Electrophoresis Forensics Lab

Wrongfully Convicted?

miniPCR bio™ Electrophoresis Forensics Lab
Wrongfully Convicted?
Instructor's and Student's Guides

Version: 1.1
Release: February 2022

Instructor's Guide Contents

At a glance

J.M. was convicted of a crime and is currently incarcerated, but he has always maintained his innocence. Use molecular techniques to determine if J.M. was wrongfully convicted.

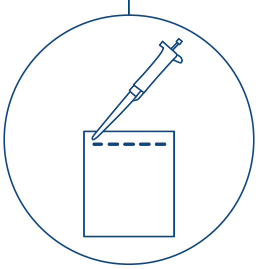
This lab offers students an introduction to forensic DNA analysis. Students will use gel electrophoresis to analyze DNA samples from a closed case to see if the new DNA evidence supports the original conviction.

TECHNIQUES	TOPICS	LEVEL	WHAT YOU NEED	AP CONNECTION
Micropipetting Gel electrophoresis	Forensics Genetics Biotechnology	Middle school through Advanced high school	Micropipettes Gel electrophoresis and visualization system	AP Biology units 5.3, 5.6, 6.8, 7.4 Skills and Practices 1.A-1.C, 2.A-2.D, 3.C-3.E, 5.A, 6.A-6.E

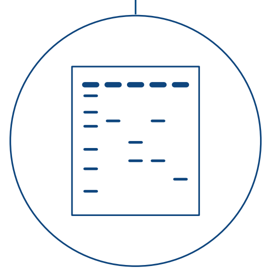
Planning your time

SINGLE CLASS PERIOD: 45 MINUTES* if gels are prepared in advance

Gel electrophoresis



Interpret results



* Allow an additional 25 minutes if students will be preparing the gels in class.

Additional Supports

Help your students achieve mastery of pipetting, PCR, and gel electrophoresis with additional instructional videos, worksheets, and activities available at: https://www.minipcr.com/tutorials/

Taking it further - extension activities, page 25
Explore population genetics and the statistical methods that forensic scientists use to analyze DNA profiles.

For answers to the lab study questions and extensions, email answers@minipcr.com. Please include the name of the lab, as well as your name, school, and title in the body of the email.

Materials needed

Supplied in Kit (KT-1504-01)

Reagents and supplies	Amount provided in kit	Amount needed per lab group	Storage	Teacher's checklist
DNA Samples • Victim DNA • J.M. DNA • DNA Evidence 1 • DNA Evidence 2	**150 µl each**	**12 µl each**	-20°C freezer	
Fast DNA Ladder 1	**100 µl**	**12 µl**	-20°C freezer	

Sold Separately in Learning Lab Companion Kit (KT-1510-01)

This lab requires reagents for running and visualizing DNA samples on a 2% agarose gel with a fluorescent DNA stain (e.g., SeeGreen™ or GelGreen®). The Learning Lab Companion Kit provides enough electrophoresis reagents for 8 groups when using the blueGel™ electrophoresis system. Gels can also be prepared with agarose tabs or agarose powder. Refer to www.minipcr.com/agarose-gel/ for detailed instructions.

Reagents and supplies	Amount provided in kit	Amount needed per lab group	Storage	Teacher's checklist
All-in-one agarose tabs	**8**	**One tab per agarose gel (2% agarose gel)**	Room temp., protected from light	
TBE electrophoresis buffer	**Supplied as liquid concentrate or powder** Sufficient to prepare 600 ml of 1X working solution	**30 ml of 1X buffer per blueGel™ system**	Room temp.	
Plastic tubes	**50 microtubes** (1.7 ml)	**5**		

Materials needed (cont.)

Supplied by teacher

Available at miniPCR.com

Reagents and supplies	Amount needed per lab group	Teacher's checklist
Horizontal gel electrophoresis apparatus: e.g. blueGel™ electrophoresis system	**1** **If sharing gels, reserve 1 lane for ladder and 4 lanes for each group**	
Blue light transilluminator *Note: A blue light transilluminator is integrated in the blueGel electrophoresis system.		
Micropipettes • 2-20 µl: one per lab group	**1**	
Disposable micropipette tips	**At least 5 per group**	
Microcentrifuge (optional; only needed to collect liquid at tube bottom)		
Distilled water for making agarose gels and diluting TBE buffer	**50 ml per gel**	
Flask or beaker to dissolve agarose		
Microwave or hot plate to dissolve agarose		
Other supplies: • Disposable laboratory gloves • Protective eyewear • Permanent marker • Cup to dispose of tips		

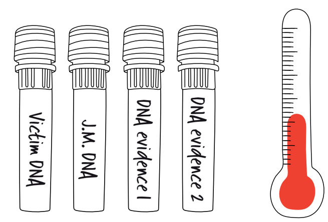
Lab setup

The following activities can be carried out by the instructor ahead of class. Reagents are sufficient to be used with 8 student groups. Reagents are stable at room temperature for several hours during setup and the lab experiment, but should otherwise remain frozen for long-term storage.

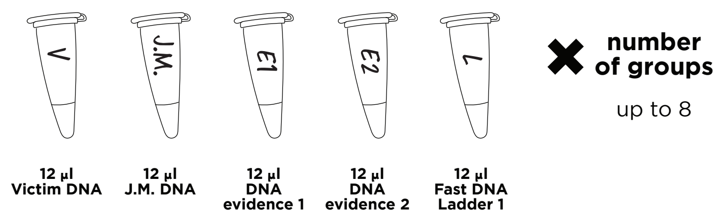
Gloves and protective eyewear should be worn for the entirety of this lab.

A. Dispense reagents and supplies

- Thaw tubes containing the DNA samples by placing them on a rack or benchtop at room temperature
- For each lab group, dispense the following reagents into five labeled 1.7 ml microtubes
    - Victim DNA 12 µl
    - J.M. DNA 12 µl
    - DNA Evidence 1 12 µl
    - DNA Evidence 2 12 µl
    - Fast DNA Ladder 1 12 µl



Defrost tubes



B. Distribute supplies and reagents to lab groups

Check	At the start of this experiment, every lab group should have:	Amount
	DNA samples for analysis • Victim DNA • J.M. DNA • DNA Evidence 1 • DNA Evidence 2	12 µl of each sample
	DNA ladder	12 µl
	2-20 µl micropipette	1
	Micropipette tips	5
	Five wells in an electrophoresis gel	

C. Prepare for gel electrophoresis

- Prepare 1X TBE buffer.
  - TBE buffer is often provided as liquid concentrate or powder.
  - Follow manufacturer's instructions to prepare 1X TBE buffer solution.
  - Volume to prepare depends on the method used to prepare gels; see "Important Note" below.
- Gels can be poured in advance of the class.
  - This lab requires running and visualizing DNA samples on a 2% agarose gel with a fluorescent DNA stain (e.g., SeeGreen™ or GelGreen®).
  - Pre-poured gels can be stored at ambient temperature, in a sealed container or wrapped in plastic wrap, and protected from light for up to three days.
- Have the banding pattern of the Fast DNA Ladder 1 handy (page 18) to help interpret the electrophoresis results.

IMPORTANT NOTE: There are several ways to prepare agarose gels.

