background sheet

Investigating influenza

photo by Sari Dennise, www.fotopedia.com/items/flickr-3483676076, CC-BY-2.0

Evolution is generally considered to be a process that occurs over very long time frames: usually millions of years. This resource explores mechanisms of viral evolution and enables students to venture into the field of bioinformatics, which uses computer technology to manage biological information. Through analysis and interpretation of data, scientists gain new insights and understandings into complex processes, such as evolution.

The activity *Influenza – an evolving problem!* involves use of the Influenza Research Database (IRD) to develop understanding that evolution is a process that can proceed rapidly. It exposes students to an authentic research tool, using the context of influenza viruses and RNA nucleotide sequences to research evolution.

Changing virus: antigenic drift

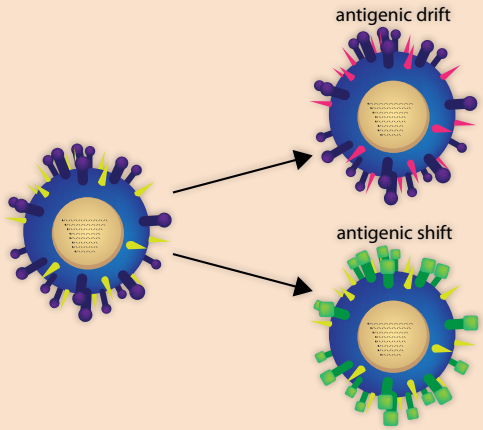
Antigenic **drift** occurs when chance mutations in RNA nucleotide sequences appear during replication. With no proof-reading ability, RNA polymerase doesn't recognise when mistakes are made. Such changes are relatively slow (but rapid in a fast reproducing virus), and happen continually over time. Mutations may result in small changes to viral proteins, particularly surface proteins H and N. Small changes to these surface proteins may render viruses less recognisable to hosts so detection by their immune system fails.

How is it possible for changes to occur in H and N proteins, yet mutated viruses still infect host cells? As H and N proteins are specific to host receptors, changes must be away from parts of proteins that attach to host cells. Parts of viruses that attach to cells' receptors are highly conserved, meaning they rarely change. Areas on viruses where host antibodies attach are non-conserved, meaning they change often.

Changing virus: antigenic shift

Antigenic **shift** occurs when genes of a single strain, or genes of two different strains, exchange genetic material. Antigenic shift may produce a novel combination of genes that has a survival advantage over contributing strains. When a new version of a virus emerges it may cause a major illness if it is not recognised by its host's immune system.

Viruses that originate from different species can take part in this process. Reassortment events, from human and avian (bird) influenza A viruses, have been the source of pandemics.



Rates of viral evolution: the need for regular immunisation

Rapidly evolving RNA viruses are estimated to have substitution rates approximately one million times greater than eukaryotic genes (Jenkins et al, 2002 in Lam et al, 2010), due to error-prone replication. Differences can develop or evolve within timeframes of months to years. Viruses can change so rapidly they escape human immunity – a vaccine formulated one year may not be effective the next. This is evolution in action.

Tracking viral evolution - the Influenza Research Database

The Influenza Research Database (IRD) is an online bioinformatics database and analysis resource that provides data for well-characterised influenza virus strains, and a suite of online analytical tools. The database is free to use. Using IRD, sequences of DNA fragments can be visualised and compared. To visualise influenza RNA as DNA it must be reverse transcribed into complementary DNA (cDNA), then amplified using a technique such as polymerase chain reaction (PCR). Once there's enough DNA to analyse, it's sequenced and compared against known variants. Comparisons focus on H and N proteins.

Various bioinformatics tools are available that use built-in algorithms to analyse relationships between genetic sequences. These tools allow large data sets to be analysed. In essence, the software aligns sequences; determines what genetic (nucleotide) differences are present; then converts genetic differences into a visual representation – a phylogenetic tree. Phylogenetic trees visualise relationships within a set of sequences.

Sequence data and alignment

Aligning sequence data enables quantification of relationships between different viral strains. To draw a phylogenetic tree that shows relationships and rate of evolution, sequences that are being compared must be carefully aligned. This is done by arranging different sequences, in rows, vertically above each other.

```
-ATGAAGGCAATACTAGTAGTTCTGCTAT
-ATGAAGGCAATACTAGTAGTTCTGCTAT
-ATGAAGGCAATACTAGTAGTTCTGCTAT
-ATGAAGGCAATACTAGTAGTTCTGCTAT
-ATGAAGGCAATACTAGTAGTTCTGCTAT
AATGAAAGTAAAACTACTTGTTCTGTTAT
```

Different computer algorithms are used to construct phylogenetic trees. IRD uses the following:

- MUSCLE (Multiple Sequence Comparison by Log-Expectation) algorithm to align sequences;
- JalView interactive alignment viewer to visualize nucleotide or amino acid sequence alignments; and
- PhyML (Guindon, S. and Gascuel, O., (2003). *Syst Biol. 52*: 696-704) to infer phylogenies based on nucleotide sequences.

If you're interested in reading more about analytical methods and algorithms, Lam et al (2010) provides a good overview.

Phylogenetic trees

Phylogenetic trees are tools that can show evolutionary relationships and rates. In this case, they show how influenza virus A H surface proteins are evolving.

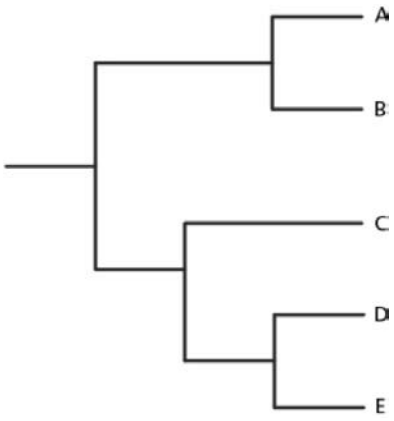
A meaningful tree has related sequences that share a common ancestor (they are homologous).

Rooted trees provide information about ancestor-descendent sequence relationships. A good way to root a tree is to include one or more sequences from a distant group, called an outgroup. Alternatively an older sequence can be included as the root (root sequences are the oldest and tip sequences the most recent). Both methods work in the case of influenza evolution as it's a form of progressive evolution, and give trees an evolutionary direction.

Horizontal distances in trees correspond to the amount of evolution. They are measured in terms of base substitutions per site.

Vertical distances have no evolutionary meaning.

A cladogram, as shown below, shows sequence ancestry but lacks any information about genetic distance.



a cladogram

Want to find out more?

Virology blog, about viruses and viral disease, retrieved 6 June 2012 from http://www.virology.ws/influenza-101/

Lam, T.T., Hon, C.C., Tang, J.W. (2010). Use of phylogenetics in the molecular epidemiology and evolutionary studies of viral infections. *Critical Reviews in Clinical Laboratory Science, 47*(1), 5-49.

Hay A.J., Gregory, V., Douglas, A.R., Lin, Y.P. (2001). The evolution of human influenza viruses. *Philosophical Transactions of the Royal Society B*. 356, 1861-1870.

Influenza (flu) viruses, Centres for Disease control and prevention, retrieved 6 June 2012 from http://www.cdc.gov/flu/about/viruses/