



# Mitochondrial DNA PCR and Sequencing

## Table of Contents Fall 2012

### **Mitochondrial DNA as a Molecular Clock**

Introduction .....	1
Mitochondrial DNA Replication .....	1
Setting the Molecular Clock .....	3
Mitochondrial DNA and its Role in Human Ancestry .....	4
Illustration of the Polymerase Chain Reaction .....	5
Laboratory Exercise: DNA Extraction .....	6
Important Laboratory Practices .....	6
DNA Preparation Using a Saline Mouthwash .....	7
Laboratory Exercise: Polymerase Chain Reaction.....	10
Laboratory Exercise: Agarose Gel Electrophoresis .....	12
Electrophoresis of Amplified DNA .....	13
Staining and Photographing Agarose Gels .....	14
Mitochondrial D-loop PCR Amplification Results .....	15
Submitting PCR samples for sequencing through CSU East Bay .....	16
Sequencing Activity One: Using the Sequence Server at DNALC .....	18
Sequence Comparisons .....	18
Sequence Alignments .....	22
Student Data .....	23
Review Questions: Molecular Clocks.....	24
Sequencing Activity Two: Determining your mtDNA Haplogroup .....	27
BLASTing your mtDNA Sequence .....	27
Using your SNPs to Determine your Haplogroup.....	29
Human mtDNA Migrations .....	30
Acknowledgements .....	31

# Mitochondrial DNA as a Molecular Clock

## Student Version

Maria Abilock  
BABEC

Frank H. Stephenson, Ph.D.  
Applied Biosystems

We gratefully acknowledge David Micklos and the staff of the Dolan DNA Learning Center at Cold Spring Harbor Laboratory for their generous help. Some materials for this exercise were adapted, by permission, from the *Genomic Biology: Advanced Instructional Technology for High School and College Biology Faculty* laboratory manual, Cold Spring Harbor Laboratory, 1999 and the *Genetic Origins* website at <http://geneticorigins.org>.

### Introduction

The DNA of every species on Earth is susceptible to change. Base pairs are lost. Base pairs are gained. One base pair can be substituted for another. We call these changes in DNA **mutations** or **single nucleotide polymorphisms (SNPs)** and they can arise by different mechanisms. By allowing organisms to adapt to environmental changes, mutation drives the steady and inevitable march of evolution.

Because a mutation within a gene can change the amino acid sequence of the encoded protein, many mutations spell disaster for that gene's function and potentially for the organism that bears it. Almost all genomes, however, from those of viruses to humans, carry segments of DNA that neither directly code for a protein nor are involved in the control of gene expression. A mutation occurring within such regions can usually be tolerated by the organism since it will most likely not impart any disadvantage to its survival nor impair its ability to reproduce.

Recently, forensic scientists, anthropologists, and evolutionary biologists have looked at mutations within the DNA of the mitochondrion to explore differences between peoples and populations. **Mitochondria** are found in all eukaryotic cells and are believed to have once been free-living bacteria that were assimilated early in evolution. Mitochondria divide independently of the cell and can be found in great numbers. They provide the cell in which they reside with the genes needed for the synthesis of the energy-carrying molecule ATP. Each mitochondrion contains several copies of its own circular genome and each cell may contain hundreds of mitochondria.

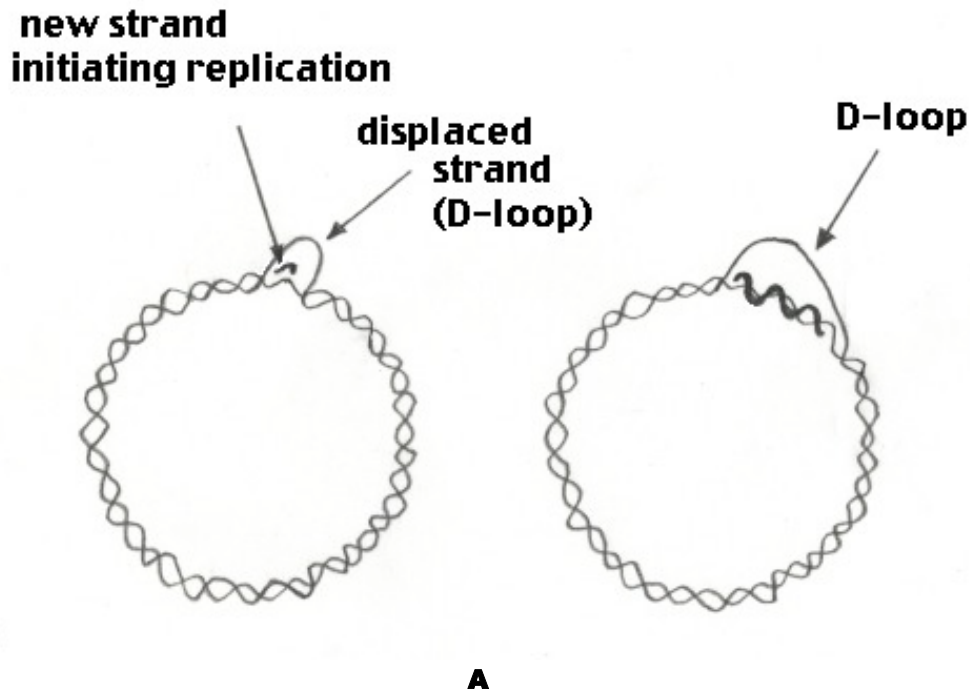
Because of its abundance, mitochondrial DNA has become a target for those scientists who do not have a ready supply of blood, bodily fluids, or tissue to work with but who still need to examine differences between people at the molecular level. Whether it is the skeletal remains of a Neandertal or a trace amount of hair left at the scene of a crime, where intact genomic DNA might be hard to come by, mitochondrial DNA can often be readily recovered.

The mitochondrial genome is 16,569 bp in length and contains 37 genes. Within its structure, however, there is a 1200 base pair non-coding segment, called the **control region**, carrying the genetic signals needed for replication and transcription. Since much of this DNA segment is not vital to the survival of the mitochondrion or the host cell, it is free to accumulate mutations. By studying the number and variety of base changes within this area, geneticists can determine the relatedness between individuals. Using the mutation rate within the mitochondrial control region as a "molecular clock," evolutionists can plot the course that hominid evolution has taken.

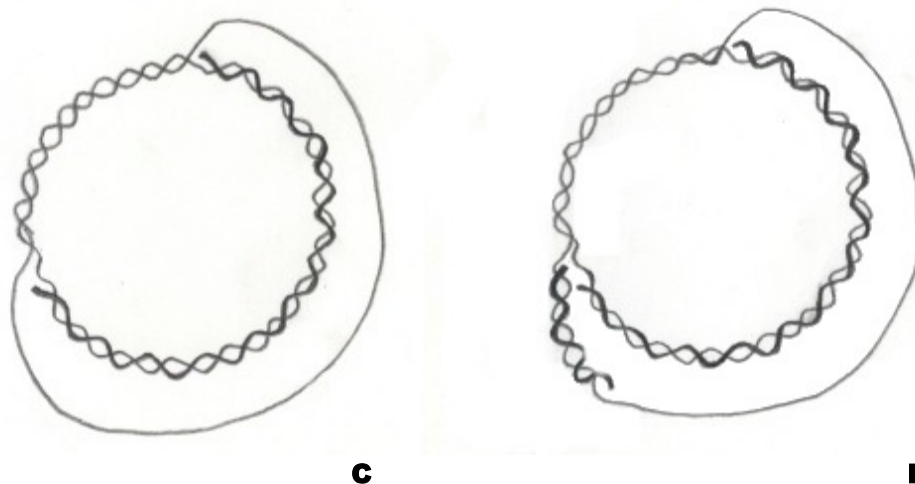
### Mitochondrial DNA Replication

When thinking about the human genome and all the traits that make us what we are tucked away within those six billion base pairs, it is easy to forget that all other cells in our body contain another genome, that of the mitochondrion. In some ways, mitochondrial DNA resembles the small circular pieces of DNA called **plasmids** found in bacterial cells. Like a plasmid, mitochondrial DNA is circular with a genome a fraction of the size of that of its host cell. Also, like a plasmid, there are multiple copies within each cell. Some cells in the human body carry thousands of mitochondria, and therefore contain thousands of copies of the mitochondrial genome. Because our bodies are composed of trillions of cells, with hundreds or thousands of mitochondria present in each cell, our bodies may contain more than 5 quadrillion (5,000,000,000,000,000) copies of the mitochondrial genome. A great deal of DNA replication has taken place to reach the adult body's full complement of mitochondrial genomes!

Replication of mitochondrial DNA proceeds in the following manner:



Replication of mitochondrial DNA begins on only one strand within the non-coding “control” region. As this strand is replicated, the opposite strand of the original DNA duplex is displaced and forms a single-stranded loop (hence the name “**D-loop**” for Displacement loop in Figure A). If the mitochondrion is not committed to the replication of its genome, copying of the non-displaced strand stops close to the protein-encoding boundary (Figure B). The displaced strand is then broken down and the replication process begins again displacing a single strand in the process. These events are repeated again and again until a signal is received that commits the entire molecule to replication. The segment you will amplify by PCR is within the D-loop sequence.



Once the signal is received to replicate the entire genome, replication continues around the circular molecule, increasing the size of the displaced strand (Figure C). When replication has proceeded approximately two-thirds of the way around the molecule, replication begins on the displaced strand (Figure D) until two new circular genomes are created.

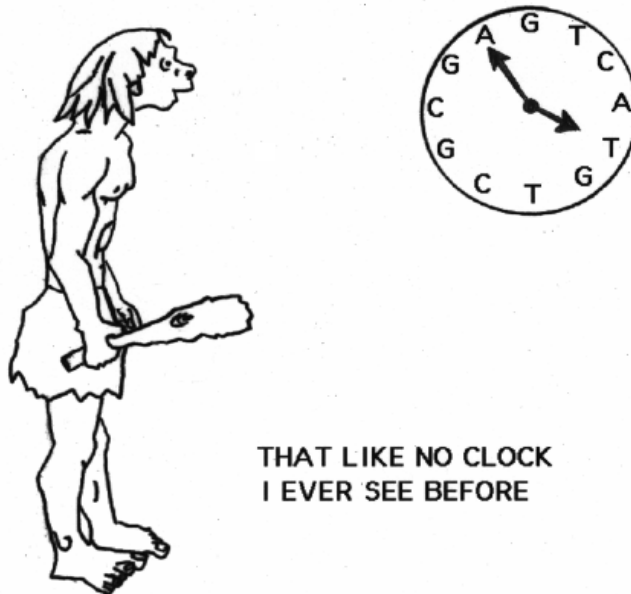
## Setting the Molecular Clock

A **species** is defined as a group of organisms that are capable of interbreeding to produce viable, reproductive offspring. New species can arise when members of a population separate to form their own breeding group within a new environment that demands of its inhabitants a unique set of survival skills. As the separated group struggles to fill a new and different ecological niche, the genes that provide individuals an advantage in the competition to flourish and mate are selected for and passed on to the next generation. Those individuals carrying genes that do not provide a selective advantage may neither survive into adulthood nor mate. Their genes are lost to the population.

