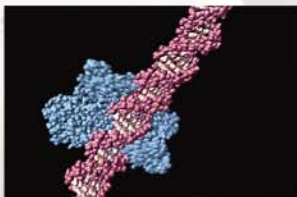




# Using Mitochondrial DNA Polymorphisms in Evolutionary and Forensic Biology

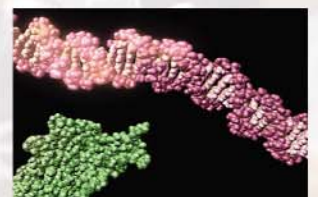
**DNA KITS**  
Learning Center



DOLAN  
  
DNA  
LEARNING  
CENTER



**CAROLINA**



# Using Mitochondrial DNA Polymorphisms in Evolutionary and Forensic Biology

## IMPORTANT INFORMATION

---

**Storage:** Upon receipt of the kit, store proteinase K, mt primer/loading dye mix, and DNA marker pBR322/*Bst*NI in a freezer (approximately  $-20^{\circ}\text{C}$ ). All other materials may be stored at room temperature (approximately  $25^{\circ}\text{C}$ ).

**Use and Lab Safety:** The materials supplied are for use with the method described in this kit only. Use of this kit presumes and requires prior knowledge of basic methods of gel electrophoresis and staining of DNA. Individuals should use this kit only in accordance with prudent laboratory safety precautions and under the supervision of a person familiar with such precautions. Use of this kit by unsupervised or improperly supervised individuals could result in injury.

**Limited License:** Polymerase chain reaction (PCR) is protected by patents owned by Hoffman-La Roche, Inc. The purchase price of this product includes a limited, non-transferable license under U.S. Patents 4,683,202; 4,683,195; and 4,965,188 or their foreign counterparts, owned by Hoffmann-La Roche Inc. and F. Hoffmann-La Roche Ltd. (Roche), 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, educational, and training activities of the purchaser when this product is used either manually or in conjunction with 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 The Perkin-Elmer Corporation, 850 Lincoln Center Drive, Foster City, California 94404 or at Roche Molecular Systems, Inc., 1145 Atlantic Avenue, Alameda, California 94501.

**Printed material:** The student instructions, pages 5-23, as well as the *Carolina*BLU™ staining protocol on page 29 may be photocopied as needed for use by your students.

## REAGENTS, SUPPLIES, AND EQUIPMENT CHECKLIST

---

### Included in the kit:

#### *DNA extraction and amplification (all kits):*

- 1.5 g Chelex® resin
- 5 mL proteinase K (100 µg/mL)
- 700 µL mt primer/loading dye mix
- 25 \*Ready-to-Go™ PCR Beads
- 5 mL mineral oil
- 130-µL tube pBR322/*Bst*NI markers (0.075 µg/µL)
- Instructor's manual with reproducible Student Lab Instructions
- MtDNA CD-ROM

#### **\*\*Electrophoresis kits with ethidium bromide staining**

##### *(Kits 21-1237 and 21-21-1237A) also include:*

- 5 g agarose
- 150 mL 20 $\times$  TBE
- 250 mL ethidium bromide, 1 µg/mL
- 4 latex gloves
- 6 staining trays

#### **\*\*Electrophoresis kits with CarolinaBLU™ staining**

##### *(Kits 21-1238 and 21-1238A) also include:*

- 5 g agarose
- 150 mL 20 $\times$  TBE
- 7 mL CarolinaBLU™ Gel & Buffer Stain
- 250 mL CarolinaBLU™ Final Stain
- 4 latex gloves
- 6 staining trays

### Needed but not supplied:

- 0.9% saline solution (NaCl), 10 mL per student in 15 mL tube
- Micropipets and tips (1 µL to 1000 µL)
- 1.5 mL microcentrifuge tubes, polypropylene, 2 per student
- Microcentrifuge tube racks
- Microcentrifuge for 1.5-mL tubes
- 0.2 mL or 0.5 mL PCR tubes, 1 per student (1.5 mL microcentrifuge tubes may also be used.)
- 0.2 mL or 0.5 mL tube adapters for microcentrifuge (can be made from 0.5 mL and/or 1.5 mL tubes)
- Thermal cycler, programmable
- Electrophoresis chambers
- Electrophoresis power supplies
- Gel staining trays
- UV transilluminator (ethidium bromide staining)
- White light box (*CarolinaBLU™* staining, optional)
- Camera or photo-documentary system (optional)
- Paper cup, 1 per student
- Permanent markers
- Container with cracked or crushed ice
- Boiling water bath (optional, see instructions)

\*Ready-to-Go™ PCR Beads incorporate *Taq* polymerase, dNTPs, and MgCl<sub>2</sub>. Each bead is supplied in an individual 0.5-mL tube or a 0.2-mL tube.

\*\*Electrophoresis reagents must be purchased separately for Kits 21-1230 and 21-1230A.

# Using Mitochondrial DNA Polymorphisms in Evolutionary and Forensic Biology

## CONTENTS

<b>STUDENT LAB INSTRUCTIONS</b> .....	5
INTRODUCTION .....	5
LAB FLOW .....	7
METHODS .....	8
BIOINFORMATICS .....	13
RESULTS AND DISCUSSION .....	17
<b>INFORMATION FOR INSTRUCTOR</b> .....	25
CONCEPTS AND METHODS .....	25
INSTRUCTOR PLANNING, PREPARATION, AND LAB FINE POINTS .....	26
<i>Carolina</i> BLU™ STAINING .....	32
BIOINFORMATICS .....	33
ANSWERS TO BIOINFORMATICS QUESTIONS .....	33
ANSWERS TO DISCUSSION QUESTIONS .....	34
CD-ROM CONTENTS .....	35



# STUDENT LAB INSTRUCTIONS

## INTRODUCTION

---

In addition to the 46 chromosomes found in the nucleus of human cells, each mitochondrion in the cell cytoplasm has several copies of its own genome. The mitochondrial (mt) genome contains only 37 genes, all of which are involved in the process of oxidative phosphorylation—the production of energy and its storage in ATP.

There is strong evidence that the mitochondrion once existed as a free-living bacterium that was taken up by a primitive ancestor of eukaryotic cells. In this endosymbiosis, the host cell provided a ready source of energy-rich nutrients, and the mitochondrion provided a means to extract energy using oxygen. This attribute was key to survival, as oxygen accumulated in the primitive atmosphere. Mitochondria are physically in the same size range as bacteria, and the mt genome retains several bacterial features. Like bacterial chromosomes and plasmids, the mt genome is a circular molecule, and mt genes are not interrupted by introns. These features are contrary to those of eukaryotic chromosomes—which are linear—and eukaryotic genes—which have numerous introns.

The entire DNA sequence of the human mt genome (16,569 nucleotides) was determined in 1981, well in advance of the Human Genome Project. Genes take up the majority of the mt genome. However, a noncoding region of approximately 1200 nucleotides contains signals that control replication of the chromosome and transcription of the mt genes. The DNA sequence of the control region is termed “hypervariable,” because it accumulates point mutations at approximately 10 times the rate of nuclear DNA. This high mutation rate results in distinctive patterns of single nucleotide polymorphisms (SNPs).

The combination of SNPs inherited by each person is termed a haplotype (or “half type”). The mt genome provides only a half set of genes, because it is inherited exclusively from the mother with no paternal contribution. The female egg is a huge cell with on the order of 100,000 mitochondria, while the tiny sperm cell is powered by fewer than 100 mitochondria at the base of the flagellum. Any male mitochondria that may enter the egg cell at conception are identified by their ubiquitin surface proteins as “foreign” and actively destroyed by enzymes in the egg cytoplasm.

In the 1980s, Allan Wilson and coworkers at the University of California at Berkeley used mtDNA polymorphisms to create a “family tree” showing ancestral relationships between modern populations. Reasoning that all human populations arose from a common ancestor in the distant evolutionary past, Wilson’s group calculated how long it would take to accumulate the pattern of mt mutations observed in modern populations. They concluded that the ancestor of all modern humans arose in Africa about 200,000 years ago. This common ancestor was widely reported as the “mitochondrial Eve.”

While each cell contains only two copies of a given nuclear DNA sequence — one on each of the paired chromosomes — there are hundreds to

thousands of copies of a given mt DNA sequence in each cell. This amplification increases the chances that enough mtDNA can be obtained for forensic analysis when tissue samples are old or badly degraded. For example, polymorphisms in the mt hypervariable region have been used to:

- Identify remains from wars and natural disasters.
- Identify the remains of the Romanov royal family assassinated during the Russian Revolution.
- Determine the relationship of Otzi, the Tyrolean ice man and the ancient hominid Neandertal to modern humans.

Because of the high copy number of the mt genome, it offers the simplest means for a person to visualize a discrete region of their own genetic material. This experiment examines a sequence within the hypervariable region of the mt genome. A sample of human cells is obtained by saline mouthwash (alternatively DNA may be isolated from hair sheaths). DNA is extracted by boiling with Chelex® resin, which binds contaminating metal ions, and the control region sequence is amplified by polymerase chain reaction (PCR).

However, since SNP polymorphisms do not change the size of the PCR product (amplicon), gel electrophoresis shows no differences between student samples. To analyze SNPs that vary from person to person, the nucleotide sequence of each student amplicon must be determined. This can be done by submitting student amplicons to the Sequencing Service of the Dolan DNA Learning Center (DNALC) of Cold Spring Harbor, which will generate student mtDNA sequences and post the results via Internet free of charge. Tools for generating mt haplotypes, comparing student sequences, and studying human evolution comparing are found at the DNALC's *BioServers* Internet site ([www.bioservers.org](http://www.bioservers.org)).

- Biesecker L.G., Bailey-Wilson J.E., Ballantyne J., Baum H., Bieber F.R., Brenner C., Budowle B., Butler J.M., Carmody G., Conneally P.M., Duceman B., Eisenberg A., Forman L, Kidd K.K., Leclair B., Niezgodna S., Parsons T.J., Pugh E., Shaler R., Sherry S.T., Sozer A. and Walsh A. (2005. Epidemiology. DNA Identifications After the 9/11 World Trade Center Attack. *Science* 310(5751):1122-1123. Cann R.L., Stoneking, M. and Wilson, A.C. (1987). Allan C. Mitochondrial DNA and Human Evolution. *Nature* 325: 31-36.
- Gill, P., Ivanov, P.L., Kimpton, C., Piercy, R., Benson, N., Tully, G., Evett, I., Haefelberg, E. and Sullivan, K. (1994). Identification of the Remains of the Romanov Family by DNA Analysis. *Nature Genetics* 6(2): 130-135.
- Handt O., Richards M., Trommsdorff M., Kilger C., Simanainen J., Georgiev O., Bauer K., Stone A., Hedges R., Schaffner W, et al. (1994). Molecular genetic analyses of the Tyrolean Ice Man. *Science* 264(5166):1775-1778.
- Krings M., Stone A., Schmitz R.W., Krainitzki H., Stoneking M. and Pääbo S. (1997). Neandertal DNA Sequences and the Origin of Modern Humans. *Cell* 90(1):19-30.
- Mullis, K. (1990). The Unusual Origin of the Polymerase Chain Reaction. *Scientific American* 262(4): 56-65.

