zunal.com

Inferring Evolutionary Relationships from Sequencing Data

WebQuest Description: Students will use authentic gene or protein sequences from pathogens of their choice to infer evolutionary relationships and inform an epidemiological investigation.

Grade Level: 9-12
Curriculum: Science
Keywords: Evolution, gene sequencing, protein sequencing, molecular relationships, epidemiology
Published On:
Last Modified: 2018-04-26 23:38:13
WebQuest URL: http://zunal.com/webquest.php?w=380365

Introduction

You will act as a molecular biologist, working as a team with epidemiologists and other researchers to identify the source of a virus that has caused a disease outbreak. You will employ authentic sequencing data, published online, and the same tools used by actual researchers to develop phylogenic trees to infer molecular relationships between strains isolated from different patients. Using background information about each subject, you will determine the source of the disease outbreak.

Tasks

Your learning objective is to develop an understanding of cladistics and phylogeny, to recognize the types of data which can reveal evolutionary relationships, and to gain hands on experience with tools scientists use to infer relationships between organisms using sequencing data. You will begin by investigating characteristics which might be used to develop a cladogram, select a virus of interest and obtain sequences using the National Institute of Health's "Virus Variation Resource," then analyze sequences using authentic tools. Virus Variation Resource (https://www.ncbi.nlm.nih.gov/genome/viruses/variation/) : Allows researchers to easily locate DNA or protein sequences of interest for a variety of viruses. Sequences can be located by host organism, data of sample collection, geographic location, or other criteria of interest.SeaView: Alignment tool to prepare sequences for comparison by matching like regions in multiple sequences. Results may be exported in a format useful for building phylogenic trees.Fig Tree: Software to develop phylogenic trees using aligned sequences. Users may apply color to highlight relationships between samples and identify patterns based on sample characteristics in the resulting trees. You will generate a phylogenic tree including at least 20 samples. You will color code your tree according to characteristics of your choosing, to highlight relationships in your data. Finally, you will report on your findings in a one page executive summary.

Process

Research cladistics and phylogeny at Bozeman Science. Review common misconceptions here.Select a virus to research. NCBI is a free data repository (one of many) where people upload their sequenced data. They also have to include information called metadata which includes things like the location, date, species, sequencing method etc. associated with the sequence. The Virus Variation Resource is a tool to easily search metadata and access virus sequences of interest. Consider what types of samples you might be interested in, and what characteristics would allow you to highlight relationships between virus strains.If you are interested in a higher level of challenge, you may consider sequences of proteins extracted from yeast or bacterial pathogens. Download sequence files of interest.Open SeaView and load your downloaded sequences.Make an alignement. An alignment does the best job at figuring out how the protein likely evolved and arranges the characters to reflect this. Once you have aligned your sequences you will build a tree from them! In the trees menu click PhyML, then click run on the menu that pops up. Once you have your tree click save as unrooted tree Trees are difficult to interpret so when presenting them its often advisable to use some color to make your data easier for the audience to digest.Load unrooted tree in Fig Tree and apply color based on information from metadata to highlight patterns.Submit your colorized tree, along with a one page executive summary identifying your pathogen of interest, describing the source of your data, and explaining molecular relationships you can infer from your tree.

Evaluation

You will generate a phylogenic tree including at least 20 samples. You will color code your tree according to characteristics of your choosing, to highlight relationships in your data. Finally, you will report on your findings in a one page executive summary. To receive credit for this assignment, you should be able to answer yes to the following five questions:Does your tree include at least 20 samples?Have you highlighted your tree according to characteristics in the metadata, such as date or location of sample collection?Does your summary identify your pathogen of interest?Does your summary describe when, where, and how samples were acquired?Does your summary explain relationships between specimens revealed by your tree?Submit your completed documents in the Google Classroom.

Category and Score					Score
				Total Score	

Conclusion

Sequencing data can often reveal patterns in microevolution, which allows us to make inferences about the spread of diseases. For some diseases, it may be possible to pinpoint the exact time or place where an outbreak began. Considering samples acquired from nonhuman sources may help us to understand and control diseases which exist in animal reservoirs and periodically spill over to infect humans. Reflect on your process of analyzing a virus in today's activity. How might biologists and epidemiologists work together to address a public health crisis? What might we need to know to better address emerging diseases? How can we apply sequencing data and tools to answer these questions?If you are interested in learning more about epidemiology, or one disease in particular, Cambridge Infectious Diseases is a source of additional learning materials and a starting place to locate current research about many diseases which impact public health worldwide.

Teacher Page

This webquest guides students through a basic phylogenetics investigation using sequences from online databases and free, open source tools. The activity is designed to be completed in a single 75 minute block. SeaView and Fig Tree should be preinstalled on computers before class begins. If time is limited, consider downloading a selection of sequence files for several virus proteins in advance.
Standards
This webquest addresses the following standards, from the AP Biology course and exam description: Essential knowledge 1.B.2: Phylogenetic trees and cladograms are graphical representations (models) of evolutionary history that can be tested. Learning Objective 1.17: The student is able to pose scientific questions about a group of organisms whose relatedness is described by a phylogenetic tree or cladogram in order to (1) identify shared characteristics, (2) make inferences about the evolutionary history of the group, and (3) identify character data that could extend or improve the phylogenetic tree.
Credits
3D models of viruses used in webquest images obtained from Visual Science.Visual Science (2014). Visual science: Visualization, communication and education. Retrieved from https://www.visual-science.com
Other