A population's ability to adapt to a new environment is driven by the process of natural selection. Mutation makes natural selection possible. Mutation alters genes, destroying or changing their function. Mutation molds the ability of members of a species to survive under a defined set of conditions. Eventually, during the process of adaptation and over many generations, enough mutations accumulate within the separated population group that its individuals are no longer capable of interbreeding with members of the original population. This marks the birth of a new species. The longer two species diverge from each other, the greater the number of mutational differences there will be between them.

A clock measures the passage of time. Assuming that mutations occur at a constant rate, the accumulation of mutations in a DNA segment can be used as a "**molecular clock**" to measure the passage of time. In this case, the greater the number of mutations, the greater the amount of time passed. For example, if a new mutation appears in a defined region of DNA at a rate of one every 100,000 years, then after 500,000 years, 5 mutations will likely accumulate.

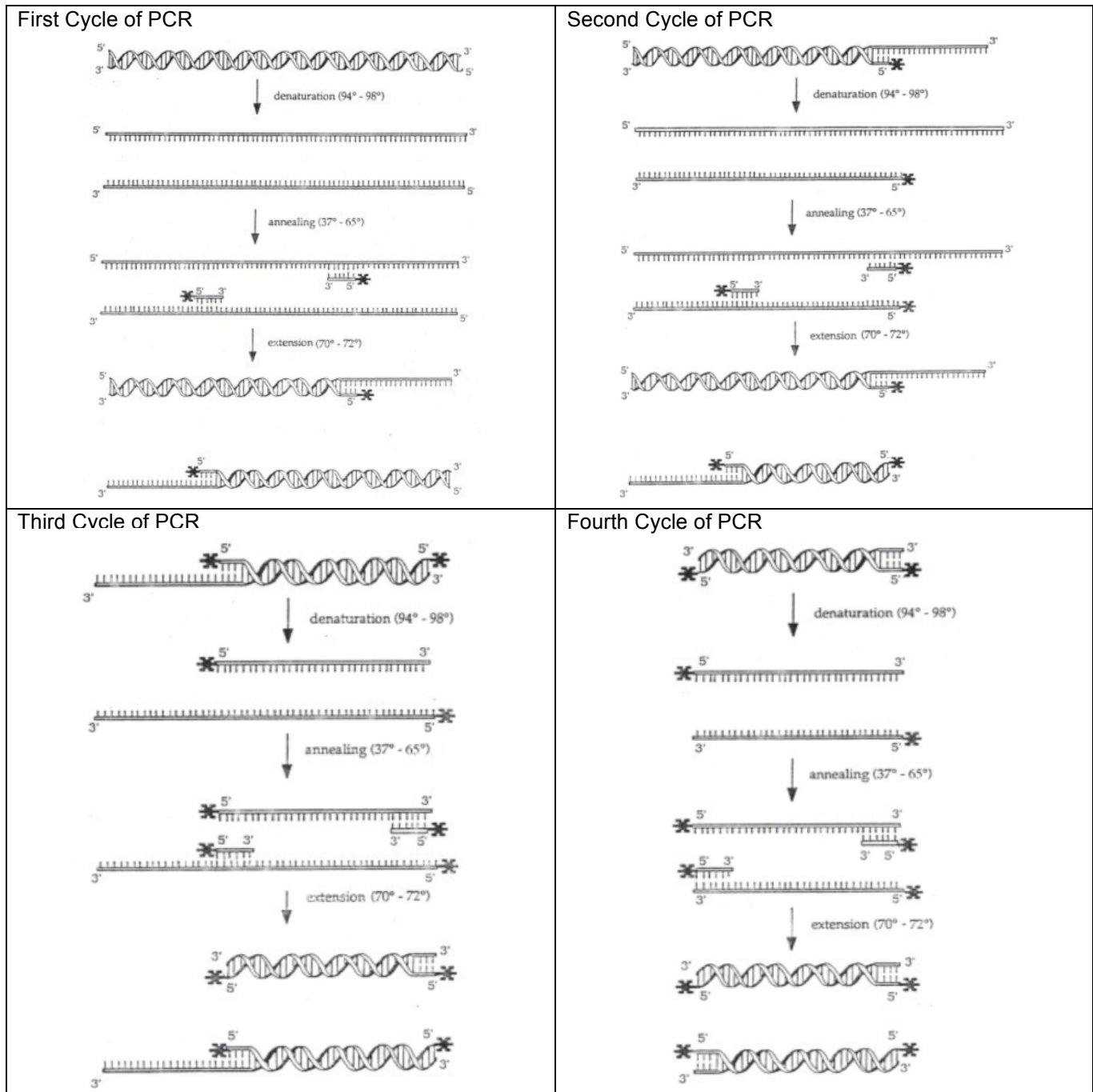
In this laboratory exercise, you will isolate mitochondrial DNA from cheek cells and amplify a 440 base pair segment of the control region by PCR. You will analyze the DNA sequence of the PCR product to reveal differences between you and the other students in your class. You can compare your sequence with those of the "Ice Man," "Lake Mungo Man," and other long-dead humans. How does your sequence compare to those of chimpanzee and Neanderthal? Could you and other modern humans have arisen from Neanderthals or did we evolve separately? If we evolved separately, at what point in time did modern humans and Neanderthals diverge on the evolutionary tree? Could Neanderthals have contributed to our gene pool? These are all questions you will investigate.





## Illustration of the Polymerase Chain Reaction

Figure 1. The First Four Cycles of the Polymerase Chain Reaction.



An excellent animated tutorial showing the steps of PCR is available at the DNA Learning Center website:  
<http://www.dnalc.org/ddnalc/resources/pcr.html>

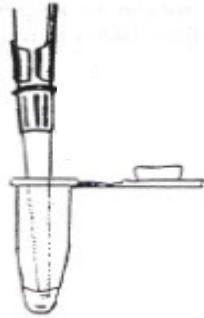



**Note:** You will need Macromedia Flash plug-in to view this online and to download the animation files to your computer.

## Laboratory Exercise

The protocol outlined below describes a procedure for isolating DNA from cheek cells. In the first step, you will rinse your mouth with a salt solution. This step typically dislodges hundreds of cells from the cheek epithelium. An aliquot of the mouthwash solution is centrifuged to collect the dislodged cells, which are then resuspended in a small volume of saline. The resuspended cells are then added to a solution of Chelex® to remove any metal ions (such as magnesium) which might promote degradation of your genomic DNA. Magnesium (and other metal ions) can act as cofactor for DNA-degrading nucleases present in saliva and the environment. The Chelex®/cell sample is then boiled to break open the cells. Since the sample is heated at a high temperature, the DNA, following this step, will be in a single-stranded form. The sample is then centrifuged briefly to collect the Chelex® and an aliquot of the supernatant containing released DNA is used for PCR.

**Objectives - student should be able to:**

1. Successfully isolate DNA from cheek cells
2. Prepare a PCR reaction for amplification of the mitochondrial D-loop.

<b>Important Laboratory Practices</b>	
<ol style="list-style-type: none"> <li>a. Add reagents to the bottom of the reaction tube, not to its side.</li> <li>b. Add each additional reagent directly into previously added reagent.</li> <li>c. <b>Do not pipet up and down, as this introduces error.</b> This should only be done only when resuspending the cell pellet and not to mix reagents.</li> <li>d. Make sure contents are all settled into the bottom of the tube and not on the side or cap of tube. A quick spin may be needed to bring contents down.</li> </ol> <div style="text-align: center; margin-top: 10px;">  </div>	<ol style="list-style-type: none"> <li>a. Pipet slowly to prevent contaminating the pipette barrel.</li> <li>b. Change pipette tips between each delivery.</li> <li>c. Change the tip even if it is the same reagent being delivered between tubes. Change tip every time the pipette is used!</li> </ol> <div style="text-align: center; margin-top: 10px;">  </div>
<p>Keep reagents on ice.</p> <div style="text-align: center; margin-top: 10px;">  </div>	<p>Check the box next to each step as you complete it.</p> <div style="text-align: center; margin-top: 10px;">  </div>

Place a check mark in the box as you complete each step.

### DNA Preparation Using a Saline Mouthwash

1. Vigorously swirl 10 mL of saline solution in your mouth for 30 seconds.

**Note: The saline solution is a 0.9% NaCl solution, the salt concentration of your blood plasma.**

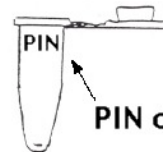


2. Expel saline into a cup and swirl to mix the cells.



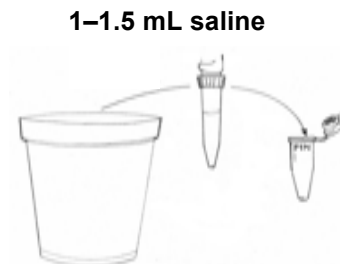
3. Label a 1.5 mL microfuge tube with your PIN.

**Note: A PIN (personal identification number) can be any combination of 2–3 numbers or letters that can uniquely identify you.**



PIN on microfuge tube

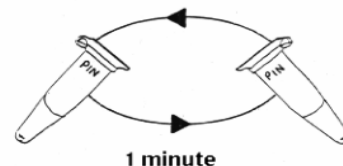
4. Transfer 1000  $\mu$ L to 1500  $\mu$ L (1 mL to 1.5 mL) of the saline/cell suspension into the labeled microfuge tube.



1-1.5 mL saline

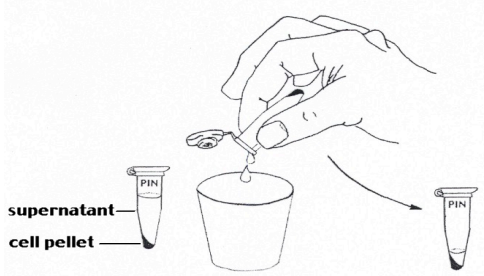
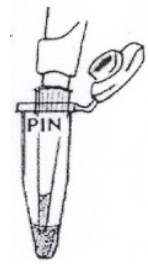
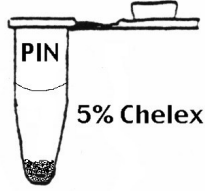
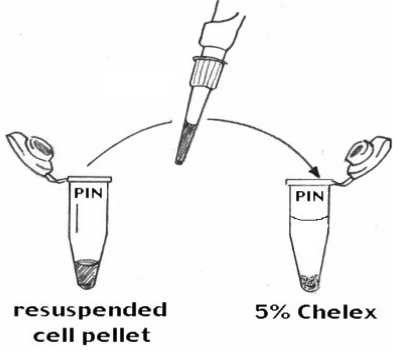
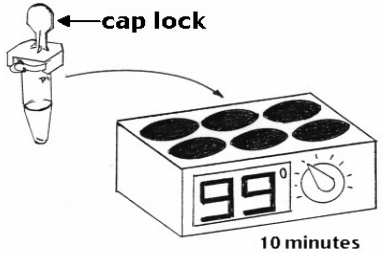
5. In a microcentrifuge, spin your saline cell suspension for 1 minute to pellet the cells. Be sure to use another student's sample as a balance.