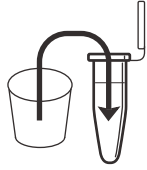
**LAB FLOW**

**I. ISOLATE DNA FROM CHEEK CELLS**

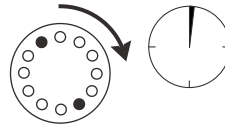
RINSE  
mouth  
with  
saline



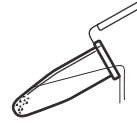
TRANSFER  
saline



CENTRIFUGE



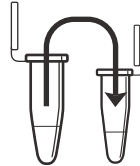
POUR OFF  
supernatant



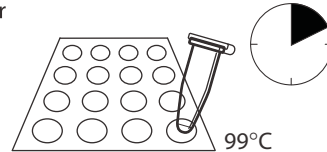
RESUSPEND



TRANSFER  
cell  
suspension



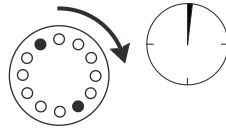
BOIL  
in thermal  
cycler



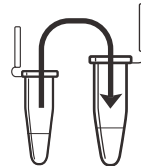
SHAKE  
vigorously



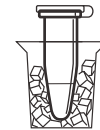
CENTRIFUGE



TRANSFER  
supernatant



STORE  
on ice

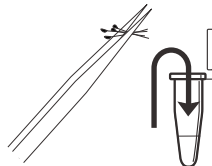


**(ALTERNATE) I. ISOLATE DNA FROM HAIR SHEATHS**

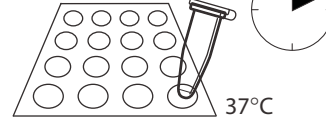
CUT  
hairs



TRANSFER  
hairs



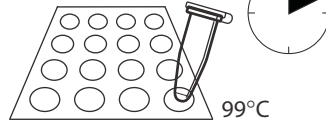
INCUBATE  
in thermal  
cycler



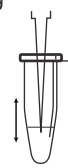
VORTEX



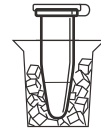
BOIL  
in thermal  
cycler



MIX  
by pipetting  
in and out



STORE  
on ice



**II. AMPLIFY DNA BY PCR**

ADD  
primer/  
loading  
dye mix



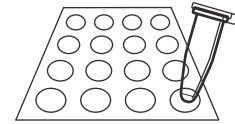
ADD  
DNA



ADD  
mineral oil  
(if necessary)

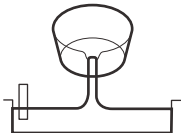


AMPLIFY  
in thermal  
cycler



**III. ANALYZE PCR PRODUCTS BY GEL ELECTROPHORESIS**

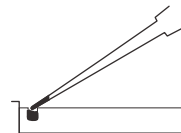
POUR gel



SET



LOAD gel



ELECTROPHORESE  
130V



## METHODS

### I. ISOLATE DNA FROM CHEEK CELLS

Reagents	Supplies and Equipment
0.9% Saline solution, 10 mL	Permanent marker
10% Chelex <sup>®</sup> , 100 $\mu$ L (in 0.2 or 0.5 mL PCR tube)	Paper cup
	Micropipets and tips (10-1000 $\mu$ L)
	1.5 mL microcentrifuge tubes
	Microcentrifuge tube rack
	Microcentrifuge adapters
	Microcentrifuge
	Thermal cycler (or water bath or heat block)
	Container with cracked or crushed ice
	Vortexer (optional)

1. Use a permanent marker to label a 1.5 mL tube and paper cup with your assigned number.
2. Pour saline solution into your mouth, and vigorously rinse your cheek pockets for 30 seconds.
3. Expel saline solution into the paper cup.
4. Swirl cup gently to mix cells that may have settled to the bottom. Use micropipet with fresh tip to transfer 1500  $\mu$ L of the solution into your labeled 1.5-mL microcentrifuge tube.
5. Place your sample tube, along with other student samples, in a balanced configuration in a microcentrifuge, and spin for 90 seconds at full speed.
6. Carefully pour off supernatant into the paper cup. Try to remove most of the supernatant, but be careful not to disturb cell pellet at the bottom of the tube. (The remaining volume will approximately reach the 0.1 mark of a graduated tube.)
7. Set micropipet to 30  $\mu$ L. Resuspend cells in the remaining saline by pipetting in and out. Work carefully to minimize bubbles.
8. Withdraw 30  $\mu$ L of cell suspension, and add to a PCR tube containing 100  $\mu$ L of Chelex<sup>®</sup>. Label the cap and side of the tube with your assigned number.
9. Place your PCR tube, along with other student samples, in a thermal cycler that has been programmed for one cycle of the following profile. The profile may be linked to a 4°C hold program.

Boiling step:      99°C    10 minutes

10. After boiling, vigorously shake the PCR tube for 5 seconds.

Before pouring off supernatant, check to see that pellet is firmly attached to tube. If pellet is loose or unconsolidated, carefully use micropipet to remove as much saline solution as possible.

Food particles will not resuspend.

Alternatively, you may add the cell suspension to Chelex in a 1.5 mL tube, and incubate in a boiling water bath or heat block.

Your teacher may instruct you to collect a sample of cell suspension to observe under a microscope.

The near-boiling temperature lyses the cell and nuclear membranes, releasing DNA and other cell contents.



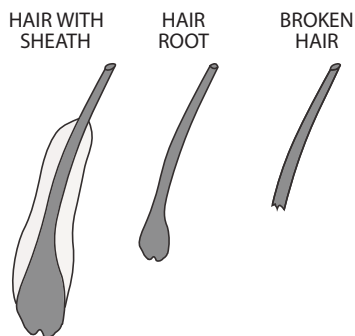
To use adapters, “nest” the sample tube within sequentially larger tubes: 0.2 mL within 0.5 mL within 1.5 mL. Remove caps from tubes used as adapters.

- Place your tube, along with other student samples, in a balanced configuration in a microcentrifuge, and spin for 90 seconds at full speed. *If your sample is in a PCR tube, one or two adapters will be needed to spin the tube in a microcentrifuge designed for 1.5 mL tubes.*
- Use a micropipet with fresh tip to transfer 30  $\mu\text{L}$  of the clear supernatant into a clean 1.5 mL tube. Be careful to avoid pipetting any cell debris and Chelex<sup>®</sup> beads.
- Label the cap and side of the tube with your assigned number. This sample will be used for setting up one or more PCR reactions.
- Store your sample on ice or at  $-20^{\circ}\text{C}$  until you are ready to continue with Part II.

### I. (ALTERNATE) ISOLATE DNA FROM HAIR SHEATHS

Reagent	Supplies and Equipment
100 mg/mL proteinase K, 100 $\mu\text{L}$ (in 0.2 or 0.5 mL tube)	Permanent marker Scalpel or razor blade Forceps or tweezers Thermal cycler (or water bath or heat block) Container with cracked or crushed ice Vortexer (optional)

Your teacher may instruct you to prepare a hair sheath to observe under a microscope.



Alternatively, you may add the hairs to proteinase K in a 1.5 mL tube, and incubate in a water bath or heat block.

- Pull out several hairs and inspect for presence of a sheath. The sheath is a barrel-shaped structure surrounding the base of the hair, and can be readily observed with a hand lens or dissecting microscope. The glistening sheath can be observed with the naked eyes by holding the hair up to a light source. (Sheaths are most easily observed on dark hair.)
- Select one to several hairs with good sheaths. Alternately, select hairs with the largest roots. *Broken hairs, without roots or sheaths, will not yield enough DNA for amplification.*
- Use a fresh razor blade or scalpel to cut off hair shafts just above the sheath.
- Use forceps to transfer hairs to a PCR tube containing 100  $\mu\text{L}$  of proteinase K. *Make sure sheath is submerged in the solution and not stuck on the test tube wall.* Label the cap and side of the tube with your assigned number.
- Place your PCR tube, along with other student samples, in a thermal cycler that has been programmed for one cycle of the following profile.  
Incubation Step: 37 $^{\circ}\text{C}$  10 minutes
- Remove sample tube to room temperature. Vortex by machine or vigorously with finger for 15 seconds to dislodge cells from hair shaft.

- Place your PCR tube, along with other student samples, in a thermal cycler that has been programmed for one cycle of the following profile. The profile may be linked to a 4°C hold program.

Boiling step:        99°C    10 minutes

- Remove sample tube to room temperature, and mix by pipetting in and out for 15 seconds.
- Store your sample on ice or in the freezer until ready to begin Part II.

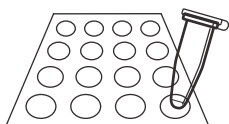
## II. AMPLIFY DNA BY PCR

Reagents (at each student station)	Supplies and Equipment
*Cheek cell DNA. 2.5 $\mu$ L (from Part I)	Permanent marker
*Mt primer/loading dye mix, 25 $\mu$ L	Micropipet and tips (1-100 $\mu$ L)
Ready-To-Go™ PCR beads (in 0.2 mL or 0.5 mL PCR tube)	Microcentrifuge tube rack
	Thermal cycler
	Container with cracked or crushed ice
<b>Shared Reagent</b>	
Mineral oil, 5 mL (depending on thermal cycler)	
*Store on ice	

The primer/loading dye mix will turn purple as the PCR bead dissolves.

If the reagents become splattered on the wall of the tube, pool them by pulsing in a microcentrifuge or by sharply tapping the tube bottom on the lab bench.

If your thermal cycler does not have a heated lid: Prior to thermal cycling, you must add a drop of mineral oil on top of your PCR reaction. Be careful not to touch the dropper tip to the tube or reaction, or the oil will be contaminated with your sample.



- Obtain a PCR tube containing a Ready-To-Go™ PCR Bead. Label with your assigned number.
- Use a micropipet with fresh tip to add 22.5  $\mu$ L of mt primer/loading dye mix to the tube. Allow the bead to dissolve for a minute or so.
- Use a micropipet with fresh tip to add 2.5  $\mu$ L of your cheek cell DNA (from Part I) *directly into* the primer/loading dye mix. Insure that no cheek cell DNA remains in the tip after pipetting.
- Store your sample on ice until your class is ready to begin thermal cycling.*
- Place your PCR tube, along with other student samples, in a thermal cycler that has been programmed for 30 cycles of the following profile. The profile may be linked to a 4°C hold program after the 30 cycles are completed.
 

Denaturing step:	94°C	30 seconds
Annealing step:	58°C	30 seconds
Extending step:	72°C	30 seconds
- After cycling, store the amplified DNA on ice or at -20°C until you are ready to continue with Part III.