- Scan the QR code for detailed instructions on how to prepare agarose gels.
- Both written and video instructions are available.

www.minipcr.com/agarose-gel/

Student's Guide Contents

Background and significance

Forensic DNA analysis

DNA analysis is considered the gold standard in forensic science. But why is this the case? When comparing two people's DNA, more than 99% will be identical. Even with so much similarity, there are more than enough differences to make every person's genetic sequence unique. This makes DNA analysis a powerful tool for identification of individuals. Scientists can isolate DNA from biological crime scene evidence, like a bloodstain or a hair follicle, and use the information in the DNA to match the sample to a specific person.

As forensic approaches have evolved, there have been several methods for matching DNA evidence to individuals. Most forensic DNA analysis looks at regions of the genome that vary in length between individuals.

Short tandem repeats

Most modern forensic DNA analysis focuses on *short tandem repeats* (STRs). STRs are short sequences, typically 2 to 5 DNA bases, that repeat several times in a row (Figure 1). The number of times the bases repeat, however, varies between people. Therefore, the length of DNA where the STRs are located also varies between people. Forensic scientists take advantage of this and measure these differences in length to compare DNA between individuals. There are many locations in the human genome where the number of copies of a specific STR varies between people. Forensics scientists give each STR a complex name that reflects its location in the genome, but for simplicity, we will refer to an example as STR 1. At this location in the human genome, the sequence GATA is repeated anywhere between 5 and 16 times (Figure 1).

Figure 1. Short tandem repeats
STRs are short DNA sequences that repeat several times in a row. For example, STR 1 is a specific location in the human genome where the sequence GATA is repeated between 5 and 16 times.

An *allele* is one of two or more versions of a DNA sequence that appears in the same location in the genome. Forensic scientists refer to STR alleles by the number of tandem repeats they contain. In the example of STR 1, if there are 5 GATA repeats, then forensic scientists would say that it is the "5" allele. Because humans have two copies of each chromosome, one inherited from each biological parent, people have two alleles for each STR region. Scientists use the term *genotype* to refer to different combinations of alleles that a person could have. For instance, someone who is "9, 9" at STR 1 inherited 9 GATA repeats from each parent. On the other hand, someone who is "5, 9" at STR 1 inherited 5 GATA repeats from one parent and 9 GATA repeats from their other parent (Figure 2).

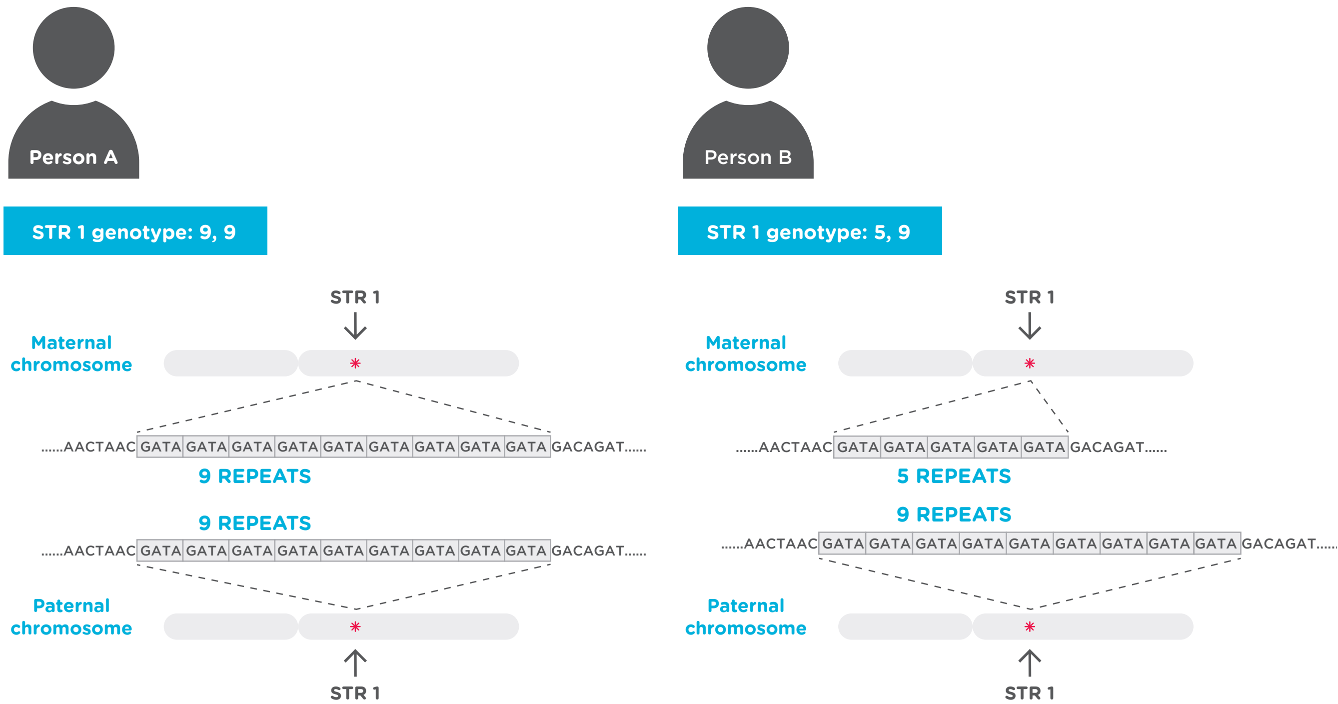


Figure 2. STR genotypes
People have two alleles for each STR region in the genome. The alleles could be the same or they could be different. Person A has two copies of the same allele for STR 1 and their genotype is 9, 9. Person B has two different alleles for STR 1 and their genotype is 5, 9.

Non-coding regions of the genome

When we think of DNA, we often think of genes, which contain the coded instructions to make RNA and proteins. However, the majority of the genome doesn't code for proteins. Scientists refer to this DNA as non-coding DNA. Non-coding DNA can have many functions. For example, non-coding DNA can be involved in gene regulation or have a structural role, like protecting the ends of chromosomes. There are also many regions of non-coding DNA where scientists don't know what the function is, or if it even has a function. Importantly, these non-coding regions are more variable between individuals than coding regions. For this reason, forensic scientists analyze non-coding regions of the genome when creating DNA profiles.

In forensic investigations, scientists examine if a person's STR genotypes match the DNA evidence. Comparing the number of repeats at multiple locations in the genome can create a unique DNA profile for genetic identification (Figure 3). The more STR regions compared in a forensic investigation, the more likely it is to find differences between people's *STR profiles*.

Police can compare STR profiles from DNA evidence with DNA from a suspect if they have one. But law enforcement agencies can also compare DNA evidence to databases of DNA profiles. The FBI's Combined DNA Index System (CODIS) is a DNA profile database for 20 standard STR locations (Figure 4). As of 2020, CODIS contains more than 19,000,000 DNA profiles from convicted offenders and individuals who have been arrested, as well as other DNA samples from forensic investigations (FBI, 2020). A centralized DNA database means that local, state, and federal authorities can all compare DNA profiles in the same way. And as of 2020, close to 500,000 criminal investigations have been aided by CODIS in the US (FBI, 2020).

