*Journal of General Virology* (2007), **88**, 875–884 DOI 10.1099/vir.0.82403-0

# West Nile virus isolates from India: evidence for a distinct genetic lineage

Vijay P. Bondre, R. S. Jadi, A. C. Mishra, P. N. Yergolkar and V. A. Arankalle

Microbial Containment Complex Unit, National Institute of Virology (ICMR), Sus Road, Pashan, Pune 411 021, India

Correspondence
Vijay P. Bondre
vpbondre@yahoo.com

The complete genomic sequence of one human isolate of *West Nile virus* (WNV) and the partial genomic sequences of 14 other strains from India isolated in the period 1955–1982 from different hosts and geographical areas were determined. Phylogenetic analyses based on complete and partial genomic sequences (921 nt of the C–prM–E region) revealed that WNV could be classified into five distinct groups that differed from each other by 20–25 % at the complete genome level and by 20–26 % using partial sequences. Of the Indian isolates, 13 formed a distinct genetic lineage, lineage 5, whereas two isolates, one from a human patient (1967) and another from a bat (1968), were related closely to lineage 1 strains. The complete genomic sequence of the Indian isolate, 804994, showed 20–22 % genetic divergence from the previously proposed lineage 1 and 2 strains and 24–25 % divergence from isolates of the newly proposed lineages 3 (Rabensburg isolate 97-103 of 1997) and 4 (Russian isolate LEIV-Krnd88-190 of 1998). Similarly, the partial genomic sequences of the Indian isolates showed 21–26 % divergence from lineage 1 and 2 strains and from the Rabensburg (97-103) and Russian (LEIV-Krnd88-190) isolates. Cross-neutralization using strain-specific polyclonal antibodies against lineage 1 strain Eg-101 and representative Indian strains suggests substantial antigenic variation. This study documents circulation of WNV strains typical to India for 27 years and the introduction of lineage 1 strains during 1967–1968. These results indicate strongly that WNV should be classified into five genetic lineages, with Indian viruses constituting the distinct genetic lineage 5.

Received 21 July 2006
Accepted 15 November 2006

## INTRODUCTION

*West Nile virus* (WNV) is a member of the Japanese encephalitis virus (JEV) serocomplex belonging to the genus *Flavivirus* of the family *Flaviviridae*. It was first isolated in the West Nile region of Uganda and has been the cause of many epidemics in different countries (Smithburn *et al.*, 1940; Tsai *et al.*, 1998; Hubalek & Halouzka, 1999; Petersen & Roehrig, 2001). In India, the existence of antibodies against WNV in the human population was recorded for the first time in 1952 (Banker, 1952). The virus was subsequently shown to be prevalent in different parts of the country (Carey *et al.*, 1968; Banerjee, 1996; Damle *et al.*, 1998; Thakare *et al.*, 2002). So far, WNV has been isolated from various mosquito species, a fruit bat (*Rousettus leschenaulti*) and sporadic human encephalitic cases in southern India (Paul *et al.*, 1970a; Banerjee *et al.*, 1979; George *et al.*, 1984; Kedarnath *et al.*, 1984). Its introduction in late 1999 and rapid spread in America with high rates of mortality and morbidity in humans, domestic animals and birds resulted in much concern globally (Centers for Disease Control and Prevention, 1999).

The genome of WNV is a single-stranded, positive-sense RNA of about 11 kb containing a single, long open reading frame flanked by non-coding regions at both ends. The 5′ region of the genome encodes structural proteins, whereas the 3′ region encodes non-structural proteins. On the basis of phylogenetic analysis, WNV has been grouped into two major genetic lineages in previously published studies. Isolates from north, west and central Africa, southern and eastern Europe, India and the Middle East have been grouped in lineage 1, whereas isolates from west, central and east Africa and Madagascar constitute lineage 2. Lineage 1 is further subdivided into three clades. Clade 1a consists of strains from Europe, Africa, the USA and Israel. The Australian Kunjin virus belongs to clade 1b, whereas isolates from India form clade 1c (Lanciotti *et al.*, 2002). Recently, Rabensburg isolate 97-103 from *Culex pipiens* mosquitoes (1997) from South Moravia, Czech Republic, and LEIV-Krnd88-190 isolated from *Dermacentor marginatus* ticks in the north-west Caucasus Mountain valley in 1998 have been proposed to be novel variants of WNV. These isolates are genetically different from lineage 1 and 2 viruses and have been proposed as members of lineages 3 (Rabensburg 97-103) and 4 (LEIV-Krnd88-190), respectively (Prilipov *et al.*, 2002; Lvov *et al.*, 2004; Bakonyi *et al.*, 2005).

Preliminary comparative studies of WNV strains from European, Asian and African countries revealed differences

in the Indian strains, thus indicating the necessity for further studies (Hammam & Price, 1966; Price & O'Learly, 1967; Gaidamovich & Sokhe, 1973; Odelola & Fabiyi, 1976; Umrigar & Pavri, 1977; Mathiot *et al.*, 1990). Comparative biological studies of some of the Indian viruses isolated from mosquitoes, humans and bat revealed different patterns with respect to pathogenicity in mice, susceptibility of Vero cell cultures and plaque size (Umrigar & Pavri, 1977). Similarly, differential neuroinvasiveness in mice has been demonstrated for some of the Indian viruses (Mathiot *et al.*, 1990). Five Indian viruses examined by different researchers using short genomic sequences of the E and NS5 regions were placed as a distinct clade (1c) within lineage 1 (Mathiot *et al.*, 1990; Porter *et al.*, 1993; Berthet *et al.*, 1997; Tsai *et al.*, 1998; Jia *et al.*, 1999; Lanciotti *et al.*, 1999, 2002; Savage *et al.*, 1999; Platonov *et al.*, 2001; Scherret *et al.*, 2001; Burt *et al.*, 2002). Analysis of the complete genome sequence of one of the Indian strains, G2266, from mosquito (1955) indicated its distinct genetic nature compared with lineage 1 and 2 viruses (Prilipov *et al.*, 2002). Considering the importance of WNV as an emerging pathogen, a systematic study of WNV strains isolated in India over a period of 27 years (1955–1982) was undertaken. Analysis of the complete genome sequence of one Indian strain and partial sequences of 14 additional strains indicates that these viruses cluster into a distinct genetic lineage.

## METHODS

**WNV isolates.** WNV isolate stocks maintained at the lowest passage level at the Virus Repository of National Institute of Virology (NIV), Pune, India, were used in this study (Table 1). These viruses were isolated from mosquito pools, and bat and human specimens by intracerebral (i.c.) inoculation in infant and adult mice. The 10 % bovalbumin phosphate saline suspensions of harvested brains of sick mice were further inoculated in infant and adult mice by the i.c. route and animals were observed for sickness. The crude suspensions of brains from the sick mice were characterized by complement fixation, haemagglutination inhibition and neutralization tests using hyperimmune sera against JEV, *Kyasanur Forest disease virus* (KFDV) and WNV (Paul *et al.*, 1970a, b; Rodrigues *et al.*, 1980; George *et al.*, 1984). Pools of respective isolates were stored at −80 °C until use.

The Institutional Animal Ethical Committee approved the work and ethical guidelines were strictly followed as per recommendations.

