

Phylogenetic and alignment analysis of West Nile Virus polyprotein amino acid sequence isolates



By: Dale Moskoff, Yifan Duan, Amelia Walsh, Sydney Berman, Micah Grubert Van Iderstine

Introduction

- West Nile Virus (WNV) was originally isolated in 1937, and has since spread globally and become an increasingly important pathogen for humans and animals³
- WNV is a mosquito-borne pathogen of the family *Flaviviridae*, genus *Flavivirus*²
- Humans are incidental hosts and don't normally show symptoms of the virus, but WNV epidemics with human deaths have been reported³
- The WNV genome is a single-stranded, positive-sense RNA encoding a polyprotein precursor that is then cleaved into 3 structural, 7 non-structural proteins²
 - The 3 structural proteins are the capsid (C) protein, precursor and membrane (prM/M) protein, and envelope (E) protein²
 - These proteins can be expected to be highly conserved among host species because they are critical for viral replication¹
- No human vaccine or therapeutic treatment has been developed for WNV²
 - Structural proteins represent viable targets for antiviral agents¹
- Recent phylogenetic analysis on nucleic acid sequence data from a portion of the E protein gene has highlighted two distinct lineages of the WNV³:
 - Lineage 1: globally distributed, virulent to humans and animals
 - Lineage 2: African distribution only, less virulent to humans

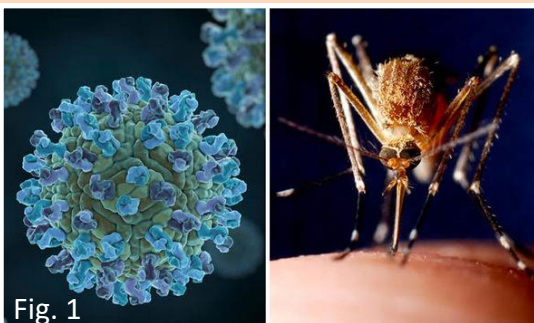


Fig. 1

Main Study Goals

- Identify highly conserved regions in structural proteins across various hosts and geographical regions.
- Investigate previous findings of two main WNV lineages.