STR location	Evidence genotype	Suspect 1 genotype	Suspect 2 genotype
STR 1	5, 8	5, 8	5, 8
STR 2	12, 13	12, 13	**14, 14**
STR 3	11, 12	11, 12	**12,** 12
STR 4	12, 12	12, 12	12, 12
STR 5	12, 18	12, 18	12, 18
	Result	**Match**	**No match**

Figure 3. Example STR profiles
Forensic scientists compare STR genotypes at multiple locations with variable STRs to determine whether DNA samples could have come from the same person. In this example, the STR genotypes show that suspect 1 could have been the source of the DNA evidence, but suspect 2's DNA does not match the evidence.

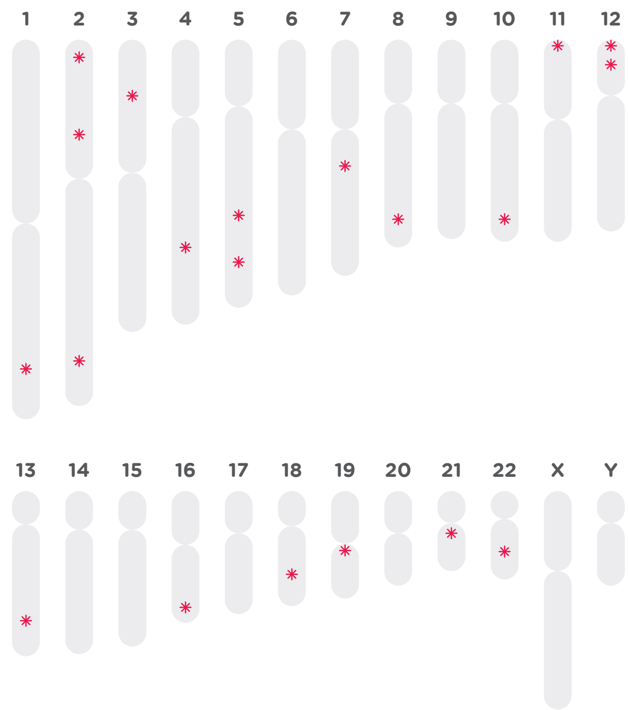
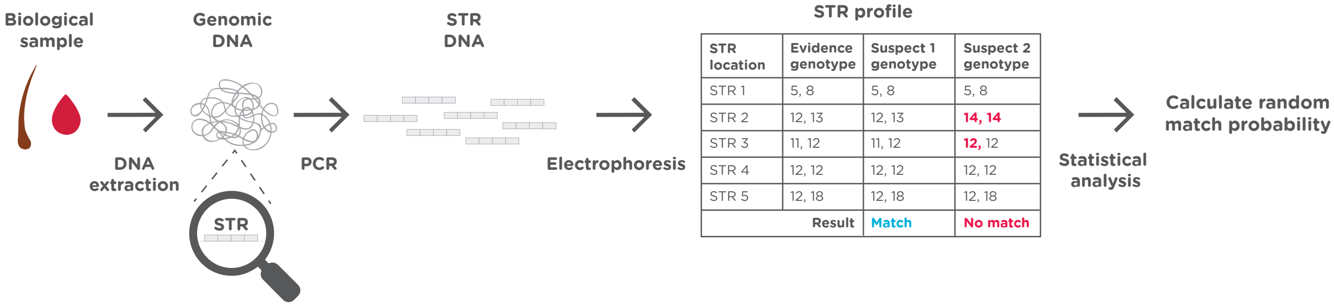


Figure 4. STR locations
Asterisks indicate the chromosomal locations of the 20 standard STR locations used in the FBI CODIS system.

FBI. "CODIS - NDIS Statistics." September 2020. https://www.fbi.gov/services/laboratory/biometric-analysis/codis/ndis-statistics. Accessed 11/20/2020.

Generating STR profiles

Forensic DNA analysis starts with a biological sample, such as a hair follicle or blood stain. Scientists break open the cells and extract the DNA for analysis (Figure 5). Luckily, forensic scientists only need a tiny amount of DNA for analysis. For example, crimes have been solved using DNA from a single eyelash or trace saliva on a cigarette butt. The key is a technique called polymerase chain reaction (PCR). PCR allows scientists to make many copies of specific regions of DNA, such as regions with a variable number of STRs.



STR location	Evidence genotype	Suspect 1 genotype	Suspect 2 genotype
STR 1	5, 8	5, 8	5, 8
STR 2	12, 13	12, 13	**14, 14**
STR 3	11, 12	11, 12	**12,** 12
STR 4	12, 12	12, 12	12, 12
STR 5	12, 18	12, 18	12, 18
	Result	Match	No match

Figure 5. Forensic DNA analysis workflow
DNA is extracted from a biological sample. Then, PCR is used to make many copies of just the locations in the genome with STRs that forensic scientists want to analyze. Electrophoresis techniques separate the copied STR DNA by length and allow forensic scientists to determine the STR genotypes. If there is a DNA match, then forensic scientists perform statistical calculations to determine the strength of the DNA evidence and the likelihood of getting a random match.

Recall that the vast majority of DNA sequences are the same in all humans and that forensic analysis focuses only on regions of DNA that tend to differ between people, like STRs. PCR allows forensic scientists to take a sample like blood that contains all of an individual's DNA and then copy just the 20 different STR locations they are interested in. After PCR, electrophoresis techniques determine the number of repeats present at each STR region (Figure 5).

Gel electrophoresis allows scientists to separate a mixture of DNA fragments based on length. Scientists use gels made of agarose, a polysaccharide extracted from seaweed, that forms a web-like mesh. An electric field causes the negatively charged DNA fragments to migrate through the gel towards the positive electrode. Smaller DNA fragments move easily through the agarose mesh and migrate quickly through the gel. Longer DNA fragments get caught in the agarose mesh and thus migrate more slowly through the gel.

Gel electrophoresis reveals "bands" of DNA fragments, with the shorter pieces of DNA being located further along the gel (Figure 6). By including a mixture of DNA fragments of known sizes, referred to as a DNA ladder, it is possible to calculate the size of unknown DNA fragments present in the samples.

Remember that STR alleles vary based on the number of repeats, and the DNA from an STR allele with more repeats will be longer than the DNA from an STR allele with fewer repeats (Figure 2). Gel electrophoresis lets scientists visualize these differences in length and differentiate between STR alleles. The gel shown in Figure 6 shows results for a single STR location. We can tell that the evidence DNA sample, suspect 1, and suspect 2 have two different alleles for this STR because the DNA samples show two bands of different sizes. On the other hand, we can tell that suspect 3 has two copies of the same allele for this STR because only one band appears on the gel.

When comparing crime scene DNA with DNA from suspects, the unique band patterns can quickly eliminate suspects whose STR profile does not match the evidence. For instance, in the gel shown in Figure 6, the DNA fragments of suspect 1 and suspect 3 do not match the evidence, so these suspects can be excluded as being the source of the DNA evidence. On the other hand, suspect 2's DNA bands do match the evidence, which means that suspect 2 cannot be excluded as being the source of the DNA evidence.

Figure 6. Interpreting gel electrophoresis results
This gel shows results examining a single STR location in the genome. By comparing the pattern of bands from the DNA evidence to the DNA of the suspects, forensic scientists can rule out suspects 1 and 3. However, suspect 2 matches the DNA evidence and cannot be excluded from further investigation.

