Identification and phylogenetic analysis of herpes simplex virus-1 from clinical isolates in India

Dimpal A. Nyayanit, Rima R. Sahay, Gajanan N. Sakpal, Anita M. Shete, Gouri C. Chaubal, Prasad Sarkale, Rashi Srivastava, Sreelekshmy Mohandas and Pragya D. Yadav*

Abstract

Human herpes simplex virus (HSV)-1 infection is acquired in childhood and persists throughout a person's lifetime. Here, we present two cases from India; one showing symptoms of postpartum haemorrhage with disseminated intravascular coagulation, and the second one showing signs of acute encephalitis syndrome. The aetiological agent in both cases was identified as HSV-1 using the PCR method. The next-generation sequencing method retrieved ~97% of the viral genome from the isolates of the clinical samples. The phylogenetic analysis of the retrieved genomes revealed that they belong to clade II of HSV-1. This study identifies a few sequence variations in the glycoprotein region of HSV-1 during two different clinical manifestations. There are a couple of papers that analyse variations in the glycoprotein region of clinical samples. Further, this study also highlights the importance of considering HSV-1 during differential diagnosis when analysing the nosocomial infection.

INTRODUCTION

Human herpes simplex virus (HSV)-1 is an enveloped DNA virus belonging to family *Herpesviridae* [1]. HSV-1 is predominantly associated with herpetic lesions in oral areas, oral cancers, sores, and is transmitted through oral contact. The HSV-1 infection can be asymptomatic or may present with recurrent lesions. In rare cases, HSV-1 infection can lead to acute encephalitis. Women are more prone to acquiring genital herpes infections, and in approximately 5% of child-infection cases transmission of HSV infection from mother-to-child is demonstrated. The transfer mainly occurs during the peripartum period (85%) or postnatally (10% cases) [2]. Neonatal herpes infection is rare but can cause high morbidity and mortality [3, 4]. Disseminated neonatal HSV infections rarely present with septic appearance, cardio-respiratory failure, hepatic failure and disseminated intravascular coagulation (DIC) [5].

HSV-1 has a 152 kb double-stranded DNA genome, which is known to encode a minimum of 84 polypeptides along with long non-coding RNAs and microRNAs [6]. HSV-1 has a GC-rich (68–70%) genome, which contains unique long and short sequences; each sequence is flanked by terminal and internal repeats, thus making it a complex genome [7]. The nucleocapsid and envelope of the virion are separated by a proteinaceous tegument [8]. The surface glycoproteins of HSV-1 are required in multiple steps; starting from the initial attachment to the fusion to the release of viral DNA into the host cell. HSV-1 encodes 12 different glycoproteins that cater for the virus entering the host cell. Glycoproteins play an important role in the virus entering the host cell and infection of the host cells [9]. They are also linked as the inducer and target for the immune system, having a role in both the adaptive and humoral immune response [10]. The effect of the single amino acid mutation in glycoprotein B has been reported to cause non-invasiveness of the virus in previous studies [11]. Studies to understand the sequence variations in the viral genes and the different glycoproteins are limited in clinical samples [11–13]. This makes it important to understand the sequence variation present on the glycoprotein sequences of the HSV-1 that could be observed in the various clinical conditions.

In this study, clinical samples of the cases demonstrating symptoms of postpartum haemorrhage (PPH) with DIC (case-1) and acute encephalitis syndrome (AES) (case-2) are

Fig. 1. Comparative HSV-1 genome structure obtained by aligning the retrieved HSV-1 genome from India (MG646679) and reference sequence (NC_0.001806.1). Sequences were aligned using the alignment program embedded in the CLC genomic workbench version 10.1. The protein coding region is depicted using the light green arrows along with the conservation in the bottom row. Red, blue and black indicate higher, lower and medium conservation amongst the two sequences.

analysed. DIC leads to the systemic activation of the blood coagulation machinery that ends up in multiple organ failure. Acute encephalitis is caused by the invasion of HSV-1 in the neural system. These are two different disease outcomes caused by HSV-1 isolated from these two clinical samples. The study aimed to identify the clade of the circulating virus along with detection of any glycoprotein-based sequence variations between the two clinical isolates. The study also explores the presence of recombination events in the isolates as well as the glycoprotein regions.

CLINICAL PROFILE
Case-1 (PPH with DIC)

A young age antenatal care (ANC) mother, second gravid (G), first para (P), one living child (L) and no history of abortion (A) (G2P1L1A0) was admitted into labour at 37 weeks of gestation for delivery in a government hospital in Telangana state, India, during February 2017. She did not have any significant medical complaints at the time of admission. She delivered a child by normal vaginal delivery. There was no prolonged or obstructed labour and episiotomy was given. The new-born child died of respiratory distress on the same day within 2 h of birth. Within 15 min of normal vaginal delivery, there was primary onset postpartum uterine bleeding for which immediate exploratory measures were taken to check for any cervical tear or retained placenta. No significant cause of bleeding was ruled out, and the patient ended up in coagulation failure, acute renal failure and circulatory failure with shock, and hence was shifted to the intensive care unit immediately. Due to circulatory failure, she was put on a mechanical ventilator. She was mildly febrile throughout treatment with a temperature of 37.2 °C. Broadspectrum antibiotics to cover sepsis, and all vasopressin agents to prevent circulatory failure were given to her. Due to profuse bleeding, her haemoglobin dropped down to 5 gm% for which she received multiple

transfusions of packed-cell volume as well as fresh-frozen plasma. Peritoneal dialysis was continuously performed to combat acute renal failure. She recovered completely within 5 days of treatment.

Case-2 (AES)

A child in the age group of 5 to 10 years was admitted to the tertiary care hospital in Maharashtra state during August 2016, with a history of two episodes of convulsions followed by altered sensorium and abnormal behaviour on the day of admission. Convulsions were not associated with projectile vomiting. The child had a history of mild to moderate fever without rash for the past 6 to 7 days for which treatment was taken from a local physician and on admission had neck rigidity. Details regarding treatment history, biochemical or haematological parameters, and brain images were not provided by the hospital.

Human ethical approval

This study describes two human clinical cases referred for diagnosis to Indian Council of Medical Research (ICMR) – National Institute of Virology (NIV) – Pune from two different states (Maharashtra and Telangana states) of India. The sample anonymity was maintained as per International Ethical Guidelines Studies, and the study was approved by the Institutional Human Ethics Committee (IHEC) at the NIV – Pune (IHEC Number-NIV/IEC/2017/D-86). The ethical committee deemed that consent was verified.