Methods

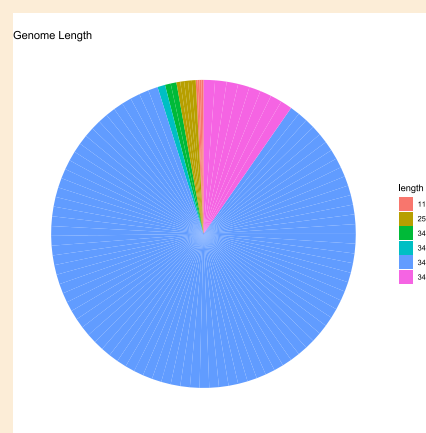


Fig. 2 Sequence Lengths

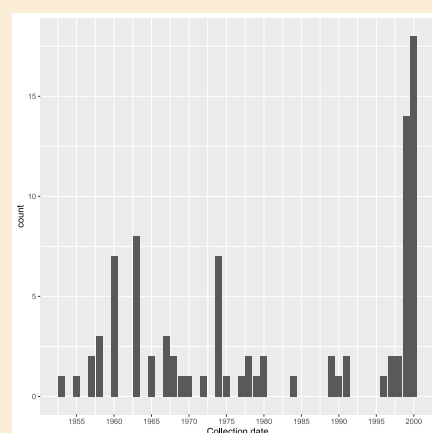


Fig. 3 WNV Sample Years

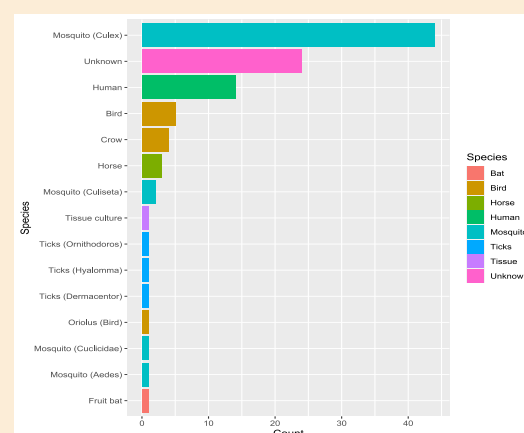


Fig. 4 WNV Vectors

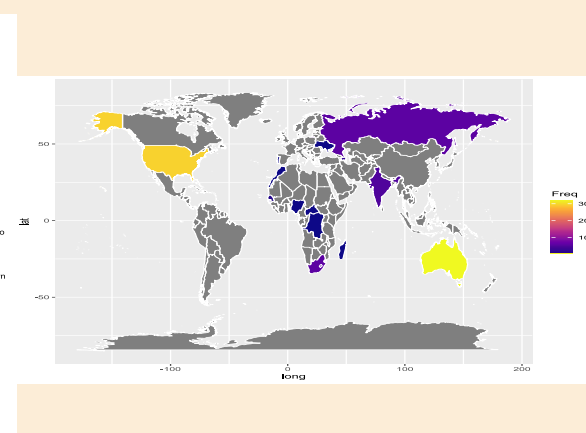


Fig. 5 WNV Sample Locations

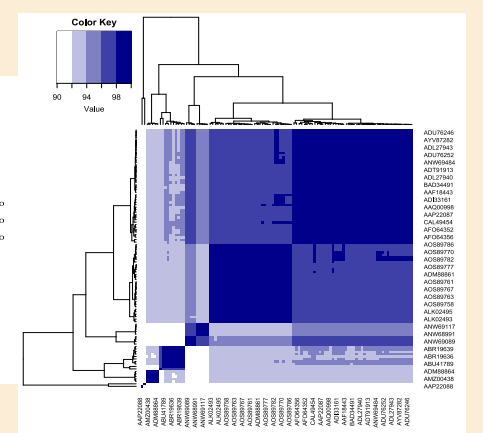


Fig. 6 Distance Matrix X Phylogeny

Main Findings - Phylogenies

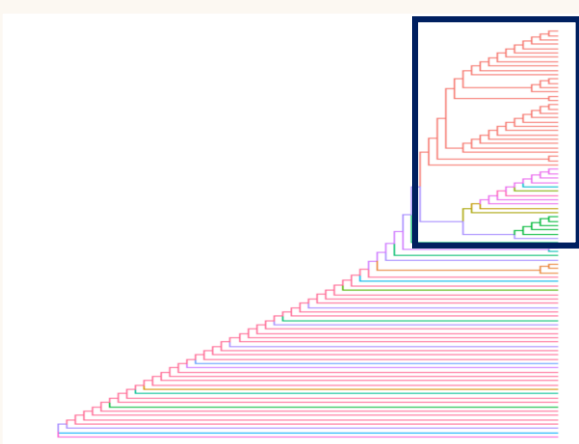


Fig. 7 C Protein Phylogeny

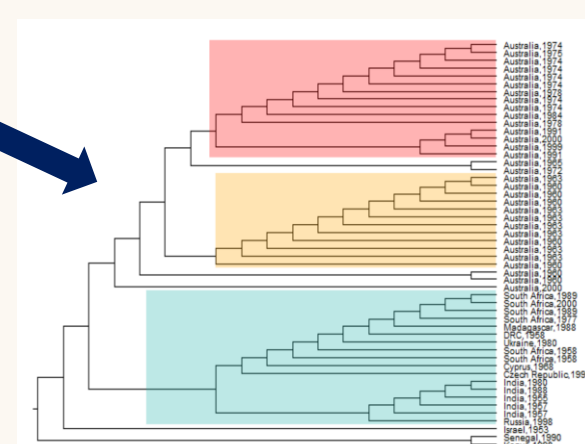


Fig. 8 Zoomed in C Protein Phylogeny

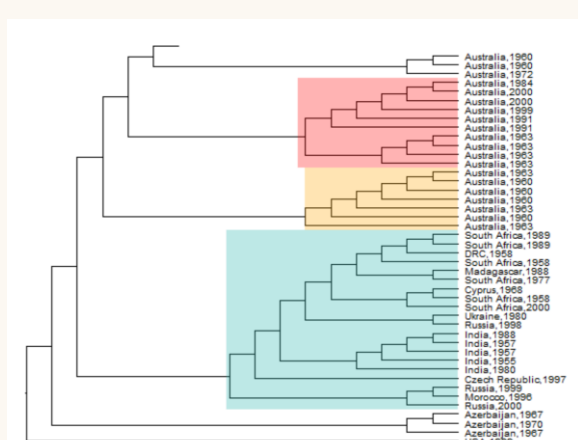


Fig. 9 Zoomed in E Protein Phylogeny

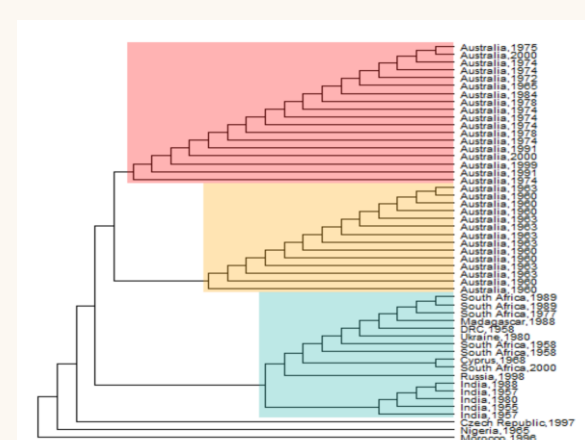
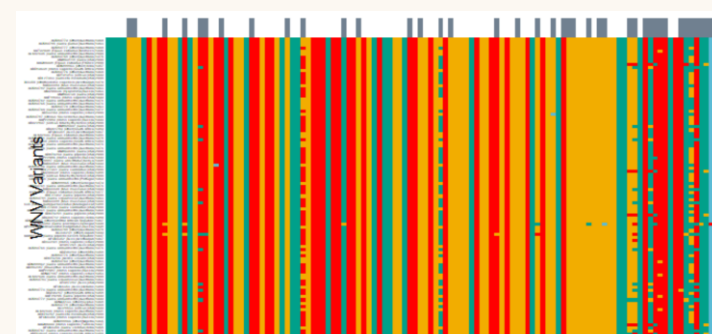


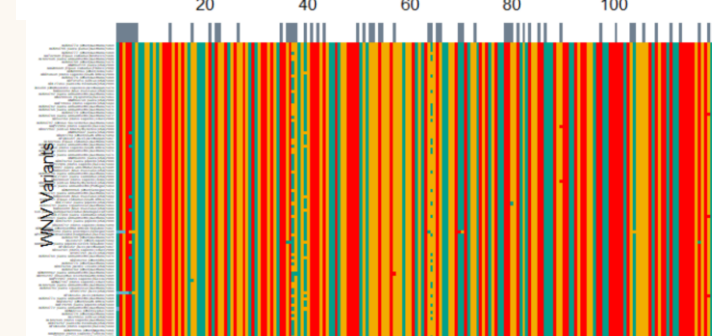
Fig. 10 Zoomed in prM Protein Phylogeny

Main Findings – Protein Alignments



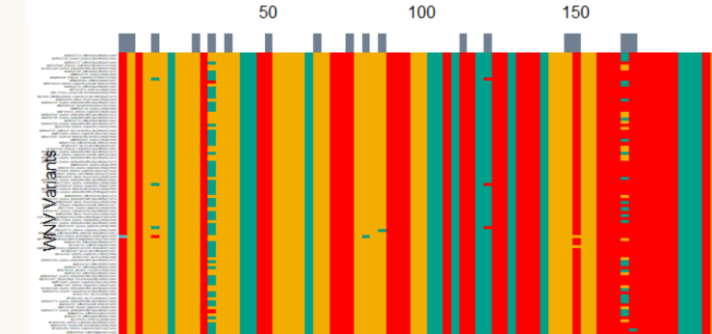
- Hydrophilic
- Small
- Hydrophobic
- Unknown

Fig. 11 C Protein Alignment



- Hydrophilic
- Small
- Hydrophobic
- Unknown

Fig. 12 E Protein Alignment



- Hydrophilic
- Small
- Hydrophobic
- Unknown

Fig. 13 prM Protein Alignment

Sequence Length

Select References

1. Brinton, M. A. 2006. Host factors involved in West Nile Virus replication. *The New York Academy of Sciences*. 951(1), 207-219.
2. Koo, Q. Y., Khan, A. M., Jung, K-O., Ramdas, S., Miotto, O., Tan, T. W., Brusic, V., Salmon, J., and August, J. T. 2009. Conservation and variability of West Nile Virus proteins. *PLoS ONE*, 4(4), e5352.
3. Lanciotti, R. S., Ebel, G. D., Deubel, V., Kerst, A. J., Murri, S., Meyer, R., Bowen, M., McKinney, N., Morrill, W. E., Crabtree, M. B., Kramer, L. D., and Roehrig, J. T. 2002. Complete genome sequences and phylogenetic analysis of the West Nile Virus strains isolated from the United States, Europe, and the Middle East. *Virology*, 298, 96-105.

