATAAGAATAAAA	+	+	-	-	-
14	ATGTTATTGTA	+	+	-	-	-
15	CATTGGGTTAC	+	+	-	-	-
16	AGAGGATAGGAGA	+	+	-	-	-
17	GAGAAAAAGAG	+	+	-	-	-
18	ATTTCCCTTTA	+	+	-	-	-
19	TAATAAATAAT	+	+	-	-	-
20	AATTCCCTTAA	+	+	-	-	-
21	GATTACATTAG	+	+	-	-	-
22	GGGATTTAGGG	+	+	-	-	-
23	TCCCTTGTTCCCT	+	+	-	-	-
24	TTTCTTTCTTT	+	+	-	-	-
25	AAGTTCTTGAA	+	+	-	-	-
26	AACCGAGCCAA	+	+	-	-	-
27	CTAGAAAGATC	+	+	-	-	-
28	AAACAGAAGACAAA	+	+	-	-	-
29	ATGAACAAGTA	+	+	-	-	-

CONCLUSION

This study identified 401 mirror repeats across the 37 genomic regions in the Nipah virus (NiV) genome of 18,246 base pairs, which includes 375 perfect and 26 imperfect repeats of different lengths. Repeated selected mirror sequences longer than 10 base pairs are present in the genomes of Nipah virus and Zika virus, but absent in Marburg virus, SARS-CoV-2, and Ebola virus, indicating a shared interspecific genomic feature between Nipah and Zikaviruses. Their widespread presence suggests potential roles in genomic organization and regulation. While their exact function remains to be determined, mirror repeats may influence transcription, translation, and other molecular processes. Further research is needed to explore their evolutionary significance and potential as therapeutic targets.

Acknowledgements:
The authors express their gratitude to the Department of Zoology, Baba Mastnath University, Asthal Bohar, Rohtak, for providing the required facilities and resources for this research.

Funding: The authors declare that no external funding was received for this research.

Competing interests: The authors do not have any competing interests.

Author contribution: All authors participated in the revision of the study design and analysis and approved the final manuscript.

REFERENCES

Banerjee, S., Gupta, N., Kodan, P., Mittal, A., Ray, Y., Nischal, N.,...& Wig, N. (2019). Nipah virus disease: A rare and intractable disease. *Intractable & rare diseases research*, *8*(1), 1-8.

Bowden, T. A., Crispin, M., Harvey, D. J., Aricescu, A. R., Grimes, J. M., Jones, E. Y., & Stuart, D. I. (2008). Crystal structure and carbohydrate analysis of Nipah virus attachment glycoprotein: a template for antiviral and vaccine design. *Journal of virology*, *82*(23), 11628-11636.

Epstein, J. H., Field, H. E., Luby, S., Pulliam, J. R., &Daszak, P. (2006). Nipah virus: impact, origins, and causes of emergence. *Current infectious disease reports*, *8*(1), 59-65.

Garbuglia, A. R., Lapa, D., Pauciullo, S., Raoul, H., &Pannetier, D. (2023). Nipah Virus: An Overview of the Current Status of Diagnostics and Their Role in Preparedness in Endemic Countries. *Viruses*, *15*(10), 2062

Gurley, E. S., Spiropoulou, C. F., & De Wit, E. (2020). Twenty years of Nipah virus research: where do we go from here?. *The Journal of infectious diseases*, *221*(Supplement 4), S359-S362.

Hafeez, M. H., Ajmal, H., Nadeem, A., Tabassum, S., &Akilimali, A. (2025). Navigating Nipah virus: Insights, challenges, and recommendations. *New Microbes and New Infections*, 101575.

Joshi, J., Shah, Y., Pandey, K., Ojha, R. P., Joshi, C. R., Bhatt, L. R., ... &Pandey, B. D. (2023). Possible high risk of transmission of the Nipah virus in South and South East Asia: a review. *Tropical medicine and health*, *51*(1), 44.

Lang, D. M. (2007). Imperfect DNA mirror repeats in the gag gene of HIV-1 (HXB2) identify key functional domains and coincide with protein structural elements in each of the mature proteins. *Virology Journal*, *4*(1), 113.