**Note: Centrifuge speed should be set to 10,000  $\times$  g (10,000 rpm).**



1 minute

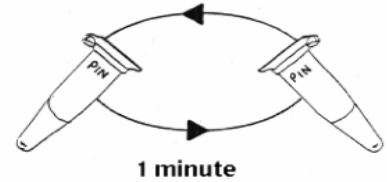


<p>6. Observe our cell pellet at the bottom of the tube. If you do not have one, you may need to start over with another 1–1.5 mL saline rinse.</p> <p>Pour off the supernatant into your cup, being careful NOT to lose your cell pellet.</p> <p><b>Note: There will be about 100 <math>\mu</math>L of saline remaining in the tube after you pour.</b></p>	 <p>supernatant</p> <p>cell pellet</p>
<p>7. Check to make sure you can see your cell pellet and that there is about 100 <math>\mu</math>L of saline covering it. You may need to add more saline to get up to about 100 <math>\mu</math>L.</p> <p>Rack or flick tube to mix, which will “resuspend” the cell and make an evenly mixed solution.</p> <p><b>Note: You can also “rack” your sample. Be sure the top of the tube is closed, hold tube firmly at the top, and pull it across a microfuge rack 2–3 times.</b></p>	<p><b>Resuspend cells in <math>\approx</math>100 <math>\mu</math>L saline</b></p> 
<p>8. Obtain a tube of Chelex from your instructor. Label with your PIN.</p>	 <p>5% Chelex</p>
<p>9. Withdraw 50 <math>\mu</math>L of your cell suspension from step 7 and add it to the tube containing Chelex.</p> <p><b>Note: Do not pipet up and down at this step, as it will clog the tip with Chelex beads.</b></p>	 <p>resuspended cell pellet</p> <p>5% Chelex</p>
<p>10. <b>Heat block version:</b> If your Chelex (with the cell suspension) is in a normal 1.5 mL microfuge tube, take your tube to a heat block station. Slide a cap lock onto the tube lid and place it in the heat block for 10 minutes. Keep track of your tube in the heat block.</p>	 <p>cap lock</p> <p>99°</p> <p>10 minutes</p>

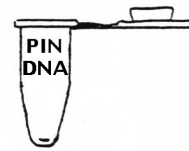
**PCR tube version:** If your Chelex (with your cell suspension) is in a tiny PCR tube, follow your teacher's instruction on placing it in a thermal cycler at 99°C for 10 minutes. Record the location of your tube.

	1	2	3	4	5	6	7	8
A	1012					0828		
B				1027				
C		0724						

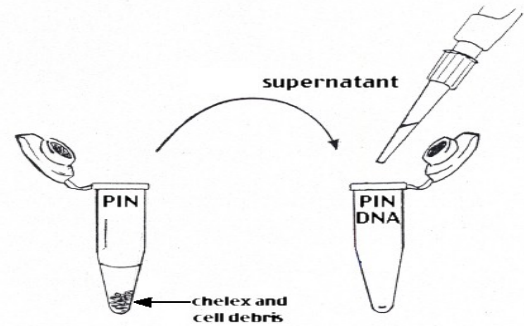
11. After heating, gently remove the cap lock and open the tube to release the pressure. Caution: the tube will be hot! Close and then rack or shake the tube well and place it in a centrifuge to spin for 1 minute.



12. Obtain another clean microfuge tube and label it with your PIN. Also write "DNA" on this tube.



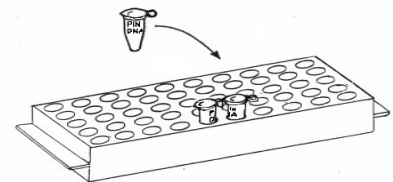
13. Holding your tube at eye level, use a P-200 to withdraw 50 µL of supernatant from the Chelex/DNA tube to the new, labeled tube. Be sure NOT to transfer any Chelex beads. **Note: This is your isolated "DNA" sample.**



14. Have someone check the "DNA" tube to be sure that no Chelex beads were transferred into it. There should be NO Chelex beads present, as they will interfere with the PCR.



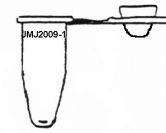
15. Place your DNA tube in the class rack. Your teacher will refrigerate your isolated DNA until you are ready to prepare your PCR amplification.



## Polymerase Chain Reaction

1. Obtain a tiny PCR tube. Label it with your PIN number, just under the lip.

**Note: Keep our PCR tube on ice when setting up the reaction.**



2. Pipet **20  $\mu$ L** of Master Mix into your PCR tube.

20  $\mu$ L of Master Mix



3. Change your pipet tip and add **20  $\mu$ L** of Primer Mix into your PCR tube.

20  $\mu$ L of Primer Mix



4. With a new pipet tip, add **10  $\mu$ L of your extracted DNA** into your PCR tube.

What is the total volume in your tube? \_\_\_\_\_  $\mu$ L

**Note: Make sure that all the liquids are settled into the bottom of the tube and not on the side of the tube or in the cap. If not, you can give the tube a quick spin in the centrifuge. Do not pipette up and down; it introduces error.**

10  $\mu$ L of DNA



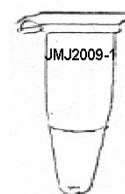
5. Setting up the controls:

- a. Two students will be asked to set up the positive control reactions (+C) for the class. They will use the positive control DNA provided in the kit. There should be enough +C PCR sample for one lane on each gel.
- b. Another two students will set up negative control reactions for the whole class (-C). They will use sterile water. There should be enough -C PCR sample for one lane on each gel.

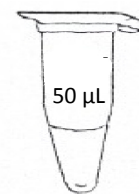
Control	Master Mix	Primer mix	DNA
+	20 $\mu$ L	20 $\mu$ L	10 $\mu$ L +C DNA
-	20 $\mu$ L	20 $\mu$ L	10 $\mu$ L sterile H <sub>2</sub> O

6. Check the volume of your PCR tube by comparing it to a reference PCR tube with 50  $\mu$ L in it. It should be near the thermal cycler, set by your teacher.

**Note: If the volume of your tube does not match, see your instructor to troubleshoot. You may need to set up the reaction again.**



PCR Tube



Reference Tube

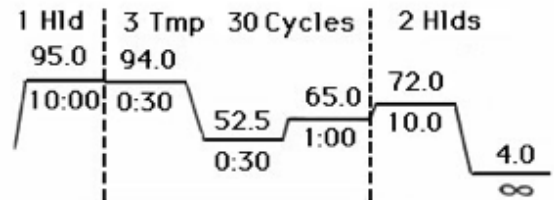
7. Place your reaction into the thermal cycler and record the location of your tube on the grid provided by your teacher.

	1	2	3	4	5	6	7
A	1123					828	
B				1027			
C		6777					9305

8. The cycling protocol for amplification of mtDNA PCR:

- 1) 95°C hold for 10 minutes
- 2) 30 cycles of:
  - 94°C for 30 seconds
  - 52.5°C for 30 seconds
  - 65°C for 1 minute
- 3) 72°C hold for 10 minutes
- 4) 4°C hold, ∞ infinity

**Thermal cycler Instrument displaying program parameters**

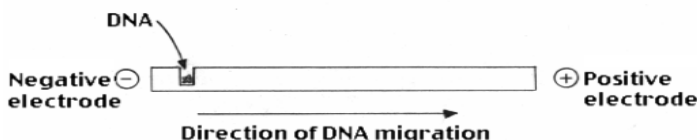


$$\begin{array}{rcl} 2 \text{ g} & = & X \text{ g} \\ 100 \text{ ml} & & 50 \text{ ml} \end{array} \quad X = 1 \text{ gram}$$

## Agarose Gel Electrophoresis

To determine whether or not the mtDNA PCR product amplified, you will need to visualize the products of your amplification. This will be done using a process called **gel electrophoresis** in which electric current forces the migration of DNA fragments through a special gel material. Since DNA is negatively charged, it will migrate in an electric field towards the positive electrode (Figure 2). When electrophoresed through a gel, shorter fragments of DNA move at a faster rate than longer ones.

**Figure 2.** Side view of an agarose gel showing DNA loaded into a well and the direction of DNA fragment migration during electrophoresis.



The gel material to be used for this experiment is called **agarose**, a gelatinous substance derived from a polysaccharide in red algae. When agarose granules are placed in a buffer solution and heated to boiling temperatures, they dissolve and the solution becomes clear. A comb is placed in the casting tray to provide a mold for the gel. The agarose is allowed to cool slightly and is then poured into the casting tray. Within about 15 minutes, the agarose solidifies into an opaque gel having the look and feel of coconut Jell-O™. The gel, in its casting tray, is placed in a buffer chamber connected to a power supply and running buffer is poured into the chamber until the gel is completely submerged. The comb can then be withdrawn to form the wells into which your PCR sample will be loaded.

**Loading dye** is a colored, viscous liquid containing dyes (making it easy to see) and sucrose, Ficoll, or glycerol (making it dense). To a small volume of your total PCR reaction, you will add loading dye, mix and then pipet an aliquot of the mixture into one of the wells of your agarose gel. When all wells have been loaded with sample, you will switch on the power supply. The samples should be allowed to electrophorese until the dye front (either yellow or blue, depending on the dye used) is 1 to 2 cm from the bottom of the gel. The gel can then be moved, stained and photographed.

### Calculations for Preparing 2% Agarose Gel

You will need a 2%, mass/volume agarose gel for electrophoresis of your PCR products. If your agarose gel casting trays holds 50 mL, then how much agarose and buffer would you need? The definition of m/v % in biology is grams (mass) / 100 mL (volume). Therefore, for 2% agarose, it will be 2 g /100 mL buffer.

Step 1: Calculate the mass of agarose needed for 50 mL total volume of agarose solution.

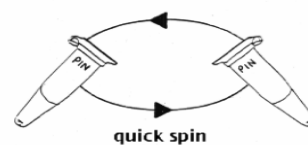
\_\_\_\_\_

Step 2: Calculate the amount of buffer needed to bring the agarose solution to 50 mL. By standard definition, 1 gram of H<sub>2</sub>O = 1 mL of H<sub>2</sub>O. The amount of buffer for the 2% agarose solution will be 49 mL (50 mL – 1 mL (1 gram of agarose)).