RESULTS
Testing for various etiological agents present in clinical samples by molecular methods

Samples were tested for RNA viruses using a different assay to exclude their role in the infection. Nasal/throat swab
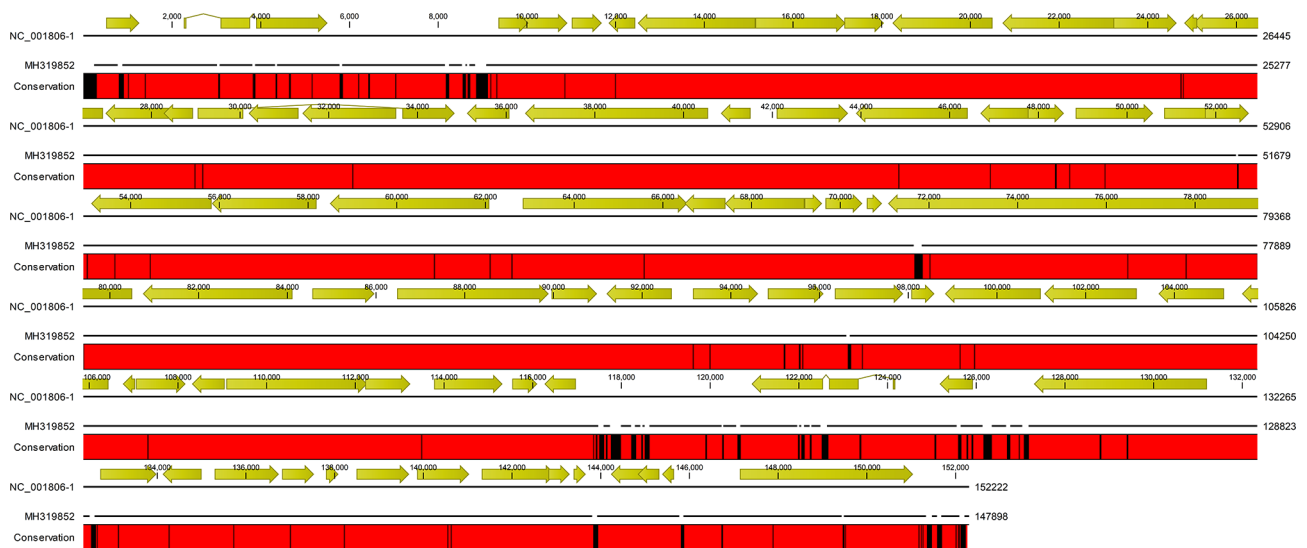
Fig. 2. Comparative HSV-1 genome structure obtained by aligning the retrieved HSV-1 genome from India (MH319852) and reference sequence (NC_0.001806.1). Sequences were aligned using the alignment program embedded in the CLC genomic workbench version 10.1. The protein coding region is depicted using the light green arrows along with the conservation in the bottom row. Red, blue and black indicate higher, lower and medium conservation amongst the two sequences.

specimen/serum samples of case-1 tested using real-time RT-PCR were found to be negative for RNA viruses (Crimean-Congo hemorrhagic fever virus [CCHFV], Kyasanur Forest disease virus [KFDV], Zika virus [ZIKV], Dengue virus [DENV], Chikungunya virus [CHIKV] and Hepatitis E [HEV]). Similarly, the CSF sample of case-2 tested negative for RNA viruses causing encephalitis (Japanese encephalitis virus [JEV], West Nile Virus [WNV] and Chandipura). Further, both the clinical samples tested positive for HSV-1 (throat swab in case-1 and CSF in case-2) and negative for Varicella Zoster Virus (VZV) using the conventional PCR method. No other DNA virus or bacteria was screened using the conventional PCR method.

Virus isolation

VeroCCL81 cells inoculated with a throat swab of case-1 showed cytopathic effect (CPE) from the second passage at the seventh day post infection (day p.i.). Serum and nasal swab samples inoculated in *VeroCCL81* cells did not show any CPE until the fourth passage. *BHK-21* cells inoculated with CSF (case-2) showed CPE at the third passage from the second day p.i.

## Next-generation sequencing of the virus isolate

Reference-based mapping retrieved a genome of 149779 and 147898 length. Figs 1 and 2 depict the bp obtained for HSV-1 in case-1 and case-2, respectively obtained using CLC genomics workbench version 10.1. The complete sequence could not be obtained and the sequences have gaps inbetween. These sequences were submitted to GenBank and have the following accession numbers: case-1 (MG646679) and case-2 (MH319852).

Phylogenetic analysis and glycoprotein characterization of the retrieved genome

Complete genome analysis of the sequences demonstrated that the two different isolates from India grouped with the East Asian and belongs to the clade II group (Fig. 3). HSV-1 from the case-1 (MG646679) HSV-1 isolate had a higher similarity (99.7 %) towards the USA strain from Texas (strain CR38) whereas the case-2 (MH319852) sample was highly similar (99.6 %) to the strain isolated from China (strain KOS).

The phylogenetic analysis of different glycoprotein's demonstrated varied clustering patterns of the clinical samples revealing some differences in the glycoprotein region (Fig. 4). It was observed that the average evolutionary nucleotide divergence for clade V glycoproteins was higher [0.39–1.89%] than the other clades (Table 1). It was observed that the gG gene has a higher divergence with respect to the other glycoprotein gene in comparison. Table 2 displays the percent nucleotide and amino acid divergence seen between the different glycoprotein compared within the study.

