

Professor Nessuno is interested in tracing how information is shared and distributed by his students. On a particular day only one student had attended his lecture on Cytochrome B, one of the proteins that accept and release electrons during oxidative metabolism in mitochondria. Only one student, Maria, heard that all eukaryotic cells possess mitochondria and thus also have cytochromes. Only she got that the DNA sequence (gene) to encode Cytochrome B might differ a little or a lot from one species to another. Later Maria let one or more other students copy her notes (by hand). Those students in turn let their friends copy their copies—also by hand. Almost inevitably errors were introduced during the copy process.

Having collected the students' notes Nessuno is interested in tracing who copied from whom. He doesn't care about the spelling errors—or even the copying—in this instance, but he is curious about its similarity to naturally occurring evolutionary processes. He decides to look at one word and track changes in its spelling. He starts with the assumptions that each copier has only a single source to copy from and that only a single change is introduced with each recopying. His assumptions might be wrong. Here are the variations of **cytochrome** and their innovators:

cidachrome (Ali)	ciderchrome (Akiko)	ciderkrome (Boris)	citachrome (Bridget)
citochrome (Carmen)	cydachrome (Dat)	cytachrome (DeEtte)	psytochrone (Fatima)
psytokrome (Guido)	sitachrome (Imelda)	sitochrome (Josef)	sydachrome (Michelle)
sydachrome (Natasha)	sytachrome (Sita)	sytochrome (Yvette)	cytochrome (Maria)

Who copied from whom? **Construct** a family tree showing the “modifications with descent” among the various essays submitted? One method might be to cut the table above along the dotted lines and rearrange the sixteen spellings based on their similarities, grouping together those with the fewest differences. **Compare** your tree with others. This one has some problems. —>



Notice that nearly all of the variations are easily recognizable and some of them are considered correct spellings in languages other than English. These “errors” are like the neutral mutations that arise over time in the genome of any organism, although this example is more like similar changes as they occur in asexually reproducing organisms which have only a single parent.

Q: What organisms reproduce only asexually?

Q: What cellular components reproduce themselves and have their own DNA separate from the nucleus?

Q: What kind of result do we usually expect when an error in copying DNA occurs?

Q: Why don't all mutations produce disastrous impacts? How many reasons could you list why a mutation you experience during your childhood won't have any effect on your children?

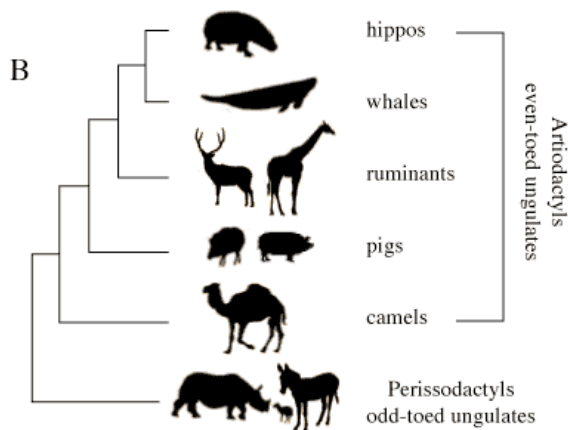
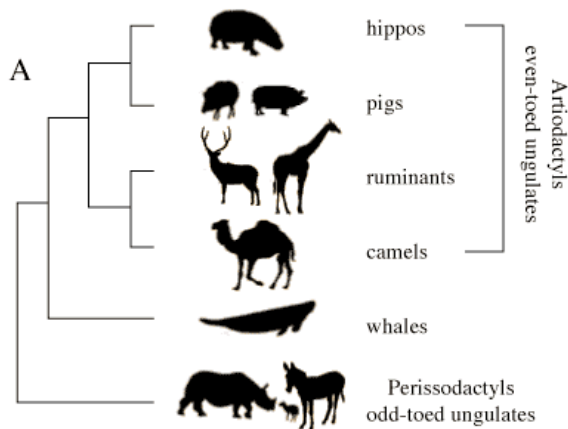
[Need insight into this? Open the Gene Explorer application, change one of the 127 base pairs (Enter New DNA Sequence) and see that changes in only some regions affect the amino acid sequence. Which ones?]



