HHMI-NMSU
MOBILE MOLECULAR BIOLOGY LAB

PHENOTYPES TO GENOTYPES
EYES ABSENT GENE

Student Instructions and Supplemental Material

Funding Provided by

HHMI
HOWARD HUGHES MEDICAL INSTITUTE

NM STATE UNIVERSITY

Student Instructions and Supplemental Material

Prepared by:
Christin Slaughter
Department of Biology
New Mexico State University

*Portions of this manual are duplicated in the **HHMI Student Instructions and Supplemental Material v.1.2** prepared by Kayce Bell and Christin Slaughter. 2007. With reference therein.*

*Literature cited: Zimmerman, J.E., Bui Q.T., Lui H. and N. Bonini. 2000. Molecular genetic analysis of Drosophila eyes absent mutation reveals an eye enhancer element. Genetics **154**: 237-246*

HHMI Mobile Lab Introduction for Students

In these lab activities you will determine the phenotype and genotype for a gene that is involved in eye formation in *Drosophila* flies (*eyes absent* – given the notation *eya*). First you will determine the color of each fly's eyes, or if they have eyes at all. That information (the fly's phenotype) will be used to predict the genotype. Then you will extract DNA from each of the flies, amplify a portion of the *eya* gene containing the *eyes absent* mutation eya^2, and determine the each fly's genotype using gel electrophoresis. Before we begin the activities you need to read through the instructions for the lab. This instruction manual also includes a vocabulary list of terms that you need to be familiar with, an explanation of the Hardy-Weinberg principle and associated exercises, instructions on how to create and keep a lab notebook, and instructions on how to use a micro-pipettor. In addition to the student instruction manual you will receive a lab notebook to record data and notes during the activities. There are pre-lab questions that you need to answer before each lab activity. All of the answers you need for those questions are contained in this workbook or will come from discussions. There are also analysis questions for you to answer after each lab activity. These questions can be answered from the activities you have just done, or they are thought questions about the activity content. Every day you should read the lab activities for the next day and answer the Pre-Lab Questions for those activities.

In addition to carefully reading and following ALL instructions, it is very important you listen to your teacher and the facilitator. If you are unsure about the instructions, please ask for clarification. These lab activities are meant to be interactive, please ask questions as they arise and don't be afraid to contribute to discussion!

Before the first day of the lab activities you should:
- Read pages 1-8.
- Write down any questions you have about the instructions or the activities.
- Answer the Pre-Lab Questions for Activity 1 on page 7.

Before the second day of the lab activities you should:
- Answer the Analysis Questions for Activity 1 on page 8.
- Answer the Pre-Lab Questions for Activity 2 on page 9, and read procedures on pages 9-11.

Before the third day of the lab activities you should:
- Answer Analysis Questions for Activity 2 on page 11.
- Answer Pre-Lab Questions for Activity 3 on page 13, and read procedures on page 13.
- Answer Pre-Lab Questions for Activity 4 on page 16.
- Read about loading agarose gels, page 19.

Before the fourth day of the lab activities you should:
- Complete and answer Analysis Questions for Activity 4 on page 17.
- Answer Pre-Lab questions for Activity 5 on page 20, and read procedures on pages 20-21.

Before the fifth day of lab activities you should:
- Complete Analysis Questions for Activity 5 page 21.
- Be prepared to calculate Hardy-Weinberg for the classroom fly population.
- Think about the Discussion Questions and answer any that you can.

MOLECULAR TECHNIQUES

The lab experiments we will be doing will utilize three techniques commonly used in molecular research: DNA extraction, polymerase chain reaction (PCR), and gel electrophoresis.

DNA extraction – The DNA we will be working with is stored in the nucleus of the cell. In order to work with the DNA we must release it from the cell and the nucleus. You will crush individual flies in tubes with a buffer solution, then incubate the tube at room temperature to release the DNA from the cells. You will then spin the tubes in a microcentrifuge to separate the DNA from the other cell material.

Polymerase chain reaction (PCR) – To determine the genotype of each of the flies we need to make many copies of the *eyeless* gene. The polymerase chain reaction makes those copies using three main ingredients (in addition to DNA): primers which are short pieces of DNA that correspond to the beginning and the end of the piece of DNA to be amplified, extra nucleotide bases, and PCR polymerase that will use the extra nucleotides to make copies of the DNA between the primers. For the PCR reaction to take place the PCR ingredients and the DNA must go through several cycles of heating and cooling. First the samples are heated to denature the double-stranded DNA so that the nucleotides of each strand are exposed, then the samples are cooled slightly so that the primers can anneal to the single stranded DNA, then the samples are heated slightly so that the PCR enzyme can extend the strand from the primers.

Gel electrophoresis – A gel matrix (we will be using agarose gels) is poured into a mold that will form a rectangular gel with depressions, or wells, at one end. The gel is submerged in a buffer solution and an electric current is applied. DNA has a slight negative charge so DNA fragments will migrate towards the positive charge on the gel. The gel matrix allows smaller fragments to move faster than larger fragments, so that over time fragments of different sizes will be separated. SYBR® Safe DNA gel stain will be used when making the agarose gels so that after electrophoresis the DNA bands can be viewed on a UV transilluminator. The stain inserts itself between DNA base pairs in double stranded DNA and fluoresces under ultra-violet light. In the lab homozygous recessive genotypes will have one band that is 367 base pairs long, the heterozygous genotypes will have one 689 base pair band, one 367 base pair band, homozygous dominant genotypes will have one 689 base pair band.

INSTRUCTIONS FOR KEEPING A LAB NOTEBOOK

In addition to the pre-lab and analysis questions you will be asked to record notes, observations and procedures in a lab notebook. Keeping a good lab notebook is important in any research laboratory - it helps others (and yourself!) understand what experiments and procedures were done, why each step was performed, and how the experiment or procedure was done. A good lab notebook contains enough detail so another person could repeat the experiment, but it's not necessarily a set of instructions like you see in the lab manual.

There are three main parts to each lab notebook entry (in addition to recording the date): **purpose**, **procedure (or methods)**, and the **outcome (or result)**. Each of these parts with examples is detailed below. A lab notebook entry is best written while the experiment is being performed and the information is still fresh in your mind, rather than after you have finished a procedure or experiment.

PURPOSE: A short 1-2 sentence entry of what you are about to do.

Purpose: Extracting DNA from Drosophila flies for later use in a PCR.

PROCEDURE: In this section you will record what you are doing and how you are doing it - **as you are doing it**. In many laboratories "recipes" for different solutions and reagents used will be listed here. You will also record notes about changes in the procedure or observations made. An entry for the fly DNA extraction may look something like this:

Procedure:

Group #: 3

labeling tubes and recording phenotype:

tube	phenotype
1-3	wild type (red)
2-3	wild type (red)
3-3	eyeless
4-3	wild type (red)
5-3	eyeless

putting fly into tube

DNA extraction continued. [*any time you have to continue your lab notebook entry onto another page you should always write a brief note indicating the procedure that is continued from the previous page*]

adding 50 ul smashing buffer (SB) to each tube

* note: tube 2-3 may not have 50 ul total - I think I messed this one up with bad pipetting

grinding each fly with individual grinders
incubating at room temperature for 30 minutes
incubation start time: 10:15 am
incubation end time: 10:50 am

*note: incubation was <u>35 minutes</u> long

putting tubes into hot water bath (95°C) 2-minutes

spinning tubes in microcentrifuge for two minutes at full speed (14, 000 rpm)

OUTCOME (RESULT): The entries for this section should contain any observed results and may include visual results like a photograph or computer outputs (like graphs or charts). An entry for the fly DNA extraction may look something like this:

<u>Result</u>: DNA extracted from cells and present in the upper layer of the buffer solution.

EXAMPLE OF A LAB NOTEBOOK ENTRY

09-16-09

Purpose: Extracting DNA from Drosophila flies for later use in a PCR.

Procedure

Group #: 4

labeling tubes and recording phenotype:

tube	phenotype	
1-4	wild type (red)	I did these tubes
2-4	wild type (red)	
3-4	eyeless	
4-4	wild type (red)	Tonia L. did these tubes
5-4	eyeless	

putting fly into tube

adding 50 ul smashing buffer (SB) to each tube

* note: tube 2-4 may not have 50 ul total - I think I messed this one up with bad pipetting

grinding each fly with individual grinders

incubating at room temperature for 30 minutes

incubation start time: 10:15 am

incubation end time: 10:50 am

*note: incubation was 35 minutes long

putting tubes into hot water bath (95°C) 2-minutes

spinning tubes in microcentrifuge for two minutes at full speed (14, 000 rpm)

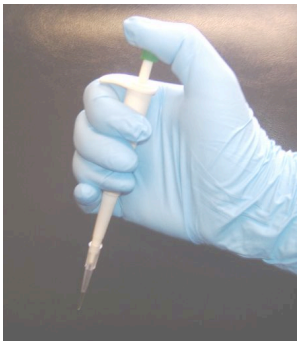
Result: DNA extracted from cells and present in the upper layer of the buffer solution.

INSTRUCTIONS FOR USING A PIPETTOR

1. Make sure you are using the pipettor with the correct volume. If you need to transfer 5 µL (µL = **micro**liters) of liquid, make sure you have the 5 µL pipettor.

2. Grasp the pipettor in your dominant hand wrapping your fingers around the barrel just below the top lip with your thumb free.

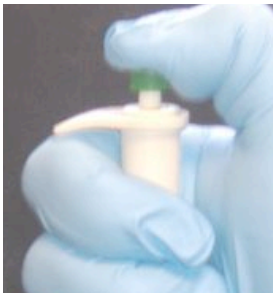
NEVER use, carry, store or hold the pipettor with the plunger toward the floor (upside-down)



3. Place the bottom of the pipettor into the top of a pipette tip and **gently** tap.

DO NOT twist or grind the pipettor into the pipette tip
DO NOT store pipettors in the tip box

4. Depress the plunger with your thumb to the first stop ***before*** you place the pipette tip into the liquid you would like to transfer. Make sure you do not depress the plunger all the way; this will not yield the correct volume.

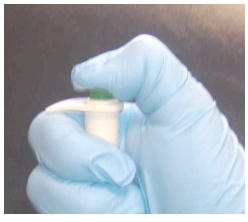


You will still be able to see some of the white on the plunger when it is held at the first stop.

5. After placing the pipette tip into the liquid *slowly* release the plunger. This action allows the pipettor to take-up the liquid.

6. Check the liquid in the pipette tip, make sure there are no bubbles.

7. Place the pipette tip in the container you are transferring the liquid to and *slowly* depress the plunger – to the first stop, then all the way if you have fluid remaining in the pipette tip. **Do not release the plunger** until you have moved the pipette tip out of the liquid.



8. Dispose of the used pipette tip into one of the disposal bottles.

Please practice using the pipettor a few times before you begin the lab, it is very important that you are able to correctly pipette liquids.

To empty out the pipette tip press the plunger down all the way.

LAB ACTIVITIES

Activity 1: *Drosophila* phenotype

Objectives

1. Determine the phenotype of *Drosophila* fly eyes.
2. Know which genotypes results in each phenotype.

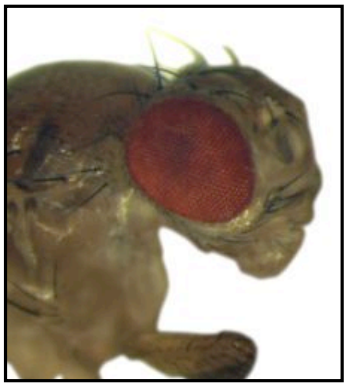
Background: In *Drosophila* flies the ***eyes absent* gene (*eya*)** is a gene that plays an important role in eye formation. This gene works together with other genes to direct eye development and is critical in not only forming an adult eye in *Drosophila*, but also plays a role in head formation and gonad development as a *Drosophila* embryo develops. There have been a number of **mutations** found for the *eya* gene with a number of effects; the mutation we are interested in is given the notation ***eya*^{2}.**

In *Drosophila* genetics, a fly's **genotype** is given by the two **alleles** it carries for the gene of interest. Flies that do not carry a mutated allele, and appear to have a "normal" phenotype, are called **wild type** with a genotype shown as (+ / +). This notation shows there is one "+" for each allele of the genotype.

There are two possible **phenotypes** when a fly is carrying the ***eya*^{2}** mutation. If the fly has a **homozygous recessive genotype** (***eya*^{2} / *eya*^{2}**) then it will have **no eyes on its head**. If a fly is **heterozygous** (***eya*$^+$ / *eya*^{2}**), having one normal (wild type) allele and one mutated allele, then it will look like a wild type fly with **red eyes**. If the fly's genotype is **homozygous dominant** (***eya*$^+$ / *eya*$^+$**) carrying no mutated *eya*^{2} allele, it is a wild type fly with **red eyes**. The flies you will examine under your hand lens will be either eyeless or wild type (having red eyes).

Pre-Lab Questions

1. How many possible phenotypes are there for the *eya*^{2} mutation of the *eyes absent (eya*) gene?
2. How many possible genotypes are there for the *eyes absent (eya*) gene?
3. What alleles could a fly that looks wild type (red eye) be carrying?



Wild type fly
***eya*$^+$ / *eya*$^+$**

Fly carrying a mutated allele - looks wild type
***eya*$^+$ / *eya*^{2}**

Eyeless fly
***eya*^{2} / *eya*^{2}**

Procedure

1. Obtain a tube with two *Drosophilia* flies inside and examine each fly under your hand lens, observing the red eyes of the wild type fly and the eyeless head of the fly homozygous for the *eya*^{2} mutation.

2. Return the flies to the tube and discard.

Analysis Questions

1. Record your phenotype and associated genotype of each of the flies you examined.

2. Using the Punnet Square determine the genotype and phenotypes of offspring from two flies mating that are (a) eyeless (***eya*^{2} / *eya*^{2}**) and wild type (***eya*$^+$ / *eya*$^+$**) and (b) eyeless (***eya*^{2} / *eya*^{2}**) and red eyed (***eya*$^+$ / *eya*^{2}**).

