

LAB MANUAL: MOLECULAR BIOLOGY AND GENETIC ENGINEERING (BIOT 275)



Victor Pham
Irvine Valley College

Lab manual: Molecular biology and Genetic
Engineering (Biot275 @ IVC)

This text is disseminated via the Open Education Resource (OER) LibreTexts Project (<https://LibreTexts.org>) and like the thousands of other texts available within this powerful platform, it is freely available for reading, printing, and "consuming."

The LibreTexts mission is to bring together students, faculty, and scholars in a collaborative effort to provide an accessible, and comprehensive platform that empowers our community to develop, curate, adapt, and adopt openly licensed resources and technologies; through these efforts we can reduce the financial burden born from traditional educational resource costs, ensuring education is more accessible for students and communities worldwide.

Most, but not all, pages in the library have licenses that may allow individuals to make changes, save, and print this book. Carefully consult the applicable license(s) before pursuing such effects. Instructors can adopt existing LibreTexts texts or Remix them to quickly build course-specific resources to meet the needs of their students. Unlike traditional textbooks, LibreTexts' web based origins allow powerful integration of advanced features and new technologies to support learning.



LibreTexts is the adaptable, user-friendly non-profit open education resource platform that educators trust for creating, customizing, and sharing accessible, interactive textbooks, adaptive homework, and ancillary materials. We collaborate with individuals and organizations to champion open education initiatives, support institutional publishing programs, drive curriculum development projects, and more.

The LibreTexts libraries are Powered by [NICE CXone Expert](#) and was supported by the Department of Education Open Textbook Pilot Project, the California Education Learning Lab, the UC Davis Office of the Provost, the UC Davis Library, the California State University Affordable Learning Solutions Program, and Merlot. This material is based upon work supported by the National Science Foundation under Grant No. 1246120, 1525057, and 1413739.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation nor the US Department of Education.

Have questions or comments? For information about adoptions or adaptations contact info@LibreTexts.org or visit our main website at <https://LibreTexts.org>.

This text was compiled on 05/03/2026

TABLE OF CONTENTS

Licensing

Activity 1-1 – Data on Human Traits

Activity 1-2 - Human Genetic Inheritance and Diseases

Activity 1-3 - Genetic evolution and Identifying Homologs

Activity 1-4 - PowerPoint Presentation on the Bioinformatