Genealogy to Trees: Inheritance and Tracing Ancestry Near and Far

You as a leaf on the Tree of Life

Because living organisms on planet earth all evolved from a common cellular ancestor, each living organism is a leaf on the enormous tree of life. Branches in the tree of life represent the connections among all species on the planet. Individuals are connected through a network of branches within species, species are connected through branches representing their history of diversification, and major groups of organisms are connected through successive branches back to the original diversification from the single-celled ancestor. Richard Dawkins in his 2009 book *The Greatest Show on Earth* provides an excellent metaphor for the connections or nodes in the Tree of Life,

> "...for any two animals there has to be a hairpin path linking them, for the simple reason that every species shares an ancestor with every other species: all we have to do is walk backwards from one species to the shared ancestor, then turn through a hairpin bend and walk forwards to the other species."

The same applies for people, considering that some individuals share with each other a more recent common ancestor(s) than with others. The webpage Genetic Connections Between Organisms from the Tree of Life Project nicely relates the genealogical connections among individuals within a family to the branching pattern within the tree of life. Because DNA contains the hereditary information for all cellular life, the branching history within the tree of life is recorded in the diversity of DNA sequences present among living organisms. Personal genomics enables an individual to discover their place within the tree of life.

Phylogenetic Trees

Biologists represent relationships among organisms using phylogenetic trees. The tips,
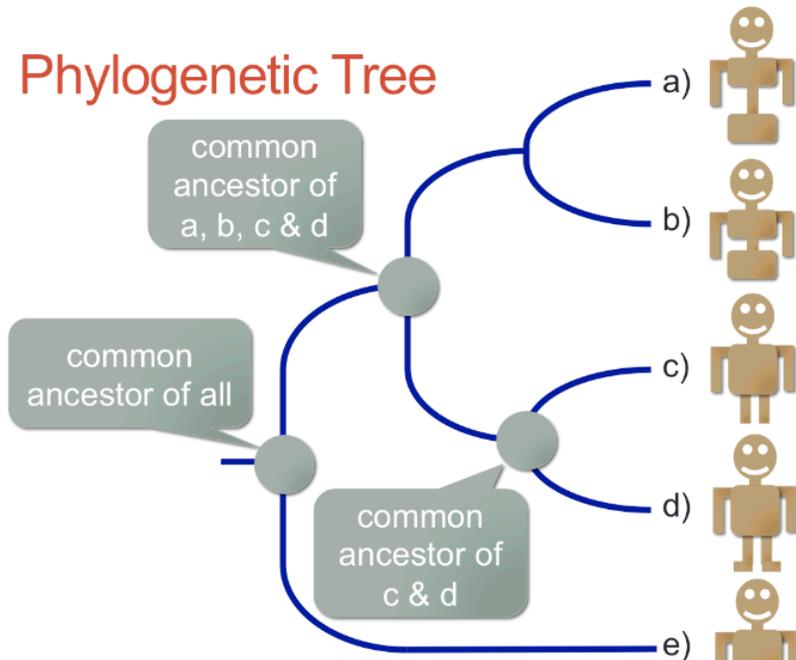
or leaves, of a phylogenetic tree can represent a variety of biological forms, such as DNA sequences, individual organisms, or species and other ta