From the time George Washington was supposedly chopping down that cherry tree, botanists and zoologists have been in the tree-building business. Across the Atlantic Ocean in Sweden Carolus Linnaeus had begun to employ two consistent approaches in naming and classifying plants and animals: using hierarchies and binomials. The binomials replaced and simplified the cumbersome, long descriptive names that had been used before. He also organized hierarchies in which species were grouped into genera, genera into orders, and so on up to kingdoms.

Much later as evolutionary ideas became widely accepted a new feature was added: besides being binomial and hierarchical, our modern system of classification is phylogenetic; it now must reflect evolutionary relationships. As fresh knowledge is gained old family trees must be revised. In the last two decades some very dramatic revisions have occurred, many driven by evidence from molecules: DNA and protein.

Consider these family trees of selected mammals taken from the BioQUEST online Whippo Problem Workspace*:



Understand:

Each of these trees is one of several possible hypotheses of how these mammals are related to one another, how they evolved.

These are “rooted” trees growing from a root or common ancestor (not named or shown) on the right. Similar trees can also be drawn from the bottom up simply to fit better on a page.

Each branch represents a point when one or more differences arose as a result of genetic changes in one group (or subpopulation) but not another.

Initially biologists could only look at physical traits (often features of skeleton, teeth or other hard body part for mammals) to find similarities and differences. Modern techniques now allow us to look at the genes themselves (or their immediate products, proteins), structures closer to the point of origins of the differences.

What is your initial impression of these two trees? (Record in your lab notebook.)

How do you interpret the differences between them? (Record in your lab notebook.)

What do you already know about these animals that would support one tree over the other? What more do you need to know? (Record in your lab notebook.)

In the next section you will be using some tools for “growing” such trees from data about genetic differences between groups of animals.

*http://www.bioquest.org/bedrock/problem_spaces/whippo/index.php





ESTEEM Home Modules About Contribute Editorial Board Search

EvolSeq

This worksheet simulates the molecular evolution of DNA sequences. The program begins with a single (random) sequence, then follows that sequence through time as it reproduces and mutates. Eventually, up to 20 evolutionary related sequences are generated. EvolSeq then calculates the genetic distances between each pair of DNA sequences, and also between the associated amino acid sequences.

Fundamental Mathematical Expression

$$p_0 = \frac{1}{2N}$$

$$\begin{aligned} \lambda &= \mu_0 p_0 \\ &= (2N\mu) \left(\frac{1}{2N} \right) \\ &= \mu \end{aligned}$$

p_0 = probability of fixation
 λ = rate of substitution
 μ = rate of mutation

Learning Objectives

Appreciate and apply general principles behind phylogenetic trees.

- Each tree is a hypothesis, one of several possible generated from same data set.
- Trees are isomeric.
- Trees based on multiple characters yield more defensible relationships.
- Branching represents differences.
- Each branch point represents the last common ancestor.
- Each branch tip represents a taxon (phylum, class, ...species, etc.).

Use EvolSeq (and graph paper) to generate a tree to represent difference distances.

Demonstrate your understanding of the general principles behind phylogenetic trees by constructing a phylogenetic tree based on seven or more sequences using the Esteem Module called EvolSeq. This exercise is designed to develop skill in solving a problem in phylogeny and not to test your knowledge of any particular animal group nor your ability to search for such information. The program generates a chosen number of DNA sequences, each representing an equivalent region of a different species. It even translates those sequences into the proteins they would encode if they were real. Your job is to draw a family tree to show their evolutionary relationships.

EvolSeq does a lot of the really tedious work involved in the task and displays the results in several interacting “sheets” called **Distance Matrices, Calculations, Sequences, and Translation**. Take a few moments to look at what each tabbed sheet displays. Only the sheet called Distance Matrices requires any input from you; the others display results generate in response to your entries or general information about transcription and translation.



