Extraction, Sequencing, and Analysis
of Mitochondrial DNA

Robert J. Kosinski[1], Donna R. Weinbrenner[2], and Melodie G. Cross[3]

Department of Biological Sciences
132 Long Hall
Clemson University
Clemson, SC 29634-0314
[1]*rjksn@clemson.edu*
[2]*donnaw@clemson.edu*
[3]*melodic@clemson.edu*

Abstract: In this laboratory, students extract mitochondrial DNA (mtDNA) from eyelash hair follicles. The instructor then uses polymerase chain reaction to amplify a 440 bp segment of the control region. After the students perform electrophoresis to verify that the segment has amplified successfully, the samples are sent to the Dolan DNA Learning Center for free sequencing. The sequence results are posted on a Dolan Web site. Students use ClustalW to determine if their mtDNA sequence is most similar to those of their own ethnic group. Finally, students analyze literature data on mtDNA in different human ethnic groups.

Introduction for the Instructor

At Clemson, we use this laboratory in the introductory biology course for majors. Prior to this exercise, the students have completed a laboratory on bioinformatics. We use this laboratory to make DNA and bioinformatics "personal." Students extract their own mitochondrial DNA (mtDNA) from hair follicles, and it is sequenced at the Dolan DNA Learning Center (DNALC) at Cold Spring Harbor Laboratory. Then the students use their bioinformatics skills to determine if their mtDNA is most similar to mtDNA of expected ethnic groups…or whether their mtDNA will surprise them.

As implemented at Clemson, this laboratory extends over several laboratory periods:

Period 1: extraction of mitochondrial DNA from hair follicles (Procedure A). This is a very short student exercise that we combine with an existing laboratory.

Between periods 1 and 2: PCR amplification of student DNA samples.

Period 2: students and instructor examine PCR products for a strong 440 bp band indicating that amplification was successful (Procedure B). Instructor photographs gels before they fade to furnish photographic proof of successful amplification to the DNALC. This is also a very short procedure inserted into an existing laboratory.

Between Periods 2 and 3: Instructor mails the successful samples and gel photographs to the DNALC. The Center sequences them for free and posts the results on a Web site. Because this takes an uncertain amount of time, Clemson submits samples at the end of the fall semester and plans to use the results only at the beginning of the spring semester.

Period 3: Students use bioinformatics tools to examine the posted sequences (Procedures C and D). At Clemson, this is a small part of a laboratory on phylogenetic taxonomy.

Student Outline
Extraction, Sequencing and Analysis of mtDNA

Human cells have six billion base pairs (bp) of chromosomal DNA, but they also have a much smaller amount of mitochondrial DNA (mtDNA). Human mitochondrial DNA is a ring consisting of only 16,569 bp. This ring contains a tiny number of genes—37, as compared with about 21,000 in the haploid human genome. Thirteen of these genes code for proteins involved in aerobic respiration. The majority of the genes code for rRNAs and tRNAs. You may remember that mitochondria have their own ribosomes; these rRNAs and tRNAs are used for protein synthesis within the mitochondrion. A small segment of the mtDNA has no genes and is called the control region because it controls the replication and protein synthesis of the rest of the mtDNA. A diagram of mtDNA appears below:
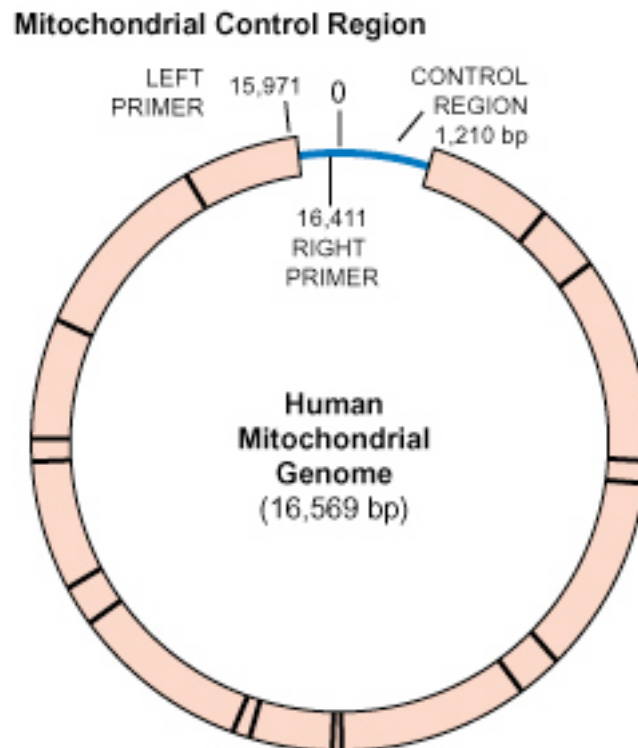
Figure 1. Human mitochondrial DNA. The radial lines in the thick part of the circle divide genes or groups of genes from one another. The thin part of the circle is the "control region." The lines at 15,971 and 16,411 show the 440-bp segment that is amplified in this laboratory. Image copied from the Dolan DNA Learning Center Web site.

This mitochondrial DNA (mtDNA) may be small, but it is important for three reasons:

First, some parts of mtDNA mutate at about 10 times the rate of chromosomal DNA, undergoing a net change of 2.5% per million years. This means that in a mere 10,000 years, we could expect 4 persisting nucleotide changes. This fast mutation rate means that mitochondrial DNA has been used to examine the differences that have developed between human ethnic groups in the

150,000-200,000 years that modern man has existed. It is thought that this high mutation rate results from the mtDNA's exposure to free radicals generated during respiration.

Second, mitochondrial DNA is inherited only from the mother. Only a few sperm mitochondria enter the zygote, and it seems they are destroyed. By contrast, the egg contains 100,000 mitochondria. Therefore, if you don't like your mitochondrial DNA, you are entirely justified in laying all the blame on your mother!

Third, while there are only two copies of any segment of chromosomal DNA per cell, there are many mitochondria per cell, so there may be hundreds or thousands of copies of the short mtDNA sequence in each cell. This large number of copies makes mtDNA relatively easy to isolate and study.

We are now starting a multi-lab project that uses student mtDNA:

a) This week, we will isolate mtDNA from hair follicles of either eyebrow or eyelash hairs.

b) The prep staff will use polymerase chain reaction (PCR) to greatly multiply the number of copies of a 440 bp segment of mtDNA called the control region.

c) After amplification, we will electrophorese each amplified mtDNA sample to make sure that it has a band that corresponds to 440 bp. Only some samples will amplify successfully.

d) The samples that show this band will be sent to the Dolan DNA Learning Center (DNALC) at Cold Spring Harbor Laboratory in New York and sequenced.

e) Once the DNALC posts the sequences on a Web site, we will use bioinformatics tools to compare the student sequences with those of different modern ethnic groups. We will determine if the student sequences are more genetically related to the expected groups (e.g., groups that have the same race and ethnic background as the students) or to unexpected ethnic groups. We will also look for signs of the evolutionary history of humans in mtDNA sequences from across the world.

## Procedure A—Extraction of mtDNA from Hair Follicles

1. Each student should obtain an Eppendorf tube that contains Chelex and an enzyme called proteinase K. The proteinase K will be used to release hair follicle cells from the hair shafts. The Chelex ties up metal ions that might interfere with the PCR reaction later. Your instructor will give you a code letter. Write this letter with a permanent pen on the top of your tube. If you are in section 4 and you are student B, then write "4B" on your tube. Keep this tube on ice as much as possible. This will keep the proteinase K enzyme working.

2. Each student should pull out one or two hairs either from the eyebrows or the eyelashes. Eyelash hairs tend to pull out more easily.

3. Use the dissecting scopes and a razor blade and cut off the hair shaft just above the bulbous, glassy end of cells from the hair follicles. Do this cutting on an index card rather than on the glass stage of the dissecting scope. Discard the hair shafts.

4. Use a forceps to put the hair follicles in the Eppendorf tubes. A *wet* forceps (perhaps wetted with the solution in the proteinase K tube) works better for picking up the follicles. Make sure that the hair follicles are immersed, not stuck to the side of the tube.

5. Put your tube (along with others) in a floating Eppendorf tube rack in the 37° C water bath for 10 minutes. During this period, the proteinase K is digesting away a membrane that holds the cells on the hair shaft.

6. After 10 minutes, vigorously vortex your tube for 15 seconds. This will disperse the cells away from the hair shaft. Keep a tight grip on the tube as you vortex it and keep the cap tightly closed.

7. Put your tube in a 100° C heat block for 8 minutes. This will boil the solution, lyse the cells, and release their DNA. The tops of the tubes tend to pop open due to the boiling. Let them pop open once, and then close them and put a weight (such as a microfuge tube rack) on top of them to keep them closed. Remove the tube with forceps and let it cool for 2 minutes.

8. Vortex the tube for 15 seconds again.

9. Give your tube to your instructor, who will collect a group of tubes and then centrifuge them at 14,000 rpm for 30 seconds.

10. When you get your tube back, get a "20/200 µL" micropipette (NOT a "2/20 µL") and make sure it is set to "150/0." Remove 150 µL of supernatant. Don't disturb the pellet. Put this 150 µL in a clean 1.5 mL centrifuge tube. Write your code letter from before plus your section number on the top of this tube. For example, if you are student B in section 4, you would write "4B" on your tube.