(a)

	***eya*^{2}**	***eya*^{2}**
***eya*$^+$**		
***eya*$^+$**		

genotype **phenotype**

(b)

	***eya*^{2}**	***eya*^{2}**
***eya*$^+$**		
***eya*^{2}**		

genotype **phenotype**

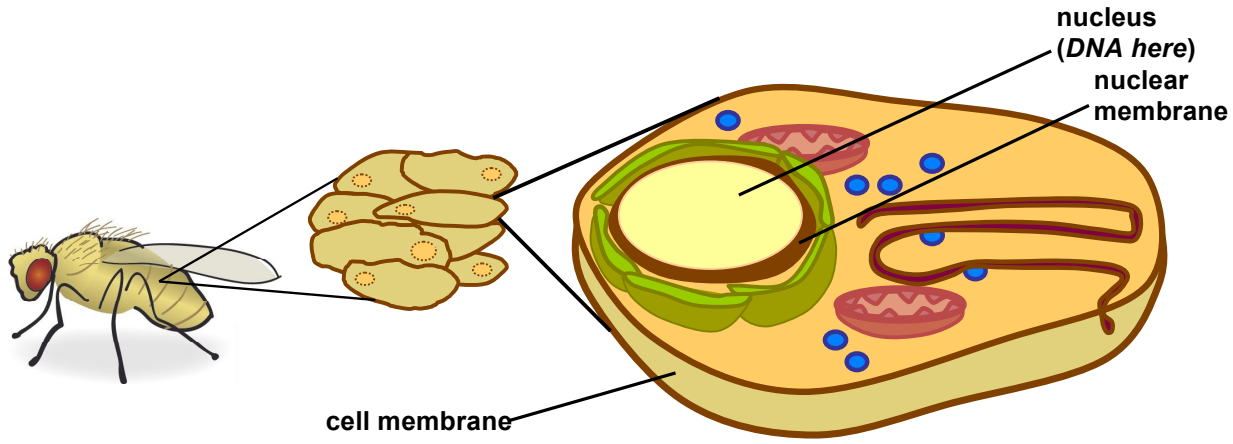
3. Using the Punnet Square (b). What is the porportion (fraction) of offspring that are homozygous recessive for the *eya*^{2} mutation? What is probability a fly will be born eyeless (homozygous recessive)?

4. Thought question. When might it be advantageous (good) to know if an organisim (human or otherwise) is carrying a mutated allele?

Activity 2: DNA Extraction

Objectives
1. Extract cellular DNA from *Drosophila* flies.
2. Understand the purpose of each step in the DNA extraction process.

Background: All nucleated cells contain DNA. To examine and analyze the DNA, cells must first be collected then the DNA must be released from the nuclei and the cells.



A buffer solution is commonly used by researchers to help release DNA from the cell and to protect the DNA strands once they have been released. You will be using a "**smashing buffer**" (**SB**) that contains chemical salts and an enzyme called **Proteinase K**.

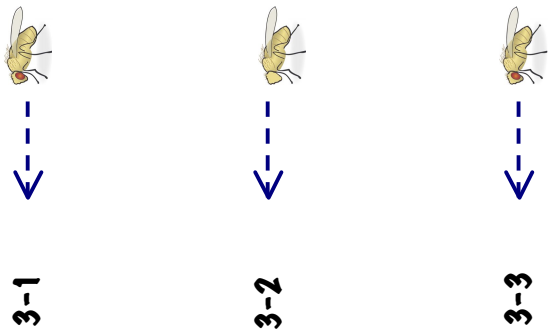
Once you have smashed each fly in its own tube with the SB, the tube will be incubated at room temperature. This incubation allows time for the cells to break down releasing the DNA and for the Proteinase K to digest proteins and remove contaminating agents released by the cells. After DNA extraction you will use PCR to make many copies of the *eya* gene.

Pre-Lab Questions
1. Where is DNA located in the cell?
2. Why is a buffer solution used during DNA extraction?
3. What is the purpose of the Proteinase K in the smashing buffer?

Procedure
Throughout the lab activities you will be asked to label several tubes with your team number and sample number. **Please make sure you are labeling as clearly as possible so you can easily identify and retrieve your samples.**

1. Use a permanent marker to label a 1.5-mL microcentrifuge tube on the **top and side** with your team number and unique sample number.

2. Your team will receive a number of *Drosophila* flies.
First identify each fly's phenotype and then place it in a tube - recording in your lab notebook the fly phenotype and tube number into which it was placed.



tube	
3-1	wild type
3-2	eyeless
3-3	wild type

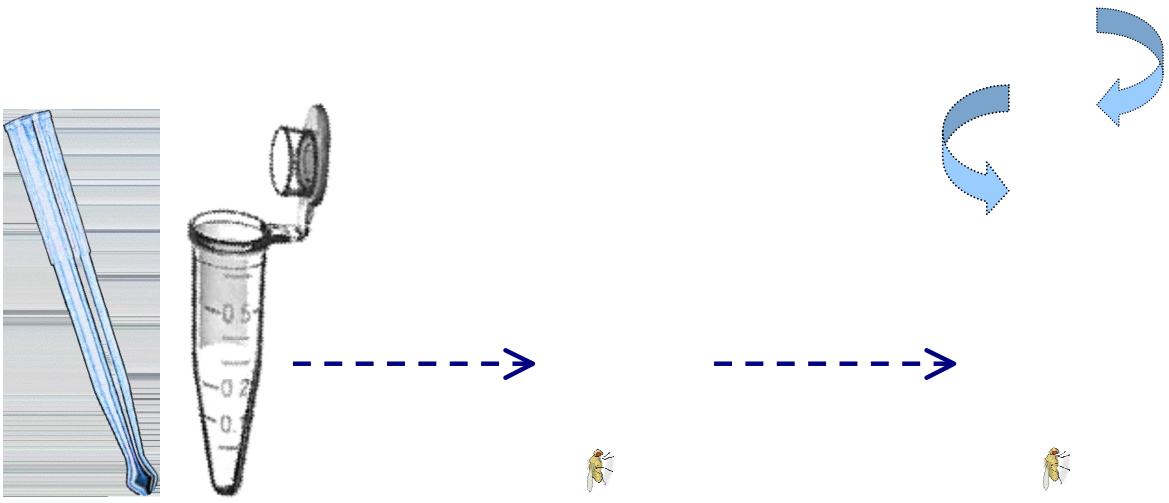
3. Add **50 µL** of smashing buffer (**SB**) into each tube containing a fly.

4. Add **5 µL** of Proteinase K (**prot K**) into each tube containing a fly and SB.

Make sure you switch tips between tubes.

5. Carefully insert a pipette tip "grinder" into each tube and smash the fly with a twisting motion until the fly body is **completely broken-down**. Discard grinders after use.

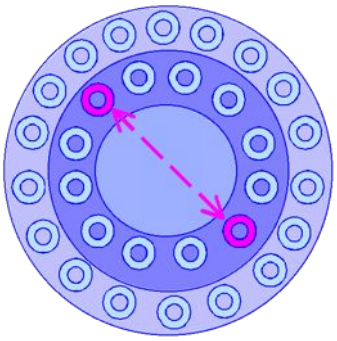
Make sure to use a different grinder for each tube.



6. Cap the 1.5 ml microcentrifuge tubes containing the buffer solution and the crushed fly. Incubate for 30 minutes at room temperature – recording in your lab notebook the incubation start time and end time.

7. After the 30-minute incubation period has ended, place your team's tubes into a **hot water bath** set at **95°C for 2 minutes**. This "heat-shock" will stop the Proteinase K activity.

8. Remove your tubes from the water bath and place your team's tubes along with other team samples, in a **balanced configuration** in the microcentrifuge. Spin the samples for **2 minutes at full speed** (14,000 rpm).



Place tubes directly across from one another to achieve a balanced configuration

8. Following the 2-minute spin, remove your team samples from the microcentrifuge and return the tubes to your large micro-centrifuge tube rack.

*carry these spun-tubes **carefully** so you don't mix the isolated DNA with the cell material*

CONGRATULATIONS - YOU HAVE EXTRACTED *DROSOPHILA* DNA!

Your team's samples contain the DNA that will be used for setting up one or more PCR reactions.

Analysis Questions

1. Why were the tubes put into a hot water bath after the 30-minute incubation?
2. Form a hypothesis about an eyeless fly phenotype and its genotype.
3. Thought question: If we can see the phenotype of each fly, why do we bother with extracting the DNA to determine the genotype?

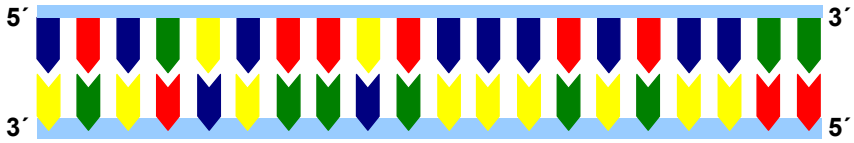
Activity 3: Polymerase Chain Reaction (PCR)

Objectives

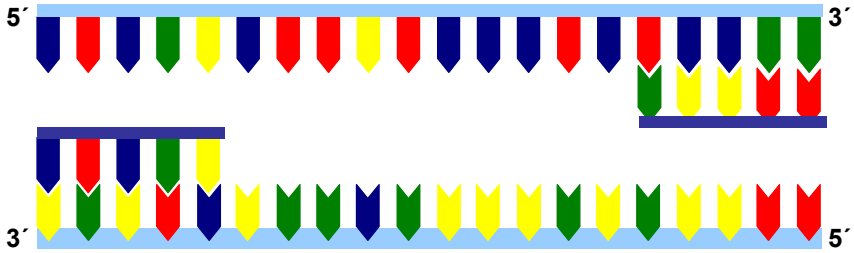
1. Use PCR to make copies of the ***eya*** gene.
2. Know the purpose of each temperature step in the thermal cycling process.
3. Know the purpose of each chemical ingredient in the PCR mix.

Background: To determine the genotype of each fly's ***eya*** gene we will need to make many copies of it. The **polymerase chain reaction (PCR)** uses three main ingredients (in addition to DNA): **primers** which are short pieces of DNA that correspond to the beginning and the end of the piece of DNA to be amplified (target DNA), **extra nucleotide bases**, and **PCR polymerase** that will use the extra nucleotides to make copies of the DNA between the primers. The PCR mix also contains a red dye for electrophoresis (Activity 5). For the PCR reaction to take place the PCR ingredients and the DNA must go through several cycles of heating and cooling.

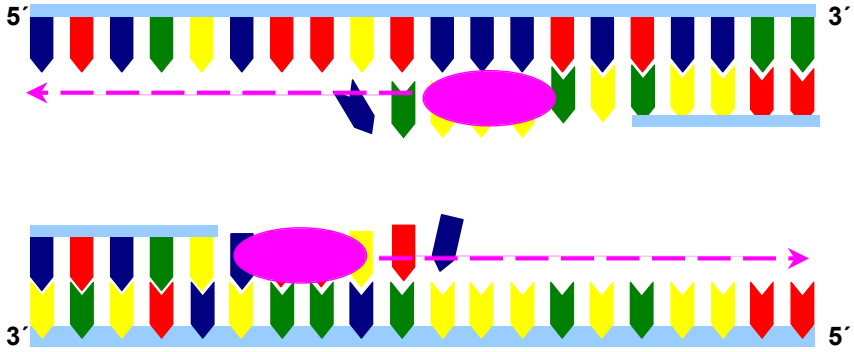
Each 3-step cycle of the PCR begins with double stranded DNA.



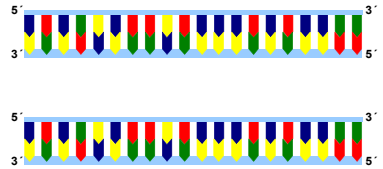
step 1. Samples are heated (94°C) to **denature** the double stranded DNA so the nucleotides of each strand are exposed.



step 2. Samples are cooled slightly (48°C) so the **primers** can **anneal** to the single stranded DNA.



step 3. Samples are heated slightly (72°C) so that the PCR enzyme (***polymerase***) can **extend** the strand from the primers



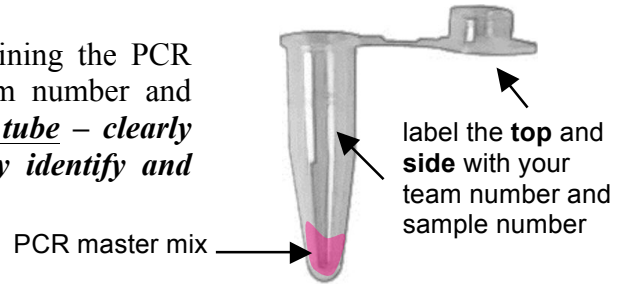
During each full cycle of the PCR (*steps 1-3*) the number of copies of the target DNA increases exponentially.

Pre-Lab Questions

1. What is the purpose of a PCR?
2. What are the chemical components of the PCR and what is the role of each of those ingredients?
3. What is the purpose of each of the three temperatures in each cycle of the PCR?