We were able to identify a few variations in gE, gG, gH, gJ and gM genes for the clinical samples analysed. The clinical sample demonstrating PPH-DIC had a change in the E gene (position: V166G) and J gene (position: P53V). The sample with AES manifestations has a variation in the M gene (position: Y414D). Deletion of single E residue was observed as compared to the other HSV-1 gG amino acid, which had 8/9 E's in stretch (position: 79–85). Further, four amino acids were deleted from the gG gene at the region ranging 131–134 of the other HSV-1 gG region. AES and PPH-DIC cases have T and I amino acid in the H gene (position: 127) as opposed
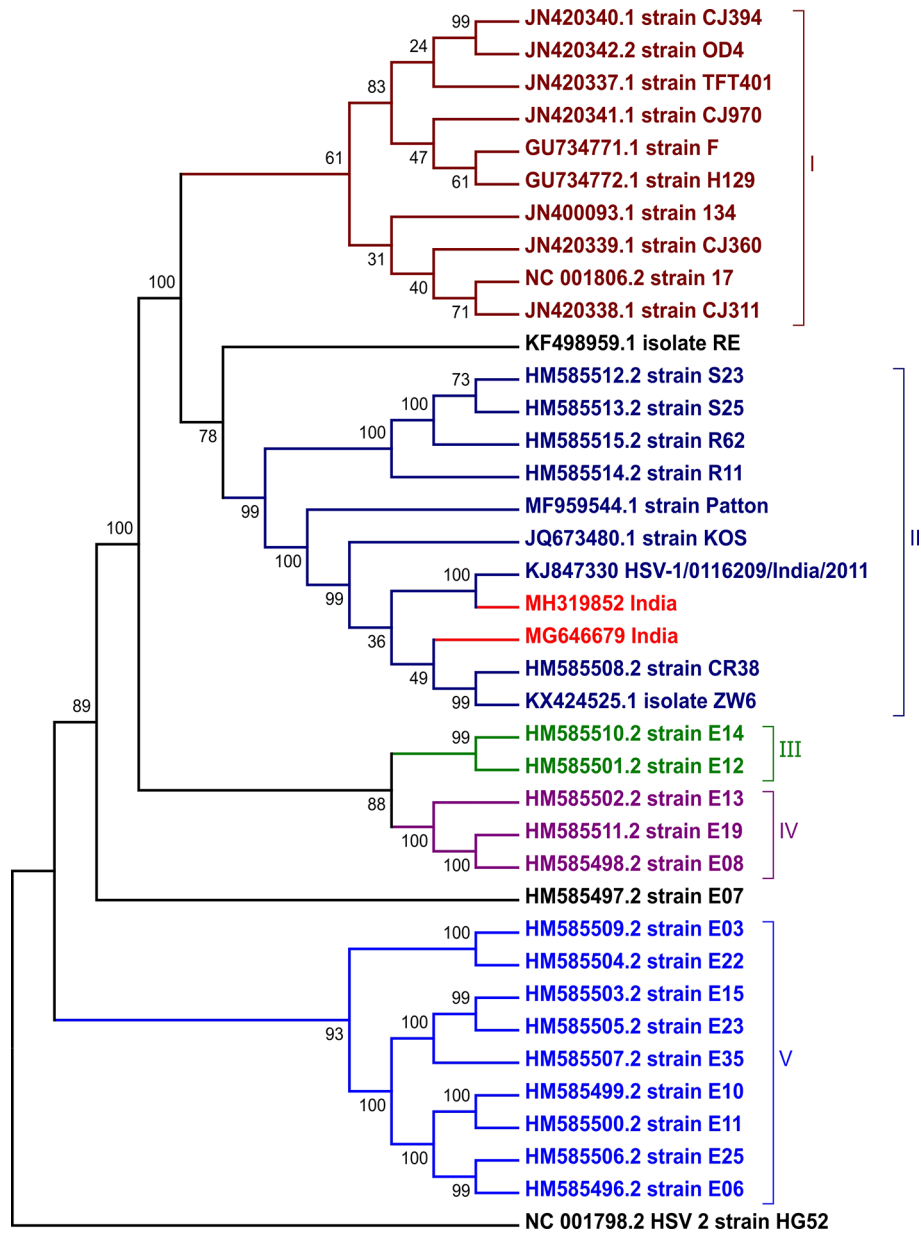
Fig. 3. Phylogenetic tree of two isolates of the HSV-1 complete genome isolated from India: a phylogenetic tree was created for 37 complete genome sequences using a general time reversible model. Gamma distribution (0.1045) was used to model the evolutionary rate along with the variation (*I*=61.33%) for some sites.

to the presence of T or I in the other reference sequences, respectively. There was no sequence variation observed in other glycoproteins that were not common in the analysed sequences.

Recombination analysis of HSV-1 sequences

Recombination analysis performed on both the HSV-1 sequences demonstrated the evidence of recombination. Fig. 5 determines the similarity plot and the bootscan plot for the isolate. The red and blue lines are for strain Kos and an E08 strain having accession number JQ673480.1 and

HM585498.2 belonging to clade II and IV, respectively. In similarity plot analysis, it was observed that the strains in comparison have approximately 97% similarity with each other. It was observed from the bootscan plot of the MG646679 isolate that the UL36 region and 3/4th of the UL37 region (pos: 74900–82900) has a higher similarity towards the HM585498.2 isolate belonging to Kenya, whereas most of the other part is similar to JQ673480.1. This leads us to conclude that recombination points exist in Indian HSV-1 (MG646679) that have higher similarity to the two strains JQ673480.1 and HM585498.2
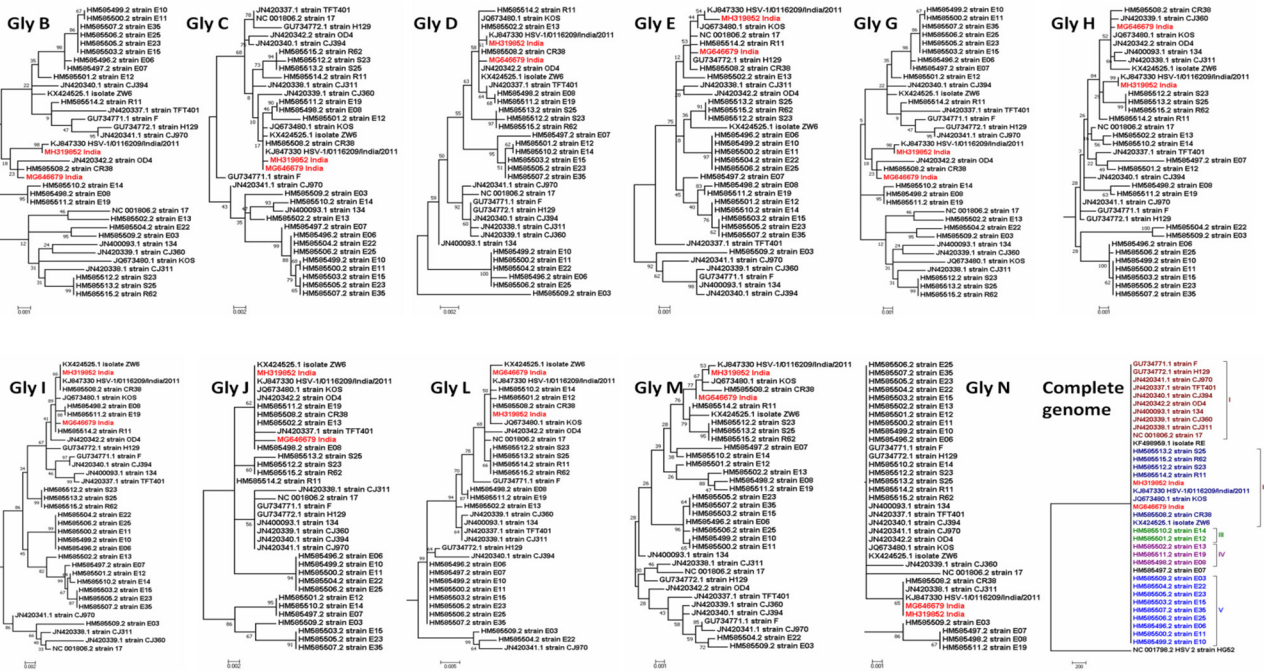
Fig. 4. Phylogenetic tree of the HSV-1 glycoprotein region. A phylogenetic tree was created for 37 complete genome sequences using a general time reversible model.

Similarity plot and the bootscan analysis of HSV-1 from case-2 (MH319852) is depicted in Fig. 6. Redline, blue line, and green are for strain KOS, CR38 and strain 134. In the similarity plot analysis, it was observed that the strains have approximately 97% similarity with each other. Further, it was noted from the boot scan plot of the MH319852 isolate that there are multiple regions of recombination observed between strains KOS, CR38and strain 134.