11. Give your tube to your lab instructor, who will arrange to have the DNA amplified by PCR.

12. Your instructor will pass around a list. Write your name next to your code letter. Also summarize your *mother's* ethnic background next to your name (e.g., "Irish," "African American," etc.). Why is only your *mother's* background relevant?

**Procedure B—Electrophoresis of Amplified DNA**

Your mtDNA samples from the last lab have been subjected to 30 cycles of PCR by the lab prep staff. However, all samples might not have amplified successfully. To determine which samples are good enough to send to the DNALC for sequencing, we must perform electrophoresis on a portion of each sample. If 440 bp segment of the mtDNA was amplified, a 440 bp band will be obvious when the sample is electrophoresed.

1. Your instructor will cast one or two gels using the agarose in the water bath. Remember that this agarose has ethidium bromide in it, and you must wear gloves when handling it or loading DNA into it. When the gel has solidified, remove the comb.

2. Your instructor also will put out a test tube rack on ice with an Eppendorf tube for everyone in the class. Each tube is marked with your code letter from the previous exercise and contains 25 µL of PCR product. You will remove only 12 µL from your tube and load it onto this gel. Put your Eppendorf tube back into the test tube rack when you've withdrawn your 12 µL, because we plan to use the 13 µL that remains. Your instructor will load one size standard lane. As you load your DNA, take care not to puncture the bottom of the well with the micropipet tip. Students will load their DNA in code letter order as follows:

negative electrode

positive electrode
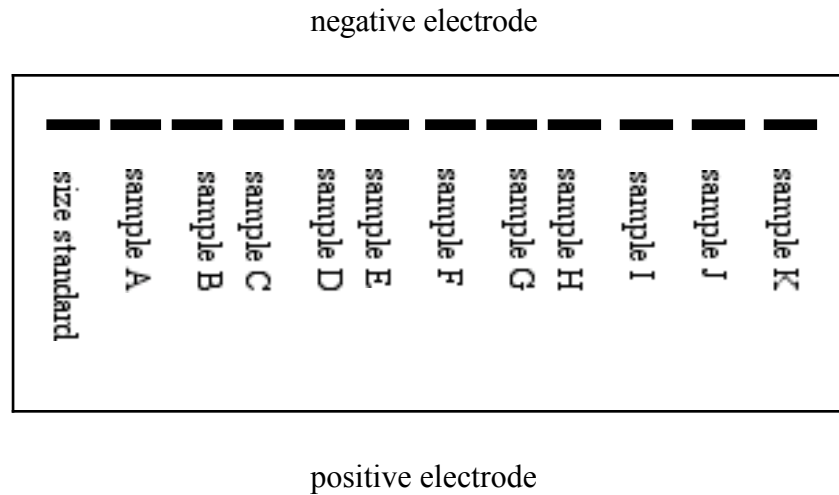
Figure 2. Pattern for loading gels with PCR products when using a 12-lane comb. The heavy lines represent the wells. A second gel would start with the letter following the last letter on the first gel ("L" in this case).

3. When all samples are loaded, your instructor will run the gel at 115-120 V for 20-30 minutes and then put it on a UV transilluminator.
4. When looking at your gel on the transilluminator, YOU MUST WEAR SAFETY GLASSES! The size standard lane should contain bands of 1,857 bp, 1,058 bp, 929 bp, 383 bp, and 121 bp. Optimally, in all lanes, we'll see bands of 440 bp, almost aligned with the fourth band (383 bp) in the size standards lane. If any lane is empty, there's no amplified DNA in that sample so we can't send that sample to the DNALC. For example, in the gel below, three samples amplified successfully, but one did not.
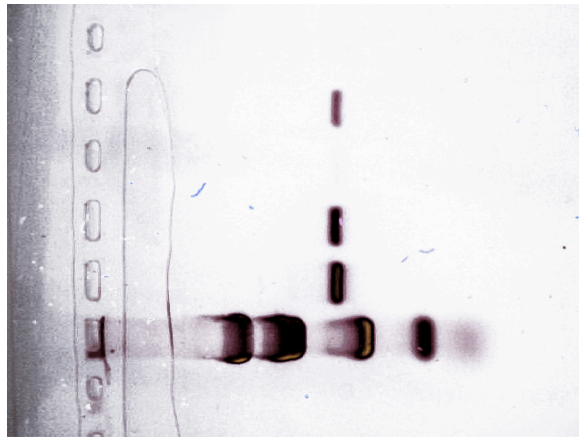
Figure 3. A gel confirming the presence of the PCR product. Size standards are in the bottom lane and samples A, B, and D show strong bands at 440 bp, indicating successful amplification. Sample C did not amplify. Photo has been color-altered to reproduce better in black and white.

5. If there's a band that's much smaller than 440 bp, this is just a PCR artifact called a "primer dimer," and, by itself, does not indicate successful amplification.

6. See if your mtDNA amplified. If it did, your instructor will note that fact on the list of samples.
7. Your instructor will place the gel in a Ziploc bag and write identifying information on the bag. He/she will also take a digital picture of the gel because the DNALC requires photographic proof that a sample contains amplified DNA before sequencing is attempted.
8. The DNALC will post the sequences on a Web site, and we will examine them in a future lab.

Procedure C—mtDNA and Human Ethnicity

     If you were assigned this exercise, the Dolan DNA Learning Center has sequenced at least some of the mtDNA samples that were extracted in a past lab. Not all mtDNA samples can be successfully sequenced. We are going to look at the results and determine how much we can tell about the ethnicity of a person from the 440 bp sequence of his/her mtDNA control region.

1. Boot up your computer and go to EBI's ClustalW site (http://www.ebi.ac.uk/clustalw/index.html). This site will be used to align and compare some mtDNA sequences that you will download in the next step. Note that the site has an empty text box. This is where you will paste your sequences.
2. Go to http://biology.clemson.edu/bpc/bp/Lab/111/mtDNAseqs.htm and download the Word document entitled, "Clemson mtDNA." The first part of this document contains a table with a maternal ethnic description of each student whose mtDNA was successfully sequenced. The bulk of the document consists of the sequences themselves.
3. From the list, choose one student with a European background, and another from a non-European background. If you are one of the students whose DNA was successfully sequenced, your instructor will tell you this and the number of the sample that contains your sequence. You'll probably want to use this as one of your two samples. If you have a choice, avoid poor-quality sequences. These will have a lot of "N" symbols, meaning that the nucleotide at that position could not be determined. They might also be longer or shorter than the good sequences.
4. Copy your first student sequence onto your clipboard. Be sure to include the FASTA format identifier line (e.g., ">Student_49"). Go back to your browser. Use your browser's back button and go back to ClustalW. Paste the first student sequence into the text box.
5. Go back to Word and copy your second student sequence, including the identifier line. Go back to ClustalW and paste in this sequence under the first one.
6. Now that you've chosen good student samples, you will determine how well the samples compare to a wide variety of mtDNA sequences from across the world. Go to http://biology.clemson.edu/bpc/bp/Lab/111/mtDNAseqs.htm again and download the file "World mtDNA." This has 65 modern human sequences from across the world, plus a fragment of a Neandertal sequence, plus sequences from a chimpanzee and a baboon. These samples have been edited to cover exactly the same mtDNA bases as our student samples (about 400 nucleotides). Most samples are labeled with their country. Some unfamiliar group names may be Kirgiz and Uzbeks (north of Afghanistan), Berbers (Algeria and Morocco), Biakas and Mbenzeles ("pygmies" from the Central African Republic), Mbuti ("pygmies" from the Congo), San and Khwe ("bushmen" from southern Africa), Chukchis, Koryaks and Inuit (native peoples of Siberia), and Pima and Navajo (native peoples of the US Southwest). These "World mtDNA" sequences come from the Human Mitochondrial DNA Database maintained by Uppsala University, Sweden, and are more accurate than the student sequences.

7. In the "World mtDNA" document, select all these samples (from ">English" to the end of the baboon sequence). Paste these into the text box following your two student samples. Just above the text box, change Output Order from "Aligned" to "Input." Press Run.

8. When the results arrive, the first screen will be a multiple alignment, starting like this:

```
CLUSTAL W (1.82) multiple sequence alignment


Student_67              -CAGTAGATCTAATTTAAACTATTCTC-TGTTCTTTCATG--GGGAAGCAGATTTGGGTA 56
Student_25              AAGTAGATTCTAATTTAAACTATTCTC-TGTTCTTTTCCC-NANAAGNGGATTTGGGTG 58
English                 GCTAAGATTCTAATTTAAACTATTCTC-TGTTCTTTCATG--GGGAAGCAGATTTGGGTA 57
French                  GCTAAGATTCTAATTTAAACTATTCTC-TGTTCTTTCATG--GGGAAGCAGATTTGGGTA 57
```
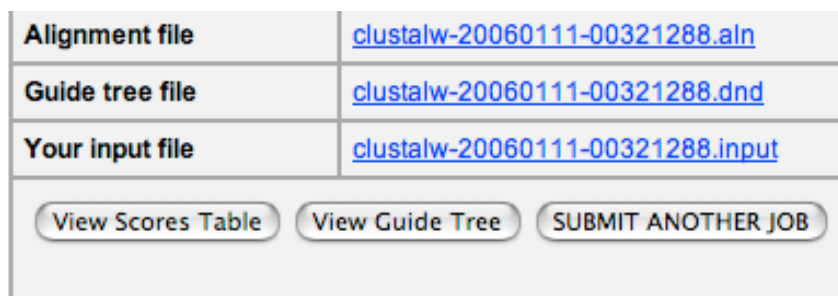
Figure 4. A part of a multiple alignment of two student samples with the "World" samples.