**Plaque reduction neutralization test.** To determine the antigenic relationship among different viruses, hyperimmune sera raised against representative WNV strains were used. The hyperimmune sera against Eg-101, 68856, G22886, 821622 and 804994 were generated by immunizing mice (6-week-old Swiss albino) with four doses of the respective viruses at 7 day intervals by the intraperitoneal (i.p.) route. The peripherally pathogenic strain 68856 was inactivated with $\beta$-propiolactone treatment (1 : 2000 final concentrations) at 4 °C for 48 h before being used for immunization. Mice were bled and serum was separated from blood by centrifugation at 3000 ***g*** for 10 min at 4 °C. The serum raised against different viruses was used for neutralization studies. An *in vitro* neutralization test was carried out in a porcine stable kidney (PS) cell line maintained in modified minimal essential medium (MEM; Gibco-BRL Life Technologies) with 10 % fetal bovine serum (FBS; Gibco-BRL Life Technologies). Fourfold serial dilutions of heat-inactivated hyperimmune sera (30 min at 56 °C) were made in MEM supplemented with 2 % heat-inactivated FBS. Serum samples (60 µl) were mixed with equal amounts of representative WNV suspensions containing 100 p.f.u. and incubated for 60 min at 37 °C in an incubator with 5 % $CO_2$. The appropriate virus strains, sera (non-immunized mouse serum and strain-specific hyperimmune sera) and uninfected PS cells were used as controls in the experiment. At the end of the incubation period, 0.1 ml virus–serum mixture was allowed to adsorb for 60 min at 37 °C in 5 % $CO_2$ on a confluent monolayer of PS cells in 24-well plates. Wells were washed with MEM and cells were overlaid

**Table 1.** Indian WNV isolates used in this study

| Isolate no. | Source of isolation | Location | Year | GenBank accession no. |
|---|---|---|---|---|
| IND804994H (complete sequence) | Human | Bangalore | 1980 | DQ256376 |
| Eg-101 | Human | NIV registry | 1951 | Not submitted* |
| WNIG2267Cv | *Culex vishnui* | Sathuperi | 1955 | AY944243 |
| WNIG16919Cv | *Culex vishnui* | Kammavanpet | 1957 | AY944245 |
| WNIG22886Cv | *Culex vishnui* | Sathuperi | 1958 | AY944241 |
| WNI672698H | Human serum | Kasoudi | 1967 | AY944238 |
| WNI68856B | *R. leschenaulti* (fruit bat) | Horabail | 1968 | AY944239 |
| WNI80245Ct | *Culex tritaeniorhynchus* | Muduuvadi | 1980 | AY639639 |
| WNI80829Cw | *Culex whitmorei* | Kodikannoor | 1980 | AY639640 |
| WNI80755Cb | *Culex bitaeniorhynchus* | Kodikannoor | 1980 | AY639641 |
| WNI804987Cv | *Culex vishnui* | Bangalore | 1980 | AY639642 |
| WNI804994H | Human | Bangalore | 1980 | AY639643 |
| WN80776AnS | *Anopheles subpictus* | Kodikannoor | 1980 | AY944244 |
| WN80235Cv | *Culex vishnui* | Arahalli | 1980 | AY944246 |
| WN80897Cw | *Culex whitmorei* | Arahalli | 1980 | AY944242 |
| WN43478Cv | *Culex vishnui* | Gownpalli | 1981 | AY944240 |
| WN821622H | Human plasma | Kolar | 1982 | AY944247 |

*The sequence generated in this study was identical to the GenBank sequence AF260968 and hence not submitted to GenBank.

with an equal amount of 2× MEM and 1 % carboxymethylcellulose supplemented with 2 % FBS. After an incubation period of 96 h, plates were prescreened under an inverted microscope. Medium was removed and plates were stained with 1 % amido black (Sigma) for 20 min at room temperature. The plates were washed with distilled water and the plaques were counted (Gore *et al.*, 1990). The serum dilution showing 80 % plaque reduction was considered as a neutralizing end point.

**RNA isolation, RT-PCR and sequencing.** Genomic RNA was prepared from WNV-infected mouse brains using Trizol reagent (Invitrogen) according to the manufacturer's instructions. Primers WN-1189R (5′-GCAATAACTGCGGACCTCTGC-3′) and WN-233F (5′-GACTGAAGAGGGCAATGTTGAGC-3′) were used for PCR amplification of the 921 nt (C–prM–E) fragment (Anderson *et al.*, 2001). Additional internal primers for sequencing (921 nt fragment) and primers used for PCR amplification and sequencing of the complete genome were designed by alignment of the WNV sequences from GenBank with CLUSTAL X 1.83. Integrated DNA Technologies synthesized all the primers used in this study. Primers used in this study are available from the author on request.

RNA (2 µl) was reverse-transcribed with Thermoscript reverse transcriptase (Invitrogen) at 52 °C. cDNA (2 µl) was amplified by PCR with Platinum *Taq* DNA Polymerase – High Fidelity (Invitrogen). The amplification was carried out by denaturing the DNA at 94 °C for 5 min, followed by 35 cycles of 94 °C for 30 s, 54 °C for 60 s and 72 °C for 60 s, and final extension at 72 °C for 7 min. PCR products were column-purified (QIAquick PCR purification kit; Qiagen) and both strands were sequenced by using a Big Dye Terminator cycle sequencing ready reaction kit (Applied Biosystems) and an automated Sequencer (ABI Prism 310 Genetic Analyzer; Applied Biosystems).

**Phylogenetic analysis.** Multiple alignments of nucleotide/amino acid sequences were carried out by using CLUSTAL X 1.83. A phylogenetic tree was constructed by using the 921 nt fragment covering the C–prM–E region of 15 Indian isolates and representative sequences of all WNVs available in GenBank. The phylogenetic status of the Indian isolates was assessed by using the software MEGA (Kumar *et al.*, 2001). For analysis in MEGA, Jukes–Cantor distance was utilized, employing the neighbour-joining algorithm. The reliability of different phylogenetic groupings was evaluated by using the bootstrap test (1000 bootstrap replications) available in MEGA. Similarly, phylogenetic analysis of genomic fragments encoding different proteins, viz. nucleocapsid, premembrane and membrane, envelope and non-structural proteins 1–5, and the 5′ and 3′ untranslated regions (UTRs) was carried out to understand the relationship between different WNV strains. A phylogenetic tree was also constructed by using the complete genome sequence of the Indian WNV isolate 804994 (this study) and complete genomic sequences (GenBank) of representative strains from different genomic groups. The JEV-GP78 sequence was used as an outgroup in the phylogenetic analysis of partial, as well as complete, genome sequences.

# RESULTS

## Plaque reduction neutralization test

As shown in Table 2, marked differences were observed in the cross-neutralization activity of polyclonal hyperimmune serum raised against Eg-101, 68856, G22886, 821622 and 804994. Highest neutralization activity was observed with homologous hyperimmune serum against the respective virus strains. A greater difference in neutralization activity was observed between Eg-101 and Indian viruses. However, 68856 showed less reactivity with other Indian viruses (G22886, 821622 and 804994) compared with the cross-reactivity between these strains; the homologous titres of these sera were always higher (Table 2).