Bootscan analysis embedded in the RDP4 software predicted recombination in gB, gG,gH and gI but these were neglected as per criteria of recombination prediction. Hence no evidence of recombination was predicted in the glycoprotein's region.

DISCUSSION

HSV-1 is differentiated into five different groups depending on its geographical location and genomic variation. Clade I circulates in America whereas clade II viruses are found in East Asian countries. Remaining clades (III–V) find their origins in East African countries. This study isolates and also characterizes two HSV-1 genomes that have a different clinical impact on humans. A phylogenetic analysis approach identified HSV-1 to be descended from clade II groups found in East Asian countries. Glycoprotein B, gD along with gH and gL present on HSV-1 are essential for the host-cell attachment, fusion, entry, and infection of the host cell. We observed a few mutations in the amino acids of the glycoprotein regions of the HSV-1 isolates from the clinical samples. There are previous reports that demonstrate the effect of amino acid mutations thereby hampering their entry and reduced pathogenesis and alter the pathogenesis of the virus [14, 15]. Also, the role of the reported variation needs to be further understood by performing future experiments.

HSV-1 is known to cause a vast spectrum of symptoms ranging from asymptomatic cases to AES manifestations,

Table 1. Average evolutionary divergence for different glycoproteins was performed over sequence pairs between clades of HSV-1 analysed in the study

Clade	gB gene	gC gene	gD gene	gE gene	gG gene	gH gene	gI gene	gJ gene	gK gene	gL gene	gM gene	gN gene
Clade I	0.60	0.27	0.31	0.67	1.44	0.37	1.36	0.54	0.11	1.52	0.55	0.32
Clade II	0.47	0.74	0.26	0.31	0.77	0.32	1.06	0.50	0.37	0.28	0.38	0.18
Clade III	0.44	0.34	0.00	0.00	0.14	0.44	0.09	0.00	0.33	0.30	0.42	0.00
Clade IV	0.25	1.30	0.11	0.28	1.48	0.48	0.74	0.00	0.20	0.30	0.33	0.72
Clade V	0.40	0.69	0.96	0.50	1.89	0.49	1.04	1.68	0.33	0.80	0.39	0.16

Table 2. Percent nucleotide and amino acid divergence between the glycoprotein with respect to the MH319852 sequence belonging to clade II from India

	Clade	gB		gC		gD		gE		gG		gH		gI		gJ		gK		gL		gM		gN	
		N*	A†	N	A	N	A	N	A	N	A	N	A	N	A	N	A	N	A	N	A	N	A	N	A
GU734771.1 strain F	I	0.41	0.55	0.90	1.46	0.51	0.51	0.12	0.18	3.12	5.13	0.44	0.48	1.59	2.67	0.72	1.09	0.69	0.00	1.93	3.13	0.91	0.85	0.36	0.00
GU734772.1 strain H129	I	0.48	0.55	0.97	1.67	1.43	0.76	0.43	0.37	3.13	4.29	0.64	0.95	1.38	0.56	1.43	2.17	0.59	0.00	1.63	3.13	0.84	0.85	0.36	0.00
JN400093.1 strain 134	I	0.44	0.55	0.83	1.46	0.93	0.51	0.43	0.74	3.13	4.29	1.03	1.19	1.47	0.83	2.15	3.26	0.59	0.59	1.63	3.13	0.98	0.85	1.45	1.10
JN420337.1 strain TFT401	I	0.44	0.55	0.28	0.42	0.25	0.00	0.49	0.74	1.41	3.40	0.75	0.95	0.28	0.00	0.00	0.00	0.49	0.30	0.59	1.34	1.05	0.85	1.45	1.10
JN420338.1 strain CJ311	I	0.55	0.55	1.11	1.88	1.27	0.76	0.43	0.37	3.13	4.29	0.68	0.95	1.38	0.56	1.43	2.17	0.59	0.00	1.63	3.13	0.91	0.85	0.36	0.00
JN420339.1 strain CJ360	I	0.55	0.55	1.11	1.88	1.27	0.76	0.43	0.37	3.13	4.29	0.68	0.95	1.38	0.56	1.43	2.17	0.59	0.00	1.63	3.13	0.91	0.85	0.36	0.00
JN420340.1 strain CJ394	I	0.37	0.33	0.62	1.26	0.76	0.51	0.49	0.55	2.83	3.85	0.52	0.72	1.47	0.83	2.15	3.26	0.49	0.00	0.15	0.00	0.70	0.21	0.36	0.00
JN420341.1 strain CJ970	I	0.81	0.66	1.11	2.09	0.08	0.00	0.37	0.55	2.54	3.83	0.56	0.48	1.01	0.56	0.00	0.00	0.20	0.30	1.04	2.23	1.13	0.85	0.36	0.00
JN420342.2 strain OD4	I	0.52	0.55	1.11	1.88	0.68	0.25	0.55	0.74	3.10	4.68	0.68	0.95	1.57	1.39	2.51	3.26	0.59	0.00	1.63	3.13	0.91	0.85	0.36	0.00
KJ847330 HSV-1/0116209/India/2011	I	0.85	0.22	0.97	1.67	1.27	0.76	0.43	0.37	3.13	4.29	1.27	1.31	1.38	0.56	1.43	2.17	0.59	0.00	3.26	4.46	1.20	0.85	0.36	0.00
NC 001806.2 strain 17	I	0.52	0.55	1.11	1.88	0.68	0.25	0.55	0.74	3.10	4.68	0.68	0.95	2.66	3.20	2.51	3.26	0.59	0.00	1.63	3.13	0.91	0.85	0.36	0.00
HM585508.2 strain CR38	II	0.52	0.55	0.97	1.67	1.27	0.76	0.43	0.37	3.13	4.29	0.68	1.07	1.38	0.56	1.43	2.17	0.59	0.00	1.63	3.13	0.91	0.85	0.36	0.00
HM585512.2 strain S23	II	0.52	0.55	1.11	1.88	0.68	0.25	0.55	0.74	3.10	4.68	0.68	0.95	1.57	1.39	2.51	3.26	0.59	0.00	1.63	3.13	0.91	0.85	0.36	0.00
HM585513.2 strain S25	II	0.18	0.11	0.07	0.21	0.08	0.00	0.18	0.18	1.41	2.98	0.44	0.48	2.17	3.40	0.00	0.00	0.10	0.00	0.00	0.00	0.28	0.42	0.00	0.00
HM585514.2 strain R11	II	0.81	0.55	0.83	1.46	1.60	1.02	1.22	0.74	2.83	3.85	1.27	1.31	4.61	6.02	2.15	3.26	0.49	0.30	3.11	4.46	1.20	0.42	1.09	1.10
HM585515.2 strain R62	II	0.44	0.55	1.04	1.67	0.76	0.51	0.49	0.55	2.97	4.27	0.56	0.48	3.74	4.19	2.15	3.26	0.49	0.30	0.15	0.45	0.56	0.21	0.36	0.00
JQ673480.1 strain KOS	II	0.44	0.55	0.28	0.42	0.25	0.00	0.43	0.55	1.41	3.40	0.60	0.60	2.52	3.40	0.00	0.00	0.49	0.30	0.59	1.34	1.20	1.06	1.45	1.10
KX424525.1 isolate ZW6	II	0.66	0.44	0.69	1.46	0.51	0.76	0.49	0.74	1.27	2.13	0.40	0.24	3.22	4.45	0.72	0.00	0.59	0.00	0.30	0.45	0.56	0.42	0.36	0.00
MG646679 India	II	0.66	0.44	0.56	1.26	0.42	0.51	0.49	0.74	1.27	2.13	0.40	0.24	3.22	4.45	1.08	1.09	0.59	0.00	0.30	0.45	0.56	0.42	0.36	0.00
HM585501.2 strain E12	III	0.41	0.33	0.62	1.26	0.25	0.00	0.18	0.37	1.27	2.56	0.28	0.24	0.09	0.00	0.36	0.00	0.10	0.00	0.30	0.45	0.35	0.21	0.36	0.00
HM585510.2 strain E14	III	0.66	0.44	0.56	1.26	0.42	0.51	0.49	0.74	1.27	2.13	0.40	0.24	3.22	4.45	0.72	0.00	0.59	0.00	0.30	0.45	0.56	0.42	0.36	0.00
HM585498.2 strain E08	IV	0.66	0.11	1.11	2.30	0.51	0.51	0.97	0.92	3.26	5.13	0.44	0.36	3.13	4.19	0.72	1.09	0.10	0.00	1.04	2.23	0.56	0.42	0.36	0.00
HM585502.2 strain E13	IV	0.66	0.55	0.49	0.63	0.17	0.00	0.67	0.92	1.13	2.14	0.48	0.48	3.04	3.93	0.36	1.09	0.49	0.00	1.04	2.23	1.13	0.63	0.36	0.00
HM585511.2 strain E19	IV	0.66	0.22	0.42	0.84	0.59	0.51	0.43	0.74	3.12	5.13	0.48	0.36	4.26	4.97	1.43	1.09	0.10	0.00	1.04	2.23	1.05	0.85	0.00	0.00
HM585496.2 strain E06	V	0.77	0.55	0.49	0.84	0.59	0.51	0.97	0.92	3.40	5.98	0.40	0.24	3.83	4.71	0.72	1.09	0.49	0.00	1.19	2.68	1.20	0.85	0.72	0.00
HM585499.2 strain E10	V	0.41	0.44	0.49	1.05	0.51	0.51	1.04	1.10	3.12	5.13	0.48	0.48	2.87	3.93	0.72	1.09	0.29	0.00	2.37	3.57	1.20	0.85	0.36	0.00
HM585500.2 strain E11	V	0.33	0.33	0.62	1.26	0.51	0.51	1.04	0.92	3.26	5.56	0.40	0.36	3.48	4.45	0.72	1.09	0.69	0.00	2.81	3.57	1.34	1.06	0.36	0.00
HM585503.2 strain E15	V	0.52	0.11	0.62	1.26	0.08	0.00	0.43	0.55	1.41	2.55	0.32	0.24	2.61	3.66	0.00	0.00	0.39	0.00	0.59	0.89	0.84	0.63	0.36	0.00