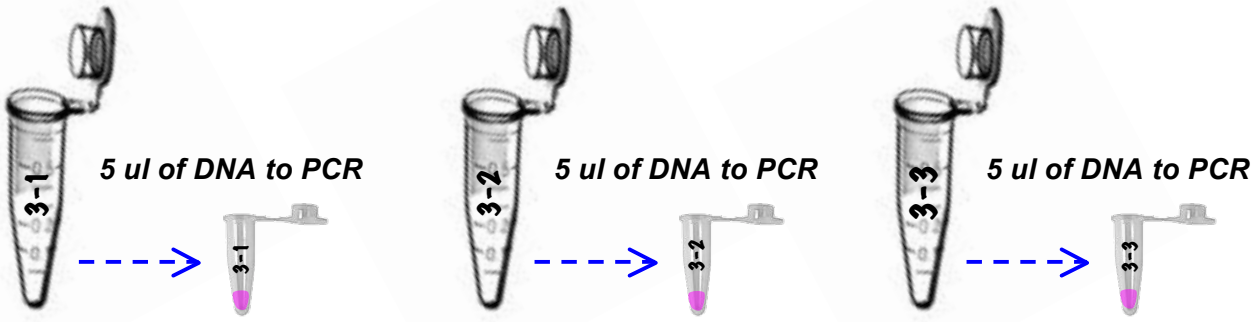
Procedure

1. You will receive several PCR tubes containing the PCR Master Mix. Label each tube with your team number and unique sample number. ***This is a very small tube – clearly label the TOP and SIDE so you can easily identify and retrieve your sample.***



2. Use a micropipettor with a fresh tip to add **5 µL** of fly DNA from 1.5 ml microcentrifuge tube (from Activity 2) directly into the small PCR tube containing the Master Mix.

make sure you are using a clean tip for each transfer



3. Cap the PCR tube and **gently** tap the tube on the tabletop to ensure all the liquid is at the bottom of the tube.

4. Place your PCR tubes into the thermalcycler.

CONGRATULATIONS – YOU HAVE SET-UP A PCR FOR THE *EYA* GENE!

Analysis Questions

1. What naturally occurring process does PCR imitate?

2. Using the single stranded DNA template provided (a) identify the complementary sequence the primers are designed to locate and (b) locate that sequence on the single stranded DNA template.

3. Thought question: What ingredient(s) of the PCR would have to change if we were targeting a gene other than the *eya* gene?

ACTIVITY 4: HARDY-WEINBERG EQUILIBRIUM

Objectives
1. Understand the Hardy-Weinberg Equilibrium Model and the model's predictions about evolutionary change.
2. Understand the terms of the Hardy-Weinberg Equation.
3. Use the Hardy-Weinberg equation to calculate allele frequency and genotype frequency.

Background: The biological sciences generally define **evolution** as the change in allele frequencies in the gene pool of a population over time. One way to measure if this change has occurred (*if the population is evolving*) is by using the Hardy-Weinberg Equilibrium model.

The **Hardy-Weinberg Principle** states that the frequency of alleles in a population will maintain a stable equilibrium (*the allele frequency is not changing*), overtime becoming fixed for the proportions of dominant and recessive alleles at a given locus.

Under the Hardy-Weinberg Principle these allele frequencies will remain the same from one generation to the next (*the population will not evolve*) only if the population meets several conditions:

- large population size (the effect of chance on allele frequencies has less of an impact)
- random mating (individuals are not mating based on a particular phenotype – *in our case the wild type or eyeless phenotype)*
- no migration (no immigration or emigration, the population is "closed")
- no mutation (no change in the DNA sequence of the alleles under investigation)
- no selection ("natural selection" – each genotype has an equal chance of surviving and reproducing; *fitness*)

The **Hardy-Weinberg Equilibrium Equation** is used to predict the *expected* frequency of alleles and/or genotypes in a population over generations. If the *expected* proportion of alleles or genotypes does not match the *observed* proportion of alleles and genotypes in a population, one or more of the Hardy-Weinberg conditions are being violated and the population is **not** in Hardy-Weinberg Equilibrium. If a significant difference between expected and observed frequencies is calculated the population is thought to be evolving.

Two equations are used to calculate expected Hardy-Weinberg frequencies:

Equation: $\mathbf{p + q = 1}$ When this equation is squared it yields: $\mathbf{p^2 + 2pq + q^2 = 1}$

p = the proportion of **alleles** in the population that are dominant (for example the ***eya***$^{+}$ allele)

q = the proportion of **alleles** in the population that are recessive (for example the ***eya*^{2}** allele)

p^2 = the proportion of the population that possesses a homozygous dominant **genotype** (***eya*$^+$/*eya*$^+$**)

2pq = the proportion of the population that possesses a heterozygous **genotype** (***eya*$^+$/*eya*^{2}**)

q^2 = the proportion of the population that possesses a homozygous recessive **genotype** (***eya*^{2}/*eya*^{2}**)

We can use the observed **phenotype** frequencies in a population and the two equations above to determine the *expected* **genotype** frequencies a population should have if it is in Hardy-Weinberg Equilibrium.

Example: A population of Snapdragon plants has a total number of 700 individuals.

450 have a Red phenotype (RR genotype)

150 have a "Red-ish" phenotype (Rr genotype)

100 have a White phenotype (rr genotype)

1. First calculate the observed **phenotype frequencies**.

the number observed/population total

Red: 450/700 = 0.64 Red-ish: 150/700 = 0.21 White: 100/700 = 0.14

This means that 64% of our population has the Red phenotype, 21% have the" Red-ish" phenotype, and 14% have the White phenotype.

2. Use the **phenotype frequency** of the **homozygous recessive** individuals to determine the recessive and dominant **allele frequency** when the heterozygous state is not known with certainty.

p + q = 1

White = rr (or r^2) = 0.14 $\sqrt{0.14}$ = **0.37** = **q** 1 – 0.37 = **0.63** = **p**

This means that the allele frequency for the recessive allele (r) is 0.37 and the allele frequency for the dominant allele (R) is 0.63.

3. Using the calculated the allele frequencies, determine the expected **genotype** frequencies.

$$\mathbf{p^2 + 2pq + q^2 = 1}$$

0.63^2 = **0.396 = 0.40 = p^2** 2(0.63)(0.37) = **0.47 = 2pq** 0.37^2 = **0.14 = q^2**

This means that 40% of our population is expected to have a homozygous dominant (RR) genotype, 47% of our population is expected to have the heterozygous genotype (Rr), and 14% is expected to have the homozygous recessive genotype (rr).

4. Use the expected genotype frequencies to estimate the number of individuals in our population that should have that particular genotype – if the population is in Hardy-Weinberg Equilibrium.

genotype frequency X population total

0.40 X 700 = **280 (RR)** 0.47 X 700 = **329 (Rr)** 0.14 X 700 = **98 (rr)**

We expect that 280 plants will be homozygous dominant, 329 will be heterozygous, and 98 will be homozygous recessive – if our population is in Hardy-Weinberg Equilibrium – not evolving, We would ***compare our observed genotypes obtained from our molecular genetics results to the expected*** *– if they are not significantly different then our population has not evolved.*

Pre-Lab Questions

1. What is the biological definition of evolution?

2. What are the five conditions a non-evolving population must meet under the Hardy-Weinberg Model?

3. Which Hardy-Weinberg equation describes **genotype** frequency? Which Hardy-Weinberg equation describes **allele** frequency?

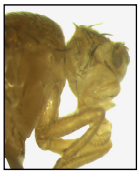
Procedure

1. Read each scenario below and answer the associated questions in order to determine the **expected genotype frequencies** for each of the populations sampled using the H-W equations.

2. Using the gel-photo documents provided;
   (a) total each of the genotypes present in the population
   (b) compare these **observed genotypes** to the **expected genotypes** to determine if each population is in Hardy-Weinberg Equilibrium.

Scenario 1 (Gel-Photo 1). Researchers are breeding *Drosophila* flies to determine is there is any selection on the ***eya*** gene. From the first population that has been bred the researchers observe the following phenotypes in **54** randomly sampled flies:

50 red eye

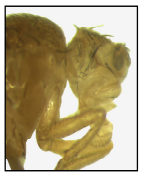


4 eyeless

Because the researchers are unsure of the true genotype of the red eyed flies you will use the frequency of the eyeless phenotype to calculate expected allele and genotype frequencies.

Scenario 2 (Gel-Photo 2). Researchers are, again, breeding *Drosophila* flies to determine is there is any selection on the ***eya*** gene (*because scientists never just do an experiment once*). From this next population that has been bred the researchers observe the following phenotypes in **54** randomly sampled flies:

52 red eye



3 eyeless

Again, the researchers are unsure of the true genotype of the red eyed flies. You will use the frequency of the eyeless phenotype to calculate expected allele and genotype frequencies.

Analysis Questions:

1. In scenario 1, was the fly population in Hardy-Weinberg Equilibrium? Was the population in scenario 2 in Hardy-Weinberg Equilibrium?

2. Which population, the population from Scenario 1 or the population from Scenario 2, is no longer evolving?

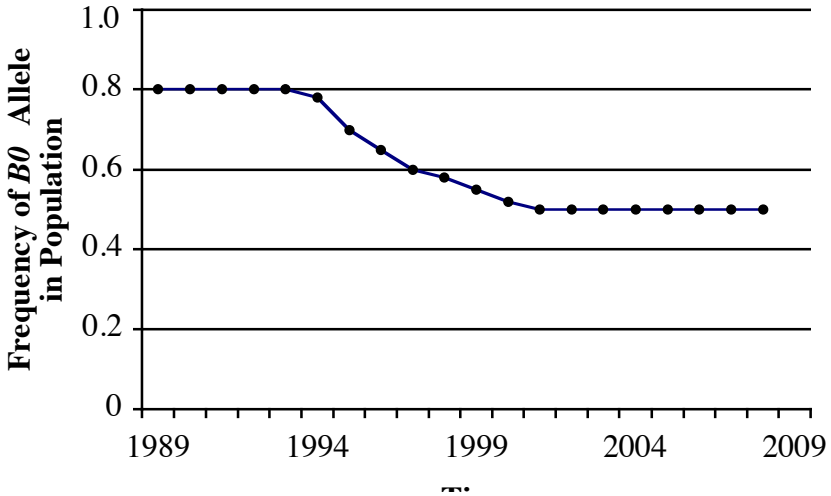
3. Thought question. Which genotype - homozygous dominant, heterozygous, or homozygous recessive – do you think plays a large role in promoting evolutionary change?

_____

_

AP Biology Questions AP1-AP5 refer to the following*

In the land snail *Cepaea nemoralis* (Grove snail) presence of bands on the shell is controlled by two alleles, B^0 and B^B at a single locus. B^0 (unbanded shell) is dominant over B^B (banded shell). A large population of snails was studied over time and the frequency of the unbanded allele B^0 was documented. The results are shown below. In 2009 a random sample of 1000 snail eggs was collected and the snails were allowed to hatch.



AP1. During which of the following time periods could the population have been in Hardy-Weinberg equilibrium for the B^0 alelle?

I. 1990-1994
II. 1995-2001
III. 2002-2009

(A) I only
(B) II only
(C) III only
(D) I and III only
(E) I, II, and III

AP2. Assuming the snail population was in Hardy-Weinberg equilibrium for the ***B*** locus, what percentage of the snails in the natural population had banded shells in 1992?

(A) 2%
(B) 4%
(C) 10%
(D) 20%
(E) 64%

AP3. Assuming the population was in Hardy-Weinberg equilibrium for the ***B*** locus, what was the frequency of the B^0 allele in the snails that were hatched in 2009?

(A) 0.33
(B) 0.50
(C) 0.67
(D) 0.75
(E) 1.00

AP4. Assuming that the population was in Hardy-Weinberg equilibrium for the ***B*** locus, what percentage of <u>unbanded</u> snails that hatched in 2009 was heterozygous?

(A) 0%
(B) 25%
(C) 33%
(D) 67%
(E) 100%

AP5. Which of the following is the most likely reason for the observed differences in the frequency of the B^0 allele between 1995-2001?

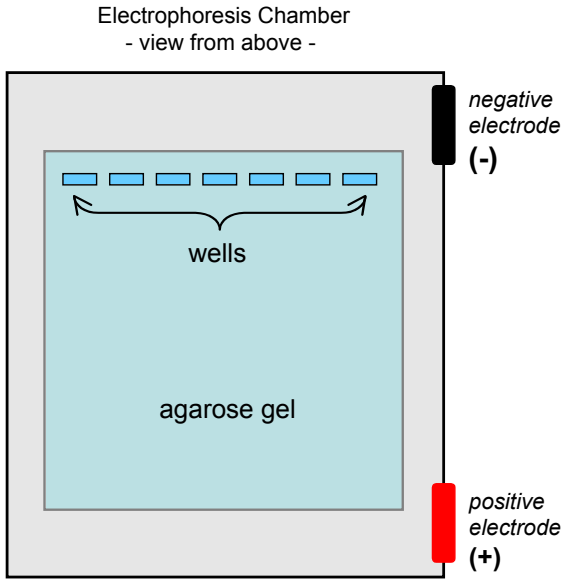
(A) Emigration of banded snails from the population.
(B) Chance
(C) Selection against unbanded phenotypes.
(D) Speciation
(E) Mutation

*modeled after 2002 College Board, Released Exam Excerpt, Section I.

During Lab Activity 4

There will be some time during Activity 4 where you will be waiting. While you are waiting you need to **practice loading agarose gels**.

The gels have been set up around the classroom. Each gel has wells at one end and is submerged in buffer. The pipettes and tips you will be using are next to each gel. A tube of dye is also with each gel, so you can see what you are pipetting and where it will go. Please practice loading gels when you have free time, it will be easier for you when you have to load samples. Below is a picture of how to load a gel. Please see the facilitator for help or if you have questions.



Tip: When you are loading the gel use two hands, one to hold the micropipettor and the other to stabilize it. Do not touch the agarose with the pipette tip. If you get agarose in the pipette tip, the sample or dye will not come out.