**Table 2.** Homologous and heterologous cross-neutralization of WNV strains

Table showing the results of *in vitro* neutralization experiments carried out using hyperimmune sera raised against the individual viruses in mice and tested against different viruses.

| Hyperimmune sera against: | Test virus (100 p.f.u.) | | | | |
|---|---|---|---|---|---|
| | Eg-101 | 68856 | G22886 | 821622 | 804994 |
| Eg-101 | 64 | 20 | 16 | 10 | 8 |
| 68856 | 32 | 128 | 8 | 8 | 10 |
| G22886 | 20 | 8 | 64 | 16 | 64 |
| 821622 | 8 | 8 | 32 | 160 | 64 |
| 804994 | 8 | 20 | 32 | 64 | 128 |

## Partial sequence analysis

The viruses examined during this study were isolated over a period of 27 years (1955–1982) from different geographical regions of southern India (Table 1). Of these, three isolates obtained in 1967, 1980 and 1982 were from human patients and one from 1968 was from a bat. Eleven isolates were obtained from different mosquito species during the period 1955–1981. Eg-101, the prototype strain isolated from a human case in 1951 in Egypt, was used as a control. The sequence obtained for Eg-101 was identical to the GenBank sequence (accession no. AF260968).

Initially, a 921 nt fragment spanning the C–prM–E region was amplified and sequenced. Phylogenetic analysis of this partial sequence showed that the different WNVs reported worldwide and those isolated in India formed five distinct lineages that differed from each other by 21.0–26.5 % (Table 3). As evident from Fig. 1, all of the previously characterized clade 1a (lineage 1) viruses, including the two Indian isolates 672698 and 68856, grouped in clade 1a with 95.23–100 % nucleotide identity (NI), whereas Kunjin virus formed clade 1b (87.87–89.87 % NI with clade 1 viruses) within lineage 1. The African viruses formed lineage 2 as reported previously (Lanciotti *et al.*, 2002). Rabensburg isolate 97-103 (1997) and Russian isolate LEIV-Krnd88-190 (1998) formed lineages 3 and 4, respectively, whereas the remaining 13 Indian isolates clustered together to form a distinct genetic lineage, lineage 5.

The NI among different WNV isolates and the relationship between different genetic lineages is shown in Table 3. Partial sequence analysis of the 15 Indian viruses showed substantial genomic divergence of isolates 672698 and 68856. These isolates showed 77.25–79.20 % NI to the other 13 Indian isolates and 96.10 ± 1.38 % NI to each other, thus indicating their distinct genetic nature. However, these two isolates showed 95.23–100 % NI to clade 1a and 89.17–89.82 % NI to

**Table 3.** NI analysis based on partial sequences (921 nt) of WNVs from five different genetic lineages formed in this study along with JEV

The mean NI ± SEM (%) between different lineages is shown. *n*, Number of partial sequences used for determination of mean NI; ND, not done; NA, not available.

| Viruses from lineage: | 1 (clade 1a) | 1 (clade 1b) | 2 | 3 | 4 | 5 | JEV-GP78 |
|---|---|---|---|---|---|---|---|
| 1 (clade 1a) (*n*=33) | 97.61±0.82 | | | | | | |
| 1 (clade 1b) (*n*=1) | 88.87±1.03 | ND | | | | | |
| 2 (*n*=3) | 81.04±1.29 | 79.58±1.32 | 94.80±0.74 | | | | |
| 3 (*n*=1) | 77.35±1.38 | 77.57±1.37 | 76.33±1.40 | NA | | | |
| 4 (*n*=1) | 74.67±1.43 | 73.48±1.46 | 73.20±1.46 | 73.26±1.46 | NA | | |
| 5 (*n*=13) | 78.38±1.35 | 76.71±1.39 | 77.41±1.37 | 75.51±1.41 | 74.13±1.44 | 98.53±0.55 | |
| JEV (*n*=1) | 67.60±1.55 | 65.87±1.56 | 66.15±1.56 | 66.74±1.55 | 66.52±1.56 | 67.39±1.54 | ND |

clade 1b viruses belonging to lineage 1. The 13 other Indian isolates showed NI of 97.07–100 % to each other. Distance analysis of the 921 nt genomic fragments of WNVs forming different genetic lineages is given in Table 3.

To determine the extent, if any, of host-dependent genomic changes in WNVs, a 921 nt fragment from eight isolates from five different mosquito species and one human patient who died of West Nile encephalitis, which were all collected during the same transmission season (1980), were compared. The 921 nt sequences of these viruses showed a maximum of 0.87 % divergence from each other. Of the eight isolates, three (80897 from *Culex whitmorei*, 80235 from *Culex vishnui* and 804994 from a human) showed two to three nucleotide changes at different locations (in C, prM and E regions). The nucleotide changes recorded were: T to C (458; amino acid V to A), A to T (653; amino acid E to V) and T to G (1037; amino acid V to G) in 80235; C to G (459), C to T (630), T to A (1037; amino acid V to E) in 80897; and G to C (1118; amino acid S to T) and G to T (1125; silent) in 804994. Sequences of the other five isolates (80776 from *Anopheles subpictus*, 80245 from *Culex tritaeniorhynchus*, 80829 from *Culex whitmorei*, 80755 from *Culex bitaeniorhynchus* and 804987 from *Culex vishnui*) were identical.

As the Indian WNVs had been placed previously within lineage 1 as a separate cluster (clade 1c) on the basis of partial E or NS5 sequences of five isolates (Scherret *et al.*, 2001; Lanciotti *et al.*, 1999, 2002), a separate analysis based on subsequences of the C–prM–E fragment encoding different proteins (C, prM, M and E) was carried out. Virtually identical results were obtained with the Indian isolates that formed a distinct genetic lineage. JEV-GP78 (outgroup) showed NI of 67.60±1.55 % (clade 1a) and 65.87±1.56 % (clade 1b) to lineage 1 viruses, 66.15±1.56 % to lineage 2 viruses, 66.74±1.55 % to Rabensburg isolate 97-103, 66.52±1.56 % to Russian isolate LEIV-Krnd88-190 and 67.39±1.54 % to Indian viruses.

## Complete genome sequence analysis

Considering the distinct status of the Indian isolates, the full genome sequence of a representative isolate (804994) was determined. Full genome-based phylogenetic analysis classified the WNV isolates in five distinct genetic lineages (Fig. 2), similar to the analysis based on partial genomic sequence analysis (Fig. 1). Lineage 1 included all the viruses previously grouped within clades 1a and 1b (Kunjin virus) of lineage 1. Lineage 2 included the viruses WNFCG, Sarafend and B956, previously grouped in lineage 2. The Rabensburg isolate 97-103 and the Russian isolate LEIV-Krnd88-190 constituted lineages 3 and 4, respectively. The Indian isolate 804994 formed a distinct genetic lineage, lineage 5.