Continued

Table 2. Continued

Clade	gB N*	gB A†	gC N	gC A	gD N	gD A	gE N	gE A	gG N	gG A	gH N	gH A	gI N	gI A	gJ N	gJ A	gK N	gK A	gL N	gL A	gM N	gM A	gN N	gN A
HM585504.2 strain E22	V	0.81	0.66	0.14	0.42	0.08	0.00	0.18	0.18	1.41	2.55	0.32	0.24	1.24	1.87	0.00	0.00	0.69	0.59	0.30	0.45	0.14	0.42	0.36	0.00
HM585505.2 strain E23	V	0.04	0.00	0.07	0.21	0.00	0.00	0.12	0.00	0.56	0.85	0.00	0.00	1.06	1.87	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.21	0.00	0.00
HM585506.2 strain E25	V	0.26	0.33	0.14	0.42	0.08	0.00	0.55	0.92	1.27	2.55	0.44	0.48	2.17	3.40	0.00	0.00	0.10	0.00	0.30	0.89	0.56	0.21	0.36	0.00
HM585507.2 strain E35	V	0.26	0.11	0.07	0.21	0.17	0.00	0.18	0.37	1.42	2.56	0.36	0.24	2.35	3.40	0.36	1.09	0.10	0.00	0.00	0.00	0.14	0.21	0.00	0.00
HM585509.2 strain E03	V	0.81	0.55	0.49	0.63	0.59	1.02	0.18	0.18	3.40	5.56	0.40	0.36	4.17	5.24	1.08	1.09	0.29	0.00	0.30	0.45	0.77	1.06	1.09	2.20

*N-Percent Nucleotide divergence.
†A-Percent Amino acid divergence.

to even death. With the level of global endemicity of this virus, it becomes difficult to understand the actual impact of HSV-1 on public health. Maternal HSV-1 is known to cause life-threatening neonatal infection. A study from New York, USA reveals increasing HSV-related neonatal death rates, indicating that an increasing number of pregnant women have a primary HSV-1 infection and are therefore at risk of acquiring the virus during pregnancy and transmitting to their infants [16]. Association of HSV-1 with sporadic acute encephalitis is widely known in India [6, 17, 18]. Also, it contributes ~10 percent of fatal sporadic AES cases [17]. Pfaff and coworkers did whole-genome sequencing of HSV-1 in 2016; they found that genotyping of HSV-1 suggested three major phylogroups corresponding to the distinct geographic location and provided reflections of human migration history [19]. Analysis of single nucleotide polymorphism (SNP) in the HSV-1 genome can help in understanding individual genetic traits [19].

It is observed that the amino acids encoded by the genes may have variation based on the clades that are analysed. Rozenberg's group reported changes in the glycoprotein B of HSV-1 during the encephalitis CSF study [11]. The sequences considered by this group for the analysis of glycoprotein B comprised of the glycoprotein B region representing different clades. Comparison of the regions between the two different clades can lead to the inappropriate mapping of the mutations and lead to reporting of amino acids that are observed in the particular clade. The present study under the light of the above factor suggests variations that might be of importance in the different clinical manifestations in these two cases. However, the changes are reported based on the single clinical isolates for each condition, and hence further studies on the isolates specific to the clinical conditions need to be done in order to concretize the above results.

Whole genomes for HSV-1 are limited, and there exists only one complete HSV-1 genome from India [6]; this sequence adds up data from India. Also, sequence divergence studies in the glycoprotein regions for the HSV-1 genome from different clinical samples unique to its clade are limited. This study increases the data for the amino acid variation present in the glycoprotein sequences of the clinical isolates. The analyses performed on the clinical samples are *in silico* in nature and require further experimental support to affirm these proposed sites.