It is crucial to keep in mind that forensic investigations typically examine at least 20 STR regions, and even if someone's complete STR profile matches the DNA evidence, it does not prove their guilt. If forensic scientists get a DNA match, then the last step of forensic DNA analysis is to calculate how common that particular STR profile is (Figure 5). The *random match probability* represents the chance that two unrelated individuals share the same STR profile. It takes into account the number of STR regions compared, the number of different alleles for each STR location, and how common each allele is in the general population. Modern STR analysis can yield genetic profiles that are statistically quite rare, on the order of one in many trillions. For more information on the statistics related to STR profiles, refer to the extension activity on probability and DNA profiles (page 25).

Remember though, that even if the statistical analysis very strongly suggests a DNA match, that still doesn't prove a suspect's guilt. It only supports that the suspect's DNA was present.

Today's lab

Only recently has DNA analysis become standard in criminal investigations. For cases that predate the use of DNA analysis, sometimes DNA can still be collected from physical evidence stored for an extended period. In many real-world law enforcement investigations, new DNA analysis of old evidence has cracked cold cases that were previously unsolvable. In other cases, new DNA analysis from closed cases has cleared wrongly convicted individuals. Today you will use gel electrophoresis to determine if DNA from an incarcerated individual matches newly extracted DNA from old crime scene evidence.

The case

In 1999, a young man was brutally attacked and strangled in an alley. The victim was found unconscious and severely injured, but he survived. The victim told the police that while the perpetrator had been wearing a ski mask, he was very tall and had a chipped front tooth. A ski mask found in a dumpster in the alley was collected as evidence.

Later that week, an officer spotted an individual, J.M., walking down the street near the crime scene. J.M. resembled the description of the perpetrator, and he was brought in for questioning. Although the victim had never seen the perpetrator's face because of the mask, he identified J.M. in a live lineup and testified in court that he was certain that J.M. was the attacker because of his distinctive chipped front tooth. In addition, several hairs were retrieved from the ski mask found in the alley. Microscopic analysis of the hair from the ski mask was said to match a hair sample collected from J.M. Based on this evidence, the jury found J.M. guilty.

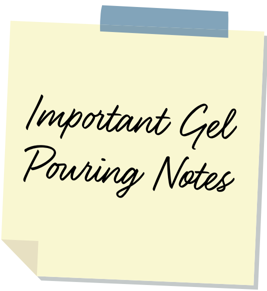
J.M. was convicted of attempted homicide and is currently serving a life sentence. J.M. has always maintained his innocence, and after years of appeals, the court has approved his request for DNA analysis of the evidence from his case. You have been sent DNA from four samples: a cheek swab from J.M., a cheek swab from the victim, and DNA from two hairs collected from the ski mask. The DNA you will test today represents the result of PCR amplification at a single STR location in the genome.

It's your job to ensure that the criminal justice system did not fail in this case. Can you prove J.M. to be innocent? Or will the evidence only implicate him further?

Laboratory guide

! Protective gloves and eyewear should be worn for the entirety of this experiment.

Gel electrophoresis - Pouring gels (before or during class period)



Gels can be prepared up to three days ahead of time and stored at ambient temperature, covered in air-tight plastic wrap and protected from light.

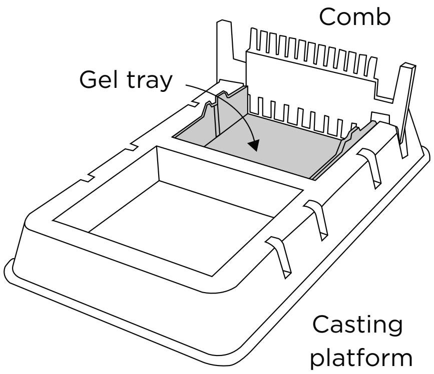
You will need 4 lanes plus one lane for ladder per group. If groups are sharing gels, a single lane for ladder is sufficient.

These instructions are designed for use with the blueGel™ electrophoresis system by miniPCR bio™. If using another electrophoresis system, these instructions may need to be adjusted according to the manufacturer's instructions.

1. **Prepare 1X TBE buffer (to be completed by teacher in advance)**
    - TBE buffer is often provided as liquid concentrate or powder.
    - Follow manufacturer's instructions to prepare 1X TBE buffer solution.

2. **Prepare a clean and dry casting platform with a gel tray and comb**
    - Place the clear gel tray in the white casting platform.
    - Place a well-forming comb at the top of the gel tray.

3. **Prepare a 2% agarose solution with a fluorescent DNA stain (e.g., SeeGreen™ or GelGreen®) using the method indicated by your instructor**



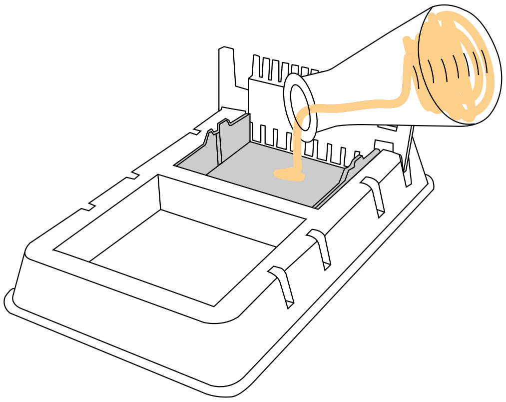
IMPORTANT NOTE: There are several ways to prepare agarose gels.

- Scan the QR code for detailed instructions on how to prepare agarose gels.
- Both written and video instructions are available.

www.minipcr.com/agarose-gel/

4. **Pour the agarose solution into the prepared casting platform with a gel tray and comb**
    - The agarose solution should cover the bottom of the gel tray and the bottom 3 mm of the comb (roughly the bottom 1/3 of the comb).

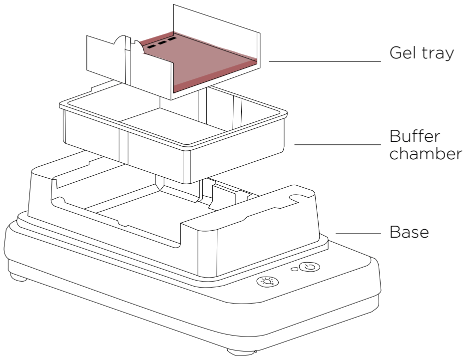
5. **Allow gel to solidify completely and remove the comb by pulling firmly upwards**
    - Gels will typically be ready in about 10 minutes.
    - Gel is ready when cool and firm to the touch.



Protective gloves and eyewear should be worn for the entirety of this experiment.

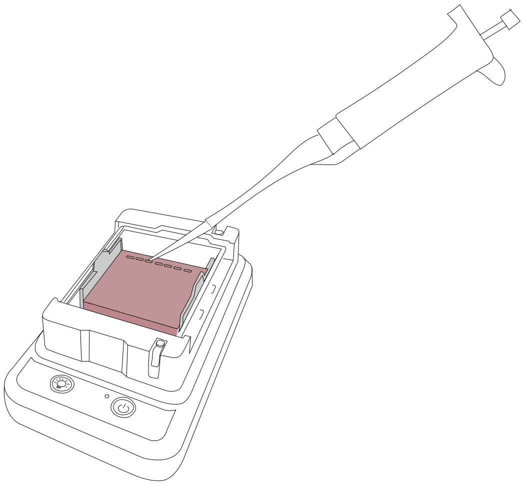
Gel electrophoresis - Running the gel

These instructions are designed for use with blueGel™ electrophoresis system by miniPCR bio™. If using another electrophoresis system, these instructions may need to be adjusted according to the manufacturer's instructions.