The NI percentages determined on the basis of complete genome sequences of viruses forming different genetic lineages are shown in Table 4. The complete genome sequence of 804994 showed 79.00–79.99 % NI to all lineage 1 viruses, including Kunjin virus, 78–79 % NI to lineage 2 viruses, 76.24±0.41 % NI to Rabensburg isolate 97-103 (lineage 3) and 75.24±0.42 % NI to isolate LEIV-Krnd88-190 (lineage 4). Among the viruses belonging to different lineages, the NI varied from 95.37 to 100 % (lineage 1, clade 1a), from 88.23 to 88.81 % (between clade 1a and 1b of lineage 1) and from 79.39 to 80.11 % between lineage 2 viruses. The viruses clustered within lineage 1 showed NI of 79.36–80.11 % to lineage 2, 77.21–77.66 % to Rabensburg isolate 97-103 (lineage 3), 75.67–76.02 % to LEIV-Krnd88-190 isolate (lineage 4) and 79.45–79.99 % to isolate 804994 (lineage 5). Similarly, the NI among lineage 2, 3, 4 and 5 viruses varied between 75 and 80 %, thus separating them from each other (Table 4). Isolate 804994 showed 75–80 % identity to different viruses forming different genetic lineages, including the newly proposed lineage 3 and 4 viruses. The changes in nucleotides of 804994 were dispersed throughout the genome and were not centred to any specific gene. Sequences for multiple viruses were available within lineages 1 (with 95–100 % NI) and 2 (with 93.5–100 % NI); however, the complete genome sequence of only one virus was available for lineage 3, 4 and 5 viruses.

Comparison of individual genes showed that, although distinct, the lineage 5 Indian isolates were closer to lineage 1 and 2 isolates than lineage 3 and 4 isolates in all regions of the genome. At the nucleotide level, the Indian viruses showed 6.51–10.99 % divergence in the 5′ UTR, 17.79–22.32 % in

the nucleocapsid, 20.22–27.89 % in the membrane, 18.05–27.64 % in the envelope, 13.04–25.95 % in NS1, 19.34–27.39 % in NS2A, 16.56–20.84 % in NS2B, 19.88–24.91 % in NS3, 21.38–27.52 % in NS4A, 21.07–26.27 % in NS4B, 11.85–19.55 % in NS5 and 14.18–23.63 % in the 3′ UTR, compared with lineage 1, 2, 3 and 4 viruses. At the nucleotide level, JEV showed 68.18–69.80 % NI to complete genomic sequences of all WNVs, whereas at the level of individual genomic fragments encoding different proteins, the NI was 66.15–67.60 %, except for both of the non-coding regions, which showed NI of 61.90–66.13 % to respective genomic fragments of WNV.

Similarly, at the amino acid level, the Indian virus showed 12.20–22.64 % divergence in nucleocapsid, 7.84–17.68 % in membrane, 7.71–21.77 % in envelope, 5.65–15.34 % in NS1, 4.56–12.18 % in NS2A, 5.31–15.21 % in NS2B, 7.64–11.07 % in NS3, 5.61–12.62 % in NS4A, 5.65–12.81 % in NS4B and 4.29–8.39 % in NS5 regions when compared with lineage 1, 2, 3 and 4 viruses. The membrane region was most divergent, whereas maximum identity was observed within the NS5 region at the amino acid level (data not shown).

## DISCUSSION

This study examined the genetic status of Indian isolates of WNV, an important emerging pathogen in the western hemisphere. Some of the Indian viruses available from other laboratories were partially sequenced to determine their phylogenetic status and were grouped within lineage 1 as clade 1c viruses in previous studies (Mathiot *et al.*, 1990; Porter *et al.*, 1993; Berthet *et al.*, 1997; Jia *et al.*, 1999; Lanciotti *et al.*, 1999, 2002; Savage *et al.*, 1999; Platonov *et al.*, 2001; Scherret *et al.*, 2001; Burt *et al.*, 2002). However, comparative studies between Indian and other lineage 1 viruses with respect to characteristics like antigenicity, neuroinvasiveness and pathogenesis in mice indicated a marked difference from other lineage 1 viruses (Hammam & Price, 1966; Price & O'Learly, 1967; Gaidamovich & Sokhe, 1973; Odelola & Fabiyi, 1976; Umrigar & Pavri, 1977; Mathiot *et al.*, 1990; Beasley *et al.*, 2002). Analysis of the complete genome sequence of one of the Indian isolates, G2266, by Prilipov *et al.*, (2002) indicated its distinct genetic nature compared with lineage 1 and 2 viruses. However, this sequence was not available in GenBank for analysis in this study.

We undertook a systematic study of the viruses isolated in India over a period of 27 years (1955–1982) from different geographical regions of southern India. Of these, three isolates obtained in 1967, 1980 and 1982 were from human patients and one isolate (1968) was obtained from a bat. The other 11 isolates were obtained from different mosquito species collected during the years 1955–1981. It was evident from the 921 nt-based analysis (Fig. 1) that 13 of the 15 Indian isolates grouped together to form a distinct genetic cluster. The degree of variation between different groups was 20–26 %, which suggests strongly that WNVs isolated worldwide can be classified into five distinct genetic lineages, with the Indian viruses forming a distinct genetic lineage. To substantiate this significant finding, the complete genome of one Indian isolate, 804994, was sequenced. Interestingly, two Indian viruses, one from a human patient isolated in 1967 and the other from a bat isolated in 1968, grouped with the lineage 1 viruses, indicating co-circulation of lineage 1 strains during 1967–1968.

The results of the full genome-based and 921 nt-based phylogenetic analyses were similar. The Indian viruses were genetically distinct from previously classified lineage 1 and 2 strains and formed a new genetic lineage. The degree of nucleotide variance between different lineages varied from 20.0 to 24.7 %. Hence, with the availability of additional complete genomic sequences of Rabensburg 97-103 (Prilipov *et al.*, 2002), LEIV-Krnd88-190, Ig-2266 (Lvov *et al.*, 2004) and 804994 (present study) viruses, the initial grouping of WNV into lineages 1 and 2 needs to be revised and modified. Our data suggest strongly that the available WNVs can be classified into five distinct genetic lineages, differing from each other by 20–25 % at the complete genome level, and also present the utility of the 921 nt genomic fragment, spanning a portion of the C–prM–E, in genetic analyses. Recent studies with Rabensburg 97-103, LEIV-Krnd88-190 and Ig-2266 isolates have shown clearly that these viruses represent new variants of WNV and belong to distinct lineages, proposed to be lineages 3 and 4 (Prilipov *et al.*, 2002; Lvov *et al.*, 2004; Bakonyi *et al.*, 2005). Considering the genetic divergence from all other known genetic lineages of WNV, the Indian WNV forms genetic lineage 5. Sequence comparison of 804994 with other isolates demonstrated that nucleotide substitutions were located throughout the genome; most substitutions were transitions. Maximum divergence among Indian viruses was found in the membrane region, whereas the greatest NI of 89.0–93.4 % was in the 5′ UTR among all viruses forming different genetic lineages. Amino acid changes unique to 804994 were recorded at 149 positions throughout the open reading frame.

Previous studies (Mathiot *et al.*, 1990; Porter *et al.*, 1993; Berthet *et al.*, 1997; Jia *et al.*, 1999; Lanciotti *et al.*, 1999, 2002; Savage *et al.*, 1999; Platonov *et al.*, 2001; Scherret *et al.*, 2001; Beasley *et al.*, 2002; Burt *et al.*, 2002) based on partial or complete envelope gene sequences of five Indian isolates (G2267, G22886, G16919, 804994 and 68856) placed all of the isolates in lineage 1. Revising these analyses showed that the Indian viruses differed from both lineage 1 and lineage 2 viruses by 22–25 %. The difference between lineage 1 and 2 viruses was 23–24 %. Analyses based on smaller fragments encoding partial envelope or NS3 regions also yielded the same results. As most of the previous studies were based on short genomic sequences, phylogenetic relationships among different lineages were derived using sequences encoding different proteins, as well as the 5′ and 3′ UTRs of WNV (data not shown). Use of smaller fragments for phylogenetic

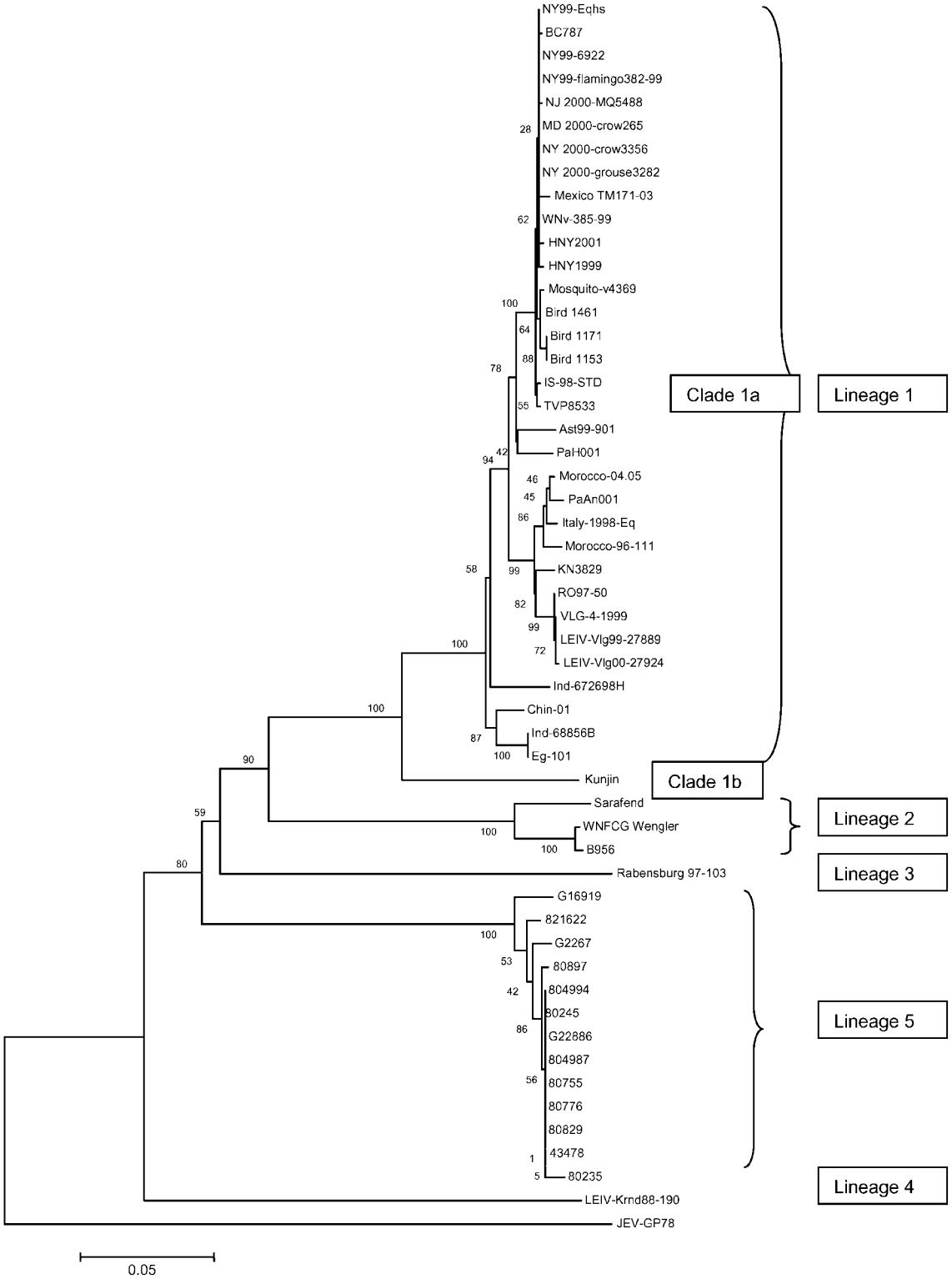
NY99-Eqhs
BC787
NY99-6922
NY99-flamingo382-99
NJ 2000-MQ5488
MD 2000-crow265
NY 2000-crow3356
NY 2000-grouse3282
Mexico TM171-03
WNv-385-99
HNY2001
HNY1999
Mosquito-v4369
Bird 1461
Bird 1171
Bird 1153
IS-98-STD
TVP8533
Ast99-901
PaH001
Morocco-04.05
PaAn001
Italy-1998-Eq
Morocco-96-111
KN3829
RO97-50
VLG-4-1999
LEIV-Vlg99-27889
LEIV-Vlg00-27924
Ind-672698H
Chin-01
Ind-68856B
Eg-101
Kunjin
Sarafend
WNFCG Wengler
B956
Rabensburg 97-103
G16919
821622
G2267
80897
804994
80245
G22886
804987
80755
80776
80829
43478
80235
LEIV-Krnd88-190
JEV-GP78
0.05
Clade 1a
Clade 1b
Lineage 1
Lineage 2
Lineage 3
Lineage 5
Lineage 4

**Fig. 1.** Phylogenetic tree based on the 921 nt sequence within the C–prM–E region of Indian WNV. JEV was used as an outgroup. The tree was constructed with the program MEGA by neighbour-joining with Jukes–Cantor parameter distances (bar). Bootstrap confidence level (1000 replicates) and a confidence probability value based on the standard error test were calculated using MEGA. GenBank accession numbers for the complete genomic sequences included in the phylogenetic analysis are: WNFCG Wengler (human, Uganda, 1937; M12294), B956 (human, Uganda, 1937; AY532665), Eg-101 (human, Egypt, 1951; AF260968), RO97-50 (*Culex pipiens*, Romania, 1996; AF260969), Morocco-96-111 (equine, Morocco, 1996; AY701412), Rabensburg isolate 97-103 (*Culex pipiens*, Czech Republic, 1997; AY765264), IS-98-STD (stork, Israel, 1998; AF481864), Italy-1998-Eq (equine, Italy, 1998; AF404757), LEIV-Krnd88-190 (*Dermacentor marginatus*, Russia, 1998; AY277251), KN3829 (*Culex univittatus*, Kenya, 1998; AY262283), NY99-Eqhs (equine, New York, 1999; AF260967), NY99-6922 (AB185915), VLG-4-1999 (human, Russia, 1999; AF317203), HNY1999 (human, New York, 1999; AF202541), NY99-flamingo382-99 (flamingo, New York, 1999; AF196835), Ast99-901 (human, Russia, 1999; AY278441), LEIV-Vlg99-27889 (human, Russia, 1999; AY277252), WNv-385-99 (dead snowy owl, USA, 1999; AY842931), NY 2000-crow3356 (crow, New York, 2000; AF404756), NY 2000-grouse3282 (ruffed grouse, New York, 2000; AF404755), NJ 2000-MQ5488 (*Culex pipiens*, New Jersey, 2000; AF404754), MD 2000-crow265 (crow, Maryland, 2000; AF404753), LEIV-Vlg00-27924 (human, Russia, 2000; AY278442), HNY2001 (human, New York, 2001; AF533540), TVP8533 (human, Texas, 2002; AY289214), Morocco-04.05 (equine, Morocco, 2003; AY701413), Sarafend (Israel; AY688948), Mexico TM171-03 (Mexico; AY660002), BC787 (AB185917), Mosquito-v4369 (*Culex quinquefasciatus*, USA; AY712948), Bird 1461 (USA; AY712947), Bird 1171 (bluejay; USA-AY712946), Bird 1153 (mourning dove, USA; AY712945), Chin-01 (China; AY490240), PaH001 (France; AY268133), PaAn001 (France-AY268132), Kunjin (strain MRM61C, *Culex annulirostris*, MRM Australia, 1980; D00246), JEV-GP78 (India; AF075723).