HSV-1 disease-control measures include the use of antivirals and topical application of microbicides. HSV-1 vaccine development is a global challenge; still, the oral and intravenous acyclovir is used as a mainstay of treatment [20]. Acyclovir, valacyclovir and famciclovir are considered as the mainstay for HSV-1 therapy [21]. Concerns are now more about acyclovir's resistance of HSV-1, though rarely reported in <1% of the population. In neonates, HSV-1 infection may present and mimic as bacterial sepsis. It is essential for in-depth maternal history as well as early initiation of treatment with acyclovir [22]. Presence of nosocomial HSV infection is well known, and clinicians
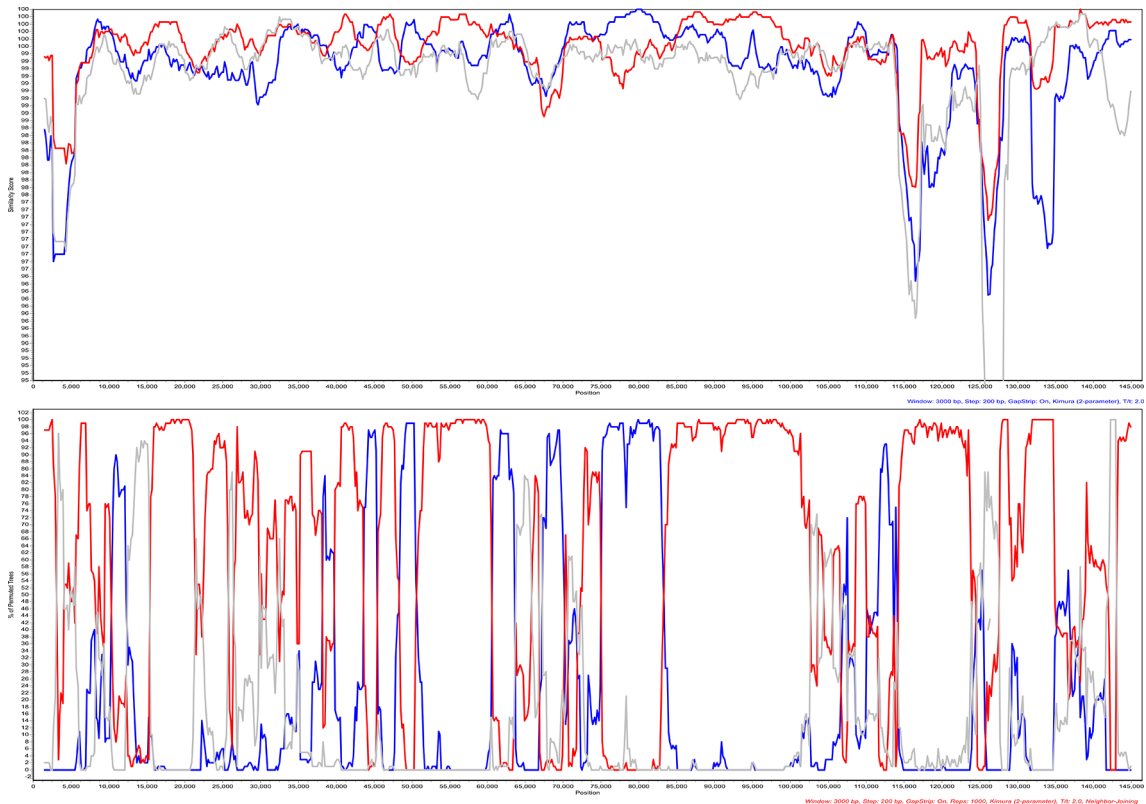
Fig. 5. Recombination analysis of the MG646679 strain of HSV-1: recombination analysis was performed using Simplot and Bootscan using a window size of 3000 bps and a step size of 200 bp. In total, 1000 pseudo-replicates were used for running Bootscan analysis.

should always keep a keen eye for ruling out and managing the HSV-1 cases having unusual presentations [23]. There are cases of nosocomial HSV-1 infection reported and also cases of disseminated HSV infection in pregnancy leading to coagulopathies [24]. Pregnant women should be screened for HSV-1 infection to prevent neonatal herpes infection. Ruling out genital herpes in pregnancy will help in planning cesarean delivery, which will reduce the chances of neonatal herpes infection. Diagnosis of the neonatal disease is challenging as only a few mothers will show symptoms of HSV infection, and most of them are asymptomatic. Hence, it is advisable for a thorough clinical examination during the antenatal period. Also, to prevent the nosocomial infection of HSV-1, it is desirable for regular fumigation of even labour rooms, avoiding the sharing of the bed sheets and linens with the patients. It is crucial for the strict adherence of hospital infection control and biomedical waste-management protocols to be followed.

In conclusion, the study identifies and characterizes HSV-1 genomes from two clinical samples having different clinical symptoms. Further, the study also determines the variation present on glycoprotein sites, which might be necessary for the manifestation of the clinical condition reported.

METHOD

Testing for various etiological agents present in clinical samples by molecular methods

Viral RNA was extracted from nasal swab, throat swab and serum samples from both the cases using the viral RNA extraction kit (Qiagen, Hilden, Germany). The extracted RNA from a serum sample of case-1 was used for performing real-time reverse transcriptase (RT)-PCR for determining the presence of RNA viruses [25–28]. The CSF sample of case-2 was tested for RNA viruses causing AES by real-time RT-PCR [29–31]. Further viral DNA was extracted from the throat swab of case-1 and CSF of case-2 and was used to diagnose HSV-1 (Genekam, Duisburg, Germany) and Varicella Zoster Virus (VZV) [32] by the conventional PCR method. The extracted DNA sample was only tested for the presence of HSV-1 and VZV, no other DNA virus or bacteria were tested.