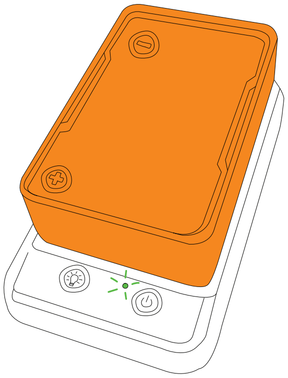
1. **Place the gel tray containing your gel in the buffer chamber**
   - Ensure that the clear buffer chamber is inside the blueGel™ electrophoresis system.
   - The wells of the gel should be on the same side as the negative electrode, away from the power button.

2. **Add 30 ml of 1X TBE electrophoresis buffer**
   - The buffer should just cover the gel and fill the wells.
   - Ensure that there are no air bubbles in the wells (shake the gel gently if bubbles need to be dislodged).



3. **Load samples onto the gel in the following sequence**
   - **Lane 1:** 10 µl Fast DNA Ladder 1
   - **Lane 2:** 10 µl victim DNA
   - **Lane 3:** 10 µl J.M. DNA
   - **Lane 4:** 10 µl DNA Evidence 1
   - **Lane 5:** 10 µl DNA Evidence 2

   Note: Samples already contain loading dye.

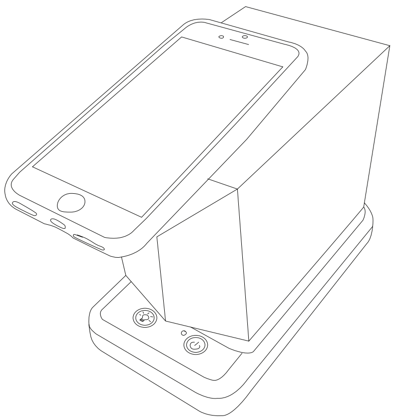


4. **Place the orange cover on the blueGel™ electrophoresis system**
   - To prevent fogging, make sure that ClearView™ spray has been evenly applied to the inside of the orange cover.
   - Match the positive and negative electrode signs on the orange lid with the corresponding positive and negative signs on the blue base.
   - The orange lid should sit flush with the blue base using little force.

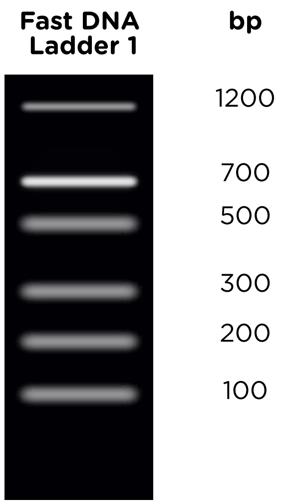
5. **Press the "Run" button**
    - Check that the green light beside the power button remains illuminated.

6. **Conduct electrophoresis for 15-25 minutes**
    - Note: Check the progress of your samples every 10 minutes to monitor the migration of your DNA samples.
    - Longer electrophoresis times will result in better separation of similar sized DNA fragments. However, if run too long, small DNA fragments can run off the end of the gel or lose fluorescence.

Gel electrophoresis – Visualizing results



1. **Press the "light bulb" button to turn on the blueGel™ transilluminator**
    - For best viewing, dim lights or use Fold-a-View™ photo documentation hood with a smartphone camera.
    - Gels may be viewed at the end of the run or periodically throughout the run.
    - If image appears hazy, wipe off the inside of the orange cover and reapply ClearView™ spray.



2. **Ensure that there is sufficient DNA band resolution in the 200-400 bp range of the Fast DNA Ladder 1**
    - Run the gel longer if needed to increase resolution.

3. **Document your results**
    - Compare the bands from the DNA samples to the ladder to obtain size estimates.
    - Place Fold-a-View™ photo documentation hood on the blueGel™ electrophoresis system to take a picture with a smartphone or other digital camera.

Study questions - pre-lab

Review

1. **Why is it essential to compare variable regions of DNA if you are trying to identify someone?**

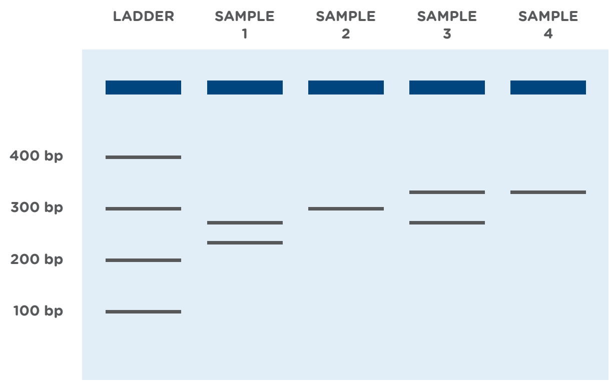
2. **Describe how you could recognize a short tandem repeat within a DNA sequence.**

3. **How do different alleles for the same STR location vary?**

4. **Why is gel electrophoresis a good tool to differentiate between alleles for the same STR region?**

Critical thinking

5. Gel electrophoresis was used to separate fragments for a single STR location, which we will call "STR 1" for simplicity.



   A. Every person has two alleles for STR 1. Which sample(s) comes from a person where the two alleles are different from each other? Explain how you can tell.

   B. Which sample(s) comes from a person where the two alleles are the same? Explain how you can tell.

   C. Do any of the samples come from people who have the same genotype? Explain how you can tell.

   D. Do any of the samples have <u>alleles</u> that are the same? If so, circle them on the gel and explain how you can tell.

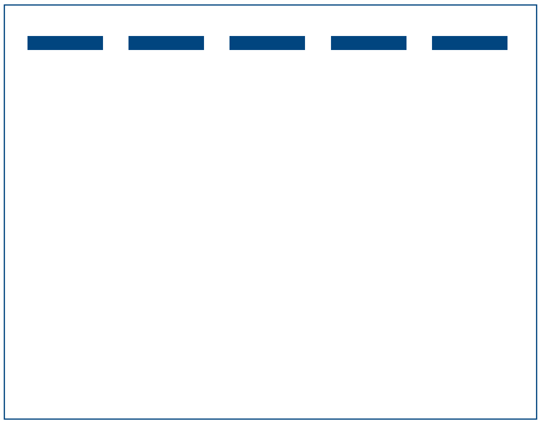
6. **When CODIS was first established it looked at 13 STR regions. Investigators have since expanded it to include 20 locations. Why do you think investigators wanted to compare more locations when looking for DNA matches?**

7. **What can you conclude if a suspect's STR profile does *not* match DNA evidence from the crime scene at a single STR location?**

8. **What can you conclude if a suspect's STR profile matches DNA evidence from the crime scene at all 20 STR locations?**

Study questions - post-lab

Interpreting results



1. Use the image on the right to illustrate your gel electrophoresis results. Be sure to label each lane with the source of the sample.

2. Compare J.M.'s DNA with the DNA evidence from the crime scene. What can you conclude?

Critical thinking

3. Why do you think it is helpful to analyze the victim's DNA if a sample is available?

4. If you were working on this case, what would your next steps be? Explain your reasoning.

CER Table

Fill in the table based on your results from the lab. Use the rubric on the next page to help your answers.