Scan up and down this alignment and notice gaps that are introduced by the inclusion of the chimpanzee and baboon, and how many times nucleotides are the same for all the human samples, but then change for the chimpanzee and baboon.

9. Return to the top and look for the button called "View Scores Table," at the bottom left in this picture. Click on this button.

Alignment file	clustalw-20060111-00321288.aln
Guide tree file	clustalw-20060111-00321288.dnd
Your input file	clustalw-20060111-00321288.input

(View Scores Table) (View Guide Tree) (SUBMIT ANOTHER JOB)

Figure 5. Where to find the button that gives you the scores table.

10. Scan down the list of similarity scores for your first student mtDNA sample. Similarity here is relative because the student samples probably have sequencing errors. Just look for the highest scores and the lowest scores for your student sample. Knowing the sample's true ethnic background, answer these questions (yes or no) for the European student sample:
    a) Do you see any tendency for people with similar ethnic background to have a higher

        similarity score with that sample? _____

    b) Do people from dissimilar backgrounds have lower similarity scores? _____

        Don't consider the chimp, baboon, and Neandertal sequences here. Consider only the

        modern human sequences.

11. Now answer the same questions for the non-European sample:

    c) Do you see any tendency for people with similar ethnic background to have a higher

        similarity score with that sample? _____

    d) Do people from dissimilar backgrounds have lower similarity scores? _____

        Again, don't consider the chimp, baboon, and Neandertal.

12. The student samples may have sequencing errors. The "World mtDNA" samples are much more accurate. Using the browser's "Back" arrow, go back to the ClustalW input screen. Delete the student samples (at the top) from the input box. Then go to the bottom and delete the Neandertal, chimp, and baboon sequences.
13. Choose an ethnic group among the remaining "World mtDNA" samples that you want to examine (e.g., "English," "Nigeria_Ibo"). Cut its identifier line and sequence from the input box and paste it at the top, where the student samples were. We do this because the scores of the first sample listed are shown the most compactly. Run ClustalW again.
14. When your results are delivered, press the "Scores" button and look at the percent similarity scores for your ethnic group. Does this group seem to have higher similarity to ethnic groups that are similar to it (e.g., to other Europeans or other East Asians)? Or does molecular similarity seem to be unrelated to ethnic group? All the samples are very similar, so in this context you must consider 100% or 99% as highly similar and 95% as very dissimilar. Fill in the blanks below:

    e) Your chosen group _____

    f) Two ethnic groups most <u>similar</u> to yours _____

                                        _____

    g) Two groups most <u>dissimilar</u> to yours   _____

                                          _____

15. Now try a different way of looking at similarity. Go to the top of the scores table and change the "Sort by Sequence Number" to "Sort by Alignment Score":
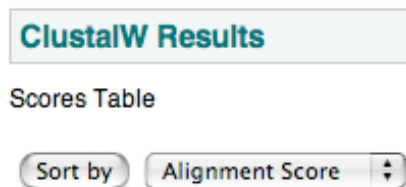
Figure 6. Location of the "Sort by" and "Alignment Score" buttons.

Press the "Sort by" button, and the pairs of samples will be resorted so that all pairs with 100% similarity come first, then pairs with 99% similarity, and so forth.

16. Look at the pairs with 100% identical sequences. Does it seem that superficially similar ethnic

    groups are 100% similar to each other? _____

17. Now scroll to the bottom. What groups often appear at the bottom of this list because they are

    most dissimilar to many other samples?

    _____

    _____

    _____

18. Do not quit from ClustalW. You will need it for the next exercise.

**Procedure D—Reconstructing Human Evolution from mtDNA**

Although so far we have only looked at the percent similarities between different mtDNA samples, ClustalW presents analysis of large data sets in a format called a phylogram. Here is a ClustalW phylogram showing the distance relationships between 8 US cities:
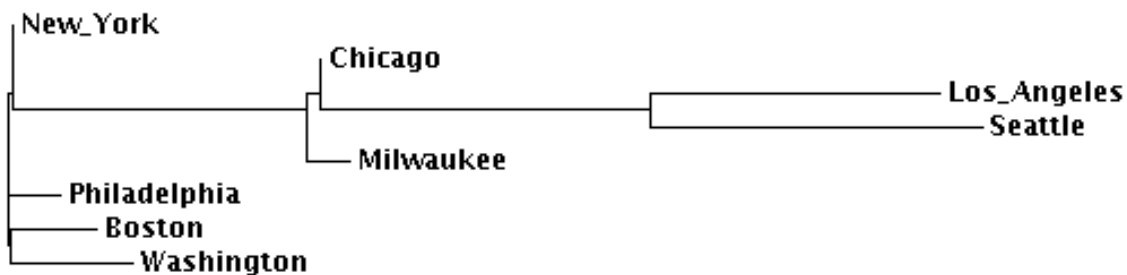

Figure 7. How a ClustalW phylogram presents the distances between certain cities.

New York, Philadelphia, Boston, and Washington are relatively close, so they are all on the left edge of the phylogram. Chicago and Milwaukee are distant from the northeastern cities, but very close to each other (less than 90 miles apart). Finally, Seattle and Los Angeles are very distant from the northeastern cities (about 2,400 miles) and distant from each other as well (about 1,000 miles).

In a phylogram, the sum of the lengths of the *horizontal* branches connecting cities is proportional to the distance between them. Vertical distance or placement doesn't matter. Two examples are shown with the highlighted branches below:
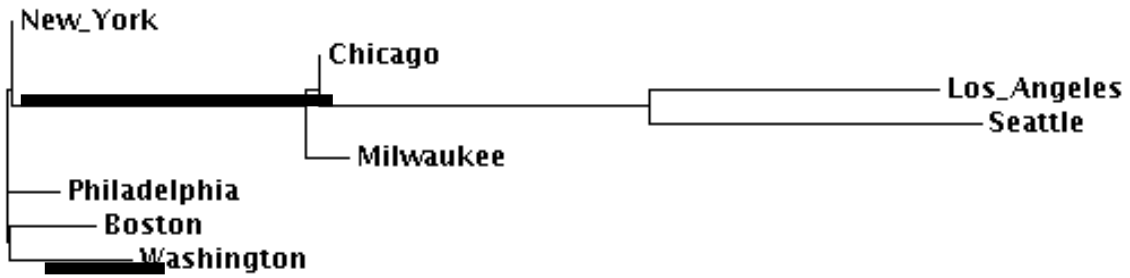
Figure 8. Highlighting shows the phylogram line segments whose summed length is proportional to the distance between Washington and Chicago.
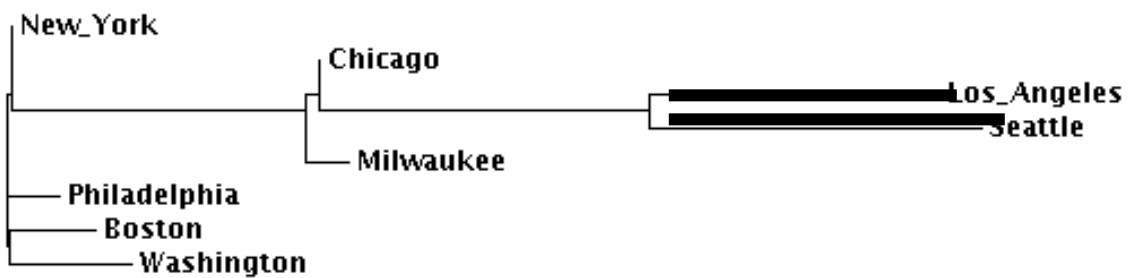

Figure 9. Highlighting shows the phylogram line segments whose summed length is proportional to the distance between Seattle and Los Angeles.

Distance between cities is analogous to genetic dissimilarity. Long horizontal branches also separate DNA samples with different sequences. You will now generate a phylogram that includes only the modern humans, and we will attempt to explain the patterns revealed.

1. Go back to ClustalW, where you were looking at the similarity scores between your chosen "World mtDNA" ethnic group and the other groups included in the "World" data set. Only the modern mtDNA samples (not the chimp, baboon and Neandertal) should have been included in your analysis.

2. Look at the top of the similarity scores, and click on the button "Back to Main Result."
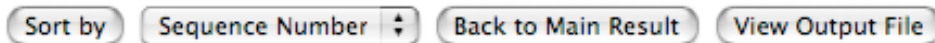
Figure 10. Buttons at the top of the ClustalW scores page.

3. Select and copy onto your clipboard everything from the label "Clustal W (1.83)" down to the end of the alignment section of the output:

```
CLUSTAL W (1.83) multiple sequence alignment


English                   AAACTATTCTCTGTTCTTTCATGGGGAAGCAGATTTGGGTACCACCCAAGTATTGACTCA 60
French                    AAACTATTCTCTGTTCTTTCATGGGGAAGCAGATTTGGGTACCACCCAAGTATTGACTCA 60
```

```
Aust_Aborigine1        GATGACCCCCCTCAGATAGGGGTCCCTTGACCACCATCC 398
Aust_Aborigine2        GATGACCCCCCTCAGATAGGGGTCCCTTGACCACCATCC 398
                       **********************  ******************
```

Figure 11. The top and bottom of the region to be selected.

4. Use your browser's back arrow to go back to the ClustalW input screen. Without disturbing the text on your clipboard, delete everything in the sequence input box. Paste in the contents of your clipboard. These mtDNA sequences have now been aligned, and a second run with ClustalW is going to allow the program to make a phylogram using these aligned DNA sequences.