## Electrophoresis of Amplified DNA

1. Retrieve your PCR tube and place it in a balanced configuration in a microcentrifuge. Spin it briefly (10 seconds) to bring the liquid to the bottom of the reaction tube.

**Note: Make sure the centrifuge adapters are in place before putting the tiny PCR tube into the centrifuge rotor.**



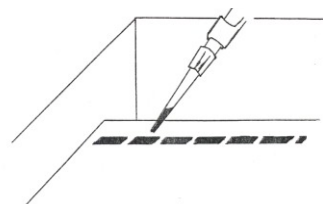
2. If you are NOT performing DNA sequencing:  
Add **5  $\mu\text{L}$**  of loading dye to your PCR tube.  
If you plan to sequence your DNA:  
Remove **20  $\mu\text{L}$**  of your PCR sample and dispense into a new tube. Add **2  $\mu\text{L}$**  of loading dye to it.

**Note: your PCR sample can't contain loading dye for sequencing.**



3. Carefully load **15 to 20  $\mu\text{L}$**  of the DNA/loading dye mixture into a well in your gel. Make sure you keep track of what sample is being loaded into each well.

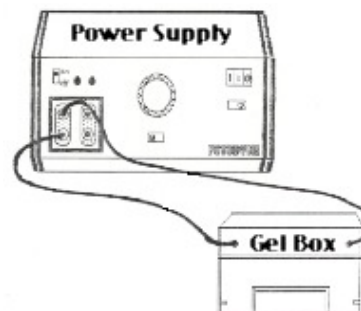
**Note: Avoid poking the pipette tip through the bottom of the gel or spilling sample over the sides of the well. Use a new tip for each sample.**



4. One student (or the instructor) should load **5-10  $\mu\text{L}$**  of 100 bp ladder (molecular weight marker) into one of the wells of **each gel**.



5. When all samples are loaded, attach the electrodes from the gel box to the power supply. Have your teacher check your connections and then electrophorese your samples at **150 Volts** for **25–40 minutes**.



6. After electrophoresis, the gels will be ready to stain and photograph.



## Staining and Photographing Agarose Gels

The PCR products separated on your agarose gel are invisible to the naked eye. If you look at your gel in normal room light, you will not be able to see the amplified products of your reaction. In order to “see” them, we must stain the gel with a fluorescent dye called **ethidium bromide (EtBr)**. Molecules of ethidium bromide are flat and can **intercalate**, or insert, between adjacent base pairs of double stranded DNA (Figure 3). When this interaction occurs, they take on a more ordered and regular configuration causing them to fluoresce under ultraviolet light (UV). Exposing the gel to UV light after staining, allows you to see bright, pinkish-orange bands where there is DNA (figure 4).

**Figure 3.** Ethidium bromide molecules intercalated between DNA base pairs.



Your teacher may stain your agarose gel and take a photograph for you so that you may analyze your PCR results.

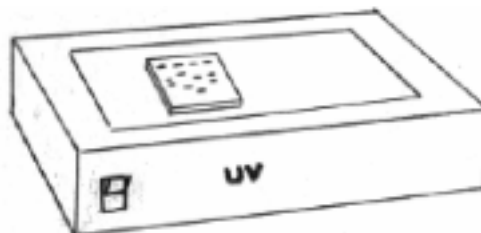
### Gel staining is done as follows:

1. Place the agarose gel in a staining tray.
2. Pour enough ethidium bromide (0.5 $\mu$ g/ mL) to cover the gel.
3. Wait **20** minutes.
4. Pour the ethidium bromide solution back into its storage bottle.
5. Pour enough water into the staining tray to cover the gel and wait 5 minutes.
6. Pour the water out of the staining tray into a hazardous waste container and place the stained gel on a UV light box.
7. Place the camera over the gel and take a photograph.
8. Check with your district on how to dispose of hazardous waste liquid and solids.

**CAUTION: Ethidium bromide is considered a carcinogen and neurotoxin. Always wear gloves and appropriate PPE (personal protective equipment) like safety glasses when handling. Students should NEVER handle EtBr.**

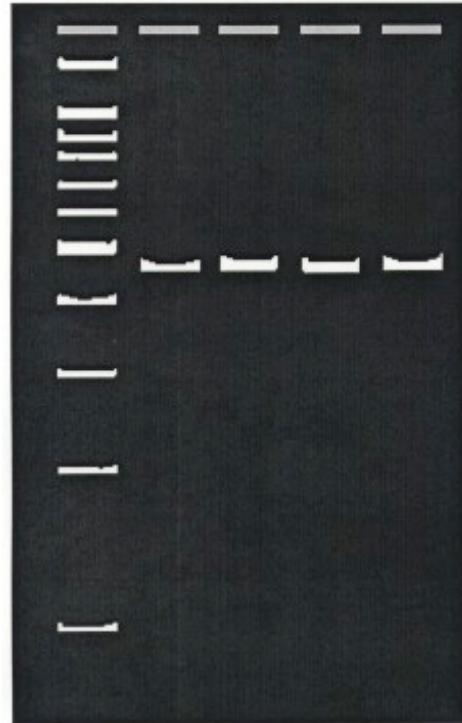
**CAUTION: Ultraviolet light can damage your eyes and skin. Always wear protective clothing and UV safety glasses when using a UV light box.**

**Figure 4.** After staining an agarose gel with ethidium bromide, DNA bands are visible upon exposure to UV light.



## Mitochondrial D-loop PCR Amplification Results

PCR amplification of the mitochondrial D-loop region using the primers for this protocol should produce a 440 bp product as shown in the figure below.



**Figure 4.** Representation of an agarose gel containing a 100 bp ladder (leftmost lane) and lanes showing 440 bp products from D-loop PCR amplification.

Tape your gel photo in the space below. Add observations and notes.



## Submitting PCR samples for sequencing through CSU East Bay

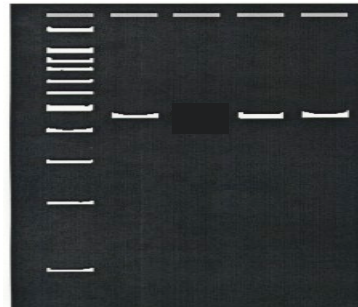
If you are submitting your PCR samples for sequencing through CSU East Bay, please continue on this page.

If you have received BigDye reagents and have made other sequencing arrangements, see the "Cycle Sequencing Supplement" on the BABEC website.

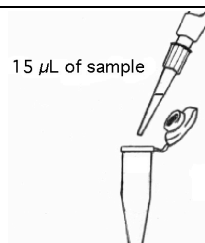
1. Obtain a clear photos of your class gel results. These photo is required for the processing of your samples at CSUEB. Label the lanes using the same numbering system that you use for your PCR tubes.

**Note to teachers:**

**If a student's sample did not amplify, then you must skip that lane and continue numbering in sequential order. Do not send in any PCR samples that did not amplify, as there will not be any material to sequence.**



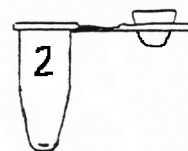
2. Spin down the original PCR tube. With a fresh tip, remove 10  $\mu$ L of the PCR product and place in a new tube.



3. Label the new tube with the corresponding number from the gel photo.

**Note to teachers:**

**Remember: do not skip numbers!**




4. The contact person at CSU East Bay is **Dr. Chris Baysdorfer**. Contact him via email at the time of shipment to confirm. Use the following email address:

[chris.baysdorfer@csueastbay.edu](mailto:chris.baysdorfer@csueastbay.edu)

Include the following class information:

- your name
- school
- number of samples
- class name (AP Bio, etc.).



<p>5. Package your samples and labeled gel photo in a zip-top plastic bag. Place in a styrofoam box with a few cooler packs.</p>	
<p>6. Ship OVERNIGHT delivery to:</p> <p><b>Professor Chris Baysdorfer</b>  Department of Biological Sciences  California State University, East Bay  Hayward, CA 94542  Phone: (510) 885-3459</p> <p><b>IMPORTANT:</b> Do not ship samples on Fridays.</p>	
<p>7. The sequences will be delivered to you via email. This should take 5-7 business days, but confirm with Dr. Baysdorfer at time of shipment. Samples are run free of charge!</p>	
<p>8. You will receive the sequences from CSU in the form of a "Trace File". They will have the same sample names that you submitted, but will end in ".ab1". The .ab1 files need a special software program to be opened. Programs can be downloaded for free:</p> <p>For MAC  <a href="http://www.mekentosj.com/science/4peaks">http://www.mekentosj.com/science/4peaks</a></p> <p>For PC:  <a href="https://products.appliedbiosystems.com/ab/en/US/adirect/ab?cmd=catNavigate2&amp;catID=600583&amp;tab=Overview">https://products.appliedbiosystems.com/ab/en/US/adirect/ab?cmd=catNavigate2&amp;catID=600583&amp;tab=Overview</a></p>	<p>"4Peaks" for MAC</p>  <p>"Peak Scanner" for PC</p> 
<p>9. Using these programs, you can visualize and edit the color chromatogram of each sequence. You can also export the sequences as text files, which will be needed for further analysis using the Sequence Server at the CSHL DNA Learning Center. See instructions below, "Mitochondrial sequence comparisons", starting on page 32.</p> <p><b>Note:</b> Feel free to contact the Manager of Education Programs at BABEC if you need help using these programs or exporting the text files. Contact information can be found at <a href="http://www.babec.org">www.babec.org</a></p>	<p>Sequence chromatogram:</p>  <p>Sequence text file:</p> <pre> BABAC1_001_E03.txt &gt;BABAC1_001_E03 GGCTTTCTAAATTTACTATTCTCTGTTCTTTCATGGGGAAGC AGATTTGGGTACCACCAAGTATTGACTCACCCATCAACAAC CGCTATGTATTTGTCATTAAGTACGACGACCATGAATATT GTACGGTACCAATAACTTGTACCACTGTAGTACATAAAAA CCCAATCCACATCAAAACCCCTCCCATGCTTCAAAGCAAG TACAGCAATCAACCCTCAACTATCACACATCAACTGCAACTC CAAAGCCACCCCTCACCACTAGGATATCAACAACCTACCC ACCCCTAACAGTACATAGCACATAAAGCCATTTACCGTACAT AGCACATFACAGTCAAAATCCCTTCTCGCCCAATGGATGACC CCCCCAGATAGGGGTCCCTTGACCACTCCTCA </pre>