### III. ANALYZE PCR PRODUCTS BY GEL ELECTROPHORESIS

#### Reagents

\*PCR product (from Part II), 15  $\mu$ L

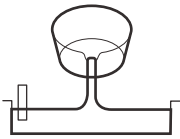
#### Shared Reagents

\*pBR322/*Bst*NI marker  
 1.5% agarose in 1X TBE, 50 mL  
 1X TBE, 300 mL  
 Ethidium bromide (1  $\mu$ g/mL), 250 mL  
 or  
*Carolina*BLU™ Gel & Buffer Stain, 7 mL  
*Carolina*BLU™ Final Stain, 250 mL

\*Store on ice

#### Supplies and Equipment

Micropipet and tips (1-100  $\mu$ L)  
 Microcentrifuge tube rack  
 Gel electrophoresis chamber  
 Power supply  
 Staining trays  
 Latex gloves  
 UV transilluminator (for use with ethidium bromide)  
 White light transilluminator (for use with *Carolina*BLU™)  
 Digital or instant camera (optional)  
 Water bath (60°C)  
 Container with cracked or crushed ice

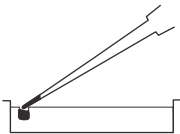


Avoid pouring an overly thick gel, which is more difficult to visualize. The gel will become cloudy as it solidifies.

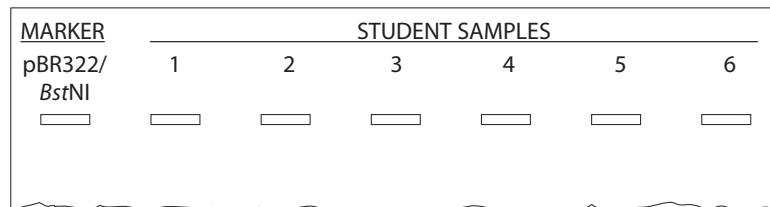
Do not add more buffer than necessary. Too much buffer above the gel channels electrical current over the gel, increasing running time.

100-bp ladder may also be used as a marker.

Expel any air from the tip before loading. Be careful not to push the tip of the pipet through the bottom of the sample well.



1. Seal the ends of the gel-casting tray with masking tape, and insert a well-forming comb.
2. Pour 1.5% agarose solution to a depth that covers about 1/3 the height of the open teeth of the comb.
3. Allow the gel to solidify completely. This takes approximately 20 minutes.
4. Place the gel into the electrophoresis chamber, and add enough 1X TBE buffer to cover the surface of the gel.
5. Carefully remove the comb, and add additional 1X TBE buffer to just cover and fill in wells – creating a smooth buffer surface.
6. Use a micropipet with a fresh tip to load 20  $\mu$ L of pBR322/*Bst*NI size marker into the far left lane of the gel.
7. CAUTION: DO NOT LOAD YOUR ENTIRE SAMPLE! Use a micropipet with a fresh tip to add 15  $\mu$ L of your sample/loading dye mixture into your assigned lane of a 1.5% agarose gel, according to the diagram below. (If you used mineral oil during PCR, pierce your pipet tip through the layer of mineral oil to withdraw the PCR sample and



leave the mineral oil behind in the original tube.)

8. Store the remaining 10  $\mu$ L of your mt PCR product on ice for subsequent use in DNA sequencing.

Destaining the gel for 5-10 minutes in tap water leaches unbound ethidium bromide from the gel, decreasing background and increasing contrast of the stained DNA.

Transillumination, where the light source is below the gel, increases brightness and contrast.

9. Run the gel at 130 V for approximately 30 minutes. Adequate separation will have occurred when the cresol red dye front has moved at least 50 mm from the wells.
10. Stain the gel using ethidium bromide or *CarolinaBLU*<sup>™</sup>:
  - a. For ethidium bromide, stain 10-15 minutes. Decant stain back into storage container for reuse, and rinse gel in tap water. *Use gloves when handling ethidium bromide solution and stained gels or anything that has ethidium bromide on it. Ethidium bromide is a known mutagen and care should be taken when using and disposing of it.*
  - b. For *CarolinaBLU*<sup>™</sup>, follow directions in the Instructor Planning section.
11. View gel using transillumination, and photograph using a digital or instant camera.

## BIOINFORMATICS

---

Biological information is encoded in the nucleotide sequence of DNA. Bioinformatics is the field that identifies biological information in DNA using computer-based tools. Some bioinformatics algorithms aid the identification of genes, promoters, and other functional elements of DNA. Others help determine evolutionary relationships between sequences.

Because of the large number of tools and DNA sequences available on the Internet, experiments done *in silico* ("in silicon," or on the computer) now complement experiments done *in vitro* (in glass, or test tube). This integration of biochemistry and computation is a key feature of modern biological research.

In Part I you will use the Basic Local Alignment Search Tool (BLAST) to find database matches to the primers used in the experiment and to determine the exact size of the fragment amplified by the primer set. In Part II you will examine the structure of the human mt chromosome, the genes it contains, and its origin. In Part III you will use MITOMAP to learn about the impact of mt gene mutations on human health and aging.

The following primer set was used in the mtDNA experiment:

Forward Primer 5'-TTAACTCCACCATTAGCACC-3'

Reverse Primer 5'-GAGGATGGTGGTCAAGGGAC-3'

### I. Use BLAST to Find DNA Sequences in Databases (Electronic PCR)

1. Initiate a BLAST search.
  - a. Open the Internet site of the National Center for Biotechnology Information (NCBI) [www.ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/).
  - b. Click on *BLAST* in the top speed bar.
  - c. Click on the link *nucleotide blast* under the heading *Basic BLAST*.
  - d. Paste or type the primer sequences into the *Enter Query Sequence* window. Omit any non-nucleotide characters from the window, because they are not recognized by the BLAST algorithm.
  - e. Under *Choose Search Set*, select the *Nucleotide collection (nr/nt) database* from the drop-down menu.
  - f. Under *Program Selection*, select *optimize* for somewhat similar sequences by selecting *blastn*.
  - g. Click on the *BLAST* button at the bottom of the page, and the query sequences are sent to a server at the National Center for Biotechnology Information in Bethesda, Maryland. There, the BLAST algorithm will attempt to match the primer sequences to the millions of DNA sequences stored in its database. A temporary page showing the status of your search will be displayed until your



results are available. This may take only a few seconds, or more than a minute if a lot of other searches are queued at the server.

- h. The results of the BLAST search are displayed in three ways as you scroll down the page:
- First, a graphical overview illustrates how significant matches, or hits, align with the query sequence. Matches of differing lengths are coded by color. What do you notice?
  - This is followed by a list of significant alignments, or hits, with *Accession links*, brief description, and statistical scores. (The accession is the unique identifier given to a sequence when it is submitted to a database, such as Genbank.)
  - Next, is a detailed view of each primer sequence (query) aligned to the nucleotide sequence of the search hit (subject). Notice that a match to the forward primer (nucleotides 1–20), and a match to the reverse primer (nucleotides 21–40) are often within the same accession.

2. What sequences does your search turn up?

- a. In the list of significant alignments, notice the scores in the *E-value* column on the right. The Expectation, or *E-value*, is the number of alignments with the query sequence that would be expected to occur by chance in the database. The lower the *E-value* the higher the probability that the hit is related to the query. (For example, an *E-value* of 1 means that a search with your sequence would be expected to turn up one match by chance.)
- b. Find a match to both primers. Do your *E-values* seem high or low? Why do you think this is so?
- c. Note the descriptions of some of significant alignments that have *E-values* less than 1. Do they make sense?
- d. To learn more, click on several of the accession links to open Genbank records with detailed information about the gene sequences that align with the primers.
  - The top part of the record contains basic information about the sequence including its basepair length, database accession number, source, and references to papers in which this sequence is published.
  - The middle part lists sequence *Features* – including regulatory region, gene, coding sequence (cds), ribosomal RNA (rRNA), and transfer RNA (tRNA) – along with their beginning and ending nucleotide positions (xx .. xx). If your Genbank record does not list any features, go back and select another hit from the list of significant alignments.
  - The bottom part lists the entire nucleotide sequence.

3. What is the predicted length of the product that this primer set would amplify in a PCR reaction (*in vitro*)?

- a. Select a hit that is labeled “complete genome,” and scroll down to the *Alignments* section to see exactly where the two primers have landed in this subject sequence.
- b. The lowest and highest nucleotide positions in the subject sequence indicate the borders of the amplified sequence. Subtracting one from the other gives the difference between the two coordinates.
- c. However, the actual length of the fragment includes both ends, so add 1 nucleotide to the result to determine the exact length of the PCR product amplified by the two primers.

## II. Examine the Structure of the mt Chromosome and Its Genes

1. What part of the mt genome do your primers amplify?
  - a. Continue working with the same hit you used in Question 3 above.
  - b. Follow the underlined link in the *Alignments* section (or the *Accession link* in the list of significant alignments). This will take you to the Genbank record for that sequence.
  - c. Print out a copy of the Genbank record or copy the report into a text document to help you with the questions below.
  - d. Scroll down to the *Features* section. Next to each feature are its nucleotide coordinates on the mt chromosome.
  - e. Find the feature that includes the nucleotide positions that define the boundaries of the forward and reverse primers (from section 3.b. above). What is it called? How large is it? Does it contain any genes? Do some research to understand each of several names used for this region.
2. What are the features of the mt chromosome? (Remember that genes can be located on either the forward or reverse strand. *Complement* denotes the location of a feature that is transcribed from the reverse strand.)
  - a. Look at the features list for your hit. List the genes that are transcribed from the forward strand and those that are transcribed from the reverse (complement) strand. (Remember that, in addition to messenger RNA (mRNA) coding for proteins, ribosomal RNAs (rRNA) are also transcribed from genes.)
  - b. On which strand are most genes found?
  - c. Judging from the genes you identified, the mt genome provides all needed components for what biological function?
  - d. Does the mt genome encode all necessary components needed for oxidative phosphorylation (energy production)?
  - e. What is the spacing like between genes – is there intergenic DNA between two adjacent genes?

- f. Focus on one of the protein coding genes. Compare the nucleotide coordinates for the gene and the CDS (coding sequence). What do you notice? What does this tell you about the structure of mt genes and their origin?

### III. Use MITOMAP to Explore the Role of mt Mutations in Human Health and Aging

MITOMAP is a database on the human mt genome.

1. Open the MITOMAP Internet site at <http://www.mitomap.org>.
2. Scroll down the page to the section *MtDNA Mutations with Reports of Disease-Associations*.
3. Follow the links *Organized by Location*:
  - a. *rRNA/tRNA Mutations*
  - b. *Coding & Control Region Mutations*
4. Explore several of the diseases:
  - a. Under the *Allele* column, correlate the location of the causative mutation with the locations of RNA and protein genes in the *Features* section of the Genbank report from Part II above.
  - b. Follow the *Disease* link to a page with links to OMIM (Online Mendelian Inheritance in Man). Click on the link for the disease, then select the link that appears on the OMIM page (usually the first one).
  - c. The OMIM report will provide detailed description of the disease.
5. Jot down notes on several disorders and share your findings with your classmates.
6. What sort of mutations generally cause mt diseases? What do many of these diseases have in common?
7. Research the mitochondrial theory of aging.

## RESULTS AND DISCUSSION

---

### I. Interpret Your Gel and Think About the mt Genome

1. Observe the photograph of the stained gel containing your sample and those from other students. Orient the photograph with the sample wells at the top. Interpret the band(s) in each lane of the gel:
  - a. Scan across the photograph to get an impression of what you see in each lane. You should notice that virtually all student lanes contain one or two prominent bands.
  - b. Now locate the lane containing the pBR322/*Bst*NI markers. Working from the well, locate the bands corresponding to each restriction fragment: 1,857 bp, 1,058 bp, 929 bp, 383 bp, and 121 bp (may be faint).
  - c. The amplification product of 440 bp should roughly align with the 383-bp marker.
  - d. It is common to see a second band lower on the gel. This diffuse (fuzzy) band is "primer dimer," an artifact of the PCR reaction that results from the primers overlapping one another and amplifying themselves. Primer dimer is approximately 50 bp, and should be in a position ahead of the 121-bp marker.
  - e. Additional faint bands, at other positions on the gel, occur when the primers bind to chromosomal loci other than mt control region and give rise to "nonspecific" amplification products.
2. How would you interpret a lane in which you observe primer dimer, but no 440-bp band?
3. The mt control region mutates at approximately 10 times the rate of nuclear DNA. Propose a biological reason for the high mutation rate of mt DNA.
4. The high mutability of the mt genome means that it evolves more quickly than the nuclear genome. This makes the mt control region a laboratory for the study of DNA evolution. However, can you think of any drawbacks to this high mutation rate?
5. There are numerous insertions of mt DNA into nuclear chromosomes. Most of these are very ancient, and are part of the genome reduction that transferred functional mt genes to the nuclear genome. However, a 540-bp insertion of the mt control region on chromosome 11 is an example of a relatively recent event that occurred about 350,000 years ago. Would you expect any difference in the mutation rates of the control region sequence in the mt genome *versus* the chromosome 11 insertion? What implication does this have in the study of human evolution?

## II. Visualize and Assess Your Own DNA Sequence

(If your DNA has not been sequenced, you can follow these directions to assess another student's mt DNA sequence.)

1. Open the *BioServers* Internet site at the Dolan DNA Learning Center [www.bioservers.org](http://www.bioservers.org).
2. Enter *Sequence Server*. You can register if you want to save your work for future reference, but it is not required.
3. The interface is simple to use: add or obtain data using the top buttons and pull-down menus, then work with the data in the workspace below.
4. Click on *MANAGE GROUPS* at the top of the page.
5. Scroll down the default *Classes* list to find your class in the list.
6. Click in the check box to select your class, and click on *OK* to move it onto the workspace.
7. Select your sample number from the pull-down menu. Click *OPEN* to view your sequence.
8. To view a chromatogram of your sequence, you will need to download the sequence data, or trace, file as well as viewer software.
  - a. Follow the link to *Trace File*, and save the .abi file to your desktop.
  - b. Follow the link to *DNA Sequencing Core* (University of Michigan), and download a chromatogram viewer. (Note that some of the programs are specific to Mac or PC platforms. *FinchTV* is a nice viewer that works on both platforms.)
9. Click on the trace file to launch your sequence in the chromatogram viewer. If this doesn't work, first open the viewer, then go to *File open*, and navigate to select your .abi trace file.
10. Inspect your chromatogram. Remember that the primers amplify a 440-nucleotide sequence, so it is physically impossible to generate a sequence (read) longer than this.
11. Each peak is the fluorescence measured at that nucleotide position. Whenever possible the software "calls" each peak as an A,T,C, or G. What does the amplitude (height) of peak represent? What do you notice about the amplitude of the peaks at the beginning of the sequence?
12. Every sequence will begin with nucleotides (A,T,C,G) interspersed with "Ns". In "clean" sequences, where experimental conditions were near optimal, the initial Ns will end within the first 25 nucleotides. The remaining sequence will have very few, if any, internal Ns. Then, at the end of the "read," the sequence will abruptly change over to Ns. Large numbers of Ns scattered throughout the sequence indicate poor quality sequence. In what part(s) of the sequence file are the most "Ns" found? What do you notice about the trace at N positions?



13. Load two different trace files into different viewers, and attempt to match (align) the DNA sequences. What do you notice? What trick did you use to align them?
14. In several percent of cases clean sequence ends abruptly about midway through the read, then gives way to poor quality sequence with many Ns. Check to see if this describes any of your classmates' sequences, and, if so, manually align this sequence with a clean sequence. What feature coincides with the abrupt transition from clean-to-poor quality sequence?

### III. Assess the Extent of Modern mt Variation

1. Determine the average number of nucleotide differences between modern humans from various population groups.
2. Click on *MANAGE GROUPS* at the top of the *BioServers* page.
3. Select *Modern Human mtDNA* from the pull-down menu of *Sequence sources* in the upper right hand corner.
4. Click in the check boxes to select *African, Asian, and European*, and click *OK* to move these groups onto the workspace.
5. Uncheck all but two individuals from among *clean* sequences from your class, as well as the reference *African, Asian, and European* sequences. (Be sure you end up with only two sequences checked.)
6. Click on *COMPARE (Align: ClustalW)*, and the sequences are sent to a server at Cold Spring Harbor Laboratory that will perform a multiple sequence alignment. The results are returned in a new window.
7. In evaluating sequence remember a grey-shaded "N" indicates a sequence error, a position in one or both sequences where a nucleotide could not be determined. A yellow-shaded box indicates a nucleotide position at which the two sequences differ.
8. The alignment typically begins with yellow-shaded dashes (-), indicating an initial stretch of nucleotides that is not present in one of the samples. (This occurs because the sequence "read" usually begins a different points in different samples.) Sequence errors (grey) and nucleotide mismatches (yellow) are also frequent at the beginning of the sequence.
9. Scan into the sequence beyond all sequence errors and mismatches until you encounter a relatively long, unbroken string of perfect matches between the two sequences. This typically occurs at position 25-40.
10. Once you are into clean sequence, count the number of yellow-shaded nucleotide differences and deletions (-). These are single nucleotide polymorphisms (SNPs). A string of internal dashes (----) likely arouse from a single event, so score it as a single mutation.
11. Do not count any Ns. Sequences with more than one or two internal

Ns or with *multiple* nucleotide differences on *every* line are not reliable. If you detect this, select other sequences to work with.

12. Stop counting when you again encounter frequent Ns and/or sequence mismatches. This should occur around position 375-400.
13. Record the names of your samples and the number of SNPs between them. Also record the number of nucleotides of clean sequence you counted (typically 325-400).
14. Use the pull-down menus and check boxes to select a different pair of individuals, and repeat Steps 6-13. (Use only the reference data if you suspect your class sequences are unreliable.)
15. Continue until you have made at least five comparisons that sample a variety of populations, and recorded your results.
16. Calculate the percent difference for each pair:
 
$$\frac{\text{\# nucleotide differences}}{\text{total \# nucleotides counted}}$$
17. Share class data to come up with the largest possible number of unique pairs. Exclude any "outliers" with suspiciously large numbers of SNPs – these comparisons likely included unreliable sequences with too many sequence errors.
18. Determine the average number of mutations, as well as the average percent difference, between any two pairs.
19. What does this number represent?
20. Mutations accumulate one at a time. If one assumes a constant mutation rate, than the number of mutations can be equated to a time. However, this "mutational clock" needs to be set by some external event. In this case, we can use anthropological data that suggests modern human arose about 140,000 years ago. If this is so, what is the approximate mutation rate for the mt control region? Why is your calculation only approximate?

#### IV. Assess Ancient Humans: "Otzi the Iceman"

mtDNA sequences from ancient humans are quite rare, but a number of several key ones aid in the study of human evolution. One sequence is from "Otzi the Iceman," a 5,300 year-old human discovered frozen in the Otztal Alps between Austria and Italy. Although Otzi had the tools and weapons of a hunter, he carried domesticated wheat. So he belonged to a agricultural society.

1. Click on *MANAGE GROUPS* at the top of the *BioServers* page.
2. Select *Prehistoric Human mtDNA* from the pull-down menu of *Sequence* sources in the upper right hand corner.
3. Click in the check box to select *Otzi the Iceman mtDNA*, and click *OK* to move this sequence onto the workspace.

4. Use a pull-down menu and check box to select a modern human from your class, Africa, Asia, or Europe to compare with Otzi.
5. Repeat Steps 6-16 in Part III above to compare the number of differences and percent difference between Otzi and each of at least five modern humans.
6. Pool class data to determine the average number of differences and average percent difference between Otzi and modern humans.
7. What does this tell you about Otzi and the DNA mutational clock?

## V. Assess Ancient Humans: Neandertal

Neandertal was an archaic member of the genus *Homo*, which lived in what is now Europe – beginning about 300,000 years ago and becoming extinct about 30,000 years ago. Clearly, during part of its span on earth, Neandertal shared its European habitat with modern humans (*Homo sapiens*). According to the multiregional model, modern humans developed concurrently from different archaic populations in Africa, Europe, and Asia – and Neandertal was the direct ancestor of modern Europeans. According to the displacement model ("Out of Africa"), the ancestors of modern humans left Africa only 50,000-70,000 years ago – and so share with Neandertal an ancient ancestor in Africa. In 1997, an international research team headed by Svante Paabo settled the debate about the relationship of Neandertal to modern humans. They extracted DNA from the humerus of the original Neandertal specimen, amplified the sample by PCR, and cloned the resulting products in *E. coli*. The cloned fragments were then used to reconstruct a 379-bp stretch of the mt control region. You can recreate their analysis:

1. Click on *MANAGE GROUPS* at the top of the *BioServers* page.
2. Select *Prehistoric Human mtDNA* from the pull-down menu of *Sequence sources* in the upper right hand corner.
3. Click in the check boxes to select several samples of *Neandertal mtDNA*, and click *OK* to move these sequences onto the workspace.
4. Use a pull-down menu and check box to select a modern human from your class, Africa, Asia, or Europe to compare with one of the Neandertal samples.
5. Repeat Steps 6-16 in Part III above to compare the number of differences and percent difference between at least five Neandertal-modern human pairs.
6. Also compare each possible pair of Neandertals to assess the extent of variation in Neandertals across their habitat range.
7. Pool class data to determine the average number of differences and average percent difference between Neandertal and modern humans – and between Neandertals.
8. What does this tell you about the relationship between Neandertals and modern humans?

9. Using 140,000 years for the divergence of modern humans, approximately how long ago did humans and Neandertal share a common mt ancestor?
10. What does the small number of available samples suggest about Neandertal diversity?

## VI. Discover What DNA Says About Human Evolution

In this exercise, you will conduct three-way sequence alignments to compare African and non-African individuals.

1. Use the pull-down menus and check boxes to select two African samples and one European or Asian sample. Record the names of the sample populations in the chart below
2. Click on *COMPARE (Align: ClustalW)* to perform a multiple sequence alignment.
3. At any yellow-shaded mismatch, one sample will be different from the other two.
4. Evaluate each mismatch and for each position determine whether one of the Africans is different or the European/Asian is different.
5. Record tick-marks for each mismatch, as illustrated below.

African 1	African 2	European or Asian	African Different	European or Asian Different

6. What do your results show? Is this what you expected?
7. Pool class results to confirm or refute your own results.
8. What does your analysis say about African mtDNA diversity? What population factors could account for this?
9. What does your analysis say about European and Asian mtDNA diversity? What population factors could account for this?

## INFORMATION FOR INSTRUCTOR

### CONCEPTS AND METHODS

---

This laboratory can help students understand several important concepts of modern biology:

- How to collect and analyze genetic information in populations.
- The use of single nucleotide polymorphisms (SNPs) in the study of human evolution.
- Identity by descent from a common ancestor.
- The movement between *in vitro* experimentation and *in silico* computation.

The laboratory uses several methods for modern biological research:

- DNA extraction and purification.
- Polymerase chain reaction (PCR).
- Gel electrophoresis.
- DNA sequencing.
- Bioinformatics.

### LAB SAFETY

---

The National Association of Biology Teachers recognizes the importance of laboratory activities using human body samples and has developed safety guidelines to minimize the risk of transmitting serious disease. ("The Use of Human Body Fluids and Tissue Products in Biology," *News & Views*, June 1996.) These are summarized below:

- Only collect samples from students under your direct supervision.
- Do not use samples brought from home or obtained from an unknown source.
- Do not collect samples from students who are obviously ill or are known to have a serious communicable disease.
- Have students wear proper safety apparel: latex or plastic gloves, safety glasses or goggles, and lab coat or apron.
- Supernatants and samples may be disposed of in public sewers (down lab drains).
- Have students wash their hands at the end of the lab period.
- Do not store samples in a refrigerator or freezer used for food.

The risk of spreading an infectious agent by this lab method is much less likely than from natural atomizing processes, such as coughing or sneezing. Several elements further minimize any risk of spreading an infectious agent that might be present in mouthwash samples:

- Each experimenter works only with his or her sample.
- The sample is sterilized during a 10-minute boiling step.
- There is no culturing of the samples that might allow growth of pathogens.
- Samples and plasticware are discarded after the experiment.



## INFORMED CONSENT AND DISCLOSURE

Student participation in this experiment raises real-life questions about the use of personal genetic data: What is my DNA sample being used for? Does my DNA type tell me anything about my life or health? Can my data be linked personally to me?

There is consensus that a human DNA sample should be obtained only with the willing consent of a donor, who understands the purpose for which it is being collected. Thus, this experiment should be explained ahead of time and students given the option to refrain from participating. (Some teachers may wish to have parents sign a consent form, such as those filled out for a field trip.) There is also consensus that a DNA sample be used only for the express purpose for which it is collected. Thus, student DNA samples should be thrown away after completing the experiment.

The PV92 polymorphism was specifically selected for this experiment because it is phenotypically neutral — it has no known relationship to any trait, disease state, or sex determination.

PV92 alleles are inherited in a Mendelian fashion and can give indications about family relationships. To avoid the possibility of suggesting inconsistent inheritance, it is best not to generate genotypes from parent-child pairs. In any event, this two-allele system would be less likely to turn up an inconsistency than the ABO blood groups. Furthermore, the chance that student samples can be mixed up when isolating DNA, setting up PCR reactions, and loading electrophoresis gels provides no certainty to any of the genotypes obtained in the experiment. (A forensic laboratory would use approved methods for maintaining “chain of custody” of samples and for tracking samples.)

## INSTRUCTOR PLANNING, PREPARATION, AND LAB FINE POINTS

The following table will help you to plan and integrate the four parts of the experiment.

Part	Day	Time	Activity
I. Isolate DNA	1	60 min.	<b>Pre-lab:</b> Prepare and aliquot saline solution Prepare and aliquot 10% Chelex® Aliquot proteinase K (alternate) Make centrifuge adapters Set up student stations <b>Lab:</b> Isolate student DNA
II. Amplify DNA by PCR	1	15 min. 15 min. 60-150 min.	<b>Pre-lab:</b> Aliquot mt primer/loading dye mix <b>Lab:</b> Set up PCR reactions <b>Post-lab:</b> Amplify DNA in thermal cycler
III. Analyze PCR Products by Gel Electrophoresis	2 3	15 min. 30 min. 15 min. 20+ min. 20+ min. 30-45 min. to overnight	<b>Pre-lab:</b> Dilute TBE electrophoresis buffer <b>Lab:</b> Prepare agarose gel solution and cast gels. Load DNA samples into gel Electrophorese samples <b>Post-lab:</b> Stain gels De-stain gels (for <i>Carolina</i> BLU∞)
IV. Label DNA by Dideoxy Chain Termination	4	30 min. 15 min. 150.	<b>Pre-lab:</b> Prepare and aliquot cycle sequencing master mix <b>Lab:</b> Set up cycle sequencing reactions <b>Post-lab:</b> Dye label DNA in thermal cycler!
V. Purify DNA and Sequence	5 6	15 min. 120 min. overnight 30 min 240-300 min	<b>Pre-lab:</b> Prepare loading dye <b>Lab:</b> Precipitate DNA <b>Post-lab:</b> Air dry DNA pellets overnight Resuspend DNA pellets Automated Sequencing
Results and Discussion	4	30-60 min.	Analyze electropherograms and perform multiple sequence alignments

## I. ISOLATE DNA FROM CHEEK CELLS

Saline mouthwash is the most reproducible of the simple methods to obtain human DNA for PCR. The mouthwash gently loosens a large number of single cells and small cluster of cheek cells. This maximizes the surface area of cells, allowing for virtually complete lysis during boiling. Cheek brushes and swabs generally yield larger clumps of cells, which are less effectively lysed by boiling.

With careful lab management, up to 90% of students should be able to “score” their *Alu* genotypes using the mouthwash method. Be especially watchful after the initial centrifugation step. Most students will have compact pellets that stay attached to the tube when the supernatant is poured off. However, about 10% of students will have diffuse or slimy masses that do not pellet well. Centrifuge these samples again, then carefully pipet out as much supernatant as possible. Surprisingly, food particles rinsed out with the mouthwash have little effect on PCR amplification. Still, it is best to avoid eating before the experiment, because food particles – especially from fruits – may block the pipet tip and make pipetting difficult.

It is worth a diversion to allow students to view their own squamous epithelial cells under a compound microscope. Add several  $\mu\text{L}$  of suspension remaining after Step I. 8. to a microscope slide, add a drop of 1% methylene blue (or other stain), and add a cover slip.

DNA is liberated from cheek cells by boiling in 10% Chelex<sup>®</sup>, which binds contaminating metal ions that are the major inhibitors of PCR. The boiling step is most easily accomplished using the same thermal cycler used for PCR. To do this, provide each student with 100  $\mu\text{L}$  of 10% Chelex<sup>®</sup> suspension in a PCR tube that is compatible with the thermal cycler you will be using: either 0.2 mL or 0.5 mL. It is not necessary to use a “thin-walled” tube. Alternatively, use 1.5 mL tubes in a heat block or a boiling water bath. *Watch out for lids opening as the tubes heat.* (Make a simple water bath by maintaining a beaker of water at a low boil on a hot plate. Place 1.5 mL tubes in a floating rack or in holes punched in a double layer of aluminum foil over the top. If using aluminum foil, insure that tubes are immersed, and add hot water as necessary to maintain water level.)

### **Pre-lab Preparation**

Prepare saline by dissolving 0.9 g NaCl in 100 mL distilled or deionized water. For each student, aliquot 10 mL into a 15-mL polypropylene tube.

Prepare 10% Chelex<sup>®</sup> by adding 15 mL distilled or deionized water to 1.5 g of Chelex<sup>®</sup>. For each student, aliquot 100  $\mu\text{L}$  of 10% Chelex<sup>®</sup> into either a 0.2 mL or 0.5 mL tube (whichever format is accommodated by your thermal cycler). Alternatively, use a 1.5 mL microcentrifuge tube if you are planning to use a heat block or water bath instead of a thermal cycler. *The Chelex<sup>®</sup> resin quickly settles, so be sure to shake the stock tube to re-suspend the Chelex<sup>®</sup> each time before pipetting a student aliquot.*

Remove caps from 1.5-mL tubes to use as adapters in which to centrifuge the 0.5-mL PCR tubes used for Chelex<sup>®</sup> extraction. Two adapters are needed to spin 0.2-mL PCR tubes—a capless 0.5-mL PCR tube is nested within a capless 1.5-mL tube.

### **Pre-lab Set Up for DNA Isolation from Cheek Cells** (per student station)

- Saline solution (0.9% NaCl) tubes, 10 mL (in 15 mL tube)
- 10% Chelex<sup>®</sup>, 100  $\mu\text{L}$  (in 0.2 or 0.5 mL tube, depending on thermal cycler)
- 2- 1.5 mL microcentrifuge tubes
- Permanent marker
- Micropipets and tips (10-1,000  $\mu\text{L}$ )
- Microcentrifuge tube rack
- Container with cracked or crushed ice
- Paper cup

### Shared Items

Microcentrifuge  
 Microcentrifuge adapters for 0.2 mL or 0.5 mL PCR tubes  
 Thermal cycler  
 Vortexer (optional)

## I. (ALTERNATE) ISOLATE DNA FROM HAIR SHEATHS

Hair roots provide the simplest source of DNA for PCR amplification; no special equipment is required for extraction. Hairs also are an extremely safe source of cells. Risk of spreading an infectious agent is minimized by "dry" collection, which does not involve any body fluid or generate any supernatant. This method also stresses the power of PCR in forensic cases – even one growing hair root provides enough DNA for excellent amplification.

HOWEVER, forensic biologists generally rate hair as a poor source of DNA for analysis, for the same reason that it can prove difficult in the classroom. Most plucked or shed hairs are broken off from the root, which is the source of cells for DNA extraction.

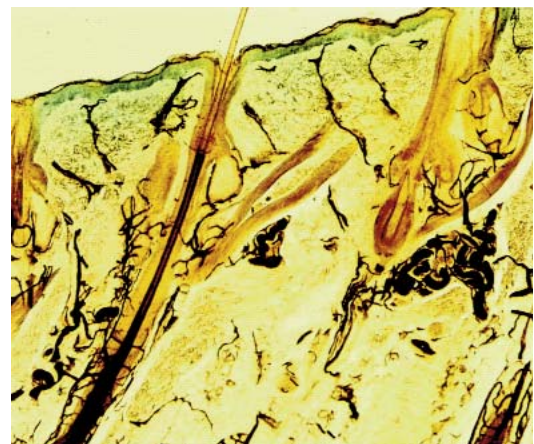
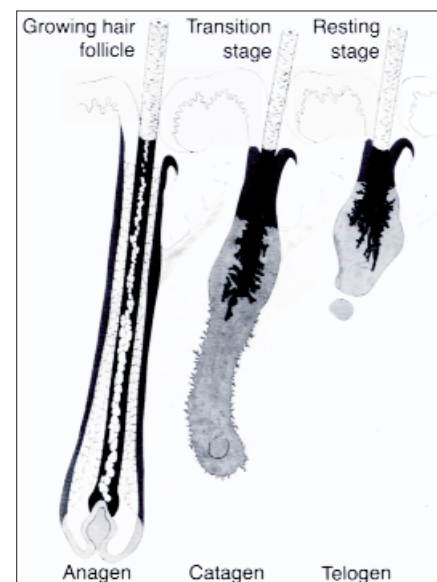
The success of this method is entirely dependent upon finding large roots from growing hairs. This can be tricky and time consuming – if often hilarious. With vigilance, up to 80% of students may find hairs with good roots from which to isolate DNA. However, it is more likely that only about 60-70% of students ultimately will be able to score their *Alu* genotypes using this method.

