Genetic analysis of West Nile New York 1999 encephalitis virus

Xi-Yu Jia, Thomas Briese, Ingo Jordan, Andrew Rambaut, Han Chang Chi, John S Mackenzie, Roy A Hall, Jacqui Scherret, W Ian Lipkin

Analysis of the genome of the flavivirus responsible for the 1999 New York City encephalitis epidemic cloned from human brain by reverse-transcription polymerase chain reaction indicates its identity as a lineage I West Nile virus (WNV; WNV-NY1999) closely related to WNVs previously isolated in the Middle East.

The *Flaviviridae* include human pathogens such as yellow fever, dengue, Japanese encephalitis, St Louis encephalitis, Kunjin (KUNV), and West Nile (WNV) viruses. WNV infections are often mild, but may result in encephalitis. WNV is transmitted by mosquitoes. Birds are maintenance hosts and may carry WNV. Epidemics have been reported in Israel, France, South Africa, and Romania where an outbreak near Bucharest in 1996-97 had an up to 5% mortality.[1] WNV infection was not recognised in the Americas until August, 1999, when there was an outbreak of encephalitis in New York City.[2] WNV-specific immunological and molecular reagents confirmed infection in 59 human beings; 103 birds; 12 horses; and in *Culex pipiens*, other *Culex* species, and *Aedes vexans* mosquitoes. The distribution of isolates indicated that the zone of infection extended beyond New York City to surrounding counties, as well as to Connecticut and Maryland. We cloned the genome of the New York WNV (WNV-NY1999) from necropsy human brain samples by RT-PCR. Due to the autolytic interval, RNA integrity was reduced: 17 overlapping amplification products were required to assemble a WNV-NY1999 virus genome sequence comprising 10 945 nucleotides.

Initial database analysis of WNV-NY1999 revealed sequence similarity to WNV-Wengler (Nigeria) and KUNV-MRM61c (Australia).[2] After cloning the WNV-NY1999 genome, sequences representing the E, NS3, NS5, and 3′-UTR regions were compared with published and unpublished flavivirus sequences. Subtypes of WNV are distinguished by antigenic variations in the E (envelope) protein and the presence of an N-glycosylation site (Asn-Tyr-Ser) at aminoacids 154-156.[3,4] Two lineages of WNVs are proposed based on signature aminoacid motifs: lineage I includes KUNV as well as WNVs from Europe; the Middle East; and North, Central and West Africa; lineage II includes WNVs from West, Central, and East Africa, and from Madagascar. Deduced WNV-NY1999 E aminoacid sequence showed an intact N-glycosylation site and the presence of lineage I signature motifs (ALa_{172}, Asn_{199}, Thr_{205}, Thr_{208}, and Thr_{210}).[4] Phylogenetic analysis of E nucleotide sequence also indicated membership in lineage I (figure 1A). Alignment of available NS3 and NS5 sequences confirmed assignment of WNV-NY1999 to the WNV and KUNV genotype lineage I defined by Berthet et al[4] based on E gene sequence (figure 1B, C). Alignment of 3′-UTR sequences, which are divergent immediately downstream of the polyprotein stop codon,[5] showed conservation between WNV-NY1999 and WNV-EGY-Eg101 (table). High levels of E nucleotide sequence conservation were also found between WN-NY1999 and WNV-EGY-Eg101,

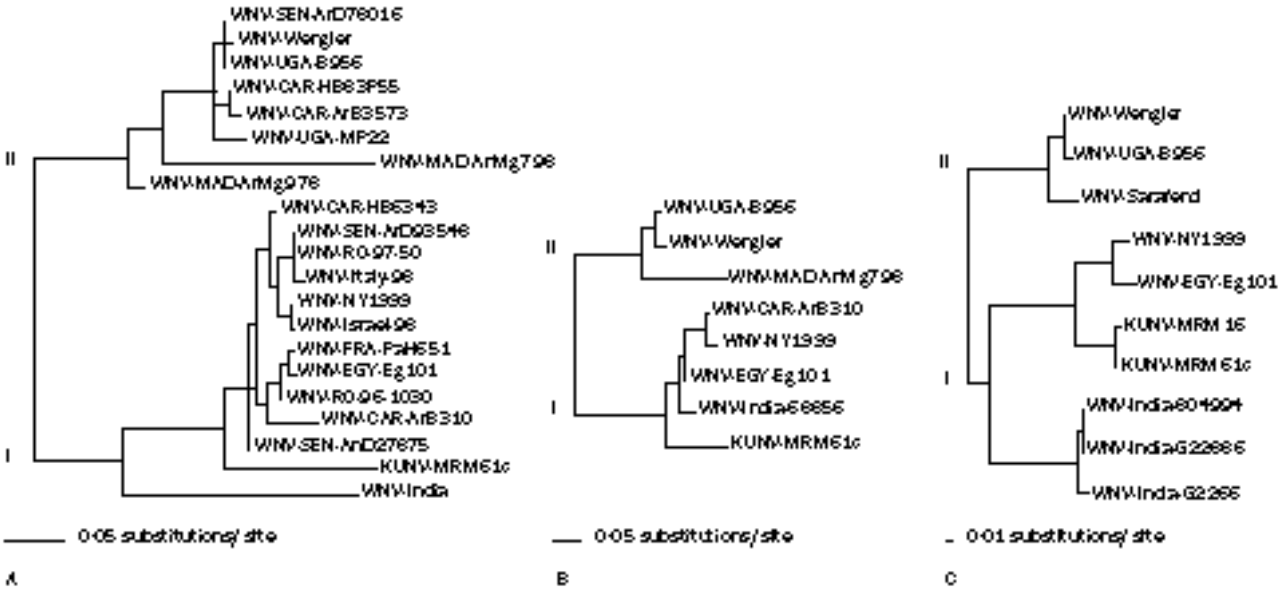

Figure 1: **Analysis of WNV-NY1999 E, NS3, NS5, and 3′-UTR sequences**

A, phylogenetic tree based on 227 nucleotides of E region sequence indicating membership of WNV-NY1999 in lineage I. B, phylogenetic tree based on 182 nucleotides of NS3 sequence indicating the relationship of WNV-NY1999 to lineage I viruses WNV-CAR-ArB310 and WNV-EGY-Eg101. C, phylogenetic tree based on 236 nucleotides of NS5 sequence indicating the relationship of WNV-NY1999 to lineage I viruses WNV-EGY-Eg101 and KUN MRM61c.

Viral sequences used in figures 1 and 2: WNV-Wengler, GenBank #M12294, KUNV-MRM61c, #D00246. Envelope: WNV-UGA-B956, this study WNV-UGA-MP22, #AF001562; WNV-MAD-ArMg978, #AF001574; WNV-MAD-AnMg798, #AF001559; WNV-SEN-ArD78016, #AF001556; WNV-SEN-AnD27875, #AF001569; WNV-SEN-ArD93548, #AF001570; WNV-CAR-HB83P55, #AF001557; WNV-CAR-ArB3573, #AF001565; WNV-CAR-HB6343, #AF001558; WNV-CAR-ArB310, #AF001566; WNV-EGY-Eg101, #AF001568; WNV-FRA-PaH651, #AF001560; WNV-RO-96-1030, #AF130363 WNV-RO-97-50, #AF130362; WNV-Romania-96, WNV-Italy–98, and WNV-Israel-98, WNV-India: V Deubel, unpublished. NS3:WNV-UGA-B956 WNV-MAD–AnMg798 (Dak-MG798), WNV-CAR-ArB310 (Dak-B310), WNV-EGY-Eg101 (E101), WNV-India-68856: Porter KR, et al. *Am J Trop Med Hyg* 1993; **48:** 440–46. NS5 and 3′-UTR: KUNV-MRM16, #L48979; WNV-UGA-B956, this study (#AF208017); WNV-EGY-Eg101, #AF017254 WNV-Sarafend, #L48977; WNV-India-804994, JSM, RAH, and JS, unpublished (#AF196540); WNV-India-G2266, JSM, RAH, and JS, unpublished (#AF196537); WNV-India-G22886, JSM, RAH, and JS, unpublished (#AF196538).