analysis gave similar results, thereby indicating their utility for analysis of strain variations in the case of WNV, especially if direct amplification from the original specimen is desirable.

Co-circulation of WNV strains belonging to two distinct genetic lineages in southern India is noteworthy. Although the majority of the isolates grouped within the new genetic lineage, two isolates, 672698 isolated in 1967 from a human patient and 68856 isolated in 1968 from *R. leschenaulti*, grouped with lineage 1 isolates. Probable reasons for the restriction of lineage 1 strain to 1967–1968 and its failure to establish transmission cycles in India remain unclear. A serological survey of 859 bird sera in the JEV endemic area of southern India indicate JEV and WNV antibody positivity, indicating a probable role of birds in the maintenance of these viruses in nature (Jamgaonkar *et al.*, 2003). However, the exact role of the birds in maintenance of WNV in other parts of India and the extent of their migration is not clear. Interestingly, isolate 68856 showed 100 % sequence similarity with Eg-101 in the 921 nt fragment encompassing partial C–prM–E regions, clearly indicating the need for further investigations of the bat isolate. However, it may be noted that, using a 182 nt sequence of the NS3 region, Porter *et al.* (1993) reported 96 % sequence similarity between Eg-101 and 68856 sequences. Further evidence of differences between 68856 and Eg-101 comes from the different biological characteristics of these viruses. The NIV had reported the isolation of WNV P-4230 from a laboratory-infected person in 1956 while working with G22886 and Eg-101 (Work, 1971). Limited sequence analysis of P-4230 showed 100 % sequence similarity with Eg-101 (not included in this study), indicating that the laboratory infection was due to Eg-101. Comparison of biological characteristics and pathogenesis in mice of several Indian isolates with P-4230 and 68856 revealed marked differences, indicating that the bat isolate and Eg-101 are in fact different strains (Umrigar & Pavri, 1977).

The geographical highlands almost at the centre of India serve as a geographical subcontinental divide. To the west of the divide, there is evidence only of WNV, whereas to the east of the divide, WNV and JEV intermingle in the southern part of the Indian Peninsula (Work, 1971). Thus, India forms a unique place where several flaviviruses like JEV, WNV, KFDV and dengue viruses (all four types) co-exist. This perhaps reflects the evolutionary significance of the region. The distinct nature of Russian and African WNVs from Indian viruses indicates that migratory birds have not played any role in the distribution of these viruses in India. Antigenic similarities between WNV strains from African countries, France, Israel, the former Soviet Union and Pakistan have been attributed to bird-mediated virus circulation between these countries, whereas southern Indian strains of WNV form a separate, geographically restricted antigenic group due to the lack of bird migratory routes linking southern India with the Middle East and Africa (Peiris & Amerasinghe, 1994).

Comparative virulence studies of lineage 1 and Indian viruses indicate that the Indian viruses may have evolved as low-virulent strains (Hammam & Price, 1966; Price & O'Learly, 1967; Gaidamovich & Sokhe, 1973; Odelola & Fabiyi, 1976; Umrigar & Pavri, 1977; Mathiot *et al.*, 1990; Beasley *et al.*, 2002). Similarly, the cross-neutralization activity among lineage 1 strain Eg-101 and Indian viruses 68856, G22886, 821622 and 804994 clearly showed the significant difference in the antigenic nature of these viruses. Eg-101 and the Indian strain 68856, which clustered with lineage 1 viruses, showed marked differences in cross-neutralization activity with other three Indian strains (Table 2). Thus, the genetic variations within lineage 1 and Indian viruses are also reflected in the antigenic and

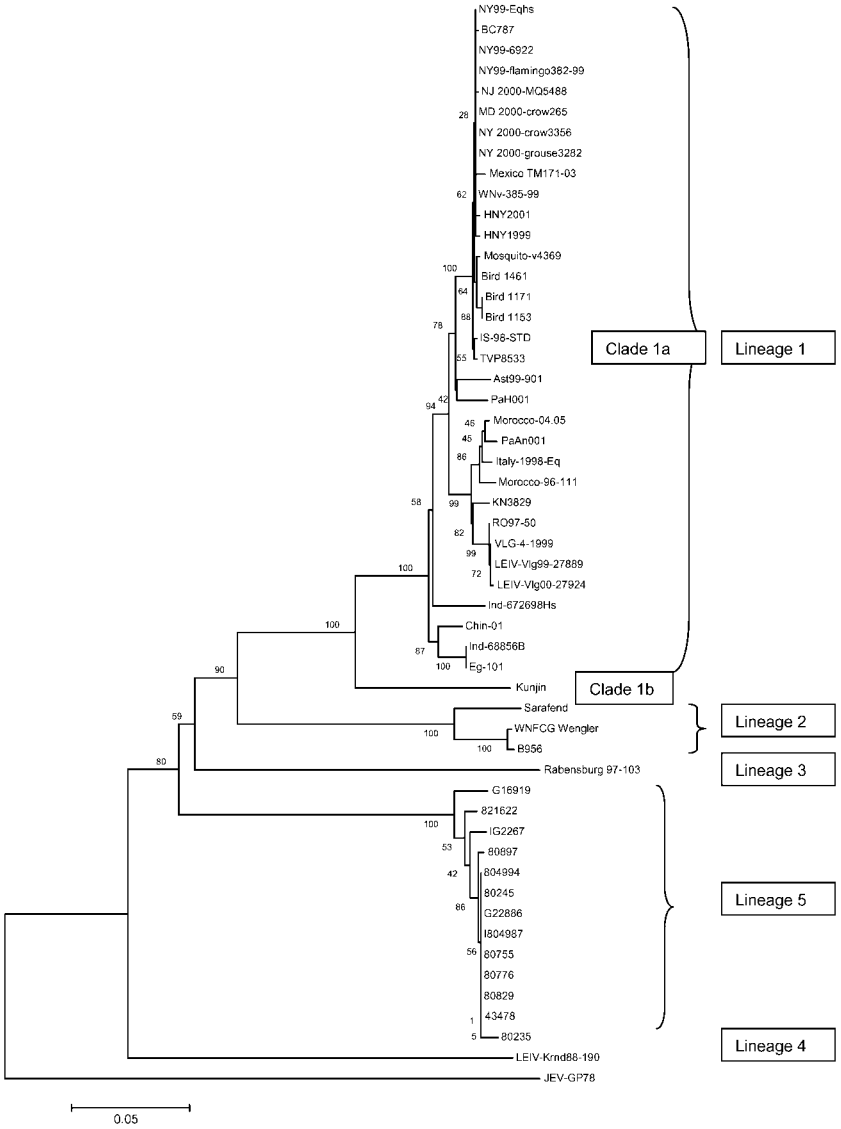


**Fig. 2.** Phylogenetic tree based on the complete genomic sequence of Indian WNV along with sequences listed in Fig. 1; GenBank accession numbers for WNV strains B-SP and 6-LP are AB185916 and AB185914, respectively. The tree was constructed with the program MEGA by neighbour-joining with Jukes–Cantor parameter distances (scale bar). Bootstrap confidence level (1000 replicates) and a confidence probability value based on the standard error test were calculated using MEGA.