A hair is anchored in the skin by a follicle, or "root," whose growing cells produce the hair shaft. Hair goes through a growth cycle with alternating periods of growth and quiescence during which the follicle increases and decreases in size. During the growth phase, the follicle extends up the hair shaft in a structure called the sheath. The sheath is a rich source of cells. The sheath membrane is easily digested by treatment with proteinase K digests, releasing squamous cells singly or in small clusters. A high percentage of these cells are lysed by boiling and release DNA.

The sheath decreases in size as the hair follicle enters a resting stage (see drawing and micrograph of growing and resting follicles). The withered bulb of a resting follicle is, in fact, what most people would consider a "root." Resting follicles usually yield little DNA for analysis. First, there are fewer cells. Second, proteinase K treatment does not effectively digest the shriveled root mass, and only cells at the edge are lysed by boiling.

Successful amplification of the PV92 locus – which is available in only two copies per cell – is closely correlated to presence of a sheath on the hair shaft. One or two hairs with long sheaths will provide plenty of DNA for PCR amplification. Three or four good sized roots will usually work, especially if they have at least small sheaths.

A good sheath is unmistakable. Especially contrasted on a dark hair, it glistens when held up to the light and extends several mm up the hair shaft. Make sure to show off the first several



good sheaths that turn up, so other students will know what to look for. Because of the hair growth cycle, most people find sheaths only on some hairs. Students whose hair grows slowly may have difficulty finding sheaths, and thin or brittle hair is likely to break off before the root. If students are having difficulty finding sheaths on hairs pulled from their scalps, have them try hairs from the eyebrow or arm.

Sheaths are the most underrated source of squamous cells for microscopic examination. Give them a try! Simply place a sheath on a microscope slide and add a drop of proteinase K (100 mg/mL). Let stand for several minutes, to allow the proteinase K to digest the sheath membrane. Then add a drop of methylene blue or other cell stain, add a cover slip, and gently press to disrupt the sheath membrane. Observe under medium power and at several time points, to see the effect of enzyme digestion. If you gently press the cover slip while the slide is on the microscope stage, you should be able to observe squamous cells squirting of tears in the sheath membrane.

### **Prelab Preparation**

For each student, aliquot 100  $\mu$ L of 100 mg/mL proteinase K into either a 0.2 mL or 0.5 mL tube (whichever format is accommodated by your thermal cycler). Alternatively, use a 1.5 mL microcentrifuge tube if you are planning to use a heat block or water bath instead of a thermal cycler.

#### **Pre-lab Set Up for DNA Isolation from Hair Sheaths** (per student station)

- 100 mg/mL proteinase K, 100  $\mu$ L (in 0.2 or 0.5 mL PCR tube)
- Permanent marker
- Scalpel or razor blade
- Forceps or tweezers

#### **Shared Items**

- Thermal cycler (or water bath or heat block)
- Container with cracked or crushed ice
- Vortexer (optional)

## **II. AMPLIFY DNA BY PCR**

The primer/loading dye mix incorporates the appropriate primer pair (0.26 picomoles/ $\mu$ L of each primer), 13.8% sucrose, and 0.0081% cresol red. The inclusion of the loading dye components, sucrose and cresol red, allows the amplified product to be directly loaded into an agarose gel for electrophoresis. Each Ready-To-Go™ PCR Bead contains reagents so that when brought to a final volume of 25  $\mu$ L, the reaction contains 2.5 units of *Taq* DNA polymerase, 10 mM Tris-HCl (pH 9.0), 50 mM KCl, 1.5 mM MgCl<sub>2</sub>, and 200  $\mu$ M of each dNTP.

The lyophilized *Taq* DNA polymerase in the bead becomes active immediately upon addition of the primer/loading dye mix and template DNA. In the absence of thermal cycling, “nonspecific priming” at room temperature allows the polymerase to begin generating erroneous products, which can show up as extra bands in gel analysis. *Therefore, work quickly. Be sure the thermal cycler is set and have all experimenters set up their PCR reactions as a coordinated effort. Add primer/loading dye mix to all reaction tubes, then add each student template, and begin thermal cycling as quickly as possible. Hold reactions on ice until all student samples are ready to load into the thermal cycler.*

PCR amplification from crude cell extracts is biochemically demanding, and requires the precision of automated thermal cycling. However, amplification of the mt control region is simplified by the large amount of available template. Therefore, the recommended amplification times and temperatures will work adequately for most common thermal cyclers, which ramp between temperatures within a single heating/cooling block. **IMPORTANT:** A different cycling profile is required for Robocycler or other brands of

thermal cyclers that physically move PCR reaction tubes between multiple temperature blocks. Because there is no ramping time between temperatures, these machines require the longer cycling times listed below:

Denaturing step:	94°C	30 seconds
Annealing step:	58°C	30 seconds
Extending step:	72°C	30 seconds

### **Pre-lab Preparation**

Aliquot 25  $\mu$ L of mt primer/loading dye mix per student. The primer/loading dye mix may collect in the tube cap during shipping; pool the reagent by spinning the tube briefly in a microcentrifuge or by sharply tapping the tube bottom on the lab bench.

#### **Pre-lab Set Up for DNA Amplification** (per student station)

Cheek cell DNA. 2.5  $\mu$ L (from Part I)  
 Mt primer/loading dye mix, 25  $\mu$ L  
 Ready-To-Go™ PCR beads (in 0.2 mL or 0.5 mL PCR tube)  
 Permanent marker  
 Micropipet and tips (1-100  $\mu$ L)  
 Microcentrifuge tube rack  
 Container with cracked or crushed ice

#### **Shared Items**

Mineral oil, 5 mL (depending on thermal cycler)  
 Thermal cycler

### **III. ANALYZE AMPLIFIED DNA BY GEL ELECTROPHORESIS**

**CAUTION:** IF you are planning to sequence student samples, on your own or through the Dolan DNA Learning Center's Sequencing Service, be sure that students electrophoresis only 15  $\mu$ L of the amplified product. The remaining 10  $\mu$ L must be retained for use in cycle sequencing.

The cresol red and sucrose in the primer mix function as loading dye, so that amplified samples can be loaded directly into an agarose gel. This is a nice time saver. However, since it has relatively little sugar and cresol red, this loading dye is more difficult to use than typical loading dyes. So, encourage students to load carefully.

Plasmid pBR322 digested with the restriction endonuclease *Bst*NI is an inexpensive marker and produces fragments that are useful as size markers in this experiment. The size of the DNA fragments in the marker are 1,857 bp, 1,058 bp, 929 bp, 383 bp, and 121 bp. Use 20  $\mu$ L of a 0.075  $\mu$ g/ $\mu$ L stock solution of this DNA ladder per gel. Other markers or a 100-bp ladder may be substituted.

View and photograph gels as soon as possible after appropriate staining/destaining. Over time, the small-sized PCR products will diffuse through the gel and lose sharpness. Refrigeration will slow diffusion somewhat, but for best results view and photograph gels as soon as staining/destaining is complete.

### **Pre-lab Preparation**

Prepare a 1X concentration of TBE by adding 75 mL of 20X concentrated stock into 1,425 mL of deionized or distilled water. Mix thoroughly.





Prepare a 1.5% agarose solution by adding 1.5 g of agarose to 100 mL of 1X TBE in a 500 mL flask or beaker. Heat the flask or beaker in a boiling water bath (approximately 15 minutes) or in a microwave oven (approximately 4 minutes) until the agarose is completely dissolved. You should no longer see agarose particles floating in solution when the beaker is swirled. Allow the agarose to cool to approximately 60°C, and hold at this temperature in a hot water bath. Cover beaker or flask with aluminum foil, and skim any polymerized “skin” off the top of the solution before pouring.

#### Pre-lab Set Up for Gel Analysis (per student station)

Amplified human DNA PCR products from Part III (store on ice)  
Container with cracked or crushed ice

#### Shared Items

pBR322/*Bst*NI markers, 20 µL per row of gel (thaw and store on ice)  
1.5% agarose in 1X TBE (hold at 60°C), 50 mL per gel  
1X TBE buffer, 300 mL per gel  
Ethidium bromide (1 µg/mL), 250 mL  
**or**  
*Carolina*BLU™ Gel & Buffer Stain, 7 mL  
*Carolina*BLU™ Final Stain, 250 mL  
Micropipet and tips (1-100 µL)  
Microcentrifuge tube rack  
Gel electrophoresis chambers  
Power supplies  
Water bath for agarose solution (60°C)  
Latex gloves  
Staining tray  
Transilluminator with digital or instant camera (optional)

## DNA SEQUENCING

Single nucleotide polymorphisms in the mt control region can only be detected by sequencing the amplicon (PCR product) obtained in Part III. This is most easily accomplished by submitting student amplicons to the *Sequencing Service* provided by the Dolan DNA Learning Center (DNALC).

### **Sequencing by the DNALC**

Sequencing of the mt control region is available free of charge to high school and college classes. *However, the DNALC can only sequence amplicons obtained using the primer sequences specified in this experiment.*

1. To insure the fastest turnaround on you samples, notify the DNALC at least one month prior to submitting.
2. Prior to the PCR experiment: Complete one online submission form for each sample set, and submit electronically at least one month prior to your expected shipment date.  
<http://www.geneticorigins.org/mito/mitoframeset.htm>
3. After the PCR experiment: Print out and complete one copy of the submission checklist at the web address above, or duplicate the form below. Carefully follow the numbering, packing, and shipping instructions.
4. You will be notified by email when sequencing is complete and student data are posted to the *BioServers* Internet site. Processing typically takes 2-3 weeks. However, demand is very high in November, March, and at the end of the school year. Submissions received near the end of the school year may not be posted until the summer. Avoid high demand periods if at all possible.
5. Contact the *Sequencing Service* at [dnalcseq@cshl.edu](mailto:dnalcseq@cshl.edu) for further information.

*Sample Submission Checklist*

- Provide a photograph a 2% agarose gel confirming amplification of student mt control region sequences from Part III of the laboratory. *Sequencing is expensive, so the DNALC can only sequence samples for which you provide evidence of amplification.*
- Provide remaining 10-15  $\mu\text{L}$  aliquots of student PCR products in clean, tightly capped 0.2 or 0.5 ml PCR test tubes.
- Samples *must not* contain bromophenol blue loading dye. Samples with bromophenol blue will not be processed.
- Use tape to *clearly* label tubes with sequential numbers beginning with "1." *Do not skip or repeat any numbers. Do not include any student names with samples.*
- Pack gel photo and student samples in zip-closure plastic bag(s) in a small Styrofoam box containing a cold pack.
- Ship by overnight mail to address below for Monday-Friday delivery. There is no one to receive packages on the weekends, so do not send packages on Fridays.