5. Change the ClustalW controls as follows: ask for an "nj" (neighbor-joining) tree and turn "Phylogenetic Tree Correct Distance" to on. The last line of controls should look like this:

OUTPUT			PHYLOGENETIC TREE		
OUTPUT FORMAT	OUTPUT ORDER	TREE TYPE	CORRECT DIST.	IGNORE GAPS
aln w/numbers	input	nj	on	off

**Figure 12.** Settings for ClustalW controls before second ClustalW run.

6. Press "Run" and wait for the output to be returned. Then go down to the *very* bottom of the output and click on the button "Show as Phylogram Tree." Scroll down to the bottom of the page to see the large phylogram.

7. After looking at the results, consider the following theory explaining the evolution of modern humans:

> Modern humans arose in Africa about 170,000 years ago, so originally, we were all Africans. Some of these Africans remained in Africa and continued their genetic diversification. Perhaps about 70,000 years ago, some other African groups left Africa and founded populations in the Middle East, Europe, and Asia. North and South America were colonized much later from eastern Asia. If this is true, we should expect that
>
> a) Africans will be the most genetically diverse modern humans because they have been evolving the longest;
> b) Europeans and Asians (founded more recently) should be less diverse than Africans;
> c) Some Africans (those who did not found populations on other continents) should be genetically different from Europeans and Asians;
> d) Other Africans (the descendants of the out-of-Africa migrants) should be markedly more similar to Europeans and Asians;
> e) Native Americans should be genetically similar to East Asians.

8. For each of the five predictions above, answer "Yes" if you think the ClustalW phylogram supports the prediction and "No" if it does not. Remember, the sum of the horizontal branch lengths between any two groups shows the genetic difference between them

a) _____ Africans are the most genetically diverse.

b) _____ Europeans and Asians are less diverse.

    c)   _____   Some Africans are markedly dissimilar to Europeans and Asians.

    d)   _____   Other Africans are similar to Europeans And Asians.

    e)   _____   Native Americans are most similar to East Asians.

9. If ClustalW draws a phylogram with the chimp and baboon included, the much different pattern in Figure 13 is seen:
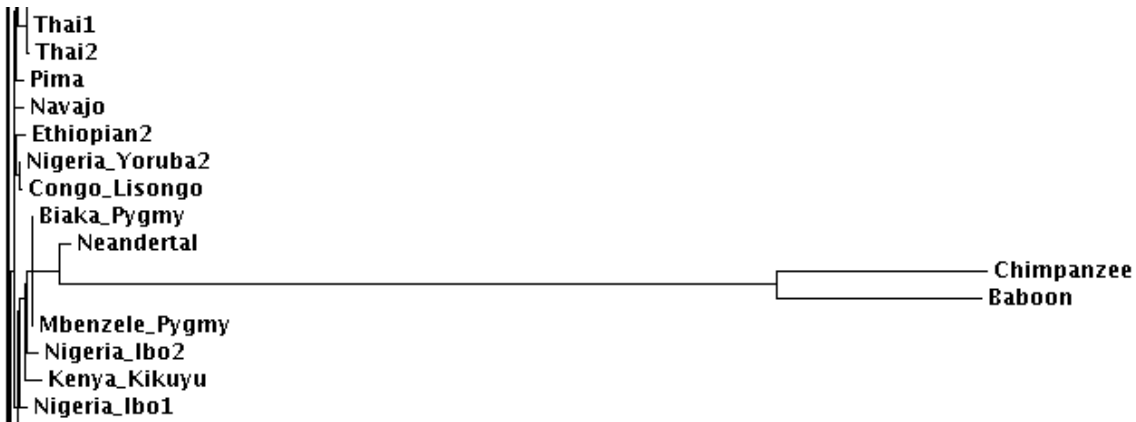
Figure 13. A fragment of a ClustalW phylogram that includes the chimpanzee and baboon.

Why do you think this happens when the two non-human primates are included?

       What does this indicate about the genetic distance between humans and their primate "neighbors" as compared with the genetic difference between human ethnic groups?

Notes for Instructor
Extraction, Sequencing, and Analysis of mtDNA

Before they do this laboratory at Clemson, our introductory biology students have completed a laboratory on bioinformatics. We use this laboratory to make DNA and bioinformatics "personal." Students extract their own mitochondrial DNA (mtDNA) from hair follicles, and it is sequenced at the Dolan DNA Learning Center (DNALC) at Cold Spring Harbor Laboratory. Then the students use their bioinformatics skills to determine if their mtDNA is most similar to mtDNA of expected ethnic groups…or whether their mtDNA will surprise them.

As implemented at Clemson, this laboratory extends over several laboratory periods:

Period 1: extraction of mitochondrial DNA from hair follicles (Procedure A). This is a very short student exercise that we combine with an existing laboratory.

Between periods 1 and 2: PCR amplification of student DNA samples.

Period 2: students and instructor examine PCR products for a strong 440 bp band indicating that amplification was successful (Procedure B). Instructor photographs gels before they fade to furnish photographic proof of successful amplification to the DNALC. This is also a very short procedure inserted into an existing laboratory.

Between Periods 2 and 3: Instructor mails the successful samples and gel photographs to the DNALC. The Center sequences them for free and posts the results on a Web site. Because this takes an uncertain amount of time (probably several weeks), Clemson submits samples at the end of the Fall semester and plans to use the results only at the beginning of the Spring semester.

Period 3: Students use bioinformatics tools to examine the posted sequences (Procedures C and D). At Clemson, this is a small part of a laboratory on phylogenetic taxonomy.

While from the student point of view, this laboratory is not too demanding, it does require much work and strong organization from the instructor. However, it is possible—at Clemson in 2006-2007, we had students prepare 240 samples. Of these, 196 amplified successfully, and all 196 were sequenced successfully by the DNALC.

I will give advice to the instructor based on the five activities list above. Recipes and further details are included in Appendix A at the end of the Instructor's Notes.

Extraction of mtDNA from Hair Follicles

This extraction procedure is derived from one published by Carolina Biological Supply Company's Kits 21-1236/1238.

I will assume that the average lab section has 24 students that are sitting at four lab benches. Recipes for all reagents are included in Appendix A. Where N is the total number of students whose DNA is going to be sequenced, the instructor or preparator should prepare N sterile, capped, 1.5 mL Eppendorf tubes and N Eppendorf tubes containing 300 µL of a 1:1 mix of 10% Chelex and 100 µg/mL proteinase-K. Store these tubes at –20° C. Each section should be provided with one Eppendorf tube rack for its sterile tubes and another rack (stored on ice) for its Chelex-Proteinase-K Eppendorf tubes. The room should have a 37° C water bath with a floating rack that will hold the Eppendorf tubes, a 100° C heating block with openings that can fit the Eppendorf tubes, a microcentrifuge capable of achieving 14,000 rpm for 30 seconds, four vortexers so students can

reach them easily, 4 micropipettes that can deliver 150 µL, five small coolers filled with ice, and four boxes of sterile tips for the micropipets. There should also be razor blades, forceps, small index cards, dissecting scopes for cutting the bulbs from the eyelash hairs, and fine-tipped Sharpies for writing the all-important code letters on the Eppendorf tubes.

One lesson that we learned the hard way in 2006: **Do not try to save primers and proteinase K from year to year. Order new supplies every year.**

Following the student outline will result in an adequate amount of DNA being released from the hair follicle cells. At Clemson, our biggest problem with this stage was that the students resisted using the dissecting scopes and razor blades to cut the bulb from the hairs. Instead they just threw in uncut eyelashes. This held the hair bulbs and their cells above the digestion mixture and resulted in a failure of amplification.

The instructor must issue code letters to the students so their samples can be tracked. At Clemson, the student who was sample B in lab section 4 would be coded as 4B. From the beginning, the instructor must keep a list of these codes and the names that correspond to them. Also have the student note the ethnicity of their *mother* next to their name. The blank form in Appendix B can be used for this record. At Clemson we've found that many students might put down just "Caucasian" or "American." At first, we thought this was a protest against recording ethnicity, but now it seems that many white Clemson students come from extremely mixed northwestern European backgrounds and simply don't know their ethnicity.

It is also worthwhile to create 2-3 different samples from yourself. Unfortunately, the DNALC sometimes reports different results for replicate samples The instructor samples can be used to check the consistency of the sequencing later. In 2006, the senior author submitted 16 samples from himself, and their similarity between pairs of these samples (which should all be 100%) ranged from 99% down to 58%.

PCR Amplification

Every student sample must be subjected to PCR to amplify the 440 bp segment of the mtDNA control region. We used Ready-To-Go PCR Beads (Amersham Biosciences 27-9559-01). Directions for PCR are given in Appendix A, and amplification directions are also given on the DNALC Web site at http://www.geneticorigins.org/geneticorigins/mito/mitoframeset.htm. Go to the second page of the "Laboratory" link. We added 22.5 µL of a primer/Cresol Red loading dye mix to a tube with a bead, and then added 2.5 µL of student DNA. Briefly spin each tube in a minicentrifuge to pool all this material in the bottom of the tube and to dissolve the bead. Write the student code (e.g., "4B") on the top of the tube with a Sharpie. Then run in a thermocycler for 30 sec at 94° C, 30 sec at 58° C, and 30 sec at 72° C. This is repeated for 30 cycles. With our thermocycler, this process lasted 1 hour and 20 minutes. Finally, pool all the tubes from one lab section in an Eppendorf tube rack and freeze them until the students will electrophorese them.