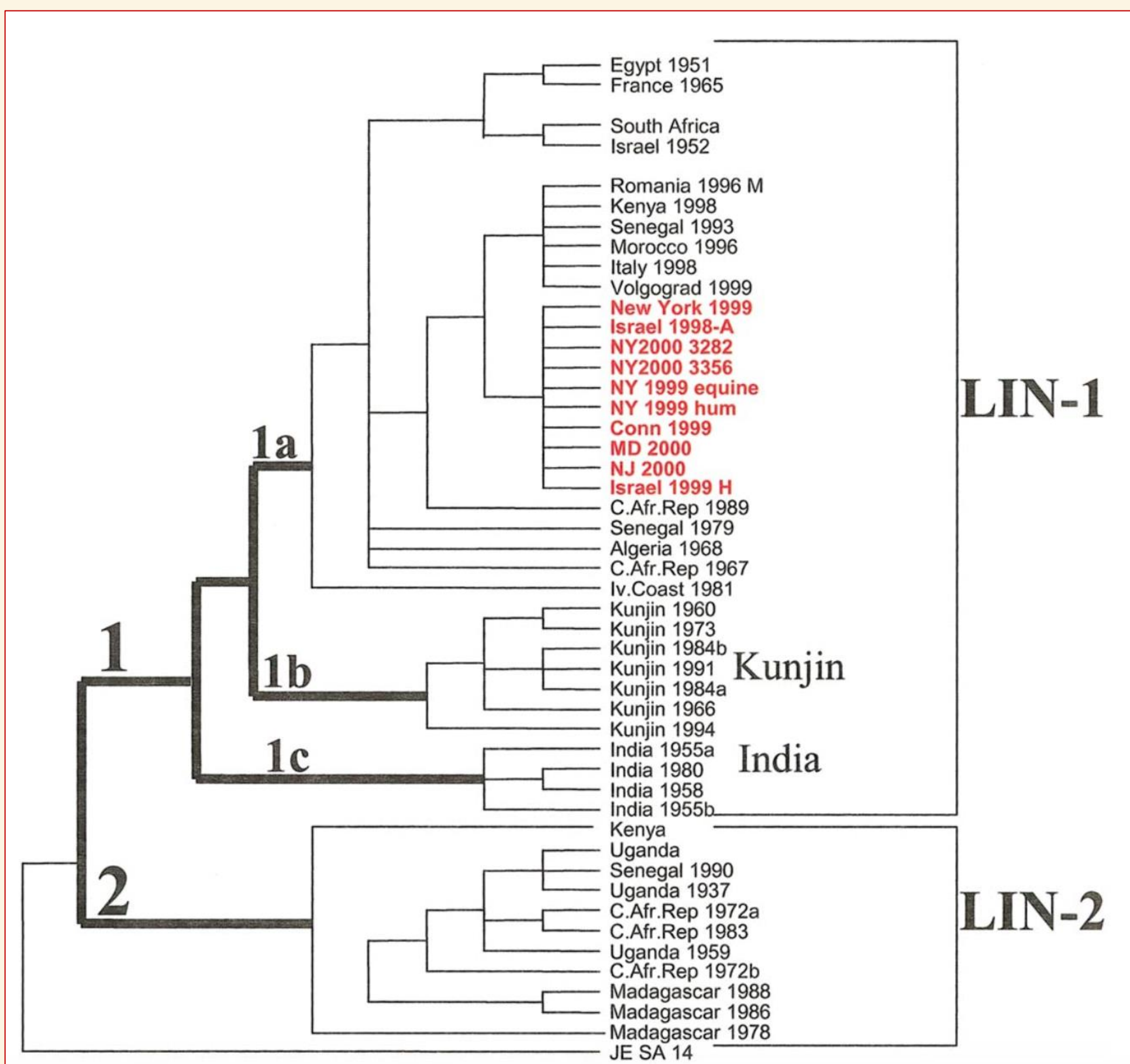


Fig. 1 Predicted lineage from literature