Virus isolation

Virus isolation attempts for samples from case-1 were made using serum, nasal swab and throat swab samples in the *VeroCCL81* cell line as described previously [32]. Passages for all three samples in the cell culture were made at an interval of 7 days (four passages of serum and nasal swab,
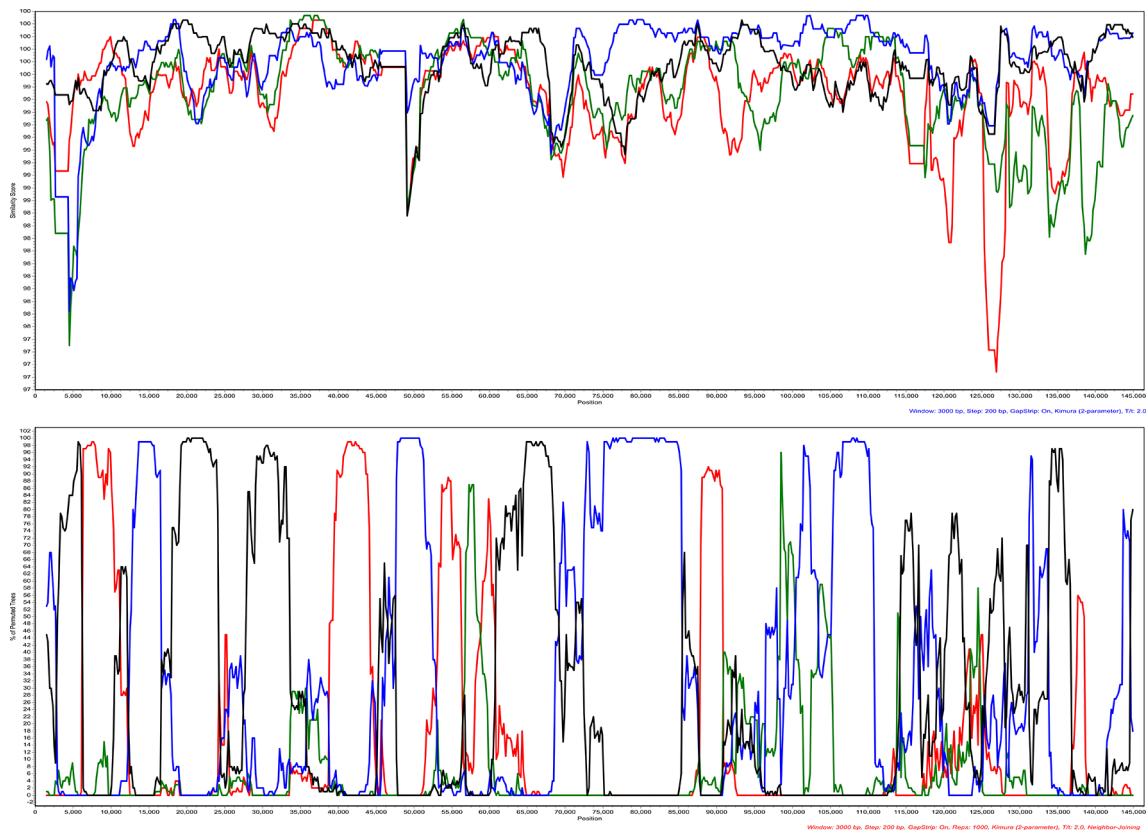
Fig. 6. Recombination analysis of the MH319852 strain of HSV-1: recombination analysis was performed using Simplot and Bootscan using a window size of 3000 bps and a step size of 200 bp. In total, 1000 pseudo-replicates were used for running Bootscan analysis.

and two passages of throat swab). Third and fourth passages for the throat swab samples were performed at an interval of 2 days. The harvested tissue culture fluid of all passages was stored at −80 °C until further use.

Virus isolation attempt from CSF of case-2 was made in BHK-21 cell lines using the protocol as described in Yadav *et al.* [33]. Three passages were done for *BHK-21* cells at an interval of 7 days. Once 80 % CPE was seen, the supernatant tissue culture fluid and infected cells were flushed entirely for single suspension and aliquots were stored at −80 °C until further use.

The cells used for passage were obtained from the American Type Culture Collection (ATCC) centre and maintained in the laboratory. The ATCC numbers for *VeroCCL81* and BHK-21 cells are ATCC CCL-81 and BHK21 [C13] (ATCC CCL10), respectively.

Next-generation sequencing of the virus isolates

The stored *VeroCCL81* and *BHK*-21 tissue culture fluid from the fourth passage of throat swab and third passage of CSF specimens, of case-1 and case-2, respectively, were reinoculated in freshly prepared *VeroCCL81* cells. The inoculum was used at a dilution of 10^{-2}. After observation of 2+CPE at the end of the second-day p.i., the cells and tissue culture

supernatant were harvested and further processed for DNA extraction.

The DNA was extracted from 1 ml of the tissue culture supernatant using a DNA extraction kit (Qiagen, Germantown, MD, USA). The DNA concentration was measured by Qubit 2.0 fluorometer using Qubit dsDNA kit (Thermofisher Scientific, MA, USA). DNA samples were normalized to a concentration of 0.2 ng μl^{-1}, and 5 μl of them were used to prepare the DNA libraries. The DNA libraries were made using the Nextera XT library preparation kit (Illumina, San Diego, CA, USA) and subsequently purified and size-selected using AgencourtAMPure XP beads (Beckman Coulter, Indianapolis, IN, USA). The prepared libraries were quantified using the KAPA library quantification kit (Roche Biosystems, Indianapolis, IN, USA) by doing real-time PCR. The libraries were normalized to 1 nM, pooled together and denatured using 0.1 N NaOH. The denatured libraries were neutralized using 0.1 M Tris (pH-7.0). The libraries were further diluted to 1.8 pM using hybridization buffer and loaded on Illumina MiniSeq mid output cartridge for paired-end sequencing. The resulting FastQ sequencing files were imported into CLC Genomics software (Qiagen, Germantown, MD, USA) and analysed further. Data were reference mapped to the sequence downloaded from the NCBI database.

Phylogenetic analysis and glycoprotein characterization of retrieved genome

The GenBank database was used to download the reference sequences for performing the comparison of the isolated viruses. Phylogenetic analysis was performed for 35 reference sequences along with the two genome sequence retrieved from this work. A total of 37 genome sequences belonging to different clades were used in the phylogenetic analysis to identify the circulating clade. The complete genome, as well as 11 glycoprotein regions for the reference sequence downloaded from Genbank, were each aligned using the ClustalW algorithm as implemented in MEGA software version 7.0 [34]. The gaps at the 5′ and 3′ ends were removed manually and the one inbetween the alignment was not disturbed during phylogenetic analysis. The best substitution matrix was identified using the best model identification program as implemented in the MEGA software. An evolutionary tree was generated using the maximum-likelihood method using the general time reversible model with Gamma+Invariable method. A bootstrap replication of 500 cycles was performed to check the statistical robustness of the obtained tree. Sequences belonging to clade II were aligned with the retrieved genome for individual glycoprotein and analysed for conservation. Average evolutionary divergence for different glycoproteins was calculated using MEGA 7.0 software.

Recombination analysis of the retrieved genome

Recombination analysis was performed using recombination detection program (RDP) software to determine evidence of recombination in both the isolates. Recombination was predicted using RDP version 4.95 [35]. The similarity plot along with the bootscan results was plotted using Simplot version 3.5.1 (http://sray.med.som.jhmi.edu/SCRoftware/SimPlot). Similarity plot and Bootscan plot were generated using the Kimura 2-parameter model and a window size of 3000 bp and step size of 200 bp. A neighbour-joining tree with a bootstrap replication of 1000 cycle was used to generate the output. A strain was determined to be recombinant if more than four programs implemented in RDP predicted recombination.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References

1. Laine RF, Albecka A, van de Linde S, Rees EJ, Crump CM *et al*. Structural analysis of herpes simplex virus by optical super-resolution imaging. *Nat Commun* 2015;6:5980.