Lou, Z., Xu, Y., Xiang, K., Su, N., Qin, L., Li, X.,...&Rao, Z. (2006). Crystal structures of Nipah and Hendra virus fusion core proteins. *The FEBS journal*, *273*(19), 4538-4547.

Lupski, J. R., & Weinstock, G. M. (1992). Short, interspersed repetitive DNA sequences in prokaryotic genomes.

Journal of bacteriology, *174*(14), 4525-4529.

Mire, C. E., Satterfield, B. A., Geisbert, J. B., Agans, K. N., Borisevich, V., Yan, L.,...&Geisbert, T. W. (2016). Pathogenic differences between Nipah virus Bangladesh and Malaysia strains in primates: implications for antibody therapy. *Scientific reports*,*6*(1), 30916.

Nicholson, B. L., & White, K. A. (2015). Exploring the architecture of viral RNA genomes. *Current opinion in virology*, *12*, 66-74.

Rahman, M. M., Miah, M., Hossain, M. E., Rahim, S., Sultana, S., Satter, S. M., ...&Jahid, I. K. (2025). Development of a culture-independent whole-genome sequencing of Nipah virus using the MinION Oxford Nanopore platform. *Microbiology spectrum*, *13*(6), e02492-24.

Rahman, M. Z., Islam, M. M., Hossain, M. E., Rahman, M. M., Islam, A., Siddika, A., ...& Gurley, E. S. (2021). Genetic diversity of Nipah virus in Bangladesh. *International Journal of Infectious Diseases*, *102*, 144-151.

Rao, S. R., Trivedi, S., Emmanuel, D., Merita, K., &Hynniewta, M. (2010). DNA repetitive sequences-types, distribution and function: A Review. *J Cell MolBiol*, *7*(2), 1-11.

Sehrawat, B., Yadav, P., Sarjeet, M., Yadav, S., Yadav, V., Goyal, N.,...&Yadav, S. (2024). In-silico evaluation of Mirror repeats in some selected genes of Candida albicans. *bioRxiv*, 2024-01.

Shapiro, J. A., & von Sternberg, R. (2005). Why repetitive DNA is essential to genome function. *Biological Reviews*, *80*(2), 227-250.

Sharma, V., Kaushik, S., Kumar, R., Yadav, J. P., &Kaushik, S. (2019). Emerging trends of Nipah virus: A review. *Reviews in medical virology*, *29*(1), e2010.

Singh, R. K., Dhama, K., Chakraborty, S., Tiwari, R., Natesan, S., Khandia, R.,...&Mourya, D. T. (2019). Nipah virus: epidemiology, pathology, immunobiology 8il4and advances in diagnosis, vaccine designing and control strategies–a comprehensive review. *Veterinary Quarterly*, *39*(1), 26-55.

Skowron, K., Bauza-Kaszewska, J., Grudlewska-Buda, K., Wiktorczyk-Kapischke, N., Zacharski, M., Bernaciak, Z., &Gospodarek-Komkowska, E. (2022). Nipah virus–Another threat from the world of zoonotic viruses. *Frontiers in Microbiology*, *12*, 811157.

Soman, P.V., Krishna, G., &Valiya, V.M. (2020). Nipah virus: past outbreaks and future containment. *Viruses*, *12*(4), 465.

Sun, B., Jia, L., Liang, B., Chen, Q., & Liu, D. (2018). Phylogeography, transmission, and viral proteins of Nipah virus. *VirologicaSinica*, *33*, 385-393.

Yadav, P., Kumari, J., Yadav, P., Yadav, R., Yadav, S., Sharma, D., ...&Yadav, S. (2023). Identification of Mirror Repeats in viral genomes using FPCB analysis. *bioRxiv*, 2023-04.

Yadav, S., Yadav, U., & Sharma, D. C. (2022). In-Silico Evaluation of 'Mirror Repeats' in HIV Genome.(2021). *Int. J. Life Sci*, *11*(5), 81-87.

Yoneda, M., Guillaume, V., Ikeda, F., Sakuma, Y., Sato, H., Wild, T. F., & Kai, C. (2006). Establishment of a Nipah virus rescue system. *Proceedings of the National Academy of Sciences*, *103*(44), 16508-16513.