Pedigrees and family trees are specialized types of phylogenetic trees commonly used to illustrate relationships in human families. Since humans reproduce sexually, each node within a pedigree consists of a pair of ancestors through which descendant lineages are produced. For example, your siblings are connected to you through your parents, your 1st cousins are connected to you through your grandparents, your 2nd cousins through your great grandparents, etc.
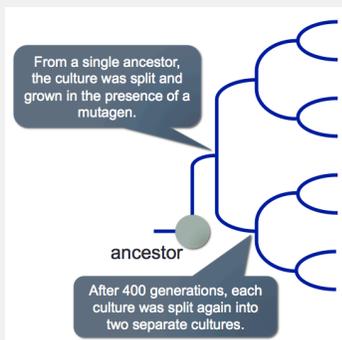
Inferring Relationships
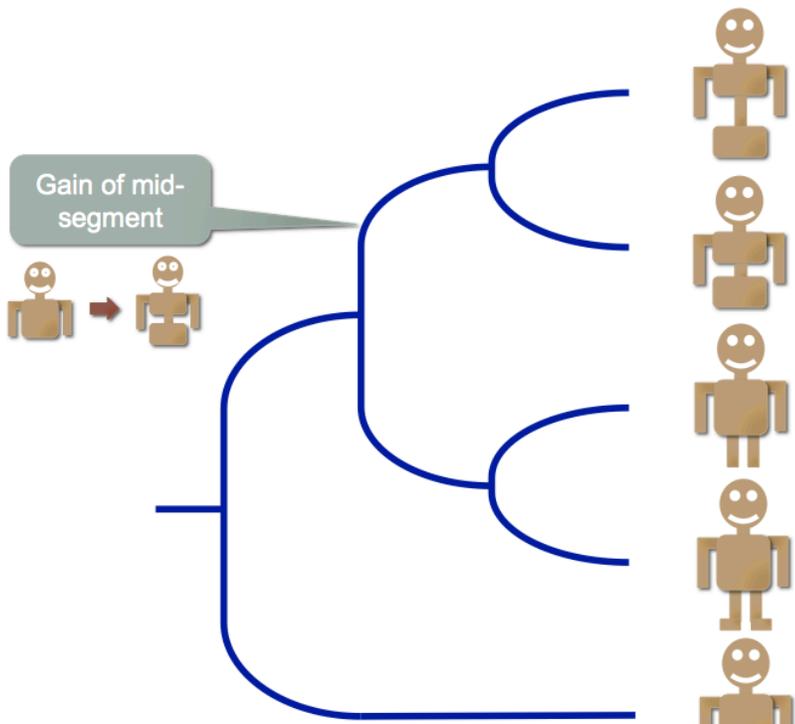
Relationships within a phylogenetic tree represent a hierarchical ordering of common ancestry. Each node connects lines of descent from a comm

Since the DNA present in each cell provides the recipe for cellular life, and this recipe is copied and transmitted as cells divide and organisms repr

Links
Tree of Life
interactive Tree of Life
Evolutionary Genealogy
BBC Tree of Life
The Tree Room
Smithsonian Human Family Tree

| An Activity on Determining Relationships |
| --- |
| A classic experiment in biology used controlled growth of a bacterial virus to evaluate the precision of methods commonly used to reconstruct evolutionary history. David Hillis, Jim Bull, and colleagues at the University of Texas conducted this example of experimental evolution. The diagram below summarizes the experimental regime. A single virus strain was used to initiate the experiment. Cultures were established and grown over successive generations, and at specific time points divided into independent cultures, so the evolutionary history of the viral lineages was controlled by the investigators. Growing the virus in the presence of a mutagen |

increased the rate of occurrence of changes in DNA sequence. At the end of the culturing phase of the experiment, eight different cultures were obtained. The DNA sequence from a representative virus from each of these cultures was obtained, compared with the DNA sequences of a representative virus from the other cultures, and the changes in DNA sequence from the single ancestral strain were used to reconstruct the evolutionary history of the viruses. Importantly, this analysis was performed without knowledge of which viral sequence came from which culture. Several different methods of phylogenetic analysis of the DNA sequence data accurately reconstructed the culture history of the viral strains.


From a single ancestor, the culture was split and grown in the presence of a mutagen.

ancestor

After 400 generations, each culture was split again into two separate cultures.

Do it yourself, this table contains the variant nucleotide positions of the viral strains. Using shared changes as an indicator of common ancestry, draw a phylogenetic tree connecting all of the evolved viral strains. Hint, the tree should be similar in shape to the culturing protocol illustrated above, tips of the tree contain a strain ID (A, B, etc), and the nodes reflect the hierarchical pattern of common ancestry inferred from the sequence data.