**Table 4.** NI analysis based on the complete sequence of WNVs from five different genetic lineages formed in this study along with JEV

The mean NI ± SEM (%) among different genetic lineages along with JEV is shown. *n*, Number of complete genomic sequences used for determination of mean NI; ND, not done; NA, not available.

| Viruses from lineage: | 1 (clade 1a) | 1 (clade 1b) | 2 | 3 | 4 | 5 | JEV-GP78 |
|---|---|---|---|---|---|---|---|
| 1 (clade 1a) ($n=31$) | 97.68 ± 0.10 | | | | | | |
| 1 (clade 1b) ($n=1$) | 88.52 ± 0.30 | ND | | | | | |
| 2 ($n=3$) | 79.75 ± 0.38 | 79.61 ± 0.38 | 96.79 ± 0.14 | | | | |
| 3 ($n=1$) | 77.55 ± 0.40 | 77.21 ± 0.40 | 77.94 ± 0.40 | NA | | | |
| 4 ($n=1$) | 75.84 ± 0.41 | 75.77 ± 0.41 | 75.96 ± 0.40 | 75.55 ± 0.42 | NA | | |
| 5 ($n=1$) | 79.88 ± 0.38 | 79.45 ± 0.39 | 78.69 ± 0.39 | 76.24 ± 0.41 | 75.24 ± 0.42 | NA | |
| JEV ($n=1$) | 69.46 ± 0.44 | 68.52 ± 0.44 | 69.80 ± 0.44 | 68.58 ± 0.44 | 68.18 ± 0.45 | 69.13 ± 0.44 | ND |

virulence activities of these viruses. Cross-neutralization using monoclonal antibodies against lineage 1 and Indian viruses will prove useful for further analysis of these viruses.

Fifteen Indian viruses isolated over 27 years from different hosts were selected to determine the impact of host species and time on evolutionary relationships. Partial sequence analysis of eight viruses isolated from different mosquito species and one human patient during 1980 did not show notable differences at the genomic level. Similarly, partial genomic sequence analysis of 13 Indian viruses that formed a distinct genetic lineage did not show notable differences with respect to time. The virus isolates of 1958, 1980 (four isolates) and 1981 had similar nucleotide sequences, whereas the isolates of 1955, 1957, 1980 (two isolates) and 1982 showed 0.5–3.5 % divergence from these isolates. Results suggest that WNV has not evolved significantly in India over the last 27 years. With the recent experience of WNV in America, special attempts are necessary to understand the current status of WNV strains circulating in India. Data from other parts of the country will need to be generated as well.

In conclusion, WNV can be grouped into five distinct genetic lineages that differ from each other by 20.0–24.7 %. The Indian viruses form lineage 5. As the last WNV available with NIV was isolated in 1982, it will be interesting to study the impact of time on evolution of WNV by characterizing currently prevalent strains in India.

## ACKNOWLEDGEMENTS

The authors are thankful to Dr K. Ray, Mr G. N. Sapakal, Rajen Lakra and Satish Ranawade for their help. Financial support by Indian Council of Medical Research to V. P. B. (project no. 02-21) is greatly appreciated.

## REFERENCES

**Anderson, J. F., Vossbrinck, C. R., Andreadis, T. G., Iton, A., Beckwith, W. I., III & Mayo, D. R. (2001).** A phylogenetic approach to following West Nile virus in Connecticut. *Proc Natl Acad Sci U S A* **98**, 12885–12889.

**Bakonyi, T., Hubálek, Z., Rudolf, I. & Nowotny, N. (2005).** Novel flavivirus or new lineage of West Nile virus, Central Europe. *Emerg Infect Dis* **11**, 225–231.

**Banerjee, K. (1996).** Emerging viral infections with special reference to India. *Indian J Med Res* **103**, 177–200.

**Banerjee, K., Gupta, N. P. & Govardhan, M. K. (1979).** Viral infections in laboratory personnel. *Indian J Med Res* **69**, 363–378.

**Banker, D. D. (1952).** Preliminary observations on antibody patterns against certain viruses among inhabitants of Bombay City. *Indian J Med Sci* **6**, 733–746.

**Beasley, D. W., Li, L., Suderman, M. T. & Barrett, A. D. T. (2002).** Mouse neuroinvasive phenotype of West Nile virus strains varies depending upon virus genotype. *Virology* **296**, 17–23.

**Berthet, F.-X., Zeller, H. G., Drouet, M. T., Rauzier, J., Digoutte, J. P. & Deubel, V. (1997).** Extensive nucleotide changes and deletions within the envelope glycoprotein gene of Euro-African West Nile viruses. *J Gen Virol* **78**, 2293–2297.

**Burt, F. J., Grobbelaar, A. F., Leman, P. A., Anthony, F. S., Gibson, G. V. F. & Swanepoel, R. (2002).** Phylogenetic relationships of Southern African West Nile virus isolates. *Emerg Infect Dis* **8**, 820–826.

**Carey, D. E., Rodrigues, F. M., Myers, R. M. & Webb, J. K. G. (1968).** Arthropod borne viral infections in children in Vellore, South India, with reference to dengue and West Nile viruses. *Indian Pediatr* **5**, 285–296.

**Centers for Disease Control and Prevention (1999).** Outbreak of West Nile-like viral encephalitis – 1999. *MMWR Morb Mortal Wkly Rep* **48**, 845–849.

**Damle, R. G., Yeolekar, L. R. & Rao, B. L. (1998).** Strain analysis and epitope mapping of West Nile virus using monoclonal antibodies. *Acta Virol* **42**, 389–395.

**Gaidamovich, S. Y. & Sokhe, J. (1973).** Studies on antigenic peculiarities of West Nile virus strains isolated in the USSR by three serological tests. *Acta Virol* **17**, 343.

**George, S., Gourie-Devi, M., Rao, J. A., Prasad, S. R. & Pavri, K. M. (1984).** Isolation of West Nile virus from the brains of children who had died of encephalitis. *Bull World Health Organ* **62**, 879–882.

**Gore, M. M., Gupta, A. K., Ayachit, V. M., Athawale, S. S., Ghosh, S. N. & Banerjee, K. (1990).** Selection of neutralization escape variant strains of Japanese encephalitis virus using monoclonal antibody. *Indian J Med Res* **91**, 231–233.

