DNA in Hollywood: Fact, Fiction and Future

Dr. David Hillis
September 14, 2007
CSI / clip
Phylogenetic analysis can be used to trace viral infections through a human population

- Origins of HIV, SARS and other viruses transmitted between animals and humans
- Global virus diversity for vaccine trials
- Epidemiological studies
- Identification of new diseases
- Forensic uses
Phylogeny Papers, 1981-2006
(with “phylogeny” or “phylogenetic” in title or abstract)
Phylogeny
Evolutionary relationships among lineages, such as genes, individuals, populations, species, etc.

Consider an ancestral lineage (e.g., descendants from one HIV virus)
Phylogeny
Evolutionary relationships among lineages, such as genes, individuals, populations, species, etc.

One lineage splits into two by mutation
Phylogeny

Evolutionary relationships among lineages, such as genes, individuals, populations, species, etc.

There are now two HIV lineages
Phylogeny
Evolutionary relationships among lineages, such as genes, individuals, populations, species, etc.

Time

Some lineages become extinct

X

Lineages continue to diversify through time
Phylogeny

Evolutionary relationships among lineages, such as genes, individuals, populations, species, etc.

The relationships among lineages produce a phylogenetic tree.
If samples are taken at this point in time...

...then the reconstructed phylogenetic tree will look like this.
Phylogeny

How is evolutionary history reconstructed?

Each of these tick marks represents a new mutation in the genome of the respective HIV lineage.
Phylogeny

How is evolutionary history reconstructed?

Lineages 1 and 2 will share these mutations.
Phylogeny
How is evolutionary history reconstructed?

Lineages 1, 2, 3, and 4 will share these mutations.
A model of HIV evolution is used to account for multiple changes to the same sites on a genome, different rates of changes among sites, and other details of HIV evolution.
How do we know phylogenetic analysis works?

Experimental evolution to test phylogenetic methods

Hillis et al., 1992 (Science 255:589-592)
HIV transmission

Viral transmission events may be traced back through time among individuals in a population.

To imagine how this is possible, start by considering the diversity of HIV within one infected individual:

Time 1: Prior to Transmission event
At the transmission event, the HIV in the recipient represents a small subset of the HIV in the source:

**Time 2: The transmission event**

HIV sequence sampled from source → To recipient
As time passes, HIV lineages in the source and recipient diversify, and other lineages become extinct.

**Time 3: Shortly after transmission event**

![Diagram showing the diversification and extinction of HIV lineages over time. The diagram includes a tree-like structure with nodes labeled as 'HIV sequence sampled from source' and 'HIV from recipient', and an annotation indicating 'Immune system has terminated the lineage'. The time of transmission is marked with a red line.](attachment:image.png)
After enough time has passed, immunoselection will result in an apparent sister-group relationship among the HIV lineages in the source and recipient:

**Time 4: Further after transmission event**
Phylogenetic evidence: *pol* sequences

Transmission from patient to victim

New sequences added

Phylogenetic evidence: *pol* sequences

Metzker et al., 2002 (PNAS 99:14292-14297)
• Schmidt was convicted of attempted murder; currently serving term of 50 years of hard labor

• First use of phylogenetic analysis in U.S. criminal case

• Phylogenetics can be used to trace infections of human pathogens among individuals
Outbreak clip?
Prediction of future evolution of virus
Phylogenetic prediction of the future of influenza:

Which current strains will lead to the epidemics of tomorrow?
Jurassic Park clip, cloning dinosaurs
Star Trek clip, showing Tricorder
You are here

The cotton in your shirt came from here

Download wall poster from: http://www.zo.utexas.edu/faculty/antisense/DownloadfilesToL.html

(see Science 300:1692-1697)
You are here

The cotton in your shirt came from here

What if, a billion years ago...