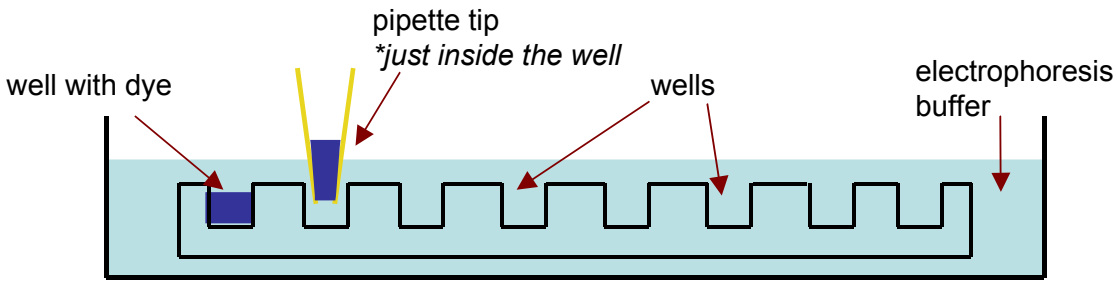
To load a well:

1. take up the dye or DNA sample into your pipette tip

2. position the tip just inside the well (*see figure below*)

3. slowly depress the plunger ejecting the sample - DO NOT release the plunger until you have removed it from the buffer - you will suck up your sample.

Electrophoresis Chamber - view from side - top end of gel



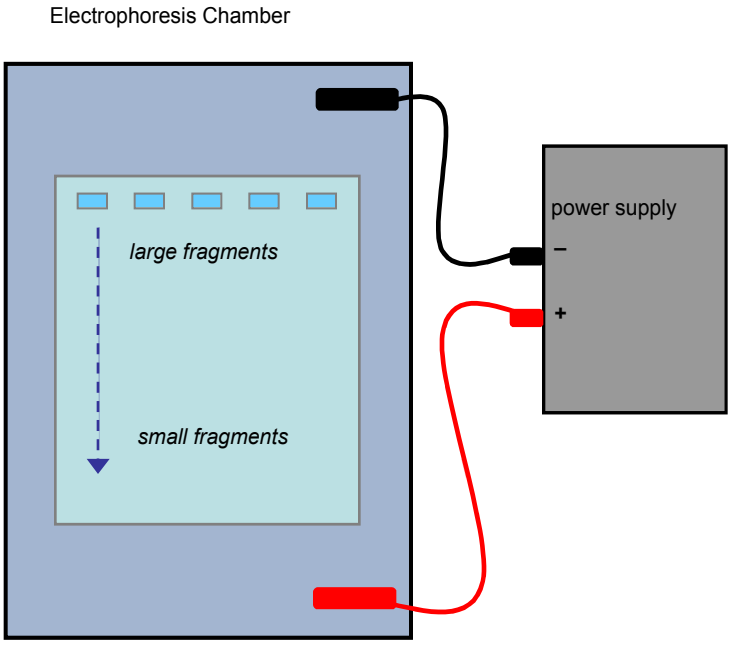
Do not "stab" the agarose gel with the micropipette tip.

Activity 5: Gel Electrophoresis

Objectives

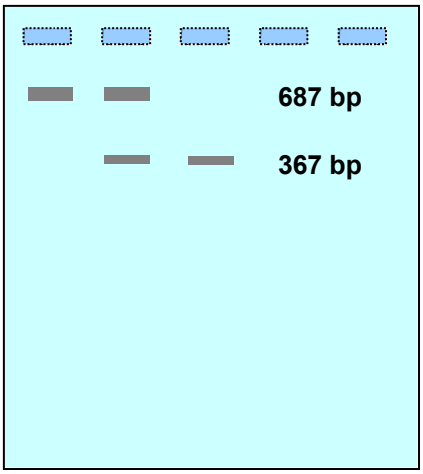
1. Use agarose gel electrophoresis to visualize the *Drosophila* PCR products
2. Know how and why DNA can be visualized on an agarose gel.
3. Understand the mechanisms of gel electrophoresis that separate DNA fragments of different sizes.

Background: DNA is often visualized using the technique of **gel electrophoresis**. First, a gel matrix (we will be using agarose gels) is heated to a liquid state then poured into a mold that will form a rectangular gel with depressions, or **wells**, at one end when the agarose cools. DNA samples, in our case PCR products, are pipetted into the wells. The gel is submerged in a **buffer** and an electric current is applied. DNA has a slight negative charge so DNA fragments will migrate towards the positive charge on the gel. The gel matrix allows smaller fragments to move faster than larger fragments, so that over time fragments of different sizes will be separated.



SYBR® Safe DNA gel stain will be incorporated into the agarose gels so that after electrophoresis the DNA bands can be viewed on a UV transilluminator. The stain inserts itself between DNA base pairs in double stranded DNA and fluoresces under ultra-violet light.

In this lab **homozygous dominant** genotypes will have one **687 base pair** (bp) band, **heterozygous genotypes** will have one **687 base pair** (bp) band and one **367 base pair** (bp) band, and **homozygous recessive genotypes** will have one band that is **367 base pairs** (bp) long.



Pre-Lab Questions

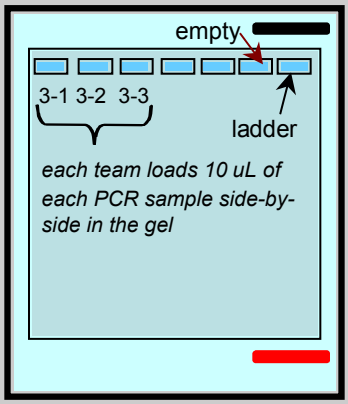
1. What material will be used to make the gel matrix?
2. Which fragment will move the fastest, a 687 base pair fragment or a 367 base pair fragment?
3. Why does a heterozygote genotype have two fragments?

*******WARNING*******

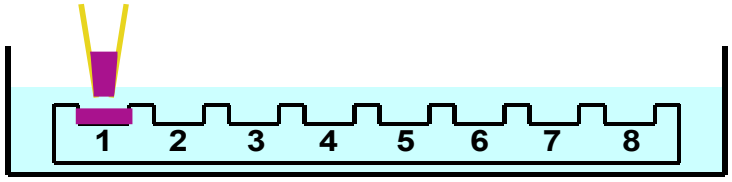
We will be using a stain in the agarose gels that inserts itself into DNA, although this stain has been shown to be fairly safe, anything that can insert itself in DNA presents a possible danger. Make sure to wear gloves at all times you are around or handling the agarose gels and once you have worked with the gel, DO NOT handle or touch anything else. Dispose of your gloves in the garbage and thoroughly wash your hands.

Procedure

Agarose gels inside electrophoresis chambers are located throughout the room. During this activity you will load your PCR and restriction digestion samples into assigned wells. The wells are numbered from 1-12 (sometimes more); you must ***count the wells*** (beginning with the one nearest to you) in order to load the ***correct sample into the correct well***.



1. Each team loads **10 µL** of each PCR product. ***Make sure to record the sample and well number in your lab notebook.*** The facilitator will load 5 µL of size marker/ladder into the gel.



each well is assigned a number - count from the well closest to you and ensure you are loading the correct sample into the correct well

2. Run the gel at 120 volts for approximately 30 minutes.

3. Remove gels carefully and carry the gel over to the UV transilluminator to view. Record the DNA fragment patterns.

Gels can be disposed of in the trash. Electrophoresis buffer can be carefully poured back into storage containers. Gel electrophoresis chambers should be rinsed with water and dried before storage.

CONGRATULATIONS – YOU HAVE PERFORMED GEL ELECTROPHORESIS!

Analysis Questions

1. Why can DNA be seen in the agarose gel under UV light?
2. Using the gel illustration, make a prediction about the fragment pattern for each of your flies. Draw the DNA fragments, and label the DNA starting point, positive and negative electrodes, and direction of DNA migration. (**use more than one lane for a red eye fly*).

AP Biology Questions

AP1*. Discuss how each of the following factors would effect the ability to visualize DNA on an agarose gel using the gel electrophoresis technique.

a. Voltage used. *For example, if high voltage is used and there is a small difference between DNA fragment sizes then the resolution is POOR? GOOD?*

b. Running time. *For example if the gel is run for a long time versus a short time; how does that impact the ability to see differences in the DNA fragments created?*

c. Amount of DNA used. *What if you load a small volume of DNA into the gel, or a large volume of DNA?*

d. Reversal of polarity. *What if you connected the electrodes so the current was applied with the DNA next to the (+) electrode?*

AP2. Two very small DNA fragments of nearly the same base pair size look like a single band, even when the DNA is run to the end of the agarose gel. What could be done to resolve (see the difference between) the two fragments? Why would it work? *Hint: think about the relationship between the agarose gel concentration (thickness) and factors (a) and (b) from above.*

AP3. How does the loading dye function in gel electrohoresis?

AP4. How can a mutation, like an insertion or deletion, be detected using gel electrophoresis?

* Pre-Lab question 2, Analysis questions 1 and 2, and AP1-4 modeled after The College Board AP Program (2001), Lab Six Molecular Biology, Exercise 6B analysis questions 1,2, 4-8.

Final Lab Activity: Data interpretation, troubleshooting and review of primary concepts

Be prepared to discuss all of the activities and all of the questions, including the questions below. This time will be used to analyze genotypic data and compare the results to your phenotype and your predictions from the first day.

Discussion Questions for Final Day

Were you able to support the hypothesis you formed on Day 1?

Were you able to correctly predict your fly's genotype based on the phenotype? What are some possible explanations for an incorrect prediction?

Was the population of flies (class) in Hardy-Weinberg Equilibrium? If not, what are the possible explanations for being out of Hardy-Weinberg Equilibrium?

Based on what you now know, do you think the information about genetics on television and in movies is accurate? What is an example of accurate and inaccurate information?

How do genetics play a role in your daily life?

HHMI-NMSU

MOBILE MOLECULAR BIOLOGY LAB

PHENOTYPES TO GENOTYPES – EYES ABSENT GENE

Supplemental Material

Vocabulary List

Bioinformatics: Phylogenetic Trees

Translation Codon Table

VOCABULARY LIST

allele: an alternative form or DNA sequence for a particular gene.

anneal (in PCR): complimentary base pair binding of primers to single stranded DNA sequence.

base pair: a complimentary pair of nitrogenous bases linked by a hydrogen bond in double stranded DNA or in DNA bound to RNA. Adenine (A) base pairs with thymine (T) in DNA and adenine base pairs with uracil (U) in RNA. Cytosine (C) base pairs with guanine (G) in DNA and RNA.

buffer (in gel electrophoresis): a liquid used to make agarose gels and once gels are solidified they are submerged in the buffer for electrophoresis. The buffer conducts electric current through the electrophoresis chamber and maintains a relatively constant pH.

clade: as in a phylogenetic tree, a monophyletic group containing a single common ancestor and all of its descendants

denature (in PCR): breaking the bonds between base pairs in double stranded DNA.

diploid: cells that have two copies of each chromosome.

dominant: the form of a trait that appears over the recessive form and is expressed in both the homozygous and heterozygous conditions.

gel electrophoresis: running electric current through a gel matrix to separate molecular fragments of differing size.

genotype: the alleles present for a particular trait (diploid cells will have two alleles).

heterozygous: a genotype where there are two different alleles present for a particular gene.

homozygous: a genotype where the two alleles for a particular gene are the same.

lyse: undergo lysis, or breaking of cell membranes.

mutation: a change in a DNA sequence such as the deletion or insertion of a nucleotide or a change from one nitrogenous base to another. Mutations may or may not result in a change in the phenotype.

nitrogenous base: a basic nitrogen containing compound that combines with a sugar to form a nucleoside; adenine, cytosine, guanine, thymine, or uracil.

nodes: as in a phylogenetic tree, a point where two (or more) taxa branches meet. Representing a common, hypothetical, ancestor.

nucleotide: a nucleoside combined with a phosphate group and forming the primary molecule of DNA and RNA.

phenotype: the physical expression or trait for a particular gene.

phylogenetic tree: a graphic representation of ancestor-descendant relationships constructed from DNA sequence data.

PCR polymerase: an enzyme necessary for the formation of a new DNA strand from an existing strand of DNA. Polymerase works during the extension step of the PCR by adding nucleotides to the primers.

polymerase chain reaction (PCR): a process used to make many copies of a target DNA sequence that generally uses three steps: heating up to denature the double-stranded DNA (generally around 93° C), cooling down to allow the primers to anneal (generally in the range of 45-65° C), and slight heating to allow polymerase to extend the DNA sequence from the primers by adding nucleotides (generally around 72° C).

primers (for PCR): a short segment of nucleotides (generally 12-28 nucleotides) that iscomplimentary to a given DNA sequence. Primers are needed to initiate replication by DNA polymerase.

recessive: the form of a trait that is only expressed when present in the homozygous form.

restriction enzyme: an enzyme found in bacteria that recognizes a particular DNA sequence (for example GGCC) and will cut the DNA strand at that site.

sequence (DNA): an ordered series of nucleotide bases (A, T, G, C) representing the DNA strand.

taxon (plural taxa): a name given to an organism or group of organisms.

tips: as in a phylogenetic tree, representing extant (living) taxa from which the DNA was sampled.

wells (in gel electrophoresis): depressions in a gel where samples to be electrophoresed are placed.

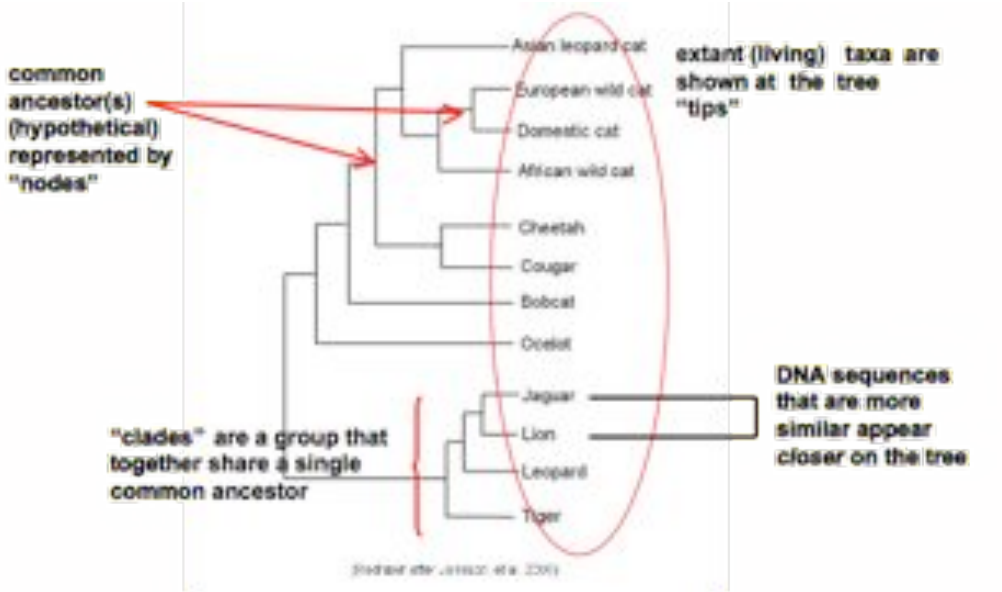
wild type: the phenotype of the "normal" or standard form of a species, often used to describe a phenotype resulting from a genotype with non-mutated alleles.