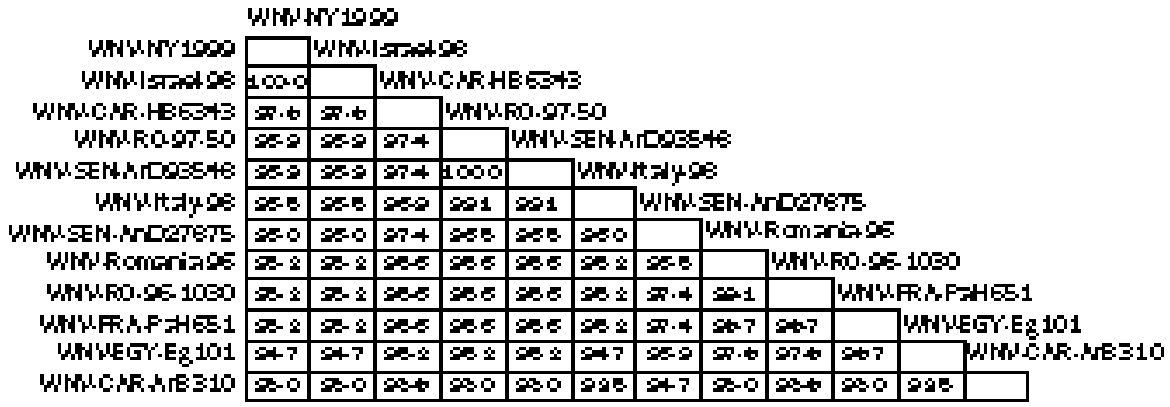

Figure 2: **Nucleotide conservation among indicated WNV E region sequences**
Percent nucleotide identity was calculated based on 227 nucleotides of E region sequence. Note 100% sequence identity between WN-NY1999 and WNV-Israel-98.

as well as other strains for which no 3′-UTR sequence is available, including WNV-Israel-98 (isolated 1998). Indeed, sequences of WNV-NY1999 and WNV-Israel-98 were identical for the 227 nucleotides available for analysis (figure 2).

Recent outbreaks of WNV infections in Italy, Czechland, Romania, Russia, and New York City suggest maintenance hosts and arthropod vectors for WNVs are widely distributed.[1] It is possible that WNVs were present in the Americas before 1999; however, the high mortality associated with infection of native corvids is more consistent with recent introduction. Similarities in sequence between WNV-NY1999 and WNV-EGY-Eg101 and WNV-Israel-98 may suggest a Middle East origin of WNV-NY1999. Potential routes for introduction of this virus to the Eastern USA include importation of infected birds, mosquitoes, or viraemic human beings. The area within New York City where WNV was prevalent includes two international airports.

Recent outbreaks of WNVs in Europe, Asia, and North America confirm their significance as emerging infectious agents and underscores the importance of global surveillance.

We thank C Calisher for advice; M A Brinton and V Deubel for access to unpublished sequences of WNVs; J Ebert, W Glass, and P Sutherland for technical assistance; R E Shope and R B Tesh for WNV RNA; and C Huang and L J Grady for necropsy brain specimens. Funding was provided by NIH grant NS29425; sequence data of Indian WNV strains were obtained under a grant from the National Health and Medical Research Council (Australia).

The assembled WNV-NY1999 sequence is submitted to GenBank under accession number AF202541. Protocol details may be found at:
http://www.ucihs.uci.edu/departments/neurovirology/

1 Hubalék Z, Halouzka J. West Nile fever: a reemerging mosquito-borne viral disease in Europe. *Emerg Inf Dis* 1999; **5:** 643–50.
2 Briese T, Jia XY, Huang C, Grady LJ, Lipkin WI. Identification of a Kunjin/West Nile-like flavivirus in brains of patients with New York encephalitis. *Lancet* 1999; **354:** 1261–62.
3 Calisher CH, Karabatsos N, Dalrymple JM, et al. Antigenic relationships between flaviviruses as determined by cross-neutralization tests with polyclonal antisera. *J Gen Virol* 1989; **70:** 37–43.
4 Berthet FX, Zeller HG, Drouet MT, Rauzier J, Digoutte JP, Deubel V. Extensive nucleotide changes and deletions within the envelope glycoprotein gene of Euro-African West Nile viruses. *J Gen Virol* 1997; **78:** 2293–97.
5 Poidinger M, Hall RA, Mackenzie JS. Molecular characterization of the Japanese encephalitis serocomplex of the flavivirus genus. *Virology* 1996; **218:** 417–21.