2. Kimberlin DW. Herpes simplex virus infections of the newborn. *Semin Perinatol* 2007;31:19–25.

3. Lee R, Nair M. Diagnosis and treatment of herpes simplex 1 virus infection in pregnancy. *Obstet Med* 2017;10:58–60.

4. Simmons A. Clinical manifestations and treatment considerations of herpes simplex virus infection. *J Infect Dis* 2002;186 Suppl 1:S71–S77.

5. James SH, Sheffield JS, Kimberlin DW. Mother-To-Child transmission of herpes simplex virus. *J Pediatric Infect Dis Soc* 2014;3:S19–S23.

6. Bondre VP, Sankararaman V, Andhare V, Tupekar M, Sapkal GN. Genetic characterization of human herpesvirus type 1: full-length genome sequence of strain obtained from an encephalitis case from India. *Indian J Med Res* 2016;144:750–760.

7. Brown JC. High G+C content of herpes simplex virus DNA: proposed role in protection against retrotransposon insertion. *Open Biochem J* 2007;1:33–42.

8. Oda S, Arii J, Koyanagi N, Kato A, Kawaguchi Y. The interaction between herpes simplex virus 1 tegument proteins UL51 and UL14 and its role in virion morphogenesis. *J Virol* 2016;90:8754–8767.

9. Reske A, Pollara G, Krummenacher C, Chain BM, Katz DR. Understanding HSV-1 entry glycoproteins. *Rev Med Virol* 2007;17:205–215.

10. Jaggi U, Wang S, Tormanen K, Matundan H, Ljubimov AV *et al*. Role of herpes simplex virus type 1 (HSV-1) glycoprotein K (gK) pathogenic CD8+ T cells in exacerbation of eye disease. *Front Immunol* 2018;9:2895.

11. Sivadon V, Lebon P, Rozenberg F. Variations of HSV-1 glycoprotein B in human herpes simplex encephalitis. *J Neurovirol* 1998;4:106–114.

12. Izumi KM, Stevens JG. Molecular and biological characterization of a herpes simplex virus type 1 (HSV-1) neuroinvasiveness gene. *J Exp Med* 1990;172:487–496.

13. Rozenberg F, Lebon P. Analysis of herpes simplex virus type 1 glycoprotein D nucleotide sequence in human herpes simplex encephalitis. *J Neurovirol* 1996;2:289–295.

14. Arii J, Wang J, Morimoto T, Suenaga T, Akashi H *et al*. A single-amino-acid substitution in herpes simplex virus 1 envelope glycoprotein B at a site required for binding to the paired immunoglobulin-like type 2 receptor alpha (PILRalpha) abrogates PILRalpha-dependent viral entry and reduces pathogenesis. *J Virol* 2010;84:10773–10783.

15. Engel JP, Boyer EP, Goodman JL. Two novel single amino acid syncytial mutations in the carboxy terminus of glycoprotein B of herpes simplex virus type 1 confer a unique pathogenic phenotype. *Virology* 1993;192:112–120.

16. Sampath A, Maduro G, Schillinger JA. Infant deaths due to herpes simplex virus, congenital syphilis, and HIV in New York City. *Pediatrics* 2016;137:e20152387

17. Beig FK, Malik A, Rizvi M, Acharya D, Khare S. Etiology and clinico-epidemiological profile of acute viral encephalitis in children of Western Uttar Pradesh, India. *Int J Infect Dis* 2010;14:e141–e146.

18. Modi A, Atam V, Jain N, Gutch M, Verma R. The etiological diagnosis and outcome in patients of acute febrile encephalopathy: a prospective observational study at tertiary care center. *Neurol India* 2012;60:168–173.

19. Pfaff F, Groth M, Sauerbrei A, Zell R. Genotyping of herpes simplex virus type 1 by whole-genome sequencing. *J Gen Virol* 2016;97:2732–2741.

20. Lan K, Luo M-H. Herpesviruses: epidemiology, pathogenesis and interventions. *Virol Sin* 2017;32:347–348.

21. James SH, Whitley RJ. Treatment of herpes simplex virus infections in pediatric patients: current status and future needs. *Clin Pharmacol Ther* 2010;88:720–724.

22. Allen UD, Robinson JL, , Canadian Paediatric Society, Infectious Diseases and Immunization Committee. Prevention and management of neonatal herpes simplex virus infections. *Paediatr Child Health* 2014;19:201–206.

23. Jouan Y, Grammatico-Guillon L, Guillon A. Nosocomial herpes simplex encephalitis: does it really exist? *J Infect Public Health* 2018;11:142.

24. Young EJ, Chafizadeh E, Oliveira VL, Genta RM. Disseminated herpesvirus infection during pregnancy. *Clin Infect Dis* 1996;22:51–58.

25. Arankalle VA, Lole KS, Deshmukh TM, Chobe LP, Gandhe SS. Evaluation of human (genotype 1) and swine (genotype 4)-ORF2-based ELISAs for anti-HEV IgM and IgG detection in an endemic country and search for type 4 human HEV infections. *J Viral Hepat* 2007;14:435–445.

26. Liu J, Ochieng C, Wiersma S, Ströher U, Towner JS *et al*. Development of a taqman array card for acute-febrile-illness outbreak investigation and surveillance of emerging pathogens, including ebola virus. *J Clin Microbiol* 2016;54:49–58.

27. Mourya DT, Yadav PD, Mehla R, Barde PV, Yergolkar PN *et al*. Diagnosis of Kyasanur forest disease by nested RT-PCR, real-time RT-PCR and IgM capture ELISA. *J Virol Methods* 2012;186:49–54.

28. Waggoner JJ, Gresh L, Mohamed-Hadley A, Ballesteros G, Davila MJV *et al*. Single-Reaction multiplex reverse transcription PCR for detection of Zika, Chikungunya, and dengue viruses. *Emerg Infect Dis* 2016;22:1295–1297.

29. Bondre VP, Jadi RS, Mishra AC, Yergolkar PN, Arankalle VA. West Nile virus isolates from India: evidence for a distinct genetic lineage. *J Gen Virol* 2007;88:875–884.

30. Gurav YK, Tandale BV, Jadi RS, Gunjikar RS, Tikute SS *et al*. Chandipura virus encephalitis outbreak among children in Nagpur division, Maharashtra, 2007. *Indian J Med Res* 2010;132:395–399.

31. Sapkal GN, Wairagkar NS, Ayachit VM, Bondre VP, Gore MM. Detection and isolation of Japanese encephalitis virus from blood clots collected during the acute phase of infection. *Am J Trop Med Hyg* 2007;77:1139–1145.

32. Weidmann M, Meyer-König U, Hufert FT. Rapid detection of herpes simplex virus and varicella-zoster virus infections by real-time PCR. *J Clin Microbiol* 2003;41:1565–1568.

33. Yadav P, Sarkale P, Patil D, Shete A, Kokate P *et al*. Isolation of Tioman virus from Pteropus giganteus bat in north-east region of India. *Infect Genet Evol* 2016;45:224–229.

34. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol* 2016;33:1870–1874.

35. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol* 2015;1.