**DNA Sequencing Service**

DNA Learning Center  
Cold Spring Harbor Laboratory  
1 Bungtown Road  
Cold Spring Harbor, NY 11724  
Phone (516) 367-5170

**Sequencing On Your Own**

Alternatively, you can have student amplicons sequenced at your own institution, or at a local university or company. You may have students perform the cycle sequencing reaction and DNA purification in Parts IV and V that follow, but check with a sequencing technician that these methods are compatible with the protocols in use at their facility.

For speed, reproducibility, and ease of tracking samples, the following procedure is optimized for a 96-well microtiter plate. This is the most common format used for DNA sequencing.

**IV. Cycle Sequence mt Amplicon by PCR****Pre-Lab Preparation**

Prepare a cycle sequencing master mix:

**Per Sample**

1  $\mu\text{L}$  Big Dye Terminator  
1  $\mu\text{L}$  5' mt DNA Primer (6pmol/ $\mu\text{l}$ )  
1.5  $\mu\text{L}$  5 Sequencing Buffer  
4.5  $\mu\text{L}$  dH<sub>2</sub>O  
8.0  $\mu\text{L}$

**Per 96-Well Plate**

105  $\mu\text{L}$  Big Dye Terminator  
105  $\mu\text{L}$  5' mt DNA Primer (6pmol/ $\mu\text{l}$ )  
157.5  $\mu\text{L}$  5x Sequencing Buffer  
472.5  $\mu\text{L}$  dH<sub>2</sub>O  
840.0  $\mu\text{L}$

Mix by vortexing or pipetting in and out.



Reagents (per reaction)	Supplies and Equipment
Cycle sequencing master mix, 8 $\mu$ L mt control region amplicon, 2 $\mu$ L	96-well microtiter plate 1-20 $\mu$ L micropipet and tips Sealing foil Thermal cycler

Seal carefully; otherwise some samples will evaporate during cycling.

1. Add 8  $\mu$ L of master mix to each well of the microtiter plate that will receive a reaction.
2. Add 2  $\mu$ L of mt amplicon (product) to each reaction.
3. Cover microtiter plate with sealing foil, and carefully seal. Press on each well of the plate with your thumb, and also run your finger around the edge of the foil to create a tight seal.
4. Store reactions on ice or in the freezer until ready to begin thermal cycling.
5. Place the microtiter plate in a thermal cycler that has been programmed according to the following profile. The program may be linked to a 4°C cycle to hold samples after completing the cycle profile.

Hold step            96°C    2 minutes

Then 34 cycles of:

Denaturing step    96°C    10 seconds

Annealing step    50°C    5 seconds

Extending step    60°C    4 minutes

## V. Precipitate Dye-Labeled DNA

### *Pre-lab Prep*

Prepare a fresh solution of 30:1 ethanol/sodium acetate (EtOH/NaAc):

3.3 mL of ethanol (EtOH)

110.0  $\mu$ L of sodium acetate (NaAc) (3M, pH 4.8)

Mix by vortexing or pipetting in and out.

Reagent	Supplies and Equipment
Cycle sequencing reactions (in 96-well microtiter plate)	20-100 $\mu$ L micropipet and tips Sealing foil
ethanol/sodium acetate (EtOH/NaAc), 30:1, 30 $\mu$ L	Clinical centrifuge (with spin trays) Freezer
Ethanol, 70%, 100 $\mu$ L	Thermal cycler (optional) Vortexer

Sample should turn a cloudy, yellowish white.

1. Add 30  $\mu$ L of 30:1 EtOH/NaAc to each well of the microtiter plate.
2. Mix by pipetting in and out several times. Do not vortex.
3. Cover microtiter plate with sealing foil, and carefully seal.
4. Incubate in freezer at  $-20^{\circ}\text{C}$  for 30 minutes.

Ethanol precipitates the extended PCR products, but leaves the primer and unincorporated nucleotides in solution. Centrifugation pellets the precipitated sequencing products on the bottom of the microtiter well.

To prevent cross-contamination between wells, move plate to a clean area of toweling between each tap.

DNA pellets are not soluble in ethanol and will not resuspend during washing.

5. Spin for 30 minutes in clinical centrifuge equipped with spin trays. Balance with 40  $\mu$ L balance plate.
6. Remove sealing foil from microtiter plate. Invert plate over sink to drain off most of the ethanol solution. Then gently tap plate on the surface of a clean paper towel to wick off as much liquid as possible.
7. Add 50  $\mu$ L of 70% ethanol into each well to wash the DNA pellet.
8. Carefully seal microtiter plate with clean foil.
9. Centrifuge microtiter plate for 15 minutes along with a 50  $\mu$ L balance plate.
10. Remove sealing foil from microtiter plate. Invert plate over sink to drain off most of the ethanol solution. Then gently tap plate on the surface of a clean paper towel to wick off as much liquid as possible.
11. Repeat steps 7–10 to wash, centrifuge and drain the DNA pellets a second time.
12. If possible, air dry the DNA pellets overnight. (Alternatively, incubate open plate in thermal cycler for 10 minutes at 37°C. Do not close thermal cycler lid!)
13. After the microtiter plate is thoroughly dry, add 20  $\mu$ L of autoclaved, distilled water to each well to resuspend the DNA. Every well on the microtiter plate – including empty ones without DNA samples – must be filled with 20  $\mu$ L of water before being loaded onto the DNA sequencer.
14. Carefully seal microtiter plate with clean foil. Label plate with identifying name and date.
15. Vortex microtiter plate at half speed. Tap plate on bench top to bring DNA droplets to bottom of wells.
16. The microtiter plate is now ready to be loaded onto an automated sequencer (ABI 3700 DNA Analyzer). Freeze microtiter plate at  $-20^{\circ}\text{C}$  until ready to sequence DNA samples.

## CarolinaBLU™ STAINING

---

### POST-STAINING

1. Cover the electrophoresed gel with the *CarolinaBLU™* Final Stain and let sit for 20–30 minutes. Agitate gently (optional).
2. After staining, pour the stain back into the bottle for future use. (The stain can be used 6–8 times.)
3. Cover the gel with deionized or distilled water to destain. Chloride ions in tap water can partially remove the stain from the DNA bands and will cause the staining to fade.
4. Change the water 3 or 4 times over the course of 30–40 minutes. Agitate the gel occasionally.
5. Bands that are not immediately present will become more apparent with time and will reach their maximum visibility if the gel is left to destain overnight in just enough water to cover the gel. Gels left overnight in a large volume of water may destain too much.

### PRE-STAINING

*CarolinaBLU™* can also be used to stain the DNA while it is being electrophoresed. Pre-staining will allow students to visualize their results prior to the end of the gel run. However, post-staining is still required for optimum viewing.

To pre-stain the gel during electrophoresis, add *CarolinaBLU™* Gel and Buffer Stain in the amounts indicated in the table below. Note that the amount of stain added is dependent upon the voltage used for electrophoresis. *Do not use more stain than recommended. This may precipitate the DNA in the wells and create artifact bands.*

Gels containing *CarolinaBLU™* may be prepared one day ahead of the lab day, if necessary. However, gels stored longer tend to fade and lose their ability to stain DNA bands during electrophoresis.

Use the table below to add the appropriate volume of *CarolinaBLU™* stain to the agarose gel:

Voltage	Agarose Volume	Stain Volume
<50 Volts	30 mL	40 µL (1 drop)
	200 mL	240 µL (6 drops)
	400 mL	520 µL (13 drops)
>50 Volts	50 mL	80 µL (2 drops)
	300 mL	480 µL (12 drops)
	400 mL	640 µL (16 drops)

Use the table below to add the appropriate volume of *CarolinaBLU™* stain to 1× TBE buffer:

Voltage	Agarose Volume	Stain Volume
<50 Volts	500 mL	480 µL (12 drops)
	3000 mL	3 mL (72 drops)
>50 Volts	500 mL	960 µL (24 drops)
	2600 mL	5 mL (125 drops)

## BIOINFORMATICS

---

Have students do the bioinformatics exercises before starting the experiment – or analyzing results. This should improve conceptual and practical understanding.

The onscreen Bio-i Guide can be played from the included CD-ROM or from the Internet site <http://bioinformatics.dnalc.org/pv92>. The default version (640 x 480 pixels) allows one to follow along with an open browser window. The full screen version (1024 x 768 pixels) is best for demonstrations.

## ANSWERS TO BIOINFORMATICS QUESTIONS

---

- I.1.h. Matches of different lengths are coded by color. What do you notice? **Typically, all or most of the significant alignments will have complete matches to the forward and reverse primers. Partial sequences will match only the forward or reverse primer.**
- I.2.b. Find a match to both primers. Do your *E-values* seem high or low? Why do you think this is so? **The lowest *E-value* obtained for a match to both primers is typically in the range of 0.02-0.03, with a larger number of matches in the 0.3-0.4 range. This might seem high for a probability, but in fact each of these values means that a match of this quality would be expected to occur by chance less than one time in this database! More precisely, a score of 0.33 would mean that a single match would be expected to occur by chance once in every 3 searches! *E-values* are based on the length of the search sequence, so the relatively short primers used in this experiment produce relatively high. Searches with longer primers or long DNA sequences return *E-values* with negative exponents (such as  $e^{-4}$ ), which look more like the probabilities one might expect. This search turns up many matches with the same *E-value*, because human mitochondrial sequences are the most common sequences in DNA databases, with tens of thousands of accessions.**
- I.2.c. Note the descriptions of some of significant alignments that have *E-values* less than 1. Do they make sense? **Yes, all are matches to human (*Homo sapiens*) mitochondrial sequences -- identified as "complete genome," D loop, "control region," or "hypervariable region."**
- I.3.b. The lowest and highest nucleotide positions in the subject sequence indicate the borders of the amplified sequence. Subtracting one from the other gives the difference between the two coordinates. **16,409 - 15,970 = 439 nucleotides. Although these are the absolute coordinates for this sequences, accessions may vary several nucleotides from these numbers.**
- I.3.c. However, the actual length of the fragment includes both ends, so add 1 nucleotide to the result to determine the exact length of the PCR product amplified by the two primers. **440 nucleotides. 439 + 1 = 440 nucleotides.**
- II.1.e. Find the feature that includes the nucleotide positions that define the boundaries of the forward and reverse primers (from section 3.b. above). What is it called? How large is it? Does it contain any genes? Do some research to understand each of several names used for this region. **This region is most frequently called the D-loop. It is approximately 1,100 nucleotides in length, and contains no genes (noncoding).**
- D-loop (displacement loop) occurs at the origin of replication, when the mt chromosome denatures into. A newly-synthesized strand displaces one of the parental strands, forming a "bubble" or loop.***

**Control region** refers to the fact that this region contains the signals that control RNA and DNA synthesis. A single promoter on each DNA strand initiates transcription in each direction, and a single origin initiates replication of each strand.