**Hammam, H. M. & Price, W. H. (1966).** Further observation on geographic variations in the antigenic character of West Nile and Japanese B encephalitis viruses. *Am J Epidemiol* **83**, 113.

**Hubalek, Z. & Halouzka, J. (1999).** West Nile fever – a reemerging mosquito-borne viral disease in Europe. *Emerg Infect Dis* **5**, 643–650.

**Jamgaonkar, A. V., Yergolkar, P. N., Geevarghese, G., Joshi, G. D., Joshi, M. V. & Mishra, A. C. (2003).** Serological evidence for Japanese encephalitis virus and West Nile virus infections in water frequenting and terrestrial wild birds in Kolar District, Karnataka State, India. A retrospective study. *Acta Virol* **47**, 185–188.

**Jia, X.-Y., Briest, T., Jordon, I., Rambout, A., Chi, H. C., MacKenzie, J. S., Hall, R. A., Scherret, J. & Lipkin, W. I. (1999).** Genomic analysis of the West Nile New York 1999 encephalitis virus. *Lancet* **354**, 1971–1972.

**Kedarnath, N., Prasad, S. R., Dandwate, C. N., Koshy, A. A., George, S. & Ghosh, S. N. (1984).** Isolation of Japanese encephalitis and West Nile viruses from peripheral blood of encephalitis patients. *Indian J Med Res* **79**, 1–7.

**Kumar, S., Tamura, K., Jakobsen, I. B. & Nei, M. (2001).** MEGA2: Molecular evolutionary genetics analysis software. *Bioinformatics* **17**, 1244–1245.

**Lanciotti, R. S., Roehrig, J. T., Deubel, V., Smith, J., Parker, M., Steele, K., Volpe, K. E., Crabtree, M. B., Scherret, J. H. & other authors (1999).** Origin of the West Nile virus responsible for an outbreak of encephalitis in the Northern US. *Science* **286**, 2333–2337.

**Lanciotti, R. S., Ebel, G. D., Deubel, V., Kerst, A. J., Murri, S., Meyer, R., Bowen, M., McKinney, M., Morrill, W. E. & other authors (2002).** Complete genome sequence and phylogenetic analysis of West Nile virus strains isolated from the United States, Europe, and the Middle East. *Virology* **298**, 96–105.

**Lvov, D. K., Butenko, A. M., Gromashevsky, V. L., Kovtunov, A. I., Prilipov, A. G., Kinney, R., Aristova, V. A., Dzharkenov, A. F., Samokhvalov, E. I. & other authors (2004).** West Nile virus and other zoonotic viruses in Russia: examples of emerging-reemerging situations. *Arch Virol Suppl* **18**, 85–96.

**Mathiot, C. C., Georges, A. J. & Deubel, V. (1990).** Comparative analysis of West Nile virus strains isolated from human and animal hosts, using monoclonal antibodies and cDNA restriction digest profiles. *Res Virol* **141**, 533–543.

**Odelola, H. A. & Fabiyi, A. (1976).** Antigenic relationships among Nigerian strains of West Nile virus by complement fixation and agar gel precipitation techniques. *Trans R Soc Trop Med Hyg* **70**, 138.

**Paul, S. D., Narasimha Murthy, D. P. & Das, M. (1970a).** Isolation of West Nile virus from a human case of febrile illness. *Indian J Med Res* **58**, 1177–1179.

**Paul, S. D., Rajgopalan, P. K. & Sreenivasan, M. A. (1970b).** Isolation of the West Nile virus from the frugivorous bat, *Rousettus leschenaulti*. *Indian J Med Res* **58**, 1169–1171.

**Petersen, L. R. & Roehrig, J. T. (2001).** West Nile virus: a reemerging global pathogen. *Emerg Infect Dis* **7**, 611–614.

**Peiris, J. S. M. & Amerasinghe, F. P. (1994).** West Nile fever. In *Handbook of Zoonoses: Section B: Viral*, 2nd edn, pp. 139–148. Edited by G. W. Beran & J. H. Steele. Boca Raton, FL: CRC Press.

**Platonov, A. E., Shipulin, G. A., Shipulina, A. Y., Tyutyunnik, E. N., Frolochkina, T. I., Lanciotti, R. S., Yazishina, S., Platonova, O. V., Obukov, I. L. & other authors (2001).** Outbreak of West Nile virus infection, Volgograd region, Russia, 1999. *Emerg Infect Dis* **7**, 128–132.

**Porter, K. R., Summers, P. L., Dubois, D., Puri, B., Nelson, W., Henchal, E., Oprandy, J. J. & Hayes, C. G. (1993).** Detection of West Nile virus by polymerase chain reaction and analysis of the nucleotide sequence variation. *Am J Trop Med Hyg* **48**, 440–446.

**Price, W. H. & O'Learly, W. (1967).** Geographic variation in the antigenic character of West Nile viruses. *Am J Epidemiol* **85**, 84–86.

**Prilipov, A. G., Kinney, R. M., Samokhvalov, E. I., Savage, H. M., Al'khovskii, S. V., Tsuchiya, K. R., Gromashevskii, V. L., Sadykova, G. K., Shatalov, A. G. & other authors (2002).** Analysis of new variants of West Nile fever virus. *Vopr Virusol* **47**, 36–41 (in Russian).

**Rodrigues, F. M., Bright Singh, P., Dandawate, C. N., Soman, R. S., Guttikar, S. N. & Kaul, H. N. (1980).** Isolation of Japanese encephalitis and West Nile viruses from mosquitoes collected in Andhra Pradesh. *Indian J Parasitol* **4**, 149–153.

**Savage, H. M., Celanu, C., Nicolescu, G., Karabatsos, N., Lanciotti, R., Vladimirescu, A., Lair, R., Ungureanu, A., Romanca, C. & Tsai, T. F. (1999).** Entomologic and avian investigations of an epidemic of West Nile fever in Romania in 1996, with serologic and molecular characterization of a virus isolate from mosquitoes. *Am J Trop Med Hyg* **61**, 600–611.

**Scherret, J. H., Poidinger, M., MacKenzie, J. S., Broom, A. K., Deubel, V., Lipkin, W. I., Briest, T. & Hall, R. A. (2001).** Studies on the relationship between West Nile and Kunjin viruses. *Emerg Infect Dis* **7**, 697–705.

**Smithburn, K. C., Hughes, T. P., Burke, A. W. & Paul, J. H. (1940).** A neurotropic virus isolated from the blood of a native of Uganda. *Am J Trop Med Hyg* **20**, 471–492.

**Thakare, J. P., Rao, T. L. & Padbidri, V. S. (2002).** Prevalence of West Nile virus infection in India. *Southeast Asian J Trop Med Public Health* **33**, 801–805.

**Tsai, T. F., Popovici, F., Cernescu, C., Campbell, G. L. & Nedelcu, N. I. (1998).** West Nile encephalitis epidemic in southern Romania. *Lancet* **352**, 767–771.

**Umrigar, M. D. & Pavri, K. M. (1977).** Comparative serological studies on Indian strains of West Nile virus isolated from different sources. *Indian J Med Res* **65**, 603.

**Work, T. H. (1971).** On the Japanese B-West Nile virus complex or an arbovirus problem of six continents. *Am J Trop Med Hyg* **20**, 169–186.