Download wall poster from: http://www.zo.utexas.edu/faculty/antisense/DownloadfilesToL.html

(see Science 300:1692-1697)
You are here

The cotton in your shirt came from here

The *E.coli* in your gut is here

Download wall poster from: http://www.zo.utexas.edu/faculty/antisense/DownloadfilesToL.html

(see *Science* 300:1692-1697)
The cotton in your shirt came from here

The E. coli in your gut is here

The fungus on your foot is here

You are here

Download wall poster from: http://www.zo.utexas.edu/faculty/antisense/DownloadfilesToL.html (see Science 300:1692-1697)
To AmphibiaTree and beyond...

Download wall poster from: http://www.zo.utexas.edu/faculty/antisense/DownloadfilesToL.html (see Science 300:1692-1697)
Number of Described Species

1.7 million species described by 2007
163,572 in GenBank on 5 July 2007
Goal: Virtually all species known; a public online database for all species; “complete” Encyclopedia of Life and Tree of Life
Number of Described Species

What needs to happen here?
Number of Described Species

What needs to happen here?

Species in GenBank from phylogenomics studies?
Is a “Biocorder of Life” Possible?

Step one: DNA isolation and amplification of a series of target genes

Step two: Rapid sequencing of amplified genes

Step three: Placement of unknown within the Tree of Life
rRNA and other nuclear genes identify unknown species as a fungus

Sample of target genes identify unknown fungus as a chytrid: *Batrachochytrium dendrobatidis*

Information added to the global database of Life

Data, locality, life history, etc. connected to the *Encyclopedia of Life*; connection to declines

Go to Fungal Tree
Hillis et al., 2001 (Herpetologica 57:266-280), and unpublished data
to other Hemidactyliini

to other Eurycea

Eurycea

Paedomolge

Notiomolge

Blepsimolge

Typhlomolge

Septentriomolge

rathbuni

robusta

naufragia

tonkawae

chisholmensis

troglodytes

nana

sosorum

New sample

Hillis et al., 2001 (Herpetologica 57:266-280), and unpublished data
New sample to other Eurycea

to other Hemidactyliini

Hillis et al., 2001 (Herpetologica 57:266-280), and unpublished data
Tree of Life (Phylogenomics) Approach

Advantages:

• Can be automated; rapid rate increase
• Cost per identification of each new species less than current methods, especially assuming development of automated and portable sequencing technology
• Puts taxonomic knowledge in public databases, rather than just ephemeral human minds
• Better uses the talents of taxonomic experts, who benefit from all the new data
• Universal databases make taxonomy much more valuable and available to all biologists
Dr. David Hillis

David Hillis was born in Copenhagen, Denmark, but spent his formative years in tropical Africa and India. In this environment, Hillis learned a love of biology, entertaining himself by making collections of butterflies, amphibians, and reptiles. Hillis received his B.S. degree (with honors) from Baylor University in 1980, and M.A., M.Ph. and Ph.D. (all with honors) from The University of Kansas in 1983, 1984, and 1985, respectively. After two years on the faculty at the University of Miami, he joined the Department of Zoology at UT Austin in 1987 and was awarded a prestigious National Science Foundation Presidential Young Investigator Award the same year. In 1992, he was appointed to the Alfred W. Roark Centennial Professorship in Natural Sciences, and in 1998 became the first Director of the School of Biological Sciences at the University of Texas-Austin. In 1999, Hillis was one of 32 people chosen to receive a prestigious MacArthur Foundation fellowship, known informally as the "genius award".

David Hillis’ research interests span much of biology, from development of statistical and computational methods for analyzing DNA sequences, to molecular studies of viral epidemiology, to studies of the diversity and phylogeny of life (particularly vertebrates), to the origin and behavior of unisexual organisms. He has published over 130 scholarly articles and two technical books, and has served as Editor or Associate Editor of a dozen scientific journals. He is an active member of many scholarly societies and national research panels, and has served as the President of the Society of Systematic Biologists. In the past decade, the 23 graduate students and 12 postdocs in his laboratory have produced an additional 120 independent scholarly articles.