Molecular Phylogeny Lab

Purpose:

To understand how similarities in mutations between species allows scientists to determine putative evolutionary relationships and construct a molecular phylogeny. To practice data analysis, including critical thinking skills, making logical inferences based on evidence, and determining the most parsimonious explanation.

Introduction:

DNA replication is a very conservative process, but occasional mutations do occur. Only mutations that occur in germ line cells can be passed on to the offspring. Scientists have observed that certain important chemicals found in many species, for example enzymes in biochemical pathways or receptors in electron transport chains, are highly conserved, that is they change little from one species to another. But there are differences, and scientists are now using these to establish putative evolutionary relationships between the species. The idea is that the mutations accumulated slowly over time, therefore species that have less differences in the DNA diverged more recently. If we can estimate how frequently mutations occur, we can also use the data to estimate how long ago the ancestors of the modern species separated, a technique called a molecular clock.

In this lab we will model the process of mutation and separation of lineages, and use the data to draw a tree showing the branching pattern of the lineages. At the beginning of the process, you will all be in one circle, representing that you are all part of one species. Over time, a combination of mutations and isolation will lead to each of you becoming a separate species. You will then reconstruct the branching pattern by which the species separated.

Data generation procedure:

1. Write your name on the molecular clock data sheet. The letters represent the order of nucleotides in one strand of DNA.

2. Stand in a circle with the rest of the class. You will need the data sheet, a pencil, and something to write on.

3. Every fifteen seconds the timekeeper will call out "mutate." One person will be chosen by the teacher to start the mutation. That person will select a base to mutate, call out the number of the base, and what letter to mutate it to, for example, "49 T". (Obviously the only valid letters are ATGC.) Each person in the group will write the change in the space directly below the letter. Be neat and make sure it is clear which space the mutation is in.

4. Each time the timekeeper calls "mutate," the next person to the right will decide on the mutation. Try to be random in the choice of number and base. Each base should only mutate once, that is there should be no mutations back to the original or to some other base.

5. As the simulation continues, the teacher will divide the class into smaller groups, until each person is a separate group. When a group is divided, each of the subgroups should pull together to form a ring. The group containing a person who just called a mutation will continue to the right, as before. The student to the right of where the ring was broken will start calling mutations in the other group. When only one person remains in a group, he or she will choose and record a mutation each time until the end of the simulation.

6. After the simulation the teacher will compile the data on an Excel spread sheet for use the next day,

Data analysis procedure:

1. Students will now work in groups to reconstruct the splitting of the class, using the base change data.

2. Compare the base changes that have occurred. First look for any shared by all the students. These took place before the first division. Next look for mutations shared by large numbers of people. This will divide the class into two groups. If you are working on a computer, select whole rows of data and move them to separate the data into those two groups. If you are not working on a computer, you may find it useful at this point to cut the paper into strips so each person is represented by one strip. Place strips with similar mutations together and keep dividing each group based on similar changes. As you go on it may help to count the number of base changes that each pair has in common to find the ones that are most similar. You may also want to use colored pencils or markers to mark mutations that are the same at one site.

3. Remember that it is possible that there are errors in the data. Someone might have marked the change in the wrong place, or marked the original base instead of the change. Write down every place you suspect there may have been a data error. Such errors occur in real data as well, and make interpretation more difficult.

4. Another thing which makes interpretation more difficult is that two people who have already been separated may choose the same mutation. Something like this also happens in real life, parallel evolution. Generally they will share one change, but differ at several others. Make note of these as well. You may find that two models are equally possible at some point. If so, record this too.

5. Once you have figured out the similarities, construct a tree showing when each of the groups split. Let each space on your paper represent one mutation. Each line represents a time between mutations when branching may have occurred. Start with the whole class together at the bottom of the page, and show the branching pattern until everyone is alone at the top, as in the example.

6. On your tree, mark the mutations that occur between each branch point. The order of mutations between branch points cannot be determined.

Lab report:

You will turn in one report per group, following the format of the standard lab report. You must also turn in the notes and rough drafts you made doing your tree, to show how you arrived at your final analysis.

Introduction
Describe the concept of molecular phylogeny, and how we are simulating it in this activity. Since this is a simulation, there is no hypothesis.

Methods
Not necessary for this lab.

Results
Draw your completed tree, with dividing points on a line and mutations shown in the spaces, as shown in the example.

Discuss any problems you noted in the data, such as data points that seem to be incorrect (person, number and letter of the mutation, and why you believe they are errors), and noteworthy results (how many large groups formed, any individuals who seem to have split off early, etc.)

Analysis:
Explain why you believe your tree is correct. Discuss any examples you found of apparent parallel evolution, identical mutations that occurred after the two had separated, and why you believe they had already separated. Make special note of any cases where two models are equally possible.

What simplifying assumptions did we make in this model? That is, in what way is this model different from mutations is nature? (This is a thought question, based on what we have discussed about DNA and mutations). Simplifying assumptions must also be made in real life when determining relatedness of species based on molecular similarities.