If electrophoresis (next step) discloses that not many samples amplified, the problem could be the amount of DNA added to the tubes. We have found that too much DNA can cause PCR failure just as readily as too little DNA can. The student tubes have 150 µL of solution, and each cycle of PCR only uses up 2.5 µL of this. Therefore, if the samples don't amplify, try adding 0.5 µL, 1.0 µL, 2 µL, and 5 µL to the PCR tubes for a small subset of the samples and try PCR again. It may be that a simple adjustment of the amount of DNA added will solve the problem.

Electrophoresis and Photographing Gels

The PCR products produced above will be electrophoresed on 2% agarose gels containing 0.5 µg/mL ethidium bromide. While using a two-comb gel seems as if it might be more efficient, for unknown reasons we had better luck in 2006 with a one-comb gel. We recommend using one comb per gel, so two gels will be needed for a lab section of 20-22 students. As soon as the students have finished loading their gels, the PCR microfuge tubes (still containing 13 µL of PCR product) should be returned to the freezer. The size standards are pBR322-BstNI, and 5 µL should be loaded into one lane of each gel, as shown in Figure 2. The electrophoresis should run for 20-30 minutes at 115-120 volts, just long enough for the cresol red dye front to move 50 mm from the wells.

The bands are visualized with a UV transilluminator. If there are two gels, it would help to have two transilluminators at opposite ends of the lab room. Set up the gels and ask the students to file past the one with their code letter and look for a band in their lane. If the 440 bp band exists in a lane, it will be obvious. Ask the students to look at Figures 2 and 3 to remind them where their band will be and how successful amplification looks. Of course, the students must wear safety glasses while they look at the gels.

In our experience, the 440 bp bands are sometimes faint, and by an hour after the electrophoresis, they will have faded to invisibility. The DNALC will not sequence a sample without photographic evidence that it has an amplified 440 bp segment. Therefore, it is important for the instructor to photograph the gels immediately. We have found that the best procedure is to do the photography in a room that can be darkened. Fix a digital camera on a tripod pointing down at a transilluminator between the legs of the tripod. Photography can be done with the gel in a Ziploc bag. First, visually decide which lanes show 440 bp bands. Then write the section number and these results on a paper towel or other paper (e.g., "Sec. 4, B C D H," meaning that this gel from section 4 had visible bands in lanes B, C, D, and H). Lay the paper towel on top of the gel bag, and take a picture of these notations with the lights on. Then remove the paper towel, turn off the lights, make sure the UV shield is in place, turn on the transilluminator, and take a picture of the bands. As long as you always take the paper towel picture first and any gel pictures later, you will be able to pair up any picture with its section, and you'll know the lanes that appeared to have bands when the gels were fresh.

It may be necessary to experiment with the exposures to find the best setting for your camera. We found that shortening the exposure (-0.3 to –0.7 stops) gave better results. Also, our transilluminators each had a transparent shield to protect users from UV light. We found that our autofocusing camera sometimes focused on the shield rather than on the gel. Before turning on the transilluminator, we removed the shield, turned it upside down, and laid it on the Ziploc bag containing the gel. This also flattened the bag against the gel.

Don't forget to take the 2-3 samples of *your* mtDNA through all these processing steps. As explained above, these can be used to check on the consistency of the sequencing.

Use these pictures to decide the samples for which you're going to claim successful amplification. Only these samples will be sent to the DNALC.

Mailing Samples to the DNALC

The DNALC Sequencing Service has a set of directions on how to submit samples at http://www.geneticorigins.org/geneticorigins/mito/mitoframeset.htm. Click the "Sequencing *Service*" (not "Sequencing *Server*") link. The DNALC asks you to submit a "Sequencing Form" online before sending the samples, and then directs you to submit the samples frozen in sealed, clean, 0.2-0.5 mL PCR test tubes. Each sample should contain 10-15 µL of PCR product. If you had N samples that successfully amplified and showed 440 bp bands, the samples must be sequentially numbered from 1 to N. This means that the old notations like "4B" are no longer relevant. The successful samples must be numbered sequentially, with no numbers skipped. The DNALC repeats this directive about strictly sequential numbers several times. IMPORTANT: You should note which numbers correspond to which samples on your original list. You will probably be sending fewer samples than you created. At Clemson, only about 50% of the student samples amplified successfully the first year we tried this exercise, and 82% amplified the second year.

Also include the photos of the gels, clearly labeled with these sequential numbers so the DNALC can tell which lanes represent which numbers.

Interpreting the Sequences

The DNALC says that it will post results in 10 days to 3 weeks after the samples are sent. Because of this uncertain schedule, Clemson submits samples at the end of the Fall semester and doesn't plan on using the results until the beginning of the Spring semester. In 2005 and 2006, almost all samples we sent resulted in useable sequences (although not necessarily error-free sequences, as explained below).

The DNALC will send you an e-mail when the sample sequences are posted on a Web site. At Clemson, we immediately download these sequences to a Word document so we don't depend on the continued availability of the site. Then the instructor should go back to the master list that gave the maternal ethnic background of every student. This information is used to produce a table that might begin:

Table 1. Maternal ethnic background of 2006 mtDNA samples.

1	German, Irish, English
2	Irish, Scottish
3	African-American
4	English, Spanish
5	Native American (Mohawk)
6	German
7	English

Then, after the list, the student sequences are listed in FASTA format as follows:

>Student_01
CAGTAGATTCTAATTTAAACTATTCTCTGTTCTTTCATGGGGAAGCAGATTTGGGTACCA
CCCAAGTATTGACTCACCCATCAACAACCGCTATGTATTTCGTACATTACTGCCAGCCA
CCATGAATATTGTACGGTACCATAAATACTTGACCACCTGTAGTACATAAAACCCAAT
CCACATCAAAACCCCCTCCTCATGCTTACAAGCAAGTACAGCAATCAACCCTCAACTAT
CACGCATCAACTGCAACTCCAAAGCCACCCCTCACCCATTAGGATACCAACAAACCTAC
CCACCCTTAACAGCACATAGTACATAAAGCCATTTACCGTACATAGCACATTACAGTCA

AATCCCTTCTCGTCCCCATGGATGACCCCCCTCAGATAGGGGTCCTGCNAANNCCNANN
TTNCNTNAA

>Student_02
CAGTAGATTCTAATTTAACTATTCTCTGTTCTTTCATGGGGAAGCAGATTTGGGTACCAC
CCAAGTATTGACTCACCCATCAACAACCGCTATGTATTTCGTACATTACTGCCAGCCAC
CATGAATATTGTACGGTACCATAAATACTTGACCACCTGTAGTACATAAAAACCCAATC
CACATCAAAACCCCCTCCCCATGCTTACAAGCAAGTACAGCAATCAACCCTCAACTATC
ACACATCAACTGCAACTCCAAAGTCACCCCTCACCCATTGGGATACCAACAAACCTACC
CACCCTTAACAGTACATAGTACATAAAGCCATTTACCGTACATAGCACATTACAGTCAA
ATCCTTCTCGTCCCCATGGATGACCCCCCTCAGATAGGGGCCTGACNANCCNNNNTTT
TCTCNAA

FASTA format is a way of identifying nucleotide or amino acid sequences in a way that can be read by bioinformatics software. The identifier follows a ">" and has no spaces in it (e.g., ">Student_01"). Then the sequence, as one long line copied from the Dolan Web site, follows on the next line. The many "N" symbols in the samples above stand for "nucleotide," and indicate positions at which A, G, T, or C could not be determined.

At Clemson, we place this document containing the ethnic backgrounds and sequences on a local Web site, from which it is downloaded during the analysis lab.

We ask the students to work with a sample from a student from a European ethnic background, and also a non-European background. If the student's own sample was sequenced successfully, the student will probably want to work with that one as one of his/her samples.

We used to ask them to run a BLAST search, which easily identifies the samples as human mitochondrial DNA. However, BLAST also gives an excessive number of "hits," few of which are identifiable as particular ethnic group. Now we ask them to use EBI's ClustalW to compare their samples with "World mtDNA," a file of 65 modern sequences from across the world. This file also includes Neandertal DNA, plus mtDNA from a chimpanzee and a baboon. The "World" sequences came from the Uppsala University (Sweden) Human Mitochondrial Genome Database, found at http://www.genpat.uu.se/mtDB/. While the sequences in the Uppsala database are all complete mitochondrial genomes, the sequences in the "World mtDNA" file have been truncated to the region sequenced by the DNALC.

We ask the students to do two analysis exercises in Procedure C:
   a)   They attempt to determine if their student sample tends to be more similar to the mtDNA of expected, related ethnic groups;
   b)   They choose a modern human sequence from the "World mtDNA" files, and see which "World" sample is most related to their chosen "World" sample.

The conclusion from the first analysis exercise will probably be that the student samples are clearly different from chimp and baboon sequences. Then things get less clear. Some students find that they are most similar to expected ethnic groups. Some find unexpected similarities. Finally (because of sequencing errors, we hope), some students are more similar to the Neandertal sequence than they are to any of the modern sequences!