Distance Matrices is normally the active sheet when EvolSeq opens. You can interact with this sheet in two ways: **specify** the number nucleotide sequences to be compared and **select** how you want the results displayed (Upper Triangle may be simpler to decipher than Symmetric).

# of sequences:	7	(must be between 2 and 20 inclusive)	Distance matrix display				
			<input checked="" type="radio"/> Upper triangular <input type="radio"/> Symmetric				
Distance matrix: Nucleotide Sequences							
	Seq. A	Seq. B	Seq. C	Seq. D	Seq. E	Seq. F	Seq. G
Seq. A	0	8	12	4	12	10	6
Seq. B		0	12	8	12	10	8
Seq. C			0	12	2	12	12
Seq. D				0	12	10	6
Seq. E					0	12	12
Seq. F						0	10
Seq. G							0

Sequences displays a number of artificially-generated DNA sequences that you specified and that EvolSeq has generated. They resemble but do not reflect real DNA sequences. Also displayed is the polypeptide chain that would be synthesized from each DNA sequence.

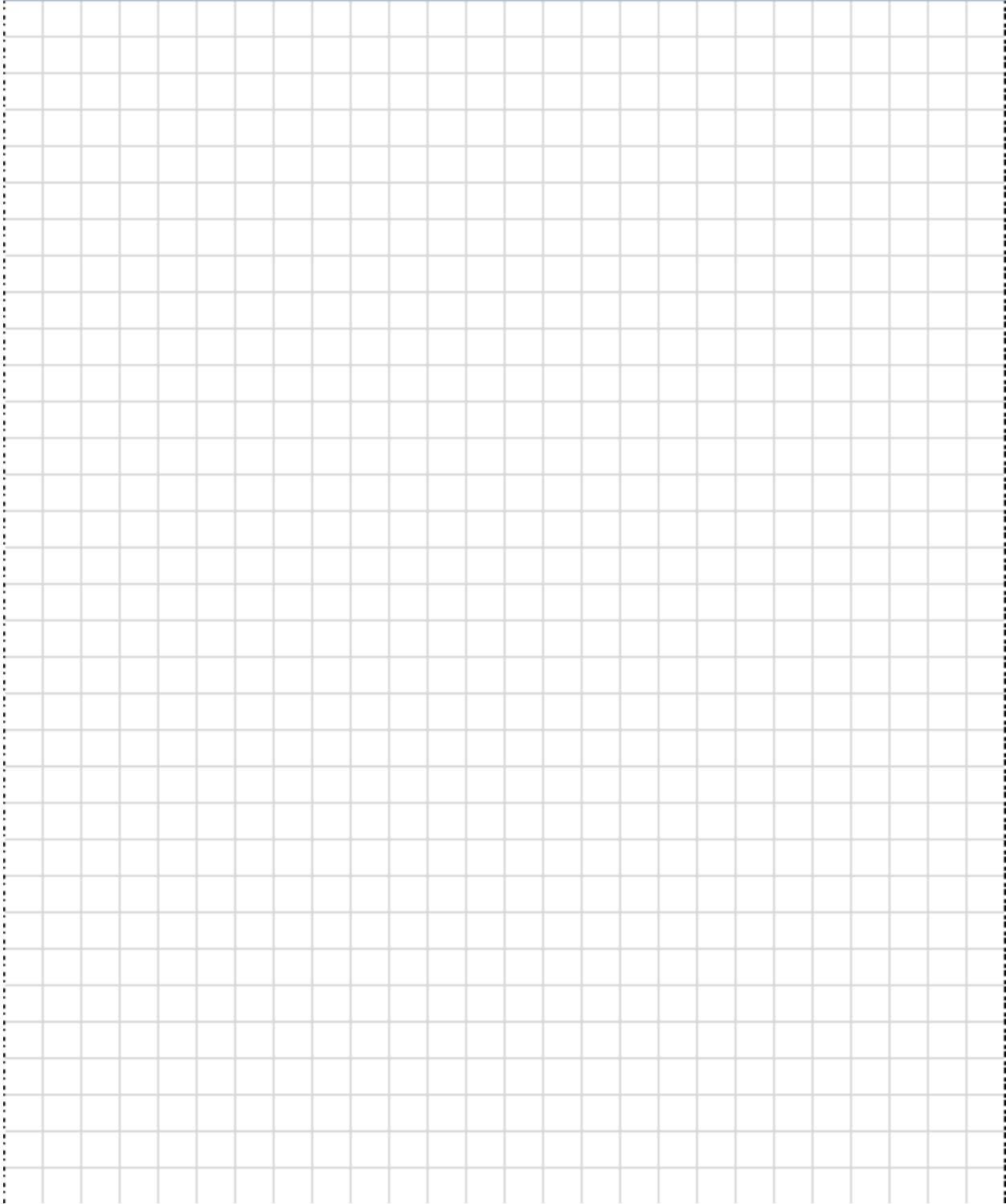
Translation shows the DNA codes for the amino acids (compare with the code table in your textbook). This sheet lists the amino acids both as 3-letter abbreviations that you could probably match with a list of their names spelled out in full and as 1-letter abbreviations. Why do you suppose there are two different sets of abbreviations for the same things?