HHMI Activity: Creating Phylogenetic Trees from DNA Sequences

Shared Change Reveals Relationships

G→C
A→T

ACCAGTTT

CCCAGTTT

AGGAATAT
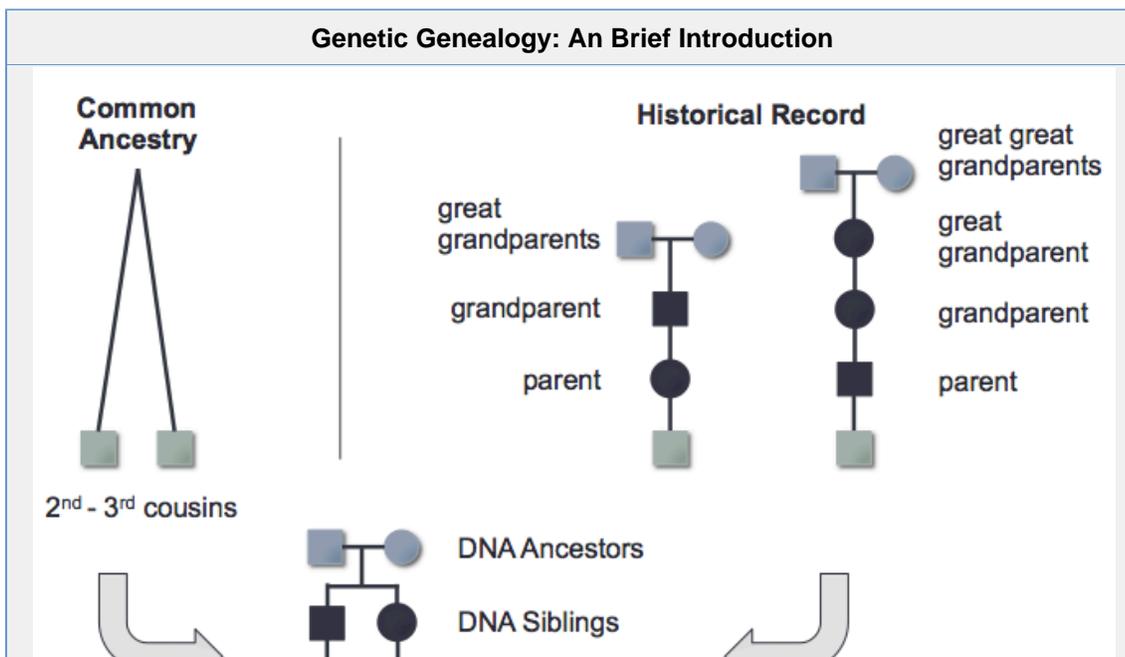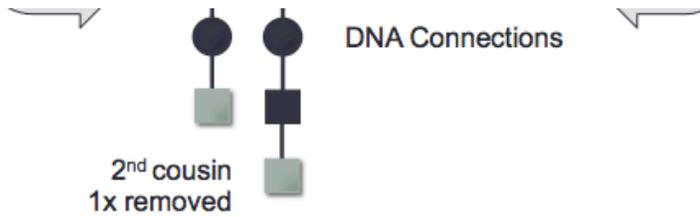
G→A

ACGA**A**A**A**T

ACG**T**GT**AA**
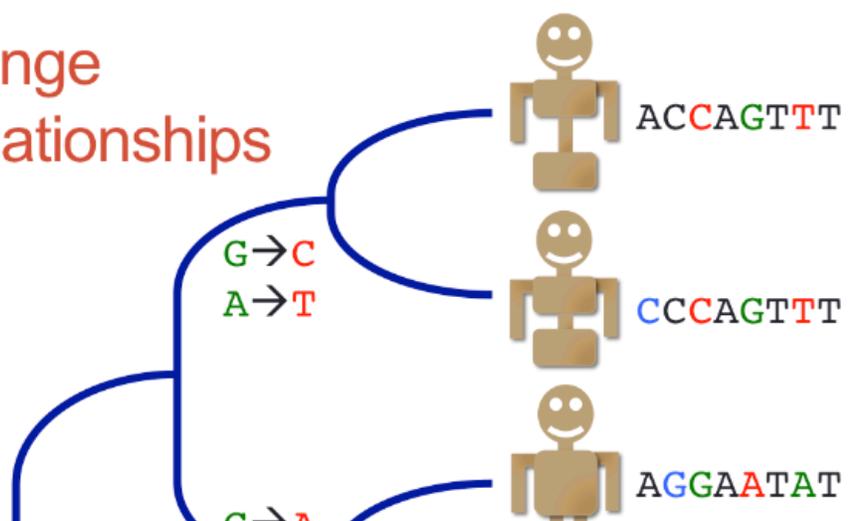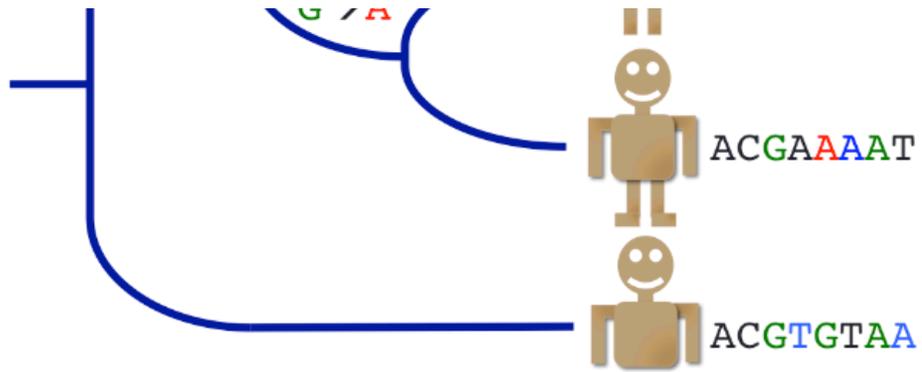
Genetic Genealogy: The Intersection of Phylogenetics and Genealogy

