

Tracking the West Nile Virus

Erica Suchmann and Mark Gallo

Video VI: Microbial Evolution

Background

West Nile virus was first isolated in 1937 from a woman in Uganda; however, scientists did not characterize the life cycle of the virus until 1951 during an outbreak in Egypt. West Nile virus, a member of the family Flaviridae, is carried by *Culex* sp. mosquitoes. These mosquitoes prefer to feed on birds, but also bite humans, horses and other mammals. In 1957, doctors identified West Nile virus as a cause of meningoencephalitis (swelling of the brain and spinal column) in humans and, in 1960, as a cause of disease in horses.

Until recently West Nile virus had never been reported in the Americas. That all changed in 1999 when illnesses due to West Nile virus infection in humans and birds were identified in New York City. Microbiologists and epidemiologists initially hoped that the New York outbreak would be limited to one summer and that the virus and the mosquitoes would be unable to overwinter in the cold New York climate. Their hopes were dashed when human cases were identified in 12 states in 2000.

By the end of 2002, 3984 human cases had been identified in 39 states, and avian, animal and mosquito cases had been identified in 44 states (Centers for Disease Control web site: <http://www.cdc.gov/ncidod/dvbid/westnile/>; January, 2003)

Most people bitten by West Nile virus-infected mosquitoes do not develop symptoms. For the 20% that do, the symptoms of West Nile virus infection include a mild fever, headache, and body chills that begin between 3-14 days after initial exposure. In most people these symptoms last for a few days before resolving with no lasting effects. Less than 1% of those infected with West Nile virus go on to develop encephalitis or meningitis, but these severe symptoms seem to be more likely to develop in people over 50.

To determine the origin of the virus isolated in the New York City outbreak, scientists at the Centers for Disease Control (CDC) constructed a database of genetic sequences of the West Nile virus from strains that have been isolated during outbreaks since 1951. Discovering how the virus travels and enters new environments greatly enhances the ability of researchers to predict where the virus may spread in future outbreaks. The same techniques for studying the movement of the virus through populations can also be used to study outbreaks of other diseases and to trace biological agents used for terrorist attacks.

Epidemiological Activity

Your task is to look at genetic sequences from West Nile virus cases in mammals and birds and to compare them in order to determine the likely origin of the strain

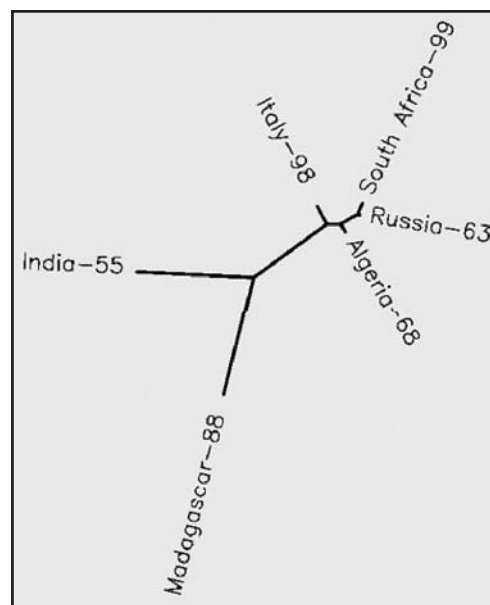


Figure 1. An unrooted tree based on aligned sequences for the envelope glycoprotein gene in six different strains of the West Nile virus. Each strain is labeled by the country and year it was collected. The tree was generated by running a ClustalW using the *Biology WorkBench*.

found first in the United States. If we assume that mutations occur continuously as the virus replicates in various hosts, then having access to the genetic sequences from several West Nile virus strains found in different locations will be helpful. We can compare these sequences and use their relatedness to hypothesize about how West Nile virus arrived in New York.

Figure 2. An example of the nucleic acid sequence data for the envelope glycoprotein gene in the West Nile virus (strain India-55).

```
>India-55
CCGACTACAGTCGAATCGCATGGCATCTACTCAACACAGCAGGGCGCCACTCAAGCAGGCCGGTTCAGCA
TAACTCCAGCCGCGCCATCATACACTCTAAAGCTTGGGGAATATGGAGAAGTCACTGTGGACTGCGAGCC
CCGCTCGGGCATAGACACAAACGCGTACTATGTGATGACAGTTGGGACGAAGACGTTTCTAGTCCACCGA
GAATGGTTCATGGACCT
```

The CDC's West Nile virus sequences are available to the public in the GenBank Viral Sequences (GBVRL) database. This database can be accessed from the *Biology Workbench* at the University of California San Diego Supercomputing Center (<http://workbench.sdsc.edu>). The bioinformatics tools you will need to look for patterns and relationships will be available there as well.

For an overview of the *Biology Workbench* and how it is organized, please see the "Orientation to the *Biology Workbench*" document on the *Microbes Count!* CD. You may also want to take a look at the "Proteins: Historians of Life on Earth" and "Tree of Life: An Introduction to Microbial Phylogeny" activities in Chapter 6 for some examples of using sequence data to create a phylogenetic tree.