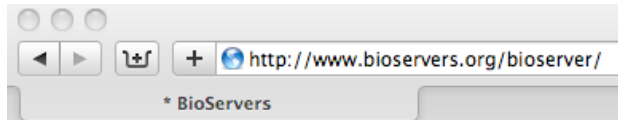
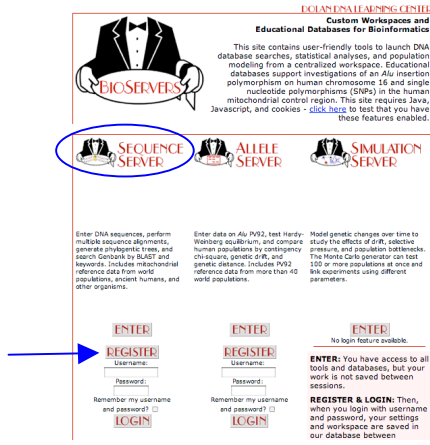
## Sequencing Activity One: Using the Sequence Server at the CSHL DNA Learning Center

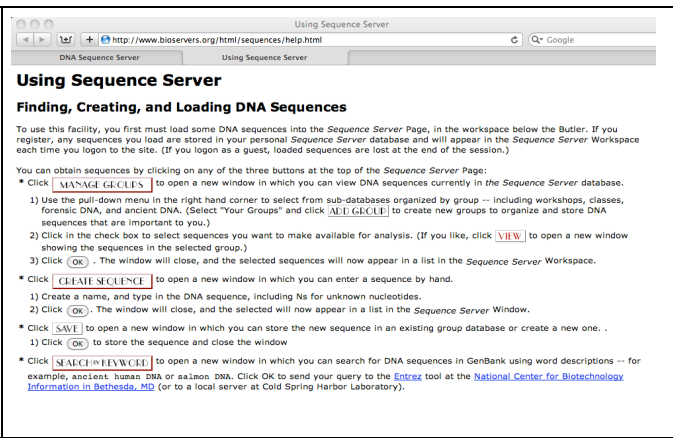
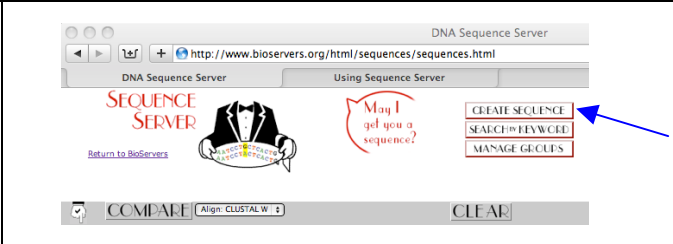
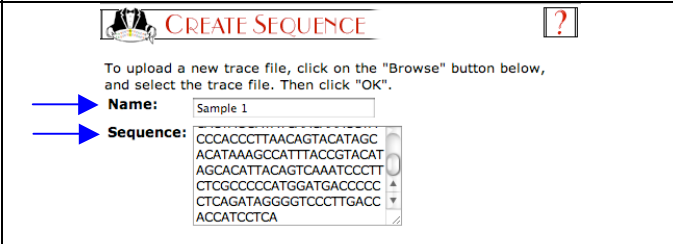

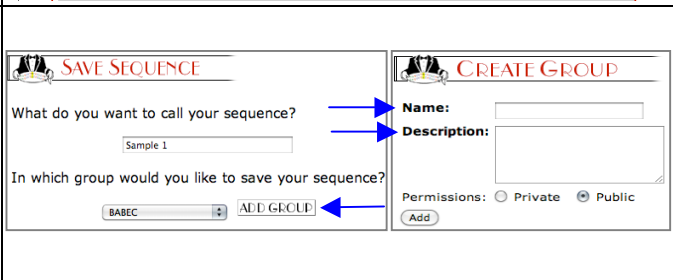
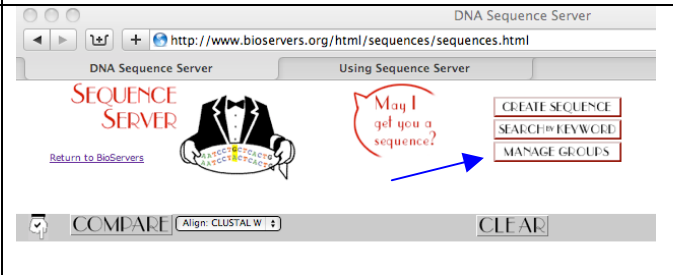
### Objectives - student should be able to:

1. View your class data in the Cold Spring Harbor Laboratory Sequence Server database.
2. Perform pair-wise sequence alignments between diverse modern humans.
3. Perform pair-wise sequence alignments between diverse modern humans and Neanderthals.
4. Set the “molecular clock” based on the number of sequence differences between modern humans.
5. Use a “molecular clock” to estimate when Neanderthals and modern humans diverged.

The DNA Learning Center at Cold Spring Harbor Laboratory has developed a number of **bioinformatics** tools for student use. Bioinformatics tools are computer programs used to help scientists make sense of biological data and solve biological problems. You will be using the Sequence Server for three different activities to help you learn more about the origins of our species.

In the following exercise, you will compare DNA sequence between individuals from several different population groups. You will first compare sequences between modern humans. This information will be used to set a “molecular clock”. You will then compare modern humans to Neanderthals to see if Neanderthals might have contributed to our gene pool. The molecular clock you derive will be used to determine when modern humans and Neanderthals diverged. In your final comparison, you will align modern human sequences to that of a chimpanzee to derive a new molecular clock. The molecular clocks will be used to estimate when modern humans first appeared.

<b>Mitochondrial DNA Sequence Comparisons</b>	
<p>1. Open up an Internet browser window. (This might be Internet Explorer, Safari, Firefox, or Netscape Navigator, etc.)</p>	
<p>2. In the address box, type in the following URL: <a href="http://www.bioservers.org/bioserver">http://www.bioservers.org/bioserver</a> and press the Enter (or Return) key on the keyboard. The DNA Learning Center’s Bioserver main page will be brought up.</p>	
<p>3. You will want to use the “Sequence Server”. Click on the “REGISTER” button if you have not previously registered with Bioservers. Fill out the required information and then hit “SUBMIT.” If you are already registered, enter your username and password, then press “LOGIN.”</p>	

<p>4. This will open the <i>main workspace window</i>. Click on the <b>red question mark</b> in the top right corner of the page. It will open up a page called "Using Sequence Server" in a separate window. The instructions contained here can be referenced if you need more information about using this site.</p>	
<p>5. Click on the "DNA Sequence Server" page to bring it forward on the desktop.</p> <p>Click on the "Create Sequence" box. This is where you can add your class data.</p>	
<p>6. You will need to upload each sequence individually. Type sequence name in "Name" box. Then cut and paste the sequence from the text file into the "Sequence" box. Then select "OK" at the bottom of the page.</p>	
<p>7. Go back to the <i>main workspace window</i> and you will see all the sequences you uploaded. Click on the "Save" box next to one of the sequences.</p>	
<p>8. Click on "Add Group" and create a name for your class in the window that pops up. Be sure to select "Public" for permissions. Public viewing will allow students and others to look at the data without creating an account.</p> <p>You will need to repeat these steps for each individual sequence until the entire class data is uploaded.</p>	
<p>9. <b>You are now ready to analyze your class data as a group.</b></p> <p>Go back to the <i>main workspace window</i> and click on the "Manage Groups" box. This will open a new window.</p>	

10. In the upper right hand corner of the “Manage Groups” window is a scroll menu. If it isn’t already showing, select “Your Groups”.

This will bring up a list of all the groups you have created in your account. Use the scroll bars on the right side of the window to locate your class, then click the box to its left to select it. Press “OK”.

**MANAGE GROUPS**

Select the appropriate period of time: 2011 (July - December) | Sequence sources: Your Groups

The following tables let you choose which of the groups of sequences we have stored in our database you would like to work with. To select a group, click on the checkbox next to the group. Clicking on the view button will show you all the sequences in the group.

Date	Your Groups	VIEW	EDIT GROUP
<input checked="" type="checkbox"/>	10/05/2011 BABEC	VIEW	EDIT GROUP

11. Go back to the “Manage Groups” window and in the upper right scroll menu, select “Ancient Human mtDNA”. Click the box to the left of one or more of the six groups on the list. Press “OK” when finished.

**MANAGE GROUPS**

Select the appropriate period of time: 2011 (July - December) | Sequence sources: Ancient Human mtDNA

The following tables let you choose which of the groups of sequences we have stored in our database you would like to work with. To select a group, click on the checkbox next to the group. Clicking on the view button will show you all the sequences in the group.

Date	Ancient Human mtDNA	VIEW
<input checked="" type="checkbox"/>	07/29/2008 Incan "Ice Maiden" (Mummy)	VIEW
<input type="checkbox"/>	01/25/2001 Lake Mungo Man mtDNA (Australia; ca. 60,000 yrs.)	VIEW
<input type="checkbox"/>	12/18/2000 "Otzi" The Iceman mtDNA (Alps, ca. 5200 yrs. old)	VIEW
<input type="checkbox"/>	12/17/2000 Egyptian mummy mtDNA (Egypt, ca. 4,000 yrs. old)	VIEW
<input type="checkbox"/>	12/16/2000 Yixi mtDNA (China, ca. 2,500 yrs. old; 3 samples)	VIEW
<input checked="" type="checkbox"/>	12/15/2000 Cahokian mtDNA (Illinois, ca. 1100 yrs. old)	VIEW

12. Your class data and several prehistoric humans' data should now be added to your workspace. Select your sequence by using the scroll menu below your class name. Deselect all the check boxes on the left except for your sample and one prehistoric human of your choice. Next to the “COMPARE” button below the Sequence Server icon, use the arrows to scroll to “Align:CLUSTAL W” then click the “COMPARE” button.

**SEQUENCE SERVER**

Return to BioServers | May I get you a sequence? | CREATE SEQUENCE | SEARCH BY KEYWORD | MANAGE GROUPS

Help window? Yes No

**COMPARE** | Align: CLUSTAL W | CLEAR | Search: BLASTN | ANALYZE

- BABEC, 2011-10-06
  - KW OPEN
  - None
- "Otzi" The Iceman mtDNA (Alps, ca. 5200 yrs. old), 2000-12-18
  - Iceman OPEN
  - None
- Egyptian mummy mtDNA (Egypt, ca. 4,000 yrs. old), 2000-12-17
  - Egyptian mummy #1 OPEN
  - None

13. Your sequence and the prehistoric human sequence you chose should align where bases are complementary. You may notice some yellow highlight regions, dashes and gray highlight regions with “N’s.” What do you think these indicate?

**Note:**

- yellow = mismatch
- gray = “N” or unknown base
- dashes = gaps due to insertions, deletions or sequencing errors

**CLUSTAL W ALIGNMENT**

Alignments are displayed in blocks of 25 nucleotides per row. Mismatches are colored in yellow. Positions containing a sequencing error (N) are colored in grey. Dashes (-) denote misalignments.