Application of DNA analysis to investigate ancestry enables an individual to support the historical record and to discover relationships that are misrepresented in the historical record or otherwise not documented. Genetic tools simply analyze shared features in DNA sequence data to infer recent common ancestry. In a database of individual genetic data, some individuals are more similar to each other due to the inheritance of DNA from a more recent common ancestor(s). Within a phylogenetic context, these related individuals are connected by a node reflecting this recent common ancestry. Estimates of whether this common ancestry is through great great grandparents, a patrilineal grandfather 10 generations removed, or a common maternal line that originated 4,000 years ago can be made and are dependent on the type of genetic data analyzed. In the ideal genetic genealogy world, all ancestral connections among all living individuals would be inferred through a network of branches connected through nodes representing all common ancestors. However, current tools are unable to provide this level of resolution beyond a few generations, so we may never be able to reveal our genetic ancestry at this level of detail. In a genealogical context, genetic data are used only to infer common ancestry, most often based on analysis of genetic data from related individuals that each descend from a set of common ancestors, some more recent than others, and not to identify these ancestors.

Genealogical research based on historical records provides the human tapestry of this ancestry with names, events, dates and places. Use of genetic tools to reconstruct common ancestry may be complementary, thus supportive, of historical records. It may also be contradictory when the history does not accurately reflect the biology. Recognize, however, that genetic data does not validate or confirm or prove the historical record except in direct comparisons between close relatives when no other relationships are plausible given the genetic data (i.e., parent-child, full siblings). Alternative historical scenarios that are not represented, or are misrepresented, in the historical record are often equally consistent with genetic data. The intersection between genetics and the historical record in the context of genealogical research is quite analogous to the intersection between molecular systematics and paleontology. Molecular systematists aim to define the relationships among species in the tree of life and paleontologists aim to describe and place fossil forms relative to common ancestral nodes. Just as the fossil record sparsely represents the biodiversity that has existed on the planet, the historical record can bias and in cases where it is missing leave much of one's genetic ancestry lacking in personality.

Genetic Genealogy: An Brief Introduction

Common Ancestry

2nd - 3rd cousins

Historical Record

great great grandparents

great grandparents — great grandparent

grandparent — grandparent

parent — parent

DNA Ancestors

DNA Siblings

DNA Connections

2nd cousin
1x removed

Genetic genealogy is the intersection of DNA-based inference of common ancestry and of historical inference about familial ancestors. Individuals may enter their search for their biological history at either side of this intersection. For example, results of DNA tests of two individuals provide evidence of recent common ancestry. In the most common genetic tests used to find relatives, the relationship between matching individuals is estimated (e. g. 2nd to 3rd cousins) from DNA data based on principles of inheritance. Comparison between the family trees of the individuals identified as having recent common ancestry is used to identify the common ancestor(s) putatively responsible for the genetic match. Alternatively, individuals that share common ancestors identified from the historical record can each submit DNA tests and determine whether the genetic match exists and is consistent with the supposed relationship.

Recognize, however, that the presence of shared common ancestors in the family trees of two individuals with matching DNA does not demonstrate that each shared individual is an ancestor. Beyond siblings, estimates of relationships do not discriminate between whether two individuals share one or two common ancestors (i.e., half- or full-sib relationships, respectively), so what is identified as a common pair of "DNA Ancestors" may reflect inheritance from only a single common ancestor. While the genetic match between descendants does not specifically provide support for each of the common ancestors, it does as singular points of genetic transmission support the historical record of the "DNA Siblings" and each "DNA Connection" in the lineages of the tested individuals. It does not, however, confirm or prove the historical record. Consider that the alternative scenarios of the nefarious brother and/or of the unwed sister as a biological parent are equally consistent with the genetic inference. Moreover, the genetic relationship may arise from a different common ancestor(s).