There are four reasons for these inconsistent results. First, even quickly changing mtDNA is remarkably similar across ethnic groups. About 79% of all bases are identical in all 65 modern ethnic groups of the "World" samples. This can be used to teach a lesson about the relative

insignificance of genetic differences even between ethnic groups that look very dissimilar. Second, we're analyzing about 2.5% of the whole mitochondrial genome. Third, while there are specific polymorphisms that are found more often in certain groups, this exercise is only looking at overall similarity. Fourth, a major reason is that the DNA Learning Center sequences include mistakes. As previously mentioned, the senior author included 16 samples from himself in 2006, and pairs of these samples ranged from 98% similar to only 58% similar. Of course, every Kosinski sample should have been 100% identical with every other Kosinski sample.

Before the potential user despairs over these errors, tests with corrupted sequences show that random errors will introduce noise but not bias into the conclusions. We took an English sequence and corrupted a random 10% of its bases to N. Then we compared this to the World mtDNA sequences. It was about 88% similar (rather than 100% similar) to the uncorrupted English sequence, but it was less than 88% similar to all the other sequences. In other words, in relative terms, it was still most similar to the most related sequence.

The second exercise, in which the students work only with the more accurate "World mtDNA" samples, gives more reliable results. Is there a tendency for mtDNA between related ethnic groups to be similar? Yes and no. One Italian sample was 100% similar to the other Italian, but 99% similar with a Japanese sample and a sample from a Siberian Koryak. A Nigerian Yoruba was 99% similar to two tribes in neighboring Cameroon, but also 99% similar to French, Spanish, and Finnish samples. A Japanese sample was most similar to other Japanese, but also to an Indian (from India) and to a sample from a native tribe of Paraguay. The results are both expected and hard to explain. Overall, mtDNA sequences do tend to be similar to sequences in related ethnic groups. However, they are also sometimes similar to samples in unrelated ethnic groups as well. At least in its coding regions, mtDNA seems to be subject to convergent evolution as independent groups get the same mutation, giving a false appearance that they had a recent common ancestor (Pakendorf and Stoneking, 2005). Ethnic group cannot be safely predicted from mtDNA control region sequence.

The Pattern of Human Evolution

The "World mtDNA" compilation can be used for a more wide-ranging exploration of human evolution. Procedure D uses ClustalW to construct a phylogram using all modern human samples. This requires two ClustalW "runs." The first run completes a multiple alignment of the samples. The second run then constructs the phylogram from the aligned sequences.

A phylogram is a diagram that portrays relative distances or differences between different entries. This is shown using cities in Figures 7-9 in the student outline. Those figures show Seattle and Los Angeles as "outliers" only because most of the cities on the phylogram are not close to Seattle and Los Angeles. If the phylogram had had several California cities plus New York, then New York would have been depicted as the outlier at the end of a long branch. The position of groups on branches that extend far to the right (e.g., the Kikuyu and San samples) is simply a statement of their dissimilarity with respect to the other groups.

Construction of the phylogram from all the modern mtDNA (Procedure D) gives the results in Figure 14.
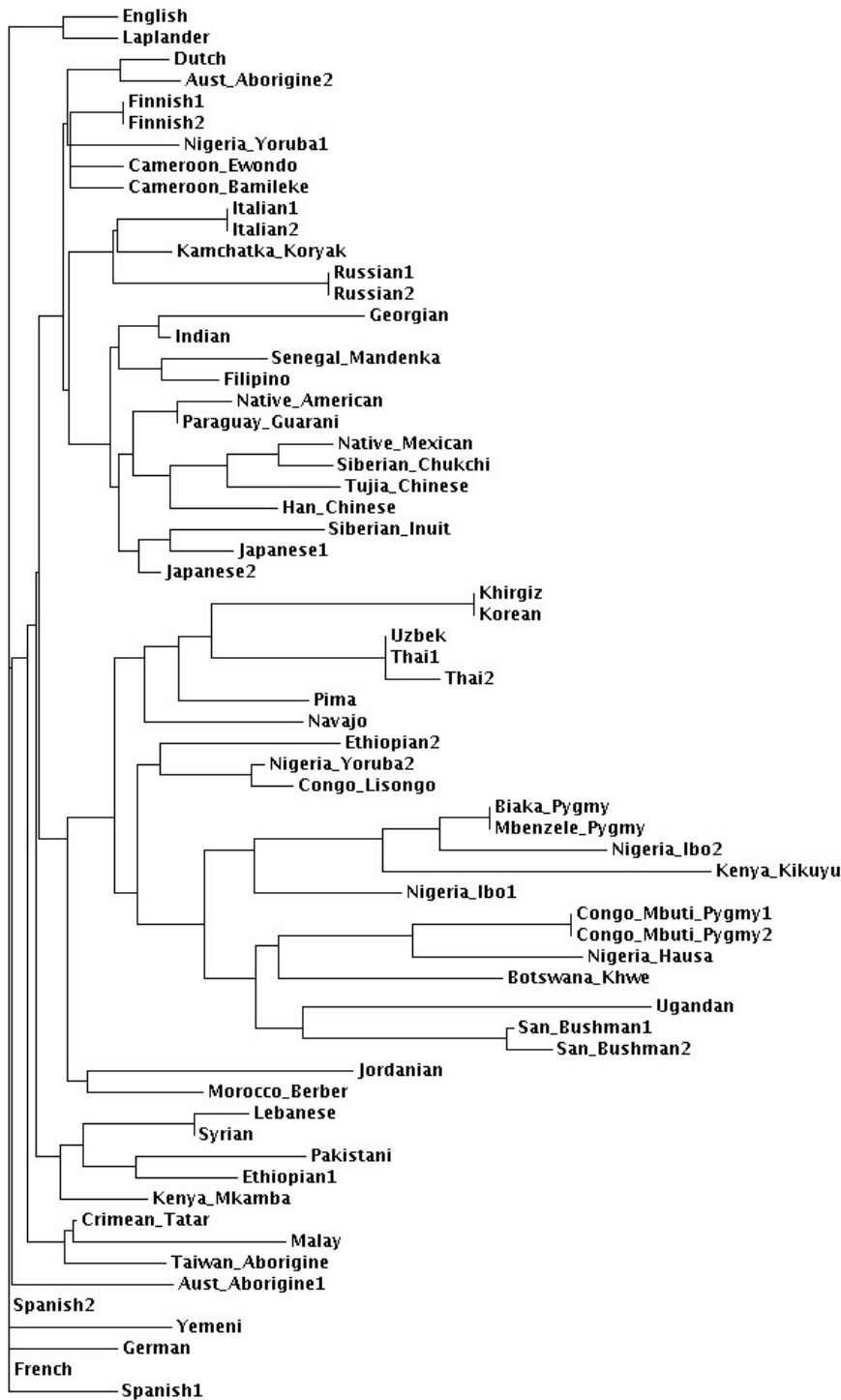
Figure 14. A ClustalW phylogram of all modern human mtDNA samples from the "World mtDNA" data set.

Again, remember that the sum of the lengths of the horizontal branches separating any two groups indicates the difference between those groups. The vertical placement means nothing, and is

an artifact of the way ClustalW's clustering algorithm works. Therefore, the English sample at the top of Figure 14 is very similar to the Laplander sample, and just as similar to the French sample at the very bottom.

The dominant feature of the phylogram is that several African groups (Kikuyus, Biakas, Mbenzeles, Hausas, Mbutis, Ugandans, Ibos, Khwes, and Sans) have long horizontal branches, showing their mtDNAs are markedly different from those of other groups. These groups are also very different from each other. At the same time, other African groups (including the Ethiopians, Lisongos, Mkambas, Mandenkas, Ewondos, and Yorubas) have shorter horizontal branches, showing more moderate differences from each other and from the other groups on the phylogram.