KW BABEC  
Egyptian mummy #1 Egyptian mummy mtDNA (Egypt, ca. 4,000 yrs. old)  
Start at position 0 of 413 base pairs (show 1000 per page)

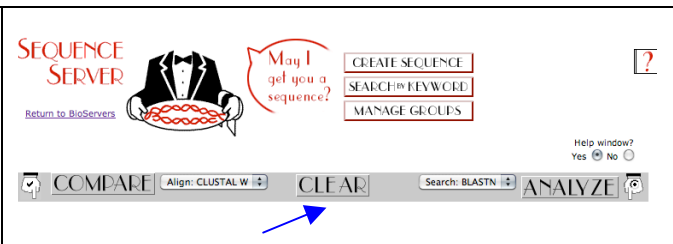
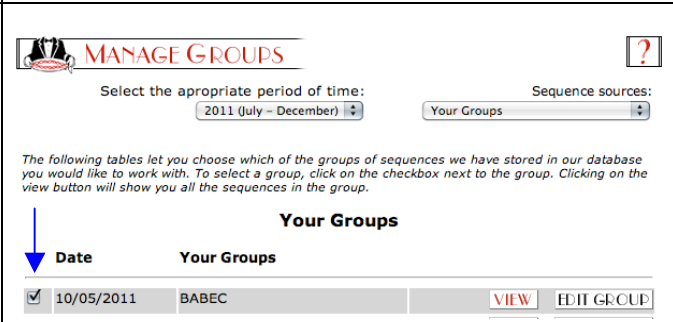
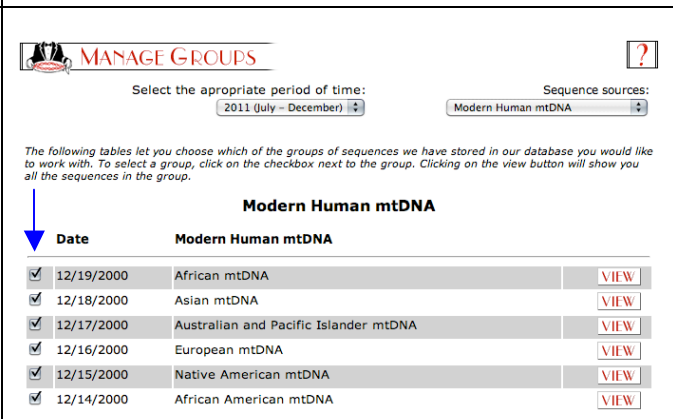
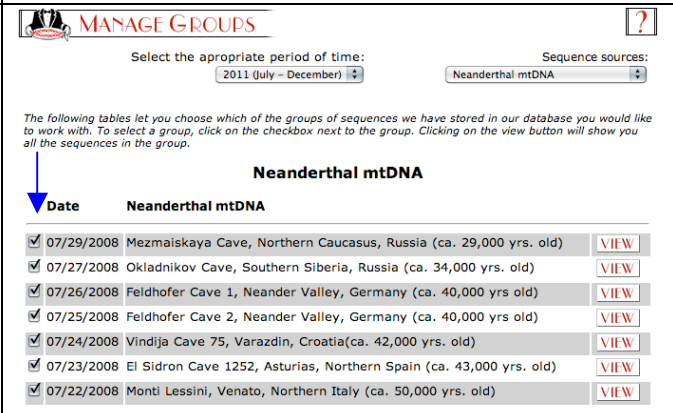
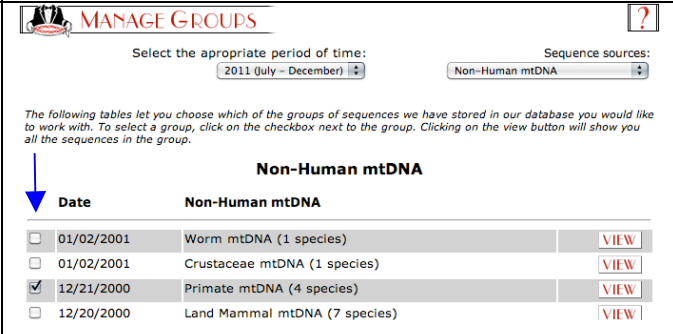
not trimmed OR trimmed

Redraw | Print this alignment...

Showing 413, starting from 0 and ending at 413

```

Egyptian_mummy_#1  ACAGTTTCATG 25
KW                GGCTTTCTAATTTACTATTCTG
Egyptian_mummy_#1  CC 50
KW                TTCATTCATGGGAAAGCAGATTGG
Egyptian_mummy_#1  AATTCATGAAATTAATTC 75
KW                ATACACCCCAAGTATTGACTCACC
Egyptian_mummy_#1  AAAATC 100
KW                ATCAACAACCGGTATGATTTCTA
  
```

<p>14. When you have examined the alignment to your satisfaction, go back to the <i>main workspace window</i> and press the “Clear” button to clear your workspace.</p> <p>Next, you will work with various diverse modern humans, Neanderthals, chimpanzee, and your classmates’ data to investigate your genetic origins.</p>	
<p>15. In the “Manage Groups” window, use the scroll menu to locate your class under the “Your Groups” option in the scroll bar. Select it by checking its box.</p>	
<p>16. Using the scroll menu again in the upper right corner of the “Manage Groups” window, locate and select “Modern Human mtDNA.” Check all boxes in this window.</p>	
<p>17. Locate and select the “Neanderthal Human mtDNA” category from the scroll bar in the “Manage Groups” window. Check all boxes in this window.</p>	
<p>18. Select the “Non-Human DNA” category from the Manage Groups window. Click the box to the left of “Primate mtDNA (4 species).” Click on the “OK” button at the bottom of the window. This will place all selected DNA sequences onto the <i>main workspace window</i>.</p> <p>Continue with the exercises listed on the next page.</p>	

## Sequence Alignments

You will perform a series of sequence alignments that will allow you to estimate a mutation rate and to calculate the timing of crucial events in human evolution. Use the following guidelines for each comparison.

- Identify a region spanning 200 bases where there is good alignment between the two sequences you are comparing. This region should contain few, if any, “N’s.”
- Excluding N’s (in gray) and dashes that may occur at the beginning or end of the alignment, count how many yellow-highlighted base positions are found in the alignment. If you find a run of three or more dashes in a row, count such a run as a single nucleotide difference.
- If you find a sequence that does not align for 200 bases, use a different sequence.

**Note to teachers:** *This is based on the assumption that it is more likely that a single event, rather than multiple, independent events, will lead to the insertion/deletion of 3 or more bases at a particular site.*

Follow the steps below to fill in the spaces in the chart that follows.

### 1. Modern Human vs. Modern Human

- a. Select any two modern humans from the groups on your workspace. Fill in the identifying information in the table.
- b. Compare these two individuals by ClustalW alignment. Count the number of mismatches, or SNPs, and record this number in the table.
- c. Repeat steps “a” and “b” using different modern humans.
- d. Now compare two students in the class and fill in all the appropriate information in the table.
- e. Calculate the average number of SNPs for this group and record in the table.

### 2. Modern Human vs. Neanderthal

- a. Select any African modern human and any Neanderthal to compare by ClustalW alignment. Fill in the identifying information in the table.
- b. Compare these two individuals and record the number of SNPs in the table.
- c. Repeat steps “a” and “b” with any Asian modern human and any Neanderthal.
- d. Repeat steps “a” and “b” again using any European modern human and any Neanderthal.
- e. Now compare your (or another student’s) sequence with any Neanderthal and record all appropriate information in the table.
- f. Calculate the average number of SNPs for this group and record in the table.

### 3. Neanderthal vs. Neanderthal

- a. Select Neanderthal #1 and Neanderthal #2 to compare by ClustalW alignment. Record the number of SNPs in the table.
- b. Do the same with the other two combinations of Neanderthals.
- c. Calculate the average number of SNPs for this group and record in the table.

### 4. Modern Human vs. Chimpanzee

- a. Select any modern human to compare with Chimp #2. Fill in the identifying information in the table.
- b. Compare these two sequences by ClustalW alignment. Count the number of SNPs and record this number in the table.
- c. Repeat steps “a” and “b” using different modern humans.
- d. Now compare your (or another student’s) sequence with Chimp #2 and record the number of SNPs along with the appropriate identifying information.
- e. Calculate the average number of SNPs for this group.

## Student Data

<b>Sequence Server Clustal W Alignments: SNPs</b>			
Modern Human vs. Modern Human	Number of SNPs	Your Average	Class Average
_____ vs. _____			~6
_____ vs. _____			
_____ vs. _____			
Student _____ vs. Student _____			
<b>Modern Human vs. Neandertal</b>			
African _____ vs. Neandertal # _____	~18		~18
Asian _____ vs. Neandertal # _____	~18		
European _____ vs. Neandertal # _____	~19		
Student _____ vs. Neandertal # _____			
<b>Neandertal vs. Neandertal</b>			
Neandertal #1 vs. Neandertal #2	~7		~5
Neandertal #1 vs. Neandertal #3	~4		
Neandertal #2 vs. Neandertal #3	~4		
<b>Modern Human vs. Chimpanzee</b>			
_____ vs. Chimp #2			~42
_____ vs. Chimp #2			
_____ vs. Chimp #2			
Student _____ vs. Chimp #2			

**Note to teachers:** Since there are so many possible combinations to choose from when making these comparisons, the values you obtain in your class may be different than those given in this table. This can bring up a good discussion topic: How might these numbers vary depending on the samples (and groups) used for comparison? You may notice how small changes in the numbers can make a huge difference in the values calculated for the evolution divergence points.

**Note to teachers:** After students complete the ClustalW alignments and calculate their averages, they will need to calculate the class averages before proceeding with the following questions.



Name \_\_\_\_\_

Date \_\_\_\_\_ Period \_\_\_\_\_

### Review Questions: Molecular Clocks

#### 1. Calculating a molecular clock

Archaeologists use a number of different techniques to estimate the age of fossils. These include radiocarbon dating, measuring changes in carbonates and tooth enamel brought about by exposure to radiation over time, and determining the age of the geological strata in which the fossil was found. By dating human fossils discovered in Africa, scientists estimate that modern humans first appeared approximately 150,000 years ago. Using this value and the class average number of differences for “Modern Humans vs. Modern Humans,” derive a **molecular clock**, or mutation rate, in years/mutation. Use the formula below:

$$\frac{150,000 \text{ years}}{\text{_____ mutations}} = \text{_____ years/ mutation}$$

#### 2. Did modern humans evolve from Neanderthals?

Neanderthal fossils have been discovered in Europe and the Middle East. Dating the fossils by radiocarbon decay suggests that Neanderthals inhabited the European continent as recently as 28,000 years ago. Estimates of when Neanderthal first appeared in Europe are far less precise but many scientists believe it may have been as long as 300,000 years ago. Although they are frequently depicted as stocky and brutish individuals, Neanderthals cared for their sick and injured, fashioned stone tools, used fire, lived and hunted in social units, and ritually buried their dead.

As far as we know, Neanderthals did not inhabit regions far outside the European continent. If modern Europeans descended from Neanderthals, you would expect that Neanderthals would be more closely related to modern European populations than to any other modern human population in the world. Based on your “Modern Human vs. Neanderthal” data, does it appear as though Europeans or any other modern world population descended from the Neanderthals? Explain.

#### 3. Human - Neanderthal divergence

How many years ago did the common ancestor of modern humans and Neanderthals live? In the equation below, use the average number of differences (mutations) you found between modern humans and Neanderthals and your calculated mutation rate to estimate this number.

$$\text{_____ mutations} \times \frac{\text{_____ years}}{\text{mutation}} = \text{_____ years}$$

Name \_\_\_\_\_

Date \_\_\_\_\_ Period \_\_\_\_\_

#### 4. Did Neanderthals contribute to the modern human mtDNA gene pool?

A **gene pool** is the collection of all genes in a population. Members of a single gene pool would be expected to have fewer differences between them than would be expected between members of different gene pools. Did Neanderthals have a separate gene pool from that of modern humans? Could Neanderthals have contributed their mitochondrial DNA to the gene pool of modern humans? Use the comparisons below (4a through 4e) to answer this question.

- a. Average difference between Neanderthals = \_\_\_\_\_
- b. Average difference between modern humans and Neandertal = \_\_\_\_\_
- c. Average difference between modern humans = \_\_\_\_\_
- d. The closest modern human/Neandertal alignment discovered by your class showed \_\_\_\_\_ differences.
- e. The two most divergent modern humans discovered by your class showed \_\_\_\_\_ differences.
- f. Do you think the Neanderthals used in this study are members of a single gene pool (assume that modern humans are of a single gene pool)? Explain.
  
- g. Do you think Neanderthals contributed their mitochondrial DNA to the modern human mtDNA gene pool? What other data would you want to answer this question?

#### 5. A molecular clock based on chimpanzee/hominid divergence

Based on the fossil record, scientists believe that chimpanzees and modern humans may have diverged 5,000,000 years ago.

- a. Would the molecular clock be different if you used the time since chimpanzees and modern humans evolved to determine the mutation rate? Calculate a new mutation rate using the formula below and the 5 million year divergence estimate.

$$\frac{5,000,000 \text{ years}}{\text{___ mutations}} = \text{_____ years/mutation}$$

Name \_\_\_\_\_

Date \_\_\_\_\_ Period \_\_\_\_\_

b. Is this value different than the one you calculated based on “Modern Human vs. Modern Human” differences? Explain.

c. Using the mutation rate you calculated in 5a, when did “Mitochondrial Eve,” the mitochondrial ancestor of all modern humans, live? Use the formula below for this calculation.

$$\underline{\hspace{2cm}} \text{ mutations} \quad \times \quad \frac{\underline{\hspace{2cm}} \text{ years}}{\text{mutation}} \quad = \quad \underline{\hspace{2cm}} \text{ years}$$

How does this estimate compare with the value you used to calculate a molecular clock in Problem 1?

d. Using the same molecular clock (calculated in 5a), when did Neanderthals and modern humans diverge and how does this estimate compare with the value you calculated in Problem 3?

$$\underline{\hspace{2cm}} \text{ mutations} \quad \times \quad \frac{\underline{\hspace{2cm}} \text{ years}}{\text{mutation}} \quad = \quad \underline{\hspace{2cm}} \text{ years}$$

e. How many mutations would you need between chimpanzee and modern humans to give the mutation rate you calculated in Problem 1? Use the equation below for your calculation.

$$\frac{5,000,000 \text{ years}}{x \text{ mutations}} \quad = \quad \underline{\hspace{2cm}} \text{ years/mutation}$$

$$x = \underline{\hspace{2cm}} \text{ mutations}$$



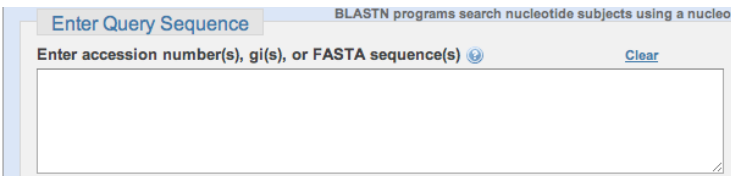

How does this number compare with the average number of SNPs your class found for the “Modern Human vs. Chimpanzee” comparisons and how can you account for any discrepancy?

f. Which mutation rate might be more accurate, that derived from the modern human/modern human comparisons or that derived from the chimpanzee/modern human comparisons? Explain.

Enter your sequence here  
 Enter **NC\_012920** here

## Sequencing Activity Two: Determining your mtDNA Haplogroup

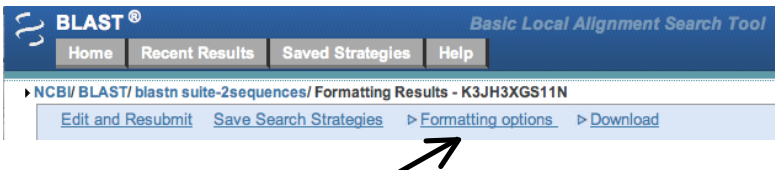
**Note:** This exercise provides information about deep ancestry and ancient human migration only. It is not meant to provide information about genealogy, family history, heredity, race, or any other classification.

<b>BLASTing your mtDNA Sequence</b>	
1. Open up an Internet browser window. (This might be Internet Explorer, Safari, Firefox, or Netscape Navigator, etc.)	
2. In the address box, type in the following URL: <a href="http://www.ncbi.nlm.nih.gov/">http://www.ncbi.nlm.nih.gov/</a> The National Center for Biotechnology Information page will open. NCBI is a database of genome sequences and biomedical research articles.	
3. On right side of home page under "Popular Resources", click on: BLAST	<div style="display: flex; justify-content: space-between;"> <div style="width: 60%;"> <p><b>Welcome to NCBI</b></p> <p>The National Center for Biotechnology Information advances science and health by providing access to biomedical and genomic information.</p> <p style="text-align: center;"> <a href="#">About the NCBI</a>   <a href="#">Mission</a>   <a href="#">Organization</a>   <a href="#">Research</a>   <a href="#">RSS Feeds</a> </p> <p><b>Get Started</b></p> <ul style="list-style-type: none"> <li>• <a href="#">Tools</a>: Analyze data using NCBI software</li> <li>• <a href="#">Downloads</a>: Get NCBI data or software</li> <li>• <a href="#">How-To's</a>: Learn how to accomplish specific tasks at NCBI</li> <li>• <a href="#">Submissions</a>: Submit data to GenBank or other NCBI databases</li> </ul> </div> <div style="width: 35%; border-left: 1px solid #ccc; padding-left: 10px;"> <p><b>Popular Resources</b></p> <ul style="list-style-type: none"> <li><a href="#">PubMed</a></li> <li><a href="#">Bookshelf</a></li> <li><a href="#">PubMed Central</a></li> <li><a href="#">PubMed Health</a></li> <li><a href="#">BLAST</a></li> <li><a href="#">Nucleotide</a></li> <li><a href="#">Genome</a></li> <li><a href="#">SNP</a></li> <li><a href="#">Gene</a></li> <li><a href="#">Protein</a></li> </ul> </div> </div>
4. Scroll down to the bottom of the BLAST page. Under <b>Specialized BLAST</b> , click on: Align two (or more) sequences using BLAST (bl2seq)	<p style="text-align: center;"><b>Specialized BLAST</b></p> <p style="text-align: center;">Choose a type of specialized search (or database name in parentheses.)</p> <ul style="list-style-type: none"> <li><input type="checkbox"/> Make specific primers with <a href="#">Primer-BLAST</a></li> <li><input type="checkbox"/> Search <a href="#">trace archives</a></li> <li><input type="checkbox"/> Find <a href="#">conserved domains</a> in your sequence (cds)</li> <li><input type="checkbox"/> Find sequences with similar <a href="#">conserved domain architecture</a></li> <li><input type="checkbox"/> Search sequences that have <a href="#">gene expression profiles</a> (GEO)</li> <li><input type="checkbox"/> Search <a href="#">immunoglobulins</a> (IgBLAST)</li> <li><input type="checkbox"/> Search using <a href="#">SNP flanks</a></li> <li><input type="checkbox"/> Screen sequence for <a href="#">vector contamination</a> (vecscreen)</li> <li><input type="checkbox"/> <b>Align</b> two (or more) sequences using BLAST (bl2seq)</li> <li><input type="checkbox"/> Search <a href="#">protein</a> or <a href="#">nucleotide</a> targets in PubChem BioAssay</li> <li><input type="checkbox"/> Search SRA <a href="#">transcript and genomic libraries</a></li> </ul>
5. In white box under <b>Enter Query Sequence</b> , copy and paste your sequence. Use the text file of your sequence, which should contain about 400 GATC letters. Be careful not to take any descriptive text other than GATC.	
6. In white box under <b>Enter Subject Sequence</b> , type the accession # for the Cambridge Reference Sequence: <b>NC_012920</b>	

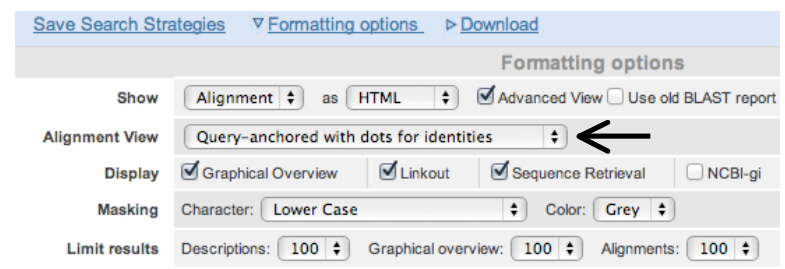
7. Click on the BLAST button at the bottom of the page



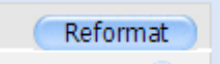
8. When the results appear, click on **Formatting Options** at the top the page.



9. Change the **Alignment View** pull-down menu from Pairwise to Query-anchored with dots for identities



10. Click on "Reformat" in the upper right corner.



11. At the bottom of the page, your sequence (Query) will appear aligned with the Cambridge Reference Sequence (NC\_012920).

Dots indicate that the sequences match. Point mutations are indicated with a letter.

For example, as position 16278, the CRS indicates that a **C** should be there. But this sequence has a **T** in that location. A **C** has been replaced by a **T**.