Once you're set up, create a Session called West Nile and then upload the CDC sequences in the file called "West Nile Sequences.doc", located in the West Nile folder on the *Microbes Count!* CD. Align the sequences using ClustalW. An unrooted phylogenetic tree will be produced. The shape of the tree, the length of the branches, and the position of the different strains of West Nile provide information about relatedness; e.g., strains within a cluster are closely related to each other.

Figure 3. The ClustalW tool aligns the sequences, enabling you to compare West Nile virus collected from various geographical areas. Only partial sequences for the envelope gene in a subset of the CDC strains are shown here.

```
South_Africa-99_ CCAACCACTGTGGATTTCGCATGGTAACCTACCCACACAGATTGGGGCCACTCAGGCAGGG
Russia-63_ CCAACCACTGTGGAGTCGCATGGAACTACCCACACAGATTGGGGCCACTCAGGCAGGG
Algeria-68_ CCAACCACTGTGGAGTCGCATGGAACTACCCACACAGATTGGGGCCACTCAGGCAGGG
Italy-98_ CCAACCACTGTGGAGTCGCATGGAACTACTCCACACAGATTGGGGCCACTCAGGCAGGG
India-55 CCGACTACAGTCGAATCGCATGGCATCTACTCAACACAGCAGGGCGCCACTCAAGCAGGC
Madagascar-88_ CCGACGACTGTTGAATCTCATGGCAATTATTCAACACAGGTTGGGGCCACCCAGGCTGGA
** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** **
```

1. What are your conclusions about the origin and path of the West Nile virus that found its way to New York in 1999? Defend your methods as well as your answers.
2. Compare your tree to the tree published in the paper "Origin of West Nile Virus responsible for an Outbreak of Encephalitis in the NE US" (Lanciotti, 1999).

Why might your tree differ from the one published in the paper?

3. Design an experiment to determine the geographic spread of West Nile virus in the USA? What data and kinds of information would you need? What factors (e.g. political borders, population density, climate, etc.) do you expect to have the largest impact? Explain.
4. The tree only shows the relationship between strains of the virus for the envelope gene. What are some reasons for why the scientists at the CDC chose the envelope glycoprotein sequences to compare?

Web Resource Used in this Activity

Biology Workbench (<http://workbench.sdsc.edu>)

Originally developed by the Computational Biology Group at the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign. Ongoing development of version 3.2 is occurring at the San Diego Supercomputer Center, at the University of California, San Diego. The development was and is directed by Professor Shankar Subramaniam.

Additional Resources

Available on the *Microbes Count!* CD

Text

A copy of this activity, formatted for printing

“West Nile Sequences.doc”, a text file for use with the activity

“Orientation to the *Biology Workbench*”

Related *Microbes Count!* Activities

Chapter 2: Searching for Amylase

Chapter 4: Molecular Forensics

Chapter 4: Exploring HIV Evolution

Chapter 6: Proteins: Historians of Life on Earth

Chapter 6: Tree of Life: Introduction to Microbial Phylogeny

Chapter 6: One Cell, Three Genomes

Chapter 7: Visualizing Microbial Proteins

Unseen Life on Earth Telecourse

Coordinates with Video VI: Microbial Evolution

Relevant Textbook Keywords

Epidemiology, Pathogen, Virus

Related Web Sites (accessed on 3/20/03)

American Society for Microbiology
<http://asmusa.org>

Centers for Disease Control, West Nile virus information
<http://www.cdc.gov/ncidod/dvbid/westnile/>

Microbes Count! Website
<http://bioquest.org/microbescount>

Unseen Life on Earth: A Telecourse
http://www.microbeworld.org/htm/mam/is_telecourse.htm

References

Lanciotti, R. S., J. T. Roehrig, V. Deubel, J. Smith, M. Parker, K. Steele, B. Crise, K. E. Volpe, M. B. Crabtree, J. H. Scherret, R. A. Hall, J. S. MacKenzie, C. B. Cropp, B. Panigrahy, E. Ostlund, B. Schmitt, M. Malkinson, C. Banet, J. Weissman, N. Komar, H. M. Savage, W. Stone, T. McNamara, and D. J. Gubler (1999). Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. *Science* 286 (5448):2333-2337.

Bibliography

Campbell G. L., A. M. Marfin, R. S. Lanciotti, and D. G. Gubler (2002). West Nile virus. *Lancet Infectious Diseases* 2:519-29.
<http://www.cdc.gov/ncidod/dvbid/westnile/resources/wnv-campbell-et-al.pdf>

Jai X-Y, T. Briese, I. Jordan, A. Rambaut, et al. (1999). Genetic analysis of West Nile New York 1999 virus. *Lancet* 354:1971-1972.

Scherret, J. H., M. Poidinger, J. S. Mackenzie, A. K. Broom, V. Deubel, W. I. Lipkin, T. Briese, E. A. Gould, and R. A. Hall (2001). The relationship between West Nile and Kunjin viruses. *Emerging Infectious Diseases* 7(4).
<http://www.cdc.gov/ncidod/eid/vol7no4/scherret.htm>

Figure and Table References

Figure 1. Modified from *Biology Workbench* (<http://workbench.sdsc.edu>)

Figure 2. Modified from *Biology Workbench* (<http://workbench.sdsc.edu>)

Figure 3. Modified from *Biology Workbench* (<http://workbench.sdsc.edu>)