The distinctness of some (but not all) African samples can be better seen if we look at European, Middle Eastern, African, and East Asian samples separately. The Kikuyu sample (the most dissimilar) is included in all of these to ensure that ClustalW's automatic scaling feature produces branches of comparable length.
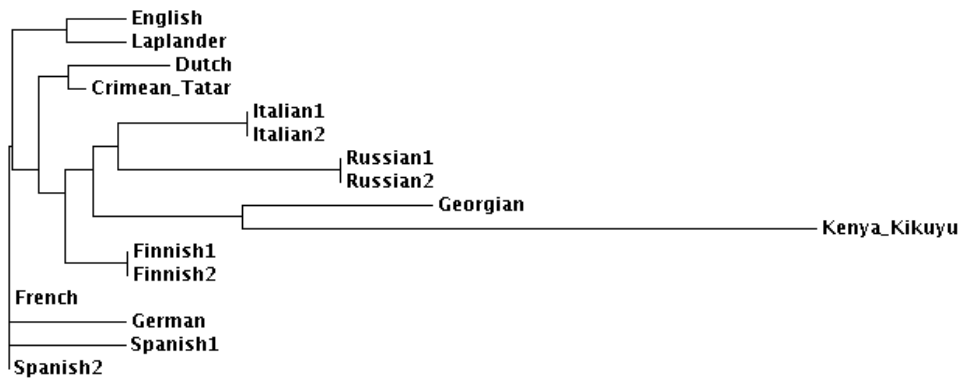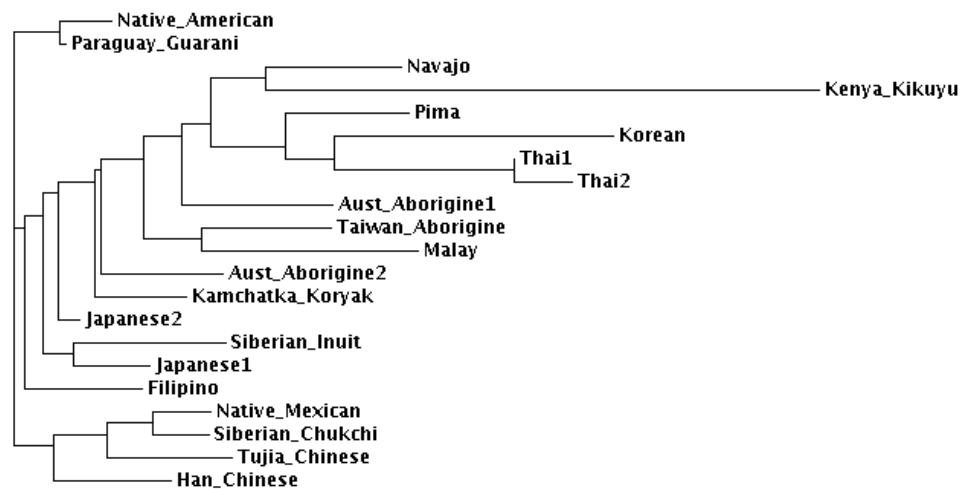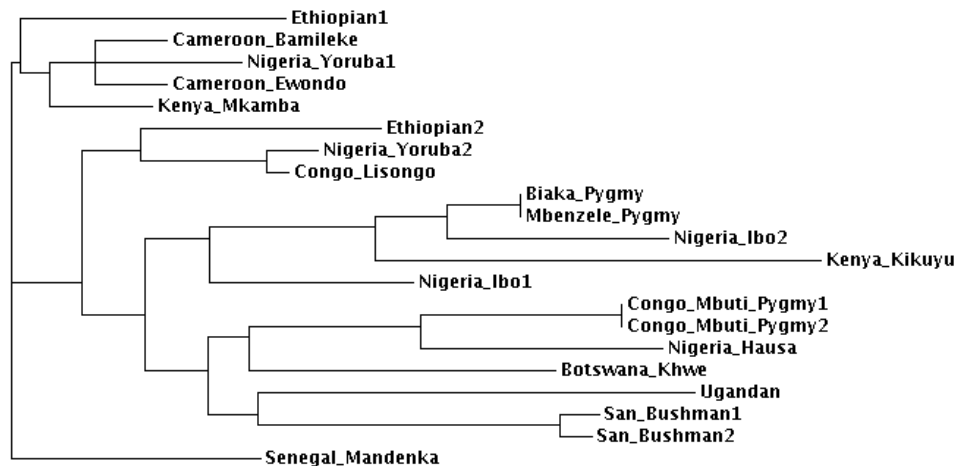

Figure 15. A phylogram of the European "World mtDNA" sequences.

Figure 16. A phylogram of the Middle Eastern "World mtDNA" sequences.

Figure 17. A phylogram of the East Asian and Native American "World mtDNA" sequences.

Figure 18. A phylogram of the African "World mtDNA" sequences.

Note the following overall trends:
  a) In the European, Middle Eastern and East Asian figures, the Kikuyu sequence extends far to the right of any of the other sequences. That is, most of the sequences are similar to each other but dissimilar from the Kikuyu sequence;
  b) In the African figure, there are several sequences that extend almost as far to the right as the Kikuyu sequence. This means that these sequences are quite different from each other (aside from pairs like the two San and the two Mbuti sequences);
  c) There are also several African groups (mostly in the upper half of Figure 18) that have short branches because they are similar to each other. Reference to Figure 14 shows that these sequences are also similar to most European, East Asian, and Middle Eastern sequences.

  A theory to explain these observations is based on the evolutionary history of modern humans. A graphic introduction to the theory is at the Mitomap Web site at http://www.mitomap.org/, and especially at a map of human migration based on mtDNA, found at http://www.mitomap.org/WorldMigrations.pdf. The theory was reviewed in detail by Forster (2004).

Basically, modern humans originated in Africa about 170,000 years ago (Forster, 2004; Ingman *et al*., 2000; Ingman and Gyllensten, 2001; Merriwether *et al*., 1991; Pakendorf and Stoneking, 2005). By about 80,000 years ago, there were three mtDNA variants present in Africa; these are called haplogroups L1, L2, and L3. As Forster (2004) summarizes the evidence, L1 is a small group that includes the San (Bushmen) and the Biaka (West Pygmies). L2 and L3 became more widespread in Africa. More detail on the diversification of the L haplogroups is given by Torroni *et al*. (2006). However, the only haplogroup to make a successful colonization of Eurasia was L3, starting about 57,000 years ago (Forster *et al*., 2001; Forster, 2004). Australia was colonized about 40,000 years ago, and the Americas were colonized from northern Siberia about 25,000 years ago. The L3 haplogroup therefore gave rise to many Africans plus all Europeans, Asians, and Americans. Ingman and Gyllensten (2001) show an mtDNA phylogram similar to Fig. 14 on which some haplogroups are identified. Further information on haplogroup distribution is given by Pafendorf and Stoneking (2005).

According to this hypothesis, Africans should be the most genetically diverse humans (Ingman *et al*., 2000; Merriwether *et al*., 1991). However, at least some Africans (the L3 haplogroup) should be similar to Europeans and Asians because they share a more recent common ancestor (Ingman *et al*., 2000). Because of their recent origin, Europeans and Asians should be less diverse than Africans, and native peoples of the Americas should be similar to each other and to East Asians.

These predictions are mostly borne out by Figures 14-18, above. Figure 14 makes the clear point that the most divergent (long-branched) mtDNA sequences in the World mtDNA data set come from Africa. Ingman *et al*. (2000) mention that the mtDNA sequence diversity of Africans is more than twice as high as among non-Africans. The 21 African sequences in Figure 18 also seem to divide into two groups—the long-branched sequences on branches 9-20 from the top, and the short-branched sequences 1-8 and 21. Using complete mtDNA sequences (not just a part of the control loop, as in this exercise), Ingman *et al*. (2000) divided Africans into almost the same two groups, and concluded that the Mkamba, Ewondo, Bamileke, Lisongo, and Yoruba were the last Africans to have a common ancestor with   non-Africans.

The prediction that Native Americans should be most related to each other and to East Asians does not enjoy so much success. For example, the "Native American" sequence had the highest similarity to the Guarani (native Paraguayan), Han Chinese, Japanese, and Indian sequences. The Native Mexican was most similar to a Siberian Chukchi. However, the Guarani, Pima, and the Navajo were equally similar to expected groups like the Koryaks and Japanese, and to unexpected groups like the French, Germans, Finns, Berbers, Ethiopians, and Yorubas.

Figure 13 and the final question of the student exercise put the relative size of these differences into perspective. The difference between the chimpanzee, baboon, and humans is so great that it dwarfs the differences between the human groups. The most dissimilar pair of "World mtDNA" samples is 95% similar; the similarity between the chimp and the modern human samples is about 8%. Clearly, all humans have a much more recent common ancestor (about 170,000 years ago) than chimps and humans (perhaps 5-6 million years ago). Whatever their minor differences, all modern humans are highly related.

The Research Utility of mtDNA Today

Pakendorf and Stoneking (2005) reviewed the current state of mtDNA research. Their review included critical examination of "dogmas" such as the assertions that mtDNA is maternally inherited

and does not recombine, as well as the nature of mitochondrial DNA variation. The approach taken in this exercise (analysis of a short sequence of the control region) was much used in the 1980s and 1990s, but now has been abandoned in favor of analysis of the whole mitochondrial genome. Evolutionary studies must also supplement these data with studies of the Y chromosome and autosomal sequences such as microsatellites. Pakendorf and Stoneking conclude that mtDNA is still valuable for studying ancient DNA, for forensic applications such as identifying crime victims and human remains from war zones, and for studying personal genetic history.

The Genographic Project

The instructor should also be aware of the Genographic Project, a large genetic survey being funded by National Geographic, IBM, and the Wiatt Family Foundation, and headed by geneticist Spencer Wells. This project (begun in 2005) is attempting to map prehistoric human migrations by obtaining DNA samples from traditional, indigenous people from throughout the world while these groups are still genetically distinct. For a fee, it also accepts DNA samples from nonindigenous peoples and will give them their basic genetic history. Participants are mailed a buccal swab kit and directions. Females get an analysis of their female lineage through mitochondrial DNA; males get analysis of their male lineage through their Y chromosome. The results take 6-8 weeks to be delivered. The fee is $99.95; with an education discount, it is $77 (October 2007).

More information about the Genographic Project can be found at
https://www3.nationalgeographic.com/genographic/.

Summary

This laboratory probably gives students their first look at their own DNA sequence, and it lets them determine how their personal sequence relates to international databases. However, the instructor should be aware that the free sequencing provided by the DNALC is not perfect. At Clemson, we bring up some of the evidence of inaccuracy above and ask the students to interpret it. Using the much more accurate sequences from the Uppsala University Database, we can begin to see some of the evidence that was used to reconstruct ancient human out-of-Africa migrations, but we can also see that mtDNA sequence is not a reliable predictor of ethnic group.

Finally, the overall lesson is that the mtDNA of all humans is very similar.

Literature Cited

Forster, P. 2004. Ice ages and the mitochondrial DNA chronology of human dispersals: a review. Philosophical Transactions of the Royal Society, London B, 359:255-264.

Ingman, M., H. Kaessmann, S. Paabo and U. Gyllensten. 2000. Mitochondrial genome variation and the origin of modern humans. Nature, 408:708-713.