<input type="checkbox"/> Query	6	TTCTAATTTAAACTATTCTCTGTTCTTTCATGGGGAAGCAGATTTGGGTACCACCCAAGT	65
<input type="checkbox"/> NC_012920	16002	.....	16061
<input type="checkbox"/> Query	66	ATTGACTCACCATCAACAACCGCTATGTATTTTCGTACATTACTGCCAGCCACCATGAAT	125
<input type="checkbox"/> NC_012920	16062	.....	16121
<input type="checkbox"/> Query	126	ATTGTACGGTACCATAAATACTTGACCACCTGTAGTACATAAAAACCCAATCCACATCAA	185
<input type="checkbox"/> NC_012920	16122	.....	16181
<input type="checkbox"/> Query	186	AACCCCTCCCATGCTTACAAGCAAGTACAGCAATCAACCTCAACTATCACACATCAA	245
<input type="checkbox"/> NC_012920	16182	.....	16241
<input type="checkbox"/> Query	246	CTGCAACTCCAAGCCACCCCTCACCCACTAGGATATCAACAAACCTACCCACCCCTTAAC	305
<input type="checkbox"/> NC_012920	16242	.....C.....	16301
<input type="checkbox"/> Query	306	AGTACATAGCACATAAAGCCATTACCGTACATAGCACATTACAGTCAAATCCCTTCTCG	365
<input type="checkbox"/> NC_012920	16302	.....T.....	16361
<input type="checkbox"/> Query	366	CCCCATGGATGACCCCTCAGATAGGGTCCCTTGACCACCATCCTC	414
<input type="checkbox"/> NC_012920	16362	T.....	16410

12. Catalog all the mismatches between your sequence and the CRS.  
 You will have to count across the rows to find the exact position of each mutation.  
 Note the nucleotide position number and the nucleotide change

In the sequence above, point mutations would be written as:

**16378T**  
**16311C**  
**16362C**

**To determine your haplogroup, proceed to the exercise on the next page.**

## Using your SNPs to Determine your Haplogroup

1. Open up an Internet browser window. (This might be Internet Explorer, Safari, Firefox, or Netscape Navigator, etc.)



2. In the address box, type in the following URL: <http://nnhgtool.nationalgeographic.com/classify/index.html>  
This is a database from the Genographic Project, sponsored by National Geographic. It contains mtDNA SNP data from hundreds of thousands of people from all around the world.



3. When the link opens, scroll to the bottom of the page and you will see a white box.

Type in the nucleotide position number and the nucleotide change for all the SNPs that you found in your mtDNA sequence.

Type one per line, then click "submit".

1. Upload the samples file:  no file selected

2. Or copy and paste its contents into the area below  
(One sample per line, separate the mutations with space)  
Example: 16039A 16188T 16189C 16223T 16290T 16319A 16356C 16362C 16519C

```
16378T
16311C
16362C
```

4. The results will appear with a determination of your haplogroup for each SNP. If more than one haplogroup appears, the program will make a call for you.

For example, haplogroup H is called here, even though one of the SNPs could belong to haplogroup HV.

The database is continuing to grow and is updated frequently. This is an active area of research.

### [Back to Haplogroup Prediction Tool](#)

Sample: Haplogroup: H  
 Sample: 16378T Haplogroup: H  
 Sample: 16311C Haplogroup: HV  
 Sample: 16362C Haplogroup: H

5. See page 32 for more information on haplogroups.

In addition, you can research your haplogroup and learn all about your deep ancestry. There are many websites available for this purpose. A few are listed here:

The Genographic Project:

[http://education.nationalgeographic.com/education/collections/genographic/?ar\\_a=1](http://education.nationalgeographic.com/education/collections/genographic/?ar_a=1)

Mitomap:

<http://www.mitomap.org/MITOMAP>

Genebase:

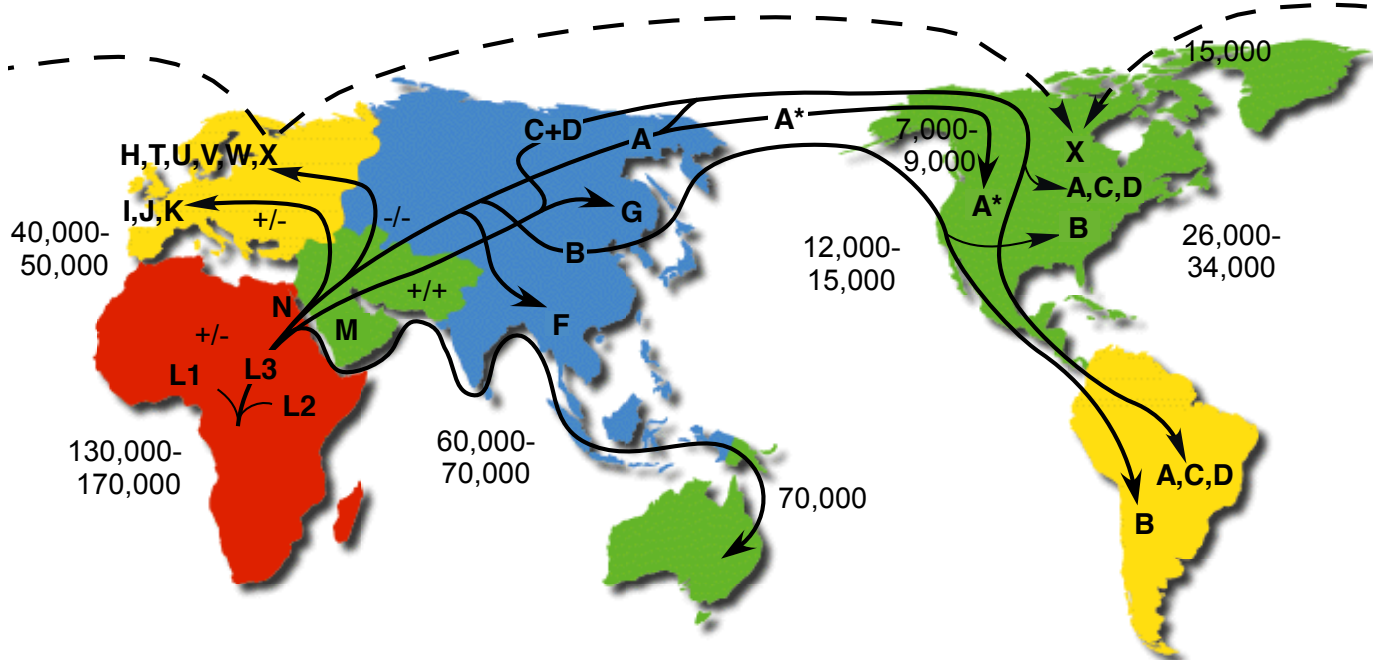
<http://www.genebase.com/learning/>

World Families:

[http://www.worldfamilies.net/reverence\\_mtDNA](http://www.worldfamilies.net/reverence_mtDNA)

## Human mtDNA Migrations

The map below shows the branch point for each haplogroup and its global spread.



<http://www.mitomap.org/pub/MITOMAP/MitomapFigures/WorldMigrations.pdf>

Content made available through a Creative Commons Attribution 3.0 License.

## Major Haplogroups

For a listing of all subgroups, see <http://www.mitomap.org>

Haplogroup	Possible time of origin	Possible place of origin (branch point)
L1, L2, L3	130,000 - 170,000 years ago	Central Africa
N	71,000 years ago	East Africa or Asia
M	60,000 years ago	North Africa or South Asia
I	30,000 years ago	Caucasus or Northeast Europe
J	45,000 years ago	Near East or Caucasus
K	16,000 years ago	Near East
H	35,000 years ago	Western Asia
T	17,000 years ago	Mesopotamia
V	15,000 years ago	Iberia and moved to Scandinavia
W	25,000 years ago	Northeast Europe or Northwest Asia
X	30,000 years ago	Northeast Europe
A	50,000 years ago	Asia
B	50,000 years ago	East Asia
C	60,000 years ago	Central Asia
D	50,000 years ago	East Asia
F	40,000 years ago	Asia
G	35,000 years ago	East Asia

My haplogroup is \_\_\_\_\_.

It branched off from haplogroup \_\_\_\_\_ in \_\_\_\_\_ (country).

This happened \_\_\_\_\_ years ago.

## Life Technologies & Applied Biosystems / BABEC Educational PCR Kits



**For Research Use Only. Not for use in diagnostic procedures.**

### NOTICE TO PURCHASER: LIMITED LICENSE

A license under U.S. Patents 4,683,202, 4,683,195, and 4,965,188 or their foreign counterparts, owned by Roche Molecular Systems, Inc. and F. Hoffmann-La Roche Ltd (Roche), for use in research and development, has an up-front fee component and a running-royalty component. The purchase price of the Lambda PCR, Alu PV92 PCR, PCR Optimization, D1S80 PCR, and Mitochondrial PCR Kits includes limited, non-transferable rights under the running-royalty component to use only this amount of the product to practice the Polymerase Chain Reaction (PCR) and related processes described in said patents solely for the research and development activities of the purchaser when this product is used in conjunction with a thermal cycler whose use is covered by the up-front fee component. Rights to the up-front fee component must be obtained by the end user in order to have a complete license. These rights under the up-front fee component may be purchased from Applied Biosystems or obtained by purchasing an authorized thermal cycler. No right to perform or offer commercial services of any kind using PCR, including without limitation reporting the results of purchaser's activities for a fee or other commercial consideration, is hereby granted by implication or estoppel. Further information on purchasing licenses to practice the PCR process may be obtained by contacting the Director of Licensing at Applied Biosystems, 850 Lincoln Centre Drive, Foster City, California 94404 or at Roche Molecular Systems, Inc., 1145 Atlantic Avenue, Alameda, California 94501.

Use of this product is covered by US patent claims and corresponding patent claims outside the US. The purchase of this product includes a limited, non-transferable immunity from suit under the foregoing patent claims for using only this amount of product for the purchaser's own internal research. No right under any other patent claim (such as the patented 5' Nuclease Process claims) and no right to perform commercial services of any kind, including without limitation reporting the results of purchaser's activities for a fee or other commercial consideration, is conveyed expressly, by implication, or by estoppel. This product is for research use only. Diagnostic uses require a separate license from Roche. Further information on purchasing licenses may be obtained by contacting the Director of Licensing, Applied Biosystems, 850 Lincoln Centre Drive, Foster City, California 94404, USA.

### TRADEMARKS:

Applied Biosystems, AB (Design), GeneAmp, and Primer Express are registered trademarks and Veriti and VeriFlex are trademarks of Applied Biosystems Inc. or its subsidiaries in the US and/or certain other countries.

AmpliTaq is a registered trademark of Roche Molecular Systems, Inc. All other trademarks are the sole property of their respective owners.

© Copyright 2001, Applied Biosystems. All rights reserved.