Question:

Based on your results, do you think J.M. is guilty?

Claim Make a clear statement that answers the above question.	
Evidence Provide data from the lab that supports your claim.	
Reasoning Explain clearly why the data you presented supports your claim. Include the underlying scientific principles that link your evidence to your claim.	

Score	4	3	2	1
CLAIM A statement that answers the original question/ problem.	Makes a clear, accurate, and complete claim.	Makes an accurate and complete claim.	Makes an accurate but incomplete or vague claim.	Makes a claim that is inaccurate.
EVIDENCE Data from the experiment that supports the claim. Data must be relevant and sufficient to support the claim.	All of the evidence presented is highly relevant and clearly sufficient to support the claim.	Provides evidence that is relevant and sufficient to support the claim.	Provides relevant but insufficient evidence to support the claim. May include some non-relevant evidence.	Only provides evidence that does not support claim.
REASONING Explain why your evidence supports your claim. This must include scientific principles/ knowledge that you have about the topic to show why the data counts as evidence.	Provides reasoning that clearly links the evidence to the claim. Relevant scientific principles are well integrated in the reasoning.	Provides reasoning that links the evidence to the claim. Relevant scientific principles are discussed.	Provides reasoning that links the evidence to the claim, but does not include relevant scientific principles or uses them incorrectly.	Provides reasoning that does not link the evidence to the claim. Does not include relevant scientific principles or uses them incorrectly.

Rubric score	3	4	5	6	7	8	9	10	11	12
Equivalent Grade	55	60	65	70	75	80	85	90	95	100

We recommend that teachers use the following scale when assessing this assignment using the rubric. Teachers should feel free to adjust this scale to their expectations.

Extension: Probability and DNA profiles

Probability and DNA profiles

Using the product rule in forensics

If you flip a coin once, there is a 50% chance that it will land heads up. In fact, any time you flip a coin, the likelihood of getting a "heads" on that specific toss is 50% regardless of whether the coin landed heads or tails on the previous tosses. We can calculate the probability of getting any combination of heads and tails over a series of tosses using the product rule. The product rule calculates the probability of a series of independent events by multiplying each event's probability. For example, if you want to know the probability of getting 3 "heads" in a row, you multiply 0.5 x 0.5 x 0.5 to get 0.125 or 12.5%.

The product rule can also estimate the probability of a given STR profile in the human population (Figure 1). Once forensic scientists have an STR profile with the genotypes for each STR location, they can calculate the *random match probability*. The random match probability estimates the frequency at which a given STR profile occurs in a population. This is a useful metric because it is the same as the probability that a randomly selected person would share the genetic profile being analyzed. To calculate the random match probability using the product rule, forensic scientists multiply the frequency of each STR genotype in the profile (Figure 1). The genotype frequency represents how prevalent the genotype is in the overall population.

STR location	Genotype	Genotype frequency
STR 1	7, 8	0.0054
STR 2	13, 13	0.0764
STR 3	11, 12	0.1772

Random match probability = genotype frequency STR 1 x genotype frequency STR 2 x genotype frequency STR 3

= 0.0054 x 0.0764 x 0.1772

= 7.31×10^{-5}

Figure 1. Calculating random match probability using the product rule

While you use the product rule for a series of coin tosses and STR profiles, STR genotypes are more complicated than a coin flip because there are more than two alleles for each STR location. Further, some STR alleles are quite common, while others are exceedingly rare. STR genotype frequencies range from more than 20% to less than 0.00003%. That means some STR genotypes are so common they are found in more than one in five people, whereas other genotypes are so rare they are found in fewer than one in thirty million people!

Interpreting random match probabilities

Multiplying the genotype frequencies for all 20 STR locations used in CODIS using the product rule gives random match probabilities that can be quite small, like one in many trillions. However, there are instances where the random match probability can be much larger. For example, if someone has several STR genotypes that are more common in the general population, then the chance of a random match will be more likely, though typically still very rare. Further, sometimes DNA evidence degrades before it is collected. In these cases, it might not be possible to analyze as many STRs. And with fewer STRs, the chance of getting a random match becomes more likely.

It is essential to note that the random match probability only represents the rarity of a specific genetic profile. Unfortunately, random match probabilities have been misinterpreted in many legal settings. For example, given a random match probability of 1 in 5 trillion, a prosecutor might claim that there is only "a 1 in 5 trillion chance that the defendant is innocent." This error in reasoning is referred to as the prosecutor's fallacy. It ignores additional factors that might make matching the DNA profile at the crime scene more or less likely. For example, the person may have been at the crime scene for reasons unrelated to the crime.

Naming of STR locations

Figure 2. STR naming conventions

Before, we referred to STR locations with simplified names like STR 1 and STR 2. In reality, forensic scientists use a complex naming system for regions of the genome with STRs. In the following questions, we will use the actual STR names because you will be using real data on the frequencies of the STR genotypes in the human population.

Most of the STR locations used in forensic analysis are in the regions of the genome that do not contain genes. These STR regions are named using a standardized system (Figure 2). Each STR name starts with a D for "DNA," followed by a number that tells you which chromosome this STR is on. Then there is an "S" followed by another number. This second number is needed because there are many different STR regions on each chromosome, so each one needs to have a unique identifier.

Study questions

1. **D7S820 can have between 5 and 16 short tandem repeats.**
   A. What is the maximum number of alleles possible for this STR?

   B. What is the maximum number of genotypes for D7S820 where both alleles contain the same number of repeats?

   C. There are 66 genotypes for D7S820 where the two alleles do not contain the same number of repeats. Give two different examples of such genotypes.

Forensic scientists use the frequency of each STR allele to determine the frequency of each STR genotype using the formulas below:

For genotypes with two copies of the same allele
(ex: “5, 5” or “11, 11”)
genotype frequency = p^2
where p = the frequency of the allele

For genotypes with different alleles (ex: “13, 14” or “6, 12”)
genotype frequency = 2pq
p = the frequency of 1st allele
q = the frequency of the 2nd allele

D7S820 alleles	Allele frequency
7 repeats	.0164
8 repeats	.1655
9 repeats	.1216
10 repeats	.2949

Table 1. Frequencies for select D7S820 alleles
National Institute of Standards and Technology. “1036 Revised U.S. Population” July 2017. Dataset. https://strbase.nist.gov/NISTpop.htm. Accessed 7/29/2020.

Table 1 shows the allele frequencies for four D7S820 alleles. Scientists have compiled STR profiles from many individuals selected to represent the entire human population and have used this information to calculate each allele’s frequency. With this information, scientists can calculate the estimated frequency of every STR genotype.

2. **Use the information in Table 1 to calculate genotype frequencies using the formulas above. Show your work.**
   - **A. D7S820 genotype 7, 7**
   - **B. D7S820 genotype 7, 9**
   - **C. D7S820 genotype 10, 10**
   - **D. D7S820 genotype 8, 9**

In the previous questions, you focused on a single STR location, but remember that forensics scientists analyze many STR regions. Table 2 shows an STR profile that contains seven STR locations. In this table, the genotype frequencies, like those you calculated in the previous problems, have been calculated for you.

