**Extension activity for “Virus Hunter: Epidemiology of Nipah Virus”**

For instructions on how to read phylogenetic trees, consider using the HHMI Biointeractive’s “[Creating Phylogenetic Trees from DNA sequences](http://www.hhmi.org/biointeractive/creating-phylogenetic-trees-dna-sequences)” Click and Learn with accompanying worksheet.

**Part 6: Evolution of Nipah virus.**

After the outbreak in Malaysia, additional outbreaks occurred in Bangladesh, India, Cambodia and Thailand. Nipah virus RNA was isolated from humans, Pteropus bats and/or pigs from these regions, sequenced, and used to create a phylogenetic tree (Figure 1).

1. Make a prediction about the relatedness of Nipah viruses isolated from humans, bats and pigs in Bangladesh, India, Malaysia, Cambodia and Thailand. Do you think all bat sequences will be closely related? Will all sequences from Bangladesh be closely related?
2. Interpret the phylogenetic tree in Figure 1. Was your prediction correct? Explain how the tree supports or disproves your prediction.



Figure 1: Phylogenetic tree of Nipah virus sequences acquired from humans, bats and pigs in Bangladesh, India, Cambodia, Malaysia and Thailand.\*

1. From your interpretation of the tree in Figure 1, draw a schematic on the map below showing the spread of Nipah virus in south Asia. Do you think pigs, bats or humans were responsible for spreading Nipah virus to humans in each location?



Figure 2. Map of South Asia. Port Dickson is indicated with a star.

Author: Holly Basta, PhD. Rocky Mountain College.

\***Phylogenetic tree methods.** Alignment was created from the nucleocapsid protein sequence of accession numbers JN808863, JN808857, FJ513078, AY988601, AJ627196, AJ564623, AJ564622, AJ564621, AY029768, AY029767, ALO75948, ACT32611, ALO75951, AEZ01374, CBM41030, ALO75942, AIS25032, AAM1340 and AEB21209 using Muscle[[1]](#endnote-2) in Mega6[[2]](#endnote-3). The tree was constructed with the UPGMA method. Bootstrap values[[3]](#endnote-4) (1000 replicates) are shown. Evolutionary distance was computed with the Poisson correction method[[4]](#endnote-5). Units are in amino acid substitutions per site. The rate of variation was computed with a gamma distribution with a shape parameter of 1. Ambiguous positions were removed in a pairwise fashion.

1. Edgar, R. C. "MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput." *Nucleic Acids Research* 32.5 (2004): 1792-797. [↑](#endnote-ref-2)
2. Tamura K., Stecher G., Peterson D., Filipski A., and Kumar S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution30: 2725-2729. [↑](#endnote-ref-3)
3. Felsenstein J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783-791 [↑](#endnote-ref-4)
4. Zuckerkandl E. and Pauling L. (1965). Evolutionary divergence and convergence in proteins. Edited in Evolving Genes and Proteins by V. Bryson and H.J. Vogel, pp. 97-166. Academic Press, New York. [↑](#endnote-ref-5)