**Hypervariable region** refers to the fact that this part of the mt chromosome accumulates point mutations at approximately 10 times the rate of nuclear DNA. This region is relatively tolerant of the high mutation rate, because binding sites for DNA and RNA polymerase are defined by only short nucleotide sequences.

- II.2.a. Look at the features list for your hit. List the genes that are transcribed from the forward strand and those that are transcribed from the reverse (complement) strand. (Remember that, in addition to proteins, ribosomal RNAs (rRNA) and messenger RNA (mRNA) are also transcribed from genes.)

**Forward Strand**

tRNA, phenylalanine  
 rRNA, 12S  
 tRNA, alanine  
 tRNA, valine  
 tRNA, 16S  
 tRNA, leucine (copy 1)  
 subunit 6)  
 gene, ND1 (NADH dehydrogenase subunit 1)  
 tRNA, isoleucine  
 tRNA, methionine  
 gene, ND2 (NADH dehydrogenase subunit 2)  
 tRNA, tryptophan  
 gene, COI (cytochrome c oxidase subunit I)  
 tRNA, aspartic acid  
 gene, COII (cytochrome c oxidase subunit II)  
 tRNA, lysine  
 gene, ATP8 (ATP synthase F0, subunit 8)  
 gene, ATP6 (ATP synthase F0, subunit 6)  
 gene, COII (cytochrome c oxidase subunit III)  
 tRNA, glycine  
 gene, ND3 (NADH dehydrogenase subunit 3)  
 tRNA, arginine  
 gene, ND4L (NADH dehydrogenase subunit 4L)  
 gene, ND4 (NADH dehydrogenase subunit 4)  
 tRNA, histidine  
 tRNA, serine (copy 2)  
 tRNA, leucine (copy 2)  
 gene, ND5 (NADH dehydrogenase subunit 5)  
 gene, CTBY (cytochrome b)  
 tRNA, threonine

**Reverse Strand (Complement)**

tRNA, glutamine  
 tRNA, asparagine  
 tRNA, cysteine  
 tRNA, tyrosine  
 tRNA, serine (copy 1)  
 gene, ND6 (NADH dehydrogenase  
 tRNA, glutamic acid  
 tRNA, proline

- II.2.b. On which strand are most genes found? **There are 29 genes on the forward strand, and eight genes on the reverse strand.**
- II.2.c. Judging from the genes you identified, the mt genome provides all needed components

for what biological function? **The mt genome encodes all components needed for protein translation: 20 tRNAs and the two ribosome subunits.**

- II.2.d. Does the mt genome encode all necessary components needed for oxidative phosphorylation (energy production)? **No. The mt chromosome encodes 13 proteins needed for oxidative phosphorylation. Additional proteins needed for oxidative phosphorylation are encoded by the nuclear genome and imported into the mitochondrion.**
- II.2.e. What is the spacing like between genes – is there intergenic DNA between two adjacent genes? **The mt chromosome has very little intergenic DNA. Many genes are immediately adjacent to one another. About a third of genes are separated by gaps of 1-70 nucleotides, and a few overlap one another by 1-7 nucleotides. Gaps and overlaps occur between adjacent genes on the same strand and on different strands.**
- II.2.f. Focus on one of the protein coding genes. Compare the nucleotide coordinates for the gene and the CDS (coding sequence). What do you notice? What does this tell you about the structure of mt genes and their origin? **Eukaryotic genes have several noncoding sequences: the 5' and 3' untranslated regions (UTRs) at either end and introns within. The coding sequence (CDS) is the set of codons that is translated into protein – beginning with a start codon, ending with a stop codon and omitting all introns. For each mt gene, the gene and coding sequences are identical, meaning that mt genes have no UTRs and no introns. These are characteristics of prokaryotic genes. The structure of mt chromosome, as well as its circular form, provide the DNA evidence that the mitochondrion evolved from free-living bacteria.**
- III.6. What sort of mutations generally cause mt diseases? What do many of these diseases have in common? **Most mt diseases are caused by point mutations. Many mt diseases affect tissues with high energy demands: nerves, skeletal muscles, and heart. Many are related to advancing age.**
- III.7. Research the mitochondrial theory of aging. **The mitochondrial theory of aging generally states that as mt genes accumulate mutations over time, the efficiency of energy production declines. Mutations in the respiratory chain proteins make them less efficient at transporting electrons. The increased “leak” of electrons leads to the production of more reactive oxygen species – which, in turn, produce more mutations in mt genes. This leads to age-related declines – especially in nerve, muscle, and heart function.**

## ANSWERS TO DISCUSSION QUESTIONS

---

- I.2. How would you interpret a lane in which you observe primer dimer, but no 440-bp band? **The presence of primer dimer confirms that the reaction contained all components necessary for amplification, but that there was insufficient template to amplify the 440-bp target sequence.**
- I.3. The mt control region mutates at approximately 10 times the rate of nuclear DNA. Propose a biological reason for the high mutation rate of mt DNA. **The mitochondrial genome is housed within the cell's energy producing factory, where it is exposed to reactive by-products of oxidative phosphorylation. Reactive oxygen species (oxygen free radicals) are potent mutagens. The number of reactive by-products, in turn, increases as enzymes involved in energy production accumulate mutations that make them function less efficiently. It is hypothesized that this decline in mitochondrial efficiency is a major contributor to aging.**
- I.4. The high mutability of the mt genome means that it evolves more quickly than the nuclear genome. This makes the mt control region a laboratory for the study of DNA evolution. However, can you think

of any drawbacks to this high mutation rate? **The mutation rate is so high that some nucleotides have mutated several times over evolutionary history. This makes it difficult to determine the actual mutation rate and to ascertain the ancestral (original) state of a particular nucleotide. These "back mutations" also make it difficult to accurately calibrate the "mutation clock."**

- I.5. Would you expect any difference in the mutation rates of the control region sequence in the mt genome *versus* the chromosome 11 insertion? What implication does this have in the study of human evolution? **Once removed from the context of the mitochondrion, the mt insertion is subject to the lower mutation rate of nuclear DNA. The mutation clock effectively slows dramatically, preserving the insertion as a "molecular fossil" from that era in DNA history. This means that one can study human evolution within oneself by comparing mt control region sequence with a nuclear insertion of the same sequence.**
- II.11. What does the amplitude (height) of peak represent? What do you notice about the amplitude of the peaks at the beginning of the sequence? **Fluorescence is proportional to the number of molecules that are terminated with dye at that position, so the height of the peak is proportional to the mass of DNA. It is important that the height of the labeled peaks is significantly greater than the background of erroneous fluorescent products. The initial peaks in the sequence tend to be very high, probably due to increased efficiency of labeling the nucleotides adjacent to the forward primer.**
- II.12. What do you notice about the trace at N positions? **At N positions, two peaks of similar amplitude overlap or intersect, or no prominent peak rises among the background "noise."**
- II.13. Load two different trace files into different viewers, and attempt to match (align) the DNA sequences. What do you notice? What trick did you use to align them? **Different chromatograms are typically out of phase by up to 30 nucleotide, because the software begins calling nucleotides different points in different file. Chromatograms can be manually aligned by matching a unique sequence of five or more nucleotides.**
- II.14. What feature coincides with the abrupt transition from clean-to-poor quality sequence? **A long "C track" coincides with the change in sequence quality. In clean reads, this sequences reads CCCCCTCCCC; in low quality sequences an internal T>C polymorphism, and often a flanking polymorphism, results in a string of 10-11 Cs. For unknown reasons, this "C track" confounds the polymerase, and it essentially falls off the template DNA at this point.**
- III.18. Determine the average number of mutations, as well as the average percent difference, between any two pairs. **This number has been experimentally determined to be about 7 mutations per 375 nucleotides (or 1.87%) within this part of the mt control region.**
- III.19. What does this number represent? **This represents the extent of mt sequence variation in modern human populations. This is equal to the average number of mutations that have accumulated since all modern humans diverged from a common maternal ancestor.**
- III.20. If this is so, what is the rough mutation rate for the mt control region? Why is your calculation only approximate? **According to this study the mutation rate for the mt**



**control region is about 1 new mutation per 20,000 years.**

$$140,000 \text{ years} / 7 \text{ mutations} = 20,000 \text{ years/mutation}$$

**This is only approximate, because it is based on the average accumulation of mutations and nothing has been done to correct for back mutations that can only be determined statistically.**

- IV.6. Pool class data to determine the average number of differences and average percent difference between Otzi and modern humans. **The average number of mutations should be close to that determined in III.18., or about 7 mutations per 375 nucleotides (1.87%). About 10% of students of European ancestry are expected to have sequences identical to Otzi over this portion of the mt control region.**
- V.7. Pool class data to determine the average number of differences and average percent difference between Neandertal and modern humans – and between Neandertals. **The average number of mutations should be about 28 mutations per 375 nucleotides (7.47%).**
- V.8. What does this tell you about the relationship between Neandertals and modern humans? **Neandertal lies outside the variation of modern humans, and, so, cannot be the direct ancestor of modern humans.**
- V.9. Using 140,000 years for the divergence of modern humans, approximately how long ago did humans and Neandertal share a common mt ancestor? **If 7 differences between modern humans = 140,000 years to a common ancestor, then 28 differences between modern humans and Neandertal = 560,000 years to a common ancestor.**
- V.10. What does the small number of available samples suggest about Neandertal diversity? **With an average of about 8 mutations per 331 nucleotides (2.41%), Neandertal had about the same level of genetic diversity as modern humans. However, Neandertal achieved this level of diversity in approximately 300,000 years of evolution as a species, compared about 150,000 years for modern humans. This suggests that Neandertal populations, like human ones, encountered several bottlenecks that reduced genetic diversity.**
- V.6. What do your results show? Is this what you expected? **In most three-way alignments, the Africans will differ at the majority of SNP locations. This is counterintuitive, since one would expect the European or Asian to be different – and for the Africans to have a greater affinity.**
- V.7. Pool class results to confirm or refute your own results. **Although students may feel that their individual results might be the results of small sample size, the trend of Africans differing from each other at most SNP locations is upheld when data is pooled.**
- VI.8. What does your analysis say about African mtDNA diversity? What population factors could account for this? **There is greater mtDNA diversity among Africans. Since mutations accumulate over time, this suggests that African populations are older than European and Asian populations. Genetic diversity also increases with population size, so this suggests that African populations have been maintained at higher levels over long periods of human evolution. This mtDNA evidence that humans arose in Africa is also supported by analyses of Y chromosome and autosomal polymorphisms.**
- VI.9. What does your analysis say about European and Asian mtDNA diversity? What population factors could account for this? **Europeans and Asians share the majority of their mt mutations with African. Most European and Asian mtDNA diversity is a subset of African diversity. This provides evidence that founders of European and Asian populations, carried a portion of the accumulated African diversity with them when they migrated out of Africa. Relatively few unique mutations have accumulated in European and Asian populations over the 50,000-70,000 years since they left Africa.**

# **Carolina Biological Supply Company**

2700 York Road, Burlington, North Carolina 27215

Phone: 800.334.5551 • Fax: 800.222.7112

Technical Support: 800.227.1150 • [www.carolina.com](http://www.carolina.com)

CB00000000