STR location	Genotype	Genotype frequency
D2S441	11, 14	0.1544
D5S818	11, 12	0.2230
D7S820	7, 8	0.0054
D10S1248	13, 13	0.0764
D13S317	11, 12	0.1772
D16S539	8, 14	.00092
D18S51	20, 21	.00069

Table 2. Genotype frequencies for select STR locations
National Institute of Standards and Technology. "1036 Revised U.S. Population" July 2017. Dataset. https://strbase.nist.gov/NISTpop.htm. Accessed 7/29/2020.

3. **Based on the genotype frequencies in Table 2:**

   A. **For which STR location(s) does this individual have a common genotype shared by more than 20% of the population?**

   B. **For which STR location(s) does this individual have a genotype so rare that less than 0.1% of the population share it?**

4. **Imagine you are testing a crime scene DNA sample. While it is standard for forensic scientists to examine at least 20 STR locations to create a DNA profile, this has not always been the case.**

   A. **Multiply the genotype frequencies in Table 2 to calculate the random match probability for this STR profile using the product rule. Show your work.**

B. Scientific notation is often used with very small numbers. Compare the random match probability that you just calculated with the scientific notation to the right. What order of magnitude is the random match probability?
    i. On the order of one in a million
    ii. On the order of one in a billion
    iii. On the order of one in a trillion

Scientific notation	In fraction form	In words
1E-6	1/1,000,000	one in a million
1E-9	1/1,000,000,000	one in a billion
1E-12	1/1,000,000,000,000	one in a trillion

Table 3. Interpreting scientific notation

C. How confident would you be in using this DNA sample to convict someone of a crime?

5. Imagine that your DNA evidence was exposed to the elements for an extended time. Because the DNA is degraded, you only get results for the first two STR locations in the table above (D2S441 and D5S818).
    A. What would the random match probability be then? Show your work.

    B. The current US population is 331,000,000 (331 million). How many individuals in the US are predicted to match your evidence DNA profile? Show your work.

    C. How confident would you be in using this DNA sample to convict someone of a crime?

6. Now assume you only get results for the last 2 STR locations in the table above (D16S539 and D18S51)
    A. What would the random match probability be then? Show your work.

    B. The current US population is 331,000,000 (331 million). How many individuals in the US are predicted to match your evidence DNA profile? Show your work.

    C. How confident would you be in using this DNA sample to convict someone of a crime?

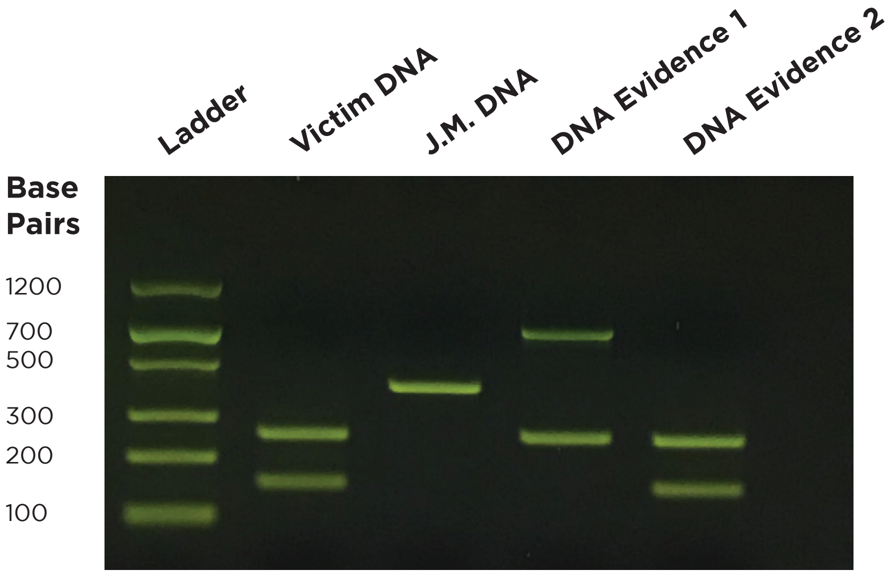
7. Explain why the random match probabilities in questions 5 and 6 are so different even though they were both calculated using two STR locations.

8. Brainstorm at least one example where you think two people are more likely to have a similar STR profile than two randomly selected individuals. Explain your reasoning.

Instructor's Guide

Expected results

Gel electrophoresis results are expected to resemble the gel image below.



- Victim DNA: should contain 2 DNA fragments - 250 bp and 150 bp
- J.M. DNA: should contain 1 DNA fragment - 400 bp
- DNA evidence 1: should contain 2 DNA fragments - 700 bp and 250 bp
- DNA evidence 2: should contain 2 DNA fragments - 250 bp and 150 bp

Interpretation

- J.M.'s DNA does not match either evidence sample. Even though only one STR was analyzed, J.M. can definitively be excluded as the source of the DNA evidence collected from the crime scene.
- The victim's DNA matches DNA evidence 2. This suggests that the second evidence sample may have come from the victim. It is not uncommon for there to be ample victim DNA after a violent crime. To be more confident of this hypothesis, more STR loci would need to be examined to generate a comprehensive STR profile.
- DNA evidence 1 represents an unknown individual since neither the victim nor J.M. has a 700 bp STR fragment. A reasonable next step in the investigation would be to search for a suspect that matches this DNA sample. If more STR loci were examined to generate a complete evidence STR profile, then a CODIS database search could be performed.

Notes on lab design

This lab serves as an introduction to forensic DNA analysis. While many forensic labs focus on using DNA to convict someone of a crime, we believe it is equally important to consider how DNA evidence can exonerate someone who was wrongly convicted. While the scenario presented in this lab is fictional, it is inspired by real cases where people have been wrongly convicted and later proven innocent through post-conviction DNA analysis.

This lab introduces the examination of short tandem repeats to generate DNA profiles. We believe our approach provides the right balance between intellectual engagement, inquiry, and accessibility. The design of this lab has simplified certain elements to achieve these goals.

- This lab uses prepared DNA to simulate the results of PCR amplification of STRs from DNA evidence.
- While forensic STR analysis typically examines at least 20 STR loci, this lab examines only one to allow for clear interpretation of the results.
- Forensic scientists use capillary electrophoresis to separate STR DNA fragments, but this technique requires expensive and specialized equipment. We substitute gel electrophoresis, which separates DNA fragments using the same principles as capillary electrophoresis.
- To allow for clear resolution of STR alleles by gel electrophoresis, the range of DNA fragments used is greater than what is observed for a given STR locus. Most STR loci have a four nucleotide repeat sequence and around 10 alleles. This means the size difference between the smallest allele and the largest allele is typically only around 40 bases. We have used size differences in the hundreds of bases as these are easily resolvable on an agarose gel.

Before carrying out this lab, students should have basic competence using a micropipette, and should understand the concept of gel electrophoresis. See the **Additional Student Supports** section of this lab for ways to scaffold this assignment for students who may be less comfortable with the aforementioned skills.

Differentiation

This lab serves as an introduction to forensic DNA analysis. With simple modifications, this activity can be used effectively in classes ranging from middle school through advanced high school.

Introductory classes: Focus on how gel electrophoresis can be used to differentiate STR profiles for genetic identification. The standard introduction and review questions for this lab take this approach.