Calculations shows what EvolSeq has done in the background to compare the DNA sequences. You probably will have little or no use for the Calculations sheet.

Activity and Challenge

On the next page carefully **compare** the Distance Matrices of 2, 3, 4 and 5 sequences and the trees made from them. Note that to get from A to B in the first, one moves down one space and then back up one space (the horizontal distance does not count since it is only there to separate A and B on the page); this is how we represent a “difference distance” of 2. On these trees the vertical distances could be understood to represent the passage of time since the last common ancestor of those groups above the split, time during which mutations have occurred in one or (probably both) groups.





Exploring Oxygen and Hemoglobin

Because oxygen is not very soluble in the watery part of blood, animals have evolved a number of strategies for acquiring, holding, and releasing oxygen in the blood. Chief among these is hemoglobin, a protein consisting of four polypeptide chains, each of which is complexed with its own molecule of heme. At the center of the heme molecule is an atom of iron which is where an oxygen molecule temporarily associates. An adult's hemoglobin molecule contains two alpha (α) chains and two beta (β) chains. Hemoglobin expressed at other life stages may have two gamma (γ) chains or two epsilon (ϵ) chains instead of β chains or have two zeta (ζ) chains, but instead two α chains.

Vertebrate muscle tissues also contain a related molecule myoglobin to store oxygen locally.

Oxygen has only one positive role, yet a very important one: to accept electrons and hydrogen ions produced during metabolism.

To become familiar with some of the wonders of hemoglobin consider the follow questions. Many of them can be answered quite adequately at <<http://en.wikipedia.org/wiki/Hemoglobin>>, although there is much more information there than will be needed in this course. If you use a resource other than Wikipedia, be sure to indicate what it was. Be aware that Wikipedia, like any source, may contain errors that haven't been caught yet.

Complete the table to show the number of kinds of polypeptides found in the hemoglobin of human beings at different stages of development:

Life Stage	alpha (α)	zeta (ζ)	beta (β)	gamma (γ)	epsilon (ϵ)
Embryo (Gower 2)					
Fetus					
Adult					

Name three organism that you were surprised to learn make hemoglobin for some reason or another.

Identify one other role of hemoglobin besides binding and releasing oxygen?

The work you've done to this point has been to establish a common set of background knowledge and skills. The next part of our exploration of hemoglobin, the animals that make it and their evolution will be to work with some additional tools and information available online used by professional biologists at universities and other research institutions. Because we access these tools using an Internet Browser you will be able to continue and extend your use of them from your home computer, putting the power of a supercomputer at your fingertips.