Ingman, M. and U. Gyllensten. 2001. Analysis of the complete human mtDNA genome: methodology and inferences for human evolution. Journal of Heredity, 92:454-461.

Merriwether, D. A., A. G. Clark, S. W. Ballinger, T. G. Schurr, H. Soodyall, T. Jenkins, S. T. Sherry, and D. C. Wallace. 1991. The structure of human mitochondrial DNA variation. Journal of Molecular Evolution, 33:543-555.

MITOMAP, 2002. The report of the committee on the human mitochondrial genome. Web site at http://www.mitomap.org/, accessed in March 2007.

MITOMAP, 2002. Human mtDNA migrations. PDF file at http://www.mitomap.org/WorldMigrations.pdf, accessed in March 2007.

Pakendorf, B. and M. Stoneking. 2005. Mitochondrial DNA and human evolution. Annual Reviews of Genomics and Human Genetics, 6:165-183.

Torroni, A., A. Achilli, V. Macaulay, M. Richards, and H. J. Bandelt. 2006. Harvesting the fruit of the human mtDNA tree. Trends in Genetics, 22: 339-345.

About the Authors

Robert J. Kosinski is a professor of Biology at Clemson University, where he is the sole lecturer in the Introductory Biology course for majors and also the coordinator of the labs for that course. He received his B.S. degree from Seton Hall University and his Ph.D. in Ecology from Rutgers University. His interests include laboratory development, investigative laboratories, and the educational use of computer simulations, all in introductory biology. He was chosen as the Alumni Master Teacher of Clemson University in 2007. He has attended every ABLE meeting since 1989, has presented at 12 of those meetings, and acted as the chair of the host committee for the 2000 ABLE meeting at Clemson University.

Donna R. Weinbrenner is a Lecturer in the Biological Sciences Department at Clemson University. She is a preparator of the General Biology Classes. She earned her B.S. and M.S. degrees from Clemson University and is now a doctoral student in Microbiology and Molecular Medicine. Her research interests include photocatalytically activated titanium dioxide nanoparticles for antimicrobial and anticancer treatments, and hydrogels and elastomers for drug delivery in vascular and orthopaedic treatments. In 1992 she earned the President's Award for Outstanding Undergraduate Research in the South Carolina Branch of the American Society for Microbiology, and the South Carolina Academy of Science and Sigma XI Award for Outstanding Research in 1993. She is an expert in TaeKwonDo, and has won numerous honors in that discipline.

Melodie G. Cross received her B.S. degree from Toccoa Falls College, Georgia, and an A.S. degree from Greenville Technical College, South Carolina. She currently is a post-baccalaureate student at

Clemson University finishing her prerequisites to begin graduate studies in Microbiology. She has aided in the prep of both Introductory Biology and Microbiology laboratory courses, and took a lead role in the successful enhancement of an introductory laboratory exercise on the hydroponic growth of tomatoes.

Appendix A
Mitochondrial DNA Prep Directions

MtDNA Extraction Exercise

Assuming 150 students…

150+ 1.5ml empty microcentrifuge tubes, sterile.

150+ 1.5ml sterilized microcentrifuge tubes containing 300 μL of Solution 3, which is a 1:1 mix of Solutions 1 and 2.

Solution 1: 10% Chelex Make in a sterile 200 mL beaker on ice while stirring continuously with a stirbar.

5g Chelex
(10% Chelex [100-200 mesh, sodium form, Sigma 95577]. To 1 g add 50 **m**M Tris [m.w.121.14] to a volume of 10 mL, adjust to **pH 11** with 4N NaOH if needed). Stores ok at room temperature ~ 3 months

Solution 2: Proteinase K   Use a sterile 200 mL beaker. Mix:

995.0 μL Tris [m.w. 121.14] (Tris HCl pH 7.2)
48.75 mL sterile ddH$_2$O
500 μL proteinase K   (thawed and vortexed)
(Protienase-K [New England Biolabs P8102s 20 **m**g/mL] must dilute to 100 **μg**/mL. Keep frozen)

Keep on ice at all times except during heating steps. As soon as solution 2 is made, add it to solution 1

Solution 3:   1:1 mix of Solution 1 and Solution 2 = 50 mL Solution 1 added to 50 mL Solution 2

Aliquot 300 μL of Solution 3 into sterile microfuge tubes on ice in a cooler. To prevent Chelex from settling out, Solution 3 must be continuously agitated or stirred during aliquoting. Store microfuge tubes containing Solution 3 at –20 $^\circ$C..

Other Equipment Needed for mtDNA Extraction

razor blades (VWR 54411-050)
12 pairs fine-tipped forceps
37°C water bath
100°C heating block with openings to fit 1.5 mL Eppendorf tubes
4+ vortexes
12 pipettors that will deliver 150 µL (don't use p1000s, tips too big to use in tubes)
12+ boxes sterile tips
12+ ultrafine tip Sharpies (Staples 642736)
microcentrifuge, set for 30 second run at 14,000 rpm
10 + microcentrifuge tube racks (one per lab section, TAs should place student extract tubes in racks
    and put in freezer.)

PCR of Extracted Samples (done by preparator after lab)

150+ PCR Tubes with Ready-To-Go PCR Beads
    (Amersham Biosciences/GE 27-9559-01)

Make primer/loading dye mix:
    Each of the tubes described below will carry out 50 PCR reactions.
    Make 10 mL of <u>1% Cresol Red</u> (0.1g Cresol Red to 10 mL ddH$_2$O = 1%)
    Use this to make <u>loading dye stock solution</u> (17g Sucrose + 49 mL Water + 1 mL 1% Cresol
        Red)
    To make <u>loading dye</u>, mix 645 µL sterile distilled water and 460 µL loading dye stock
        solution in a 1.5 mL Eppendorf.
    Finally, to make <u>primer/loading dye</u>, add to the loading dye tube:
        10 µL 5'Primer (15 pM/µL)
    5'-TTAACTCCACCATTAGCACC-3'. Vortex.
        10 µL 3'Primer (15 pM/µL**)**
    5'-GAGGATGGTGGTCAAGGGAC-3'. Vortex. Freeze.
Order primers from Sigma-Genosys: 1-800-234-5362 Contact Jennifer Carlson (ext. 1211) or Dr.
Ashely Heath (ext. 1147)

When student DNA becomes available, add 22.5µl Primer/Loading Dye to a PCR tube with bead.
    Then add 2.5µL student DNA, tap to mix, spin to pool everything in bottom of tube.
Make sure to use a new tip for each sample!

Store on ice if PCR is not done immediately.

Label PCR tubes with student numerical codes (e.g., 4B).

Run samples in thermocycler programmed with step file:

94° C-30 seconds
58° C-30 seconds
72° C-30 seconds
30 cycles

Freeze product at -20^oC when done until electrophoresis is carried out.

Electrophoresis of amplified samples (done by either students or preparator)

2% Agarose containing Ethidium Bromide (EtBr) - **(CAUTION EtBr is Carcinogenic!)**
Wear proper SAFETY attire!

Assuming 150 students with 11 student lanes per gel, and therefore 14 gels…

1. Dissolve 11.2 g Agarose (OmniPar 2120, EMD, Baltimore, MD 21211) per 560 mL dd-H_2O.
2. Microwave approximately 5 minutes, to a rolling boil.
3. Move flask to chemical hood. Add 28 µL ethidium bromide (10 mg/mL) (OmniPar, Baltimore, MD 21211), swirl to mix. Final concentration of EtBr = 0.5 µg/mL.
4. Pour 140 mL of agarose-ethidium bromide mixture into each of four 250 mL square glass bottles.
5. Temper for 15-30 Minutes in a 55^oC water bath
6. Cast gel using 40 mL of agarose mix per gel, using a 12 well comb; (Gel Box: Model MGU-202T, C.B.S. Scientific Co., Del Mar, CA). Let the gel solidify and remove the comb.
7. Pour 1x-TAE Buffer (OmniPar 8720, EMD, Baltimore, MD 21211) carefully on each side of gel until covered, approximately 200-250 mL.
8. Add 5µL of size standard ladder (pBR322 DNA-BstNI Digest) (New England BioLabs 1000 ug/mL; # N303L) into the 1st lane only.
9. Add 12 µL of student sample DNA from the first PCR tube to lane 2 in the gel. Continue until the 10 other lanes have a different sample in each lane.
10. Run the gel electrophoresis (Power Pack Models EPS-250 and EPS-300II, C.B.S. Scientific Co., Inc., Del Mar, CA) at 115-120 volts for 20-30 Minutes; adequate separation occurs when the cresol red dye front has traveled 50 mm from the wells.
11. Remove the gel with a spatula and insert into a Zip Lock plastic bag.
12. Place gel with ladder on left, on a Mini Benchtop UV Transilluminator (UVP, Model TM-10E, PN# 95-0180-01 ($841.93 state price) VWR# 21476-040, Atlanta, GA). Wear safety goggles/glasses and make sure the UV shield is in place prior to turning on the transilluminator.
13. Take digital pictures and decide which lanes show 440 bp bands.
14. Send only those samples, numbered consecutively, to the Dolan DNA Learning Center. Ship using cold freezer packs, not dry ice.

Appendix B—Form for Recording Student Information

Section:		
Code	Name	Mother's Ethnic Background
A		
B		
C		
D		
E		
F		
G		
H		
I		
J		
K		
L		
M		
N		
O		
P		
Q		
R		
S		
T		
U		
V		
W		
X		
Y		
Z		