Emerging Diseases Laboratory, Departments of Microbiology and Molecular Genetics, Anatomy and Neurobiology, Neurology, University of California, Irvine, CA 92697-4292, USA (Yi-Yu Jia MD, T Briese PhD, I Jordan PhD, Prof W I Lipkin MD); **Department of Zoology, University of Oxford, South Parks Road, Oxford, UK** (A Rambaut DPhil); **Department of Biological Chemistry, University of California, Irvine, CA** (H C Chi PhD); **Department of Microbiology and Parasitology, University of Queensland, Brisbane, Australia** (J S Mackenzie PhD, R A Hall PhD, J Scherret MSc)

Correspondence to: Prof W Ian Lipkin (e-mail: ilipkin@uci.edu)

KUNV	MRM16	ACAGTATTGT	**AAA**TACTTTG	TTAATTGTAA	ATAAATAT- -	----------	--------TG	TTATTATGTG	TAGAAGTTTA
KUNV	MRM61c	ACAGTATTGT	**AAA**TACTTTG	TTAATTGTAA	ATAAATAT- -	----------	--------TG	TTATTATGTG	TAGAAGTTTA
WNV	EGY-Eg101	ACAGTACTGT	**AAA**TACTTTA	TTAATTGTAA	ATAGACAA- -	----------	--------TG	TAAGCATGTG	TAAAAGTATA
			. . . . .	. . . . .	. . . . .		.	. . . . .	. . . . .
WNV	**NY1999**	ACAGTACTGT	**A**GATATTTAA	TCAATTGTAA	ATAGACAA- -	----------	--------TA	TAAGTATGCA	TAAAAGTGTA
WNV	India-804994	TCAGTGTTGT	**AAA**TAGTAAC	AGTTAA- - - -	----------	----------	---------GG	TATGTGTATA	GATTAGTGTT
WNV	Wengler	ACTGTTTTGT	**AA**- - - - - - - - -	----------	----------	----------	----------	----------	----------
WNV	UGA-B956	ACTGTTTTGT	**AA**AAAATAAA	GCTGTATTGA	GTAGTTGTAT	AGTTGTAGTG	TTCATAGCAA	TTTGAATTAT	GATTAATTAT
KUNV	MRM16	GCTTTGTAAT	AGTGTTTAGT	G- - - - - - TG	TTTAGAGTTA	GGAAAATTTT	AGT-GAGGAA	GTCAGGCCGG	AAAATTCCCG
KUNV	MRM61c	GCTTTATAAT	AGTGTTTAGT	G- - - - - - TG	TTTAGAGTTA	GAAAAATTTT	AGT-GAGGAA	GTCAGGCCGG	AAAATTCCCG
WNV	EGY-Eg101	GTTTTATAGT	AGCATTTAGT	GATGTTAGTG	TAAATGGTTA	AGAAAATTTT	AAG-GAGGAA	GTCAGGCCGG	AAAGTTTCCG
		. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .
WNV	**NY1999**	GTTTTATAGT	AGTATTTAGT	GGTGTTAGTG	TAAATAGTTA	AGAAAATTTT	GAG-GAGAAA	GTCAGGCCGG	GAAGTTCCCG
WNV	India-804994	GT-AAATAGG	ATTAGCTAAA	GTATGCATGT	AGGTTAGTGT	TGAGAATTTT	GTTAGAGGAA	GTCAGGCCGG	AGCATTTCCG
WNV	Wengler	- - - - - - - - AA	GATAGTATTA	TAGTTAGTTT	AGTGTAAATA	GGA-TTTATT	GAGAATGGAA	GTCAGGCCAG	ATTAATGCTG
WNV	UGA-B956	TTAGGCTTAA	GATAGTACTA	TAGTTAGTTT	AGTGTAAATA	GGA-TTTATT	GAGAATGGAA	GTCAGGCCAG	ATTAATGCTG

A maximum of 148 nt following the polyprotein stop codon (bold, WN-UGA-B956) were aligned. Dashes indicate gaps; points indicate nucleotides conserved between WNV-NY1999 and WNV-EGY-Eg101.

Alignment of 3′-UTR sequences