PHYLOGENETIC TREES

Phylogenetic trees are a commonplace application of DNA information and can be used in a number of ways. Molecular Biologists may be interested in reconstructing ancestor-descendant relationships in a study organism, or determining "relatedness" among a number of sampled individuals. Phylogenetic trees are constructed using DNA sequence data that a researcher generates and inputs into a computer software program.

The main components of a phylogenetic tree are the **tips**, **nodes**, and **clades**. The tips represent living taxa (or individuals) from which DNA has been sampled. Nodes join the tips and represent a common ancestor. Clades are groups of sampled taxa (individuals) that together share a common ancestor. Because phylogenetic trees are generated from DNA sequence data, DNA sequences that are more similar appear closer on the tree.

For the DNA sequence data above there are 2 differences between Sequence 1 and Sequence 2 and 4 differences between Sequence 1 and Sequence 3. Sequence 1 and 2 would be closer on a phylogenetic tree than Sequence 1 and 3.



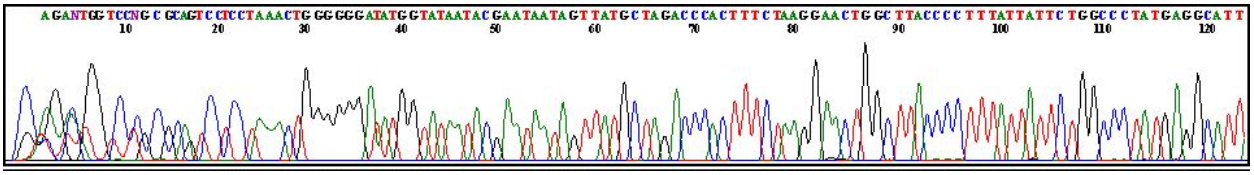
The phylogenetic tree above shows the evolutionary relationship among different cat species. Tips, nodes and clades are highlighted. The DNA sequences that are more similar will appear closer on the tree; for example Jaguars and Lions have similar DNA sequence and are closer to each other than either cat is to a Tiger.

BIOINFORMATICS

Bioinformatics is a growing discipline within the life sciences where information technology is applied to the field of molecular biology. In order to better understand biological processes, bioinformatics focuses on creating databases, mathematical algorithms, statistical techniques, and theory development. One of the primary areas of bioinformatics research is **DNA sequence alignment techniques**.

DNA sequence data provides ordered **nuceotide-by-nucelotide information** for a region of target DNA. During "dye terminator sequencing" DNA sequence data is generated through a process of DNA extraction and PCR amplification, where the PCR product that is generated is then used a template in another PCR reaction – a **sequencing reaction**. During the sequencing reaction each nucleotide is labeled with a fluorescent dye that will fluoresce at different wavelengths when processed in a machine called an "automated sequencer."

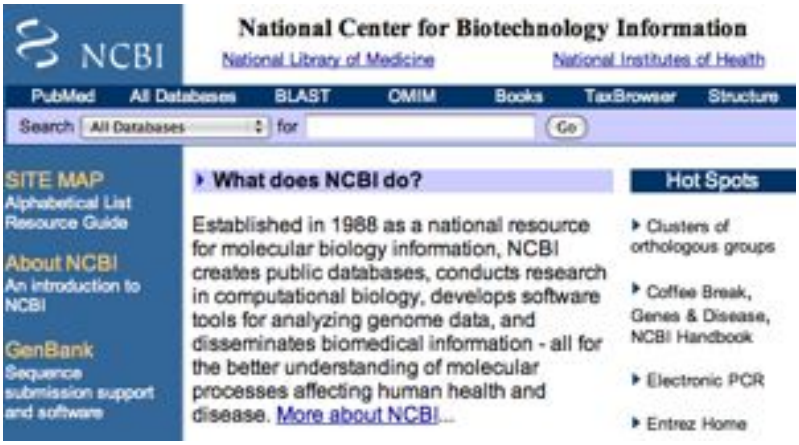
The output of automated sequencing is a dye-terminator read or "**electropharagram**" that displays the nucleotide-by-nucleotide order of the DNA strand.



In the above electropharagram, adenine (A) is labeled with a green dye, thymine (T) is labeled with a red dye, guanine (G) is labeled with a black dye and cytosine (C) is labeled with a blue dye. Notice the nucleotide bases are ordered from 5' to 3' and assigned a position along the strand (e.g. there is a G at base pair position 40).

Once generated, the DNA sequence data can be used with bioinformatics alignment techniques to compare sequence information. The aligned sequences can also be used to construct phylogenetic trees displaying the relationships among the DNA sequences generated.

One source for obtaining DNA sequence data is the **National Center for Biotechnology Information** (NCBI) located on the web at http://www.ncbi.nlm.nih.gov/. Here researchers can submit DNA sequence data they have generated, as well as access data submitted by other researchers. There are also a variety of bioinformatics tools available on the website for use in DNA analysis.

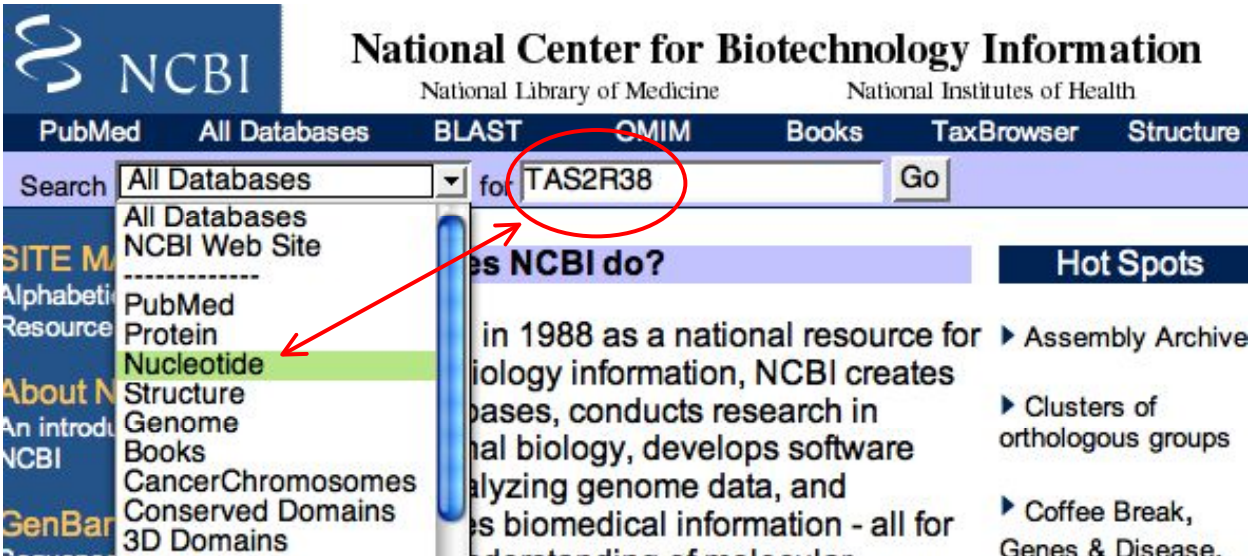


BIOINFORMATICS – PTC GENE IN MAMMALS

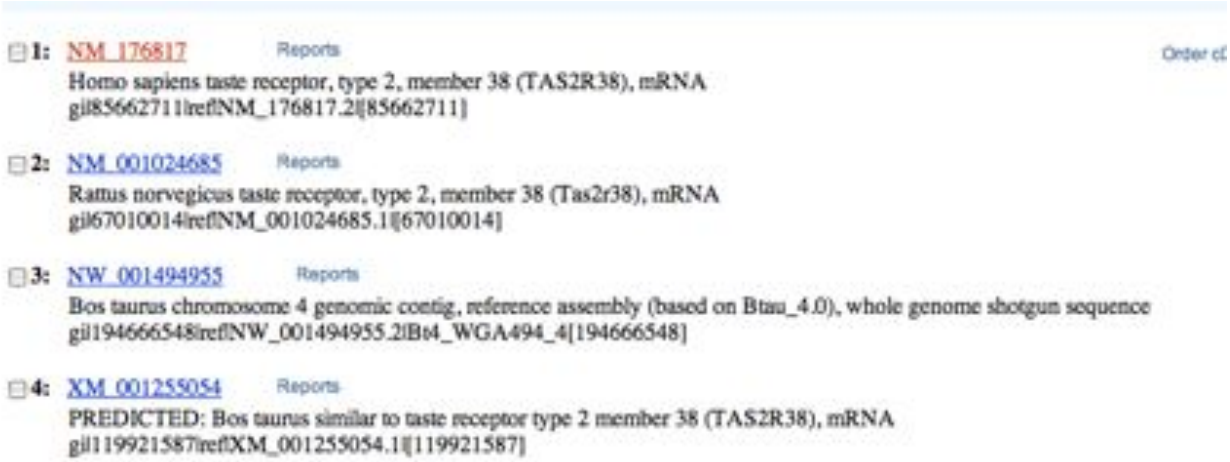
Research shows that human food preference may be linked to a genetically predisposed sensitivity to flavors like that of the bitter chemical Phenylthiocarbamide (PTC). Human and other mammals carry the PTC gene, referred to as the **TAS2R38** gene. Using the **National Center for Biotechnology Information** (NCBI) and associated bioinformatics techniques you will locate a DNA sequence for the TAS2R38 gene (PTC gene) and determine the evolutionary relationship among some mammals for this gene.

1. Access the NCBI home page by entering the web address **http://www.ncbi. nlm.nih.gov/** or doing a search for **NCBI** or **GenBank.**

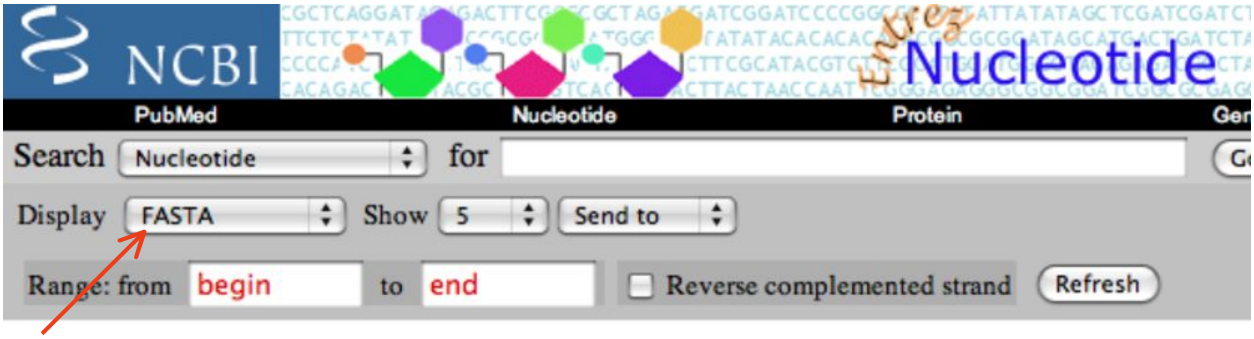
2. On the NCBI home select **Nucleotide** from the Search dropdown menu. Enter **TAS2R38** in the search term box. Click "**Go**"



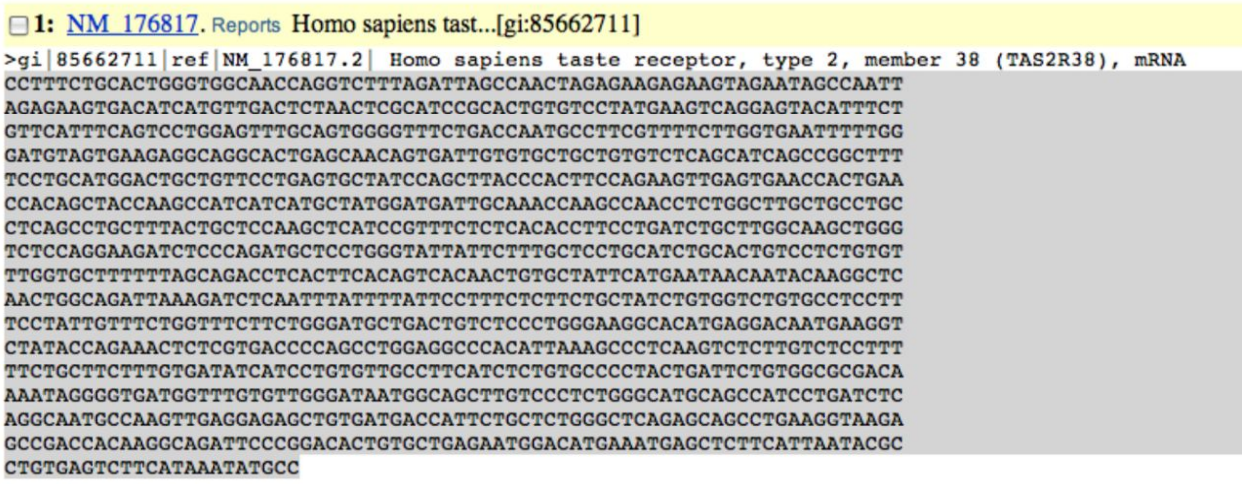
3. The results of your search will be displayed. Scroll down and look for the result for **Homo sapiens taste receptor, type 2, member 38 (TAS2R38), mRNA.** Click on it.