Advanced classes: Use the extension *Probability and DNA profiles* (page 25) to discuss population genetics and how statistical analysis is used to determine the strength of a given DNA profile for personal identification.

Additional student supports

At miniPCR bio™, we are committed to preparing students to be successful in the laboratory through high quality curriculum and training. We have created an extensive set of resources to help your students succeed in molecular biology techniques, all of which are available for free download at the miniPCR bio™ tutorials page of our website.

https://www.minipcr.com/tutorials/

Activities most relevant to this lab are listed below.

Micropipetting: Video and activity resources to train students in the basic use of a micropipette.

Gel electrophoresis: Video and worksheet activity instructing students on the fundamentals and practice of agarose gel electrophoresis.

PCR: While students do not perform PCR in this lab, the samples they analyze represent PCR products. If you want to discuss PCR in more detail with your students, we have a video and worksheet activity instructing students on the fundamentals and practice of PCR.

Extension activities

The following optional extension activities are provided for students to explore topics more deeply.

Probability and DNA profiles (page 25): Getting a DNA match is not the last step in a forensic investigation. Certain STR genotypes are more common, so forensic scientists use statistics to help determine the strength of the DNA evidence. This extension introduces basic probability calculations used by forensic scientists when analyzing DNA profiles and can serve as an entry point to discuss population genetics.

DNA fingerprinting: This lab gives students an introduction to the STR-based DNA analysis currently used in forensic science. Learn more about the past and future of DNA analysis in forensics. Link includes article and classroom questions.

https://dnadots.minipcr.com/dnadots/dna-fingerprinting

miniPCR in Forensics: See how miniPCR® machines are used in actual forensic investigation. Instead of bringing evidence to the lab, some forensic scientists are bringing miniPCR® machines to the evidence, reducing the risk of contamination or evidence mishandling.

https://www.minipcr.com/case-studies/forensic-dna-analysis-evidence/

Placement in unit

Forensics
Use this lab as an introduction to how forensic scientists use STR analysis. Once students have mastered their molecular biology skills with this lab, continue with the miniPCR Forensics Lab (https://www.minipcr.com/product/minipcr-forensics-lab-d1s80-vntr/), which allows students to sample their own DNA to solve a forensic mystery using PCR and gel electrophoresis.

Genetics
This lab offers a hands-on way to gain familiarity with population genetics and alleles using the extension activity *Probability and DNA profiles* (page 25).

Biotechnology
This lab can be used as a culminating activity to demonstrate mastery of micropipetting and gel electrophoresis. For this approach, we recommend spending more time on some of the activities discussed in the **Additional student supports** section of this lab (page 36) before beginning the experiment.

Learning goals and skills developed

Student Learning Goals:

- Describe the presence of variable number STRs in the human genome
- Understand the use of STR profiles for identification in forensics
- Apply basic probability rules to forensic DNA analysis
- Analyze DNA profiles to solve real-world scenarios

Scientific Inquiry Skills:

- Students will use experimental results to make conclusions and solve a real-world problem
- Students will follow detailed experimental protocols
- Students will make a claim based in scientific evidence
- Students will use reasoning to justify a scientific claim
- Students will follow laboratory safety protocols

Molecular Biology Skills:

- Micropipetting
- Preparation of agarose gels
- Agarose gel DNA electrophoresis
- Staining, visualization, and molecular weight analysis of DNA fragments

Standards alignment

Next Generation Science Standards

Students who demonstrate understanding can:

HS-LS3-3.	Apply concepts of statistics and probability to explain the variation and distribution of expressed traits in a population.

Science and Engineering Practice	Disciplinary Core Ideas	Crosscutting Concepts
• Asking Questions and Defining Problems • Developing and Using Models • Planning and Carrying Out Investigations • Analyzing and Interpreting Data • Using Mathematics and Computational Thinking • Constructing Explanations and Designing Solutions • Engaging in Argument from Evidence • Obtaining, Evaluating, and Communicating Information	LS1.A: From Molecules to Organisms: Structures and Processes LS3.A: Inheritance of Traits	• Patterns • Cause and Effect • Systems and System Models • Structure and Function • Stability and Change • Interdependence of Science, Engineering, and Technology • Influence of Engineering, Technology, and Science on Society and the Natural World

Common Core ELA/Literacy Standards

RST.9-10.1	Cite specific textual evidence to support analysis of science and technical texts, attending to the precise details of explanations or descriptions.
RST.9-10.3	Follow precisely a complex multistep procedure when carrying out experiments, taking measurements, or performing technical tasks, attending to special cases or exceptions defined in the text.
RST.9-10.4	Determine the meaning of symbols, key terms, and other domain-specific words and phrases as they are used in a specific scientific or technical context relevant to grades 9-10 texts and topics.
RST.9-10.5	Analyze the structure of the relationships among concepts in a text, including relationships among key terms (e.g., force, friction, reaction force, energy).
RST.9-10.9	Compare and contrast findings presented in a text to those from other sources (including their own experiments), noting when the findings support or contradict previous explanations or accounts.
WHST.9-10.1	Write arguments focused on discipline-specific content.
WHST.9-10.2	Write informative/explanatory texts, including the narration of historical events, scientific procedures/ experiments, or technical processes.
WHST.9-10.9	Draw evidence from informational texts to support analysis, reflection, and research.

* This activity has been aligned to high school NGSS and grades 9-10 Common Core standards. For information aligning this activity to middle school or other grade levels, please contact support@minipcr.com.

Ordering information

To order miniPCR™ Electrophoresis Forensics Lab kits, you can:

Call (781)-990-8PCR

email us at orders@minipcr.com

visit https://www.minipcr.com

miniPCR™ Electrophoresis Forensics Lab (catalog no. KT-1504-01) contains the following reagents:

- Victim DNA
- J.M. DNA
- DNA evidence 1
- DNA evidence 2
- Fast DNA Ladder 1

Materials are sufficient for 8 lab groups, or 32 students

All components should be kept frozen at -20°C for long-term storage

Reagents must be used within 12 months of shipment

Other reagents needed

- Agarose (electrophoresis grade)
- Fluorescent DNA stain (e.g., SeeGreen™ or GelGreen®)
- Gel electrophoresis buffer (e.g., 1X TBE)
- Distilled or deionized water (to dilute TBE buffer concentrate)

Note: Agarose, DNA stain, and TBE buffer are available at minipcr.com as part of the Learning Lab Companion Kit (KT-1510-01)

About miniPCR bio Learning Labs™

This Learning Lab was developed by the miniPCR bio™ team in an effort to help more students understand concepts in molecular biology and to gain hands-on experience in real biology and biotechnology experimentation.

We believe, based on our direct involvement working in educational settings, that it is possible for these experiences to have a real impact in students' lives. Our goal is to increase everyone's love of DNA science, scientific inquiry, and STEM. We develop Learning Labs™ to help achieve these goals, working closely with educators, students, academic researchers, and others committed to science education.

The guiding premise for this lab is that a 45-minute electrophoresis-based experiment can recapitulate a real-life forensics application and provide the right balance between intellectual engagement, inquiry, and discussion.

Starting on a modest scale working with Massachusetts public schools, miniPCR bio Learning Labs™ have been well received, and their use is growing rapidly through academic and outreach collaborations across the world.