4. The GenBank information page for the TAS2R38 data you selected opens. It contains all the information associated with the DNA data that was submitted by the researching scientists. Select **FASTA** from the Display dropdown menu.

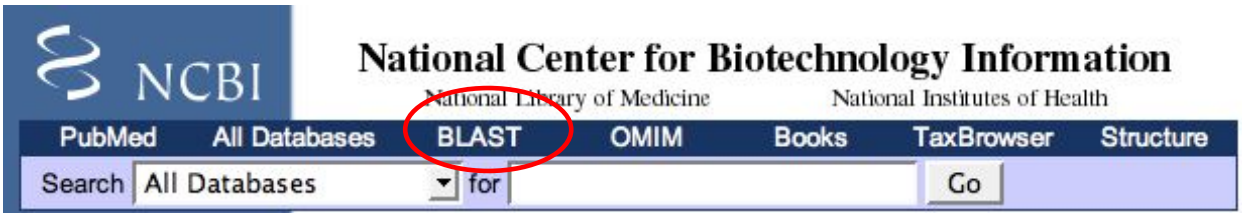


5. The entire DNA sequence for the TAS2R38 gene is displayed. Select the DNA sequence only (not the <gi|85662711|ref|NM_176817.2| Homo sapiens taste receptor, blah blah stuff) and **copy** it. Return to the **NBCI homepage** by clicking on the NCBI logo in the top-left side of the page.

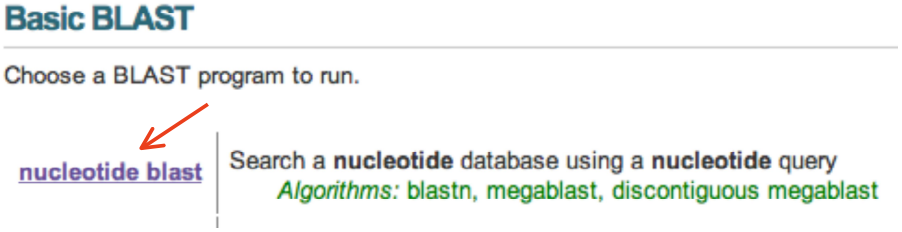

1: NM_176817. Reports Homo sapiens tast...[gi:85662711]

```
>gi|85662711|ref|NM_176817.2| Homo sapiens taste receptor, type 2, member 38 (TAS2R38), mRNA
CCTTTCTGCACTGGGTGGCAACCAGGTCTTTAGATTAGCCAACTAGAGAAGAGAAGTAGAATAGCCAATT
AGAGAAGTGACATCATGTTGACTCTAACTCGCATCCGCACTGTGTCCTATGAAGTCAGGAGTACATTTCT
GTTCATTTCAGTCCTGGAGTTTGCAGTGGGGTTTCTGACCAATGCCTTCGTTTTCTTGGTGAATTTTTGG
GATGTAGTGAAGAGGCAGGCACTGAGCAACAGTGATTGTGTGCTGCTGTGTCTCAGCATCAGCCGGCTTT
TCCTGCATGGACTGCTGTTCCTGAGTGCTATCCAGCTTACCCACTTCCAGAAGTTGAGTGAACCACTGAA
CCACAGCTACCAAGCCATCATCATGCTATGGATGATTGCAAACCAAGCCAACCTCTGGCTTGCTGCCTGC
CTCAGCCTGCTTTACTGCTCCAAGCTCATCCGTTTCTCTCACACCTTCCTGATCTGCTTGGCAAGCTGGG
TCTCCAGGAAGATCTCCCAGATGCTCCTGGGTATTATTCTTTGCTCCTGCATCTGCACTGTCCTCTGTGT
TTGGTGCTTTTTTAGCAGACCTCACTTCACAGTCACAACTGTGCTATTCATGAATAACAATACAAGGCTC
AACTGGCAGATTAAAGATCTCAATTTATTTTATTCCTTTCTCTTCTGCTATCTGTGGTCTGTGCCTCCTT
TCCTATTGTTTCTGGTTTCTTCTGGGATGCTGACTGTCTCCCTGGGAAGGCACATGAGGACAATGAAGGT
CTATACCAGAAACTCTCGTGACCCCAGCCTGGAGGCCCACATTAAAGCCCTCAAGTCTCTTGTCTCCTTT
TTCTGCTTCTTTGTGATATCATCCTGTGTTGCCTTCATCTCTGTGCCCCTACTGATTCTGTGGCGCGACA
AAATAGGGGTGATGGTTTGTGTTGGGATAATGGCAGCTTGTCCCTCTGGGCATGCAGCCATCCTGATCTC
AGGCAATGCCAAGTTGAGGAGAGCTGTGATGACCATTCTGCTCTGGGCTCAGAGCAGCCTGAAGGTAAGA
GCCGACCACAAGGCAGATTCCCGGACACTGTGCTGAGAATGGACATGAAATGAGCTCTTCATTAATACGC
CTGTGAGTCTTCATAAATATGCC
```

6. Select **BLAST** from the NCBI homepage menu.



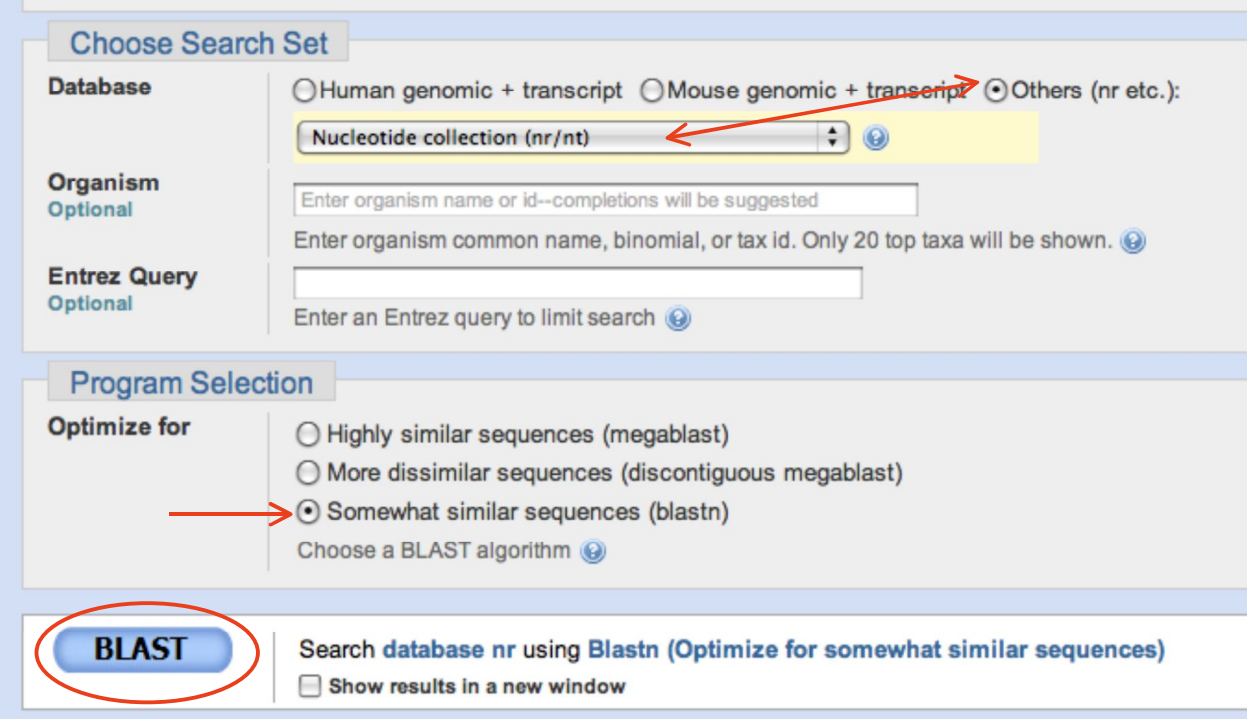
7. On the BLAST search page, select **nucleotide blast**.



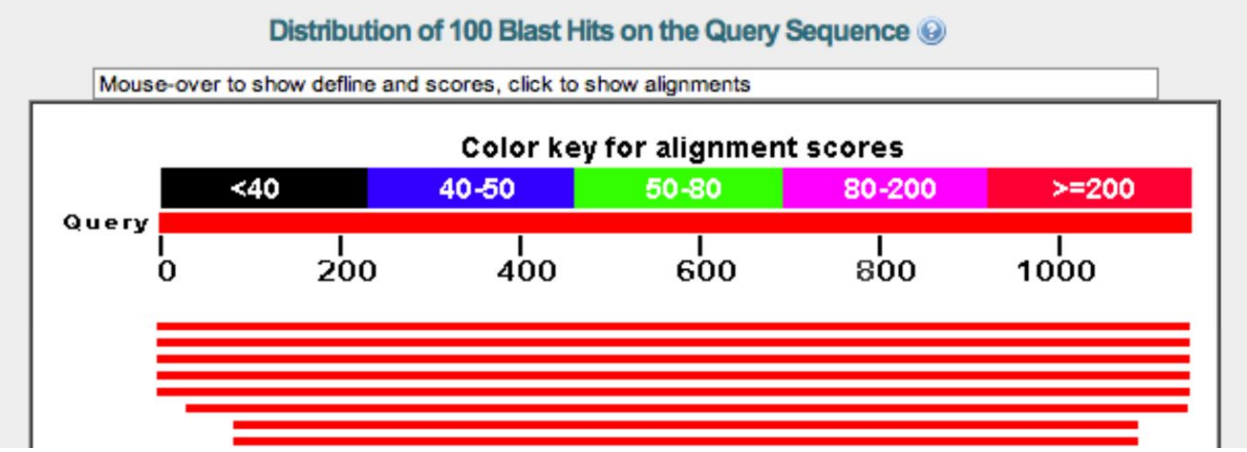
F

8. Paste the TAS2R38 sequence you copied into the **Enter Query Sequence** window.

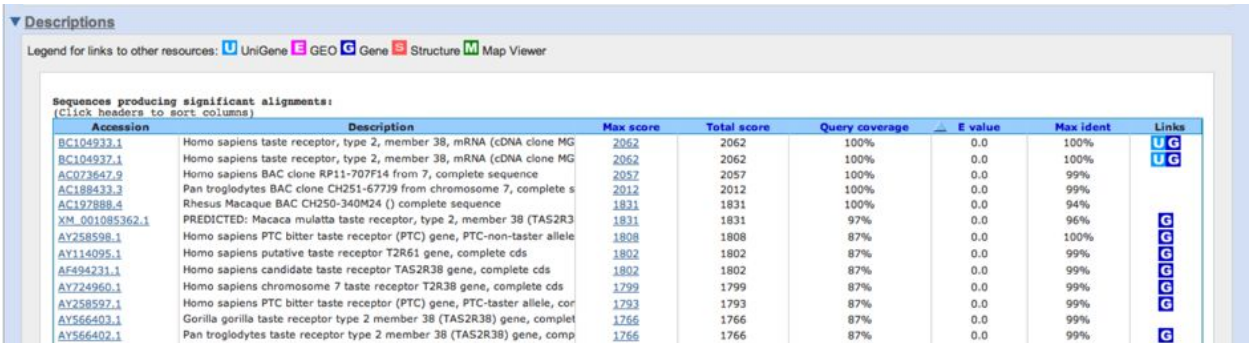
9. Scroll down and select **others (nr etc)** and **Nucleotide collection (nr/nt)** from the **Choose Search Set** options. And **Somewhat similar sequences (blastn)** from the **Program Selection** options. Then hit **BLAST**.



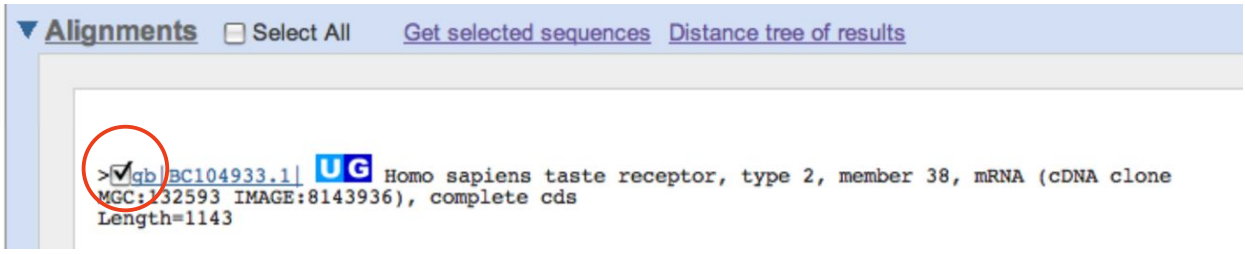
10. The BLAST results appear after the NCBI database is searched for DNA sequences that are similar to the DNA sequence used in the BLAST search. These results appear as "Hits" – highly similar sequences are shown as continuous red bars.



11. Scroll down the page and you will first see the **Descriptions** frame that contains details about each of the BLAST hits shown graphically above. Do you recognize some of the Genus and species names listed?



12. Scroll down further to the **Alignments** frame. It is in this area that you will **select DNA sequences** to use in construction of a phylogenetic tree to determine the evolutionary relationship among mammals for the PTC gene. Place a checkmark next to DNA sequences you will use – these are listed below. You will have a total of **8 DNA sequences** to compare (*do not use the sequence pictured below, it just shows where to place the checkmark to select the sequence*).



>gb|AY258598.1| Homo sapiens PTC bitter taste receptor (PTC) gene, PTC-non-taster allele, complete cds Length=1002

>gb|AY258597.1| Homo sapiens PTC bitter taste receptor (PTC) gene, PTC-taster allele, complete cds Length=1002

>gb|AY566403.1| Gorilla gorilla taste receptor type 2 member 38 (TAS2R38) gene, complete cds Length=1002

>gb|AY566402.1| Pan troglodytes taste receptor type 2 member 38 (TAS2R38) gene, complete cds Length=100

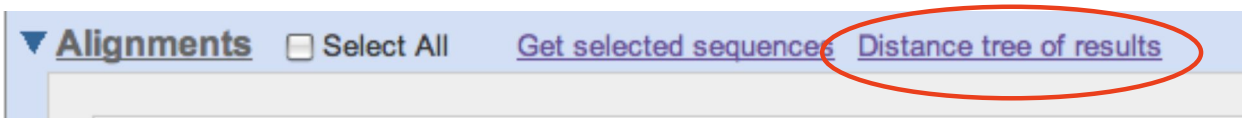
>gb|AY566405.1| Hylobates klossii taste receptor type 2 member 38 (TAS2R38) gene, complete cds Length=1002

>dbj|AB249694.1| Canis familiaris DNA for bitter taste receptor, complete cds, clone: Cafa-T2R38 Length=951

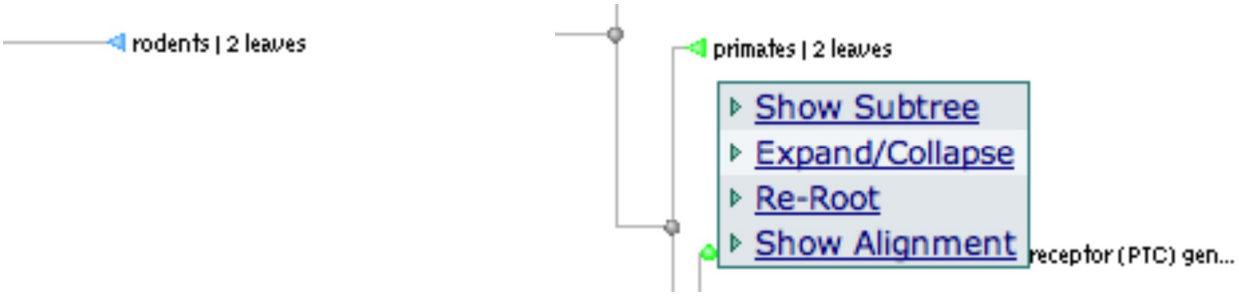
>gb|BC128016.1| Mus musculus taste receptor, type 2, member 138, mRNA (cDNA clone MGC:157613 IMAGE:40136088), complete cds Length=1000

>gb|AY362746.1| Rattus norvegicus putative taste receptor T2R26 gene, complete cds Length=99

13. Once you have selected the 8 DNA sequences you will use, scroll back to the beginning of the **Alignments** frame and click on **Distance tree of results**.



14. A new window will open that displays the phylogenetic tree constructed for the DNA sequences you have selected. The **tip** highlighted in yellow was your Query sequence. You will have to expand the **clade** that contains the **rodents** and the **primates** by scrolling over the blue or green arrow next to the clade name and selecting **Expand/Collapse** from the dropdown menu.



15. Using the worksheet provided, or on your own sheet of paper, draw the phylogenetic tree listing the **taxon** at the tips. Based on the tree you have constructed, what can you say about the evolutionary relationship among mammals for the PTC gene?

BIOINFORMATICS – DNA IN THE COURTROOM

This exercise examines molecular evidence used in a criminal trial involving the accusation of deliberate HIV-1 infection. You may be familiar with the forensic use of human DNA sequences to establish whether or not two DNA samples came from the same source (e.g. crime scene evidence and suspect). This approach has been used successfully to exonerate death row inmates (for examples see the Innocence Project website: http://www.innocenceproject.org/). As human DNA sequences mutate quite slowly, we can use such DNA profiles to definitively establish matches.

In contrast, HIV-1 mutates very rapidly. Because of its high mutation rate, the virus will continue to change (evolve) after a person is infected. Thus, within an infected individual, there may be multiple variants of the virus, all of which diverged from the same strain since the time of infection. Similarly, if many people were all infected by a common source (the same infected individual), over time we would expect to see different sequence variants arise in each infected individual, but for all those variants to be genetically related to one another. We can use the genetic sequences to generate a phylogenetic tree and test hypotheses about the genetic (& evolutionary) relationships between the different viral strains.

An Ill-Fated Argument

In August of 2004, a nurse and her then-boyfriend (a gastroenterologist) got into a serious argument. During the argument, the boyfriend stuck her in a muscle with some kind of needle/syringe. Prior to that time, the nurse had had several HIV tests (each time she gave blood, and one after having the saliva of an infected patient splash on her skin), and she had always tested negative. Her most recent blood donation was in April of 2004, and her blood tested negative at that time. In January of 2005 she tested positive for HIV. At that time, she accused her ex-boyfriend of deliberately infecting her during the argument back in August. He was brought to trial on charges of attempted second-degree murder.

You can imagine that the defense team posed alternative means by which she could have become infected. What are some other possibilities? What kinds of tests or information could be used to rule out these alternative hypotheses for her infection with HIV?

Other sources of infection include her prior sexual contacts and occupational exposure, given that she is a nurse. All seven of the men that she had been in sexual contact with (including her former boyfriend) were tested, and found to be HIV-negative.

Her employment records were examined, and there were no reports of any accidental or occupational exposures, other than the saliva that was splashed on her skin sometime in the mid-1980's. Her file did not have any documentation of any needle sticks at work.

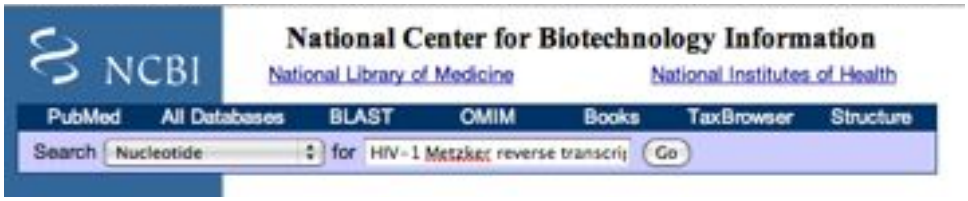
As the investigation proceeded, it was found that an HIV-positive patient under the care of the ex-boyfriend had blood drawn at the physician's offices on August 4, 2004. The paperwork for this procedure was both deliberately hidden (the log book was found at the physician's residence) and was not filled out in a manner that was consistent with the office practices.

Based on the circumstantial case against the physician, the reverse transcriptase (RT) sequences from the victim (the nurse/ex-girlfriend) and from the physician's HIV-positive patient (the putative source of the nurse's infection, via the needle stick during the ill-fated argument) were analyzed.

As HIV-1 mutates rapidly, we don't expect to find identical sequences in the victim and patient. Instead, we expect to find related sequences that share a common ancestor. We can investigate this by using patient and victim HIV RT sequences to generate a phylogenetic tree and look at the clustering of the sequences.

Here's how:

1. Go to the NCBI homepage (http://www.ncbi.nlm.nih.gov/). On the top toolbar, search NUCLEOTIDE (from the pulldown menu that will open with "all databases" selected) for: HIV-1 Metzker reverse transcriptase

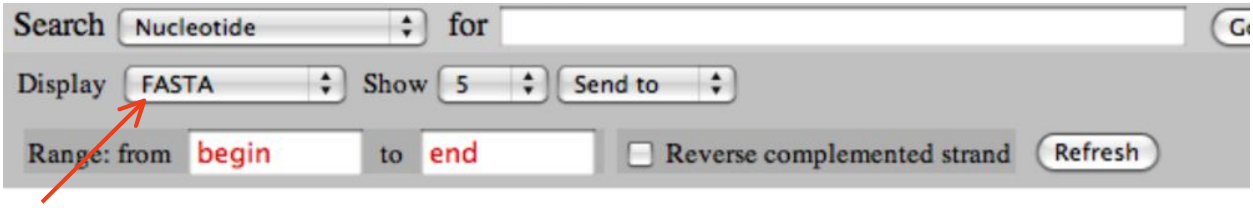


Make sure you enter **HIV-1 Metzker reverse transcriptase** with correct spelling to ensure you retrieve the proper result.

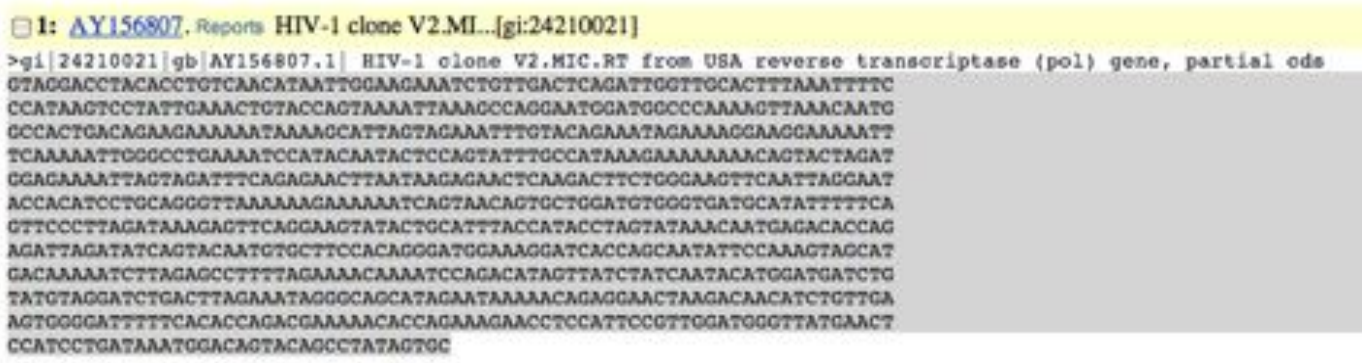
2. You may get a highlighted link of some number of "core nucleotide sequences". Click on the highlighted link.

3. You will get a list of 42 patient (P) and victim (V) sequences. Click on the link for one of the victim (HIV-1 clone V) sequences.

4. When you get to the page that opens, look near the top of the page and change the Display menu from GenBank to FASTA



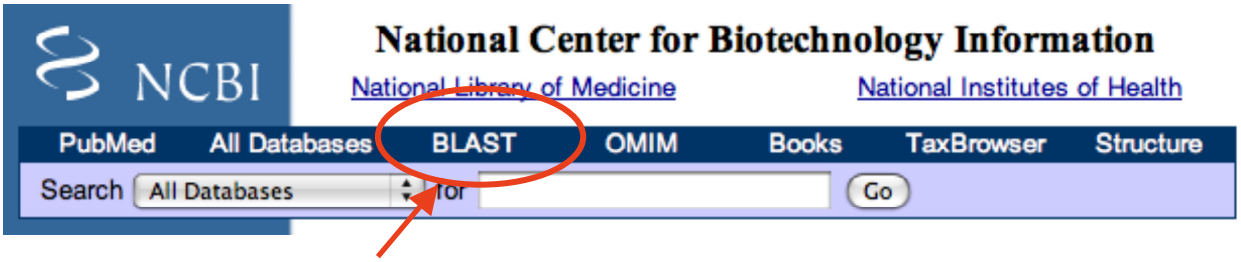
5. You will get the complete nucleotide sequence of that particular sequence. Highlight it and copy it (ONLY the sequence, not the blah blah on the first line).


1: AY156807. Reports HIV-1 clone V2.MI...[gi:24210021]

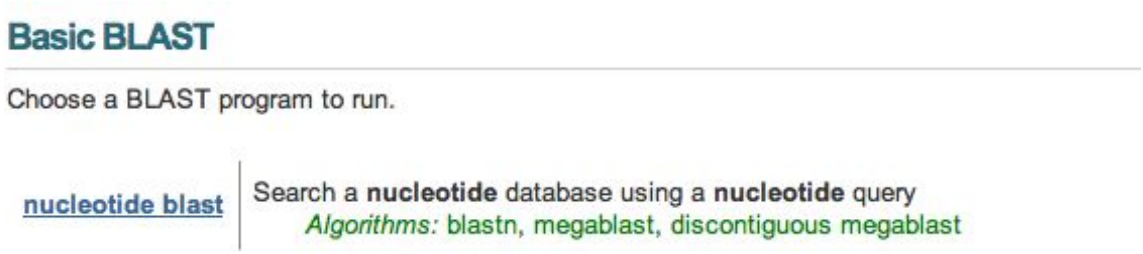
>gi|24210021|gb|AY156807.1| HIV-1 clone V2.MIC.RT from USA reverse transcriptase (pol) gene, partial cds

GTAGGACCTACACCTGTCAACATAATTGGAAGAAATCTGTTGACTCAGATTGGTTGCACTTTAAATTTTC
CCATAAGTCCTATTGAAACTGTACCAGTAAAATTAAAGCCAGGAATGGATGGCCCAAAAGTTAAACAATG
GCCACTGACAGAAGAAAAATAAAAGCATTACTAGAAATTTGTACAGAAATAGAAAAGGAAGGAAAAATT
TCAAAAATTGGGCCTGAAATCCATACAATACTCCAGTATTTGCCATAAAGAAAAAAAACAGTACTAGAT
GGAGAAAATTAGTAGATTTCAGAGAACTTAATAAGAGAACTCAAGACTTCTGGGAAGTTCAATTAGGAAT
ACCACATCCTGCAGGGTTAAAAAAGAAAAAATCAGTAACAGTGCTGGATGTGGGTGATGCATATTTTTCA
GTTCCCTTAGATAAAGACTTCAGGAAGTATACTGCATTTACCATACCTAGTATAAACAATGAGACACCAG
AGATTAGATATCAGTACAATGTGCTTCCACAGGGATGGAAAGGATCACCAGCAATATTCCAAAGTAGCAT
GACAAAAATCTTAGAGCCTTTTAGAAAACAAAATCCAGACATAGTTATCTATCAATACATGGATGATCTG
TATGTAGGATCTGACTTAGAAATAGGGCAGCATAGAATAAAAACAGAGGAACTAAGACAACATCTGTTGA
AGTGGGGATTTTTCACACCAGACGAAAAACACCAGAAAGAACCTCCATTCCGTTGGATGGGTTATGAACT
CCATCCTGATAAATGGACAGTACAGCCTATAGTGC

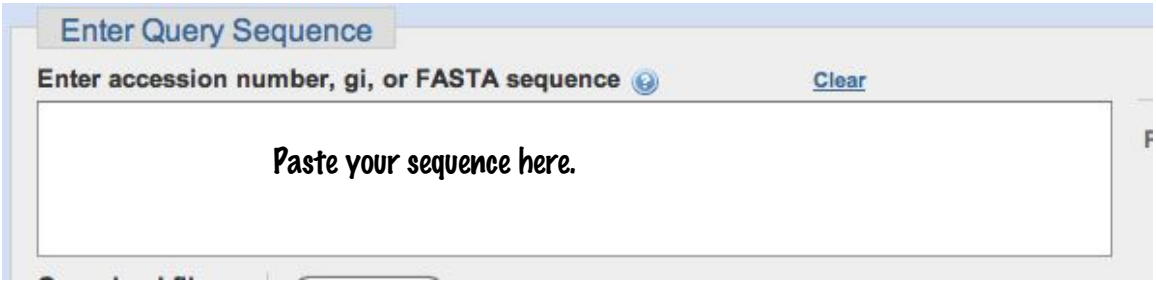
6. Now go back to the NCBI homepage and click on BLAST on the top toolbar.



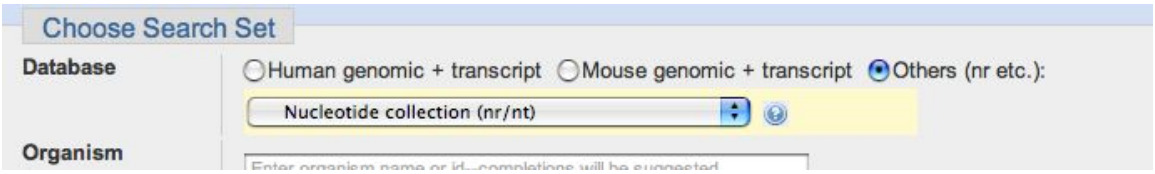
7. When you get to the BLAST homepage, click on the nucleotide blast link (left-side, about half-way down).



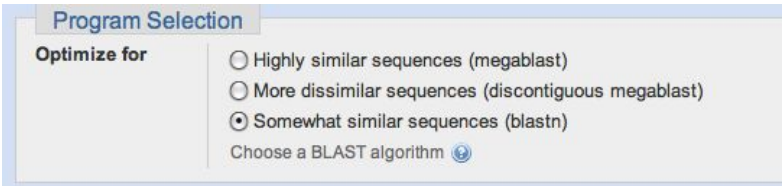
8. When you get to the nucleotide blast page, paste your sequence into the top box (Enter Query Sequence).



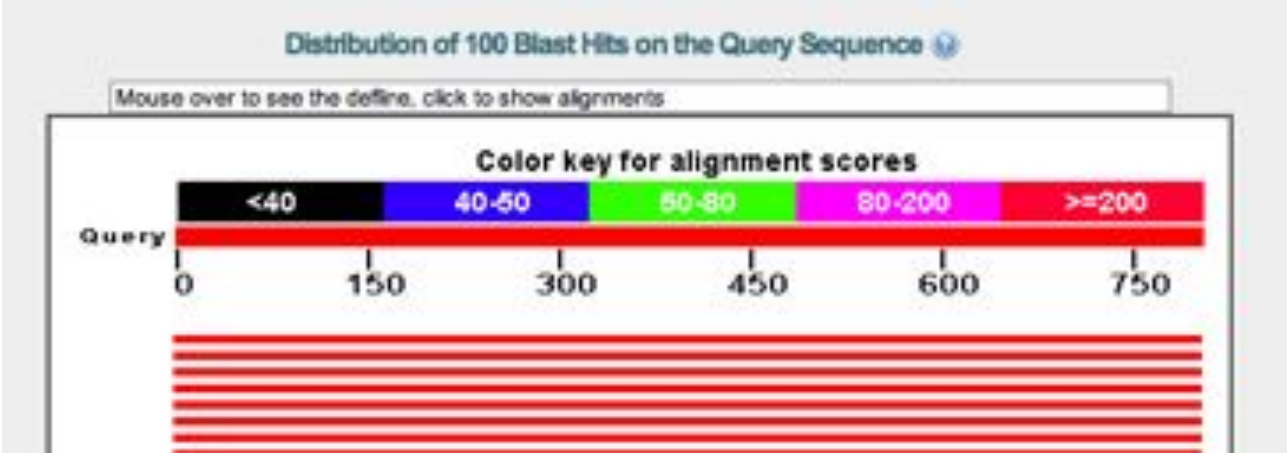
9. Under "Choose Search Set", select "Others" for the database and "nucleotide collection" from the database dropdown menu.



10. Under Program Selection, optimize for somewhat similar sequences (blastn), then:



11. After a few moments, you will get a list of "hits" that have nucleotide similarities to your Query sequence (from the victim).

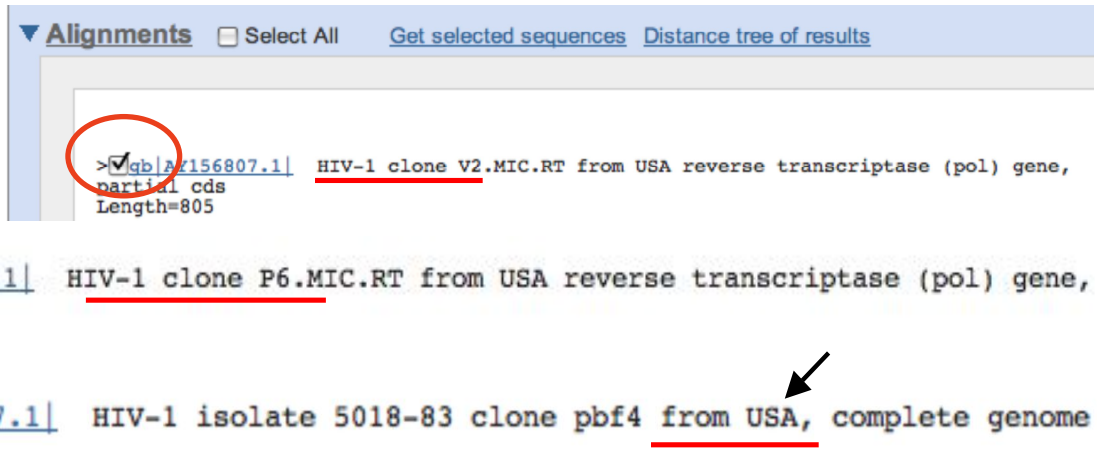


12. Not surprisingly, the top hits are patient (P) and victim (V) sequences from this case.

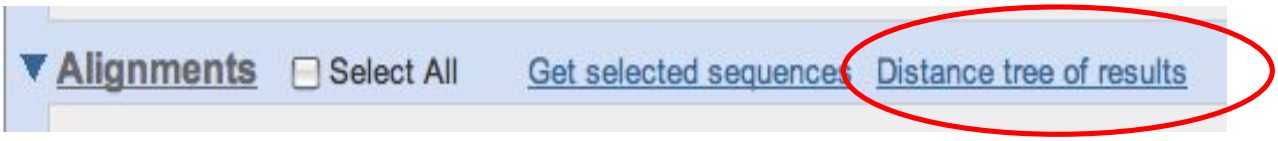
There are a couple victim sequences (V1 and V2) and six patient sequences (P1-P6) because both the victim (the girlfriend) and the (doctor ex-boyfriend's) patient had their HIV strains tested more than once.

13. Scroll down to the listing of individual sequences. Click in the boxes for all of the victim (**V1 and V2**) and patient (**P1 – P6**) sequences. Also click on at least **two** more sequences with "USA" in the title.

To select an individual sequence, mark it by clicking in the box next to the file number



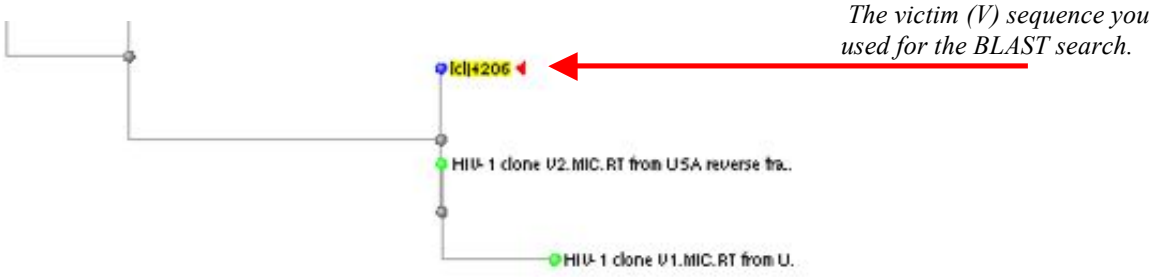
14. Once you have checked the sequences you want to compare, scroll back up to the start of the Alignments section (right between the table of sequences and the actual alignments of sequences):



You are looking for the "Distance Tree of Results" link. Click on it.

14. A new window with a phylogenetic tree will open. The sequences should be labeled. You may have to click on the arrows next to some sequence names to see all the sequences in the group. Remember that V represents HIV reverse transcriptase from the victim and P represents HIV reverse transcriptase sequences from the patient.

The victim (V) sequence you used for the BLAST search.



Answer the following questions on a separate sheet of paper:

1. Describe the tree (using phylogenetic terms). Draw a quick sketch of the tree.

2. Does there appear to be a relationship between the patient and victim sequences? Do they appear to diverge from a common ancestor?

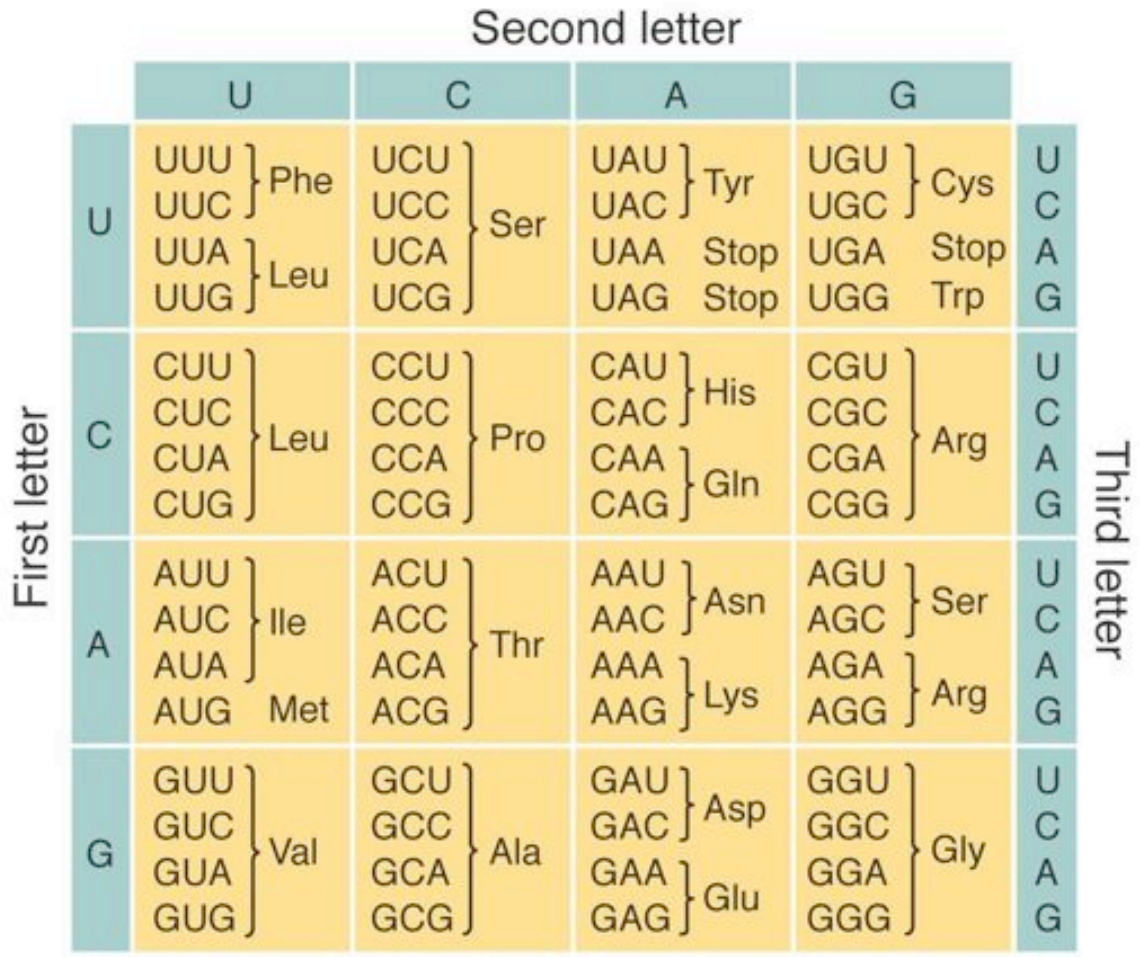
3. What conclusion can you draw from this tree?

4. Given the circumstantial evidence and the phylogenetic evidence, what do you think the verdict was in this case?

From instructional materials created by:
Dr. Michèle Schuster
Department of Biology
New Mexico State University

Reference:
Metzker et al. 2002. Molecular Evidence of HIV-1 Transmission in a Criminal Case. PNAS 99(22):14292-14297.

TRANSLATION CODON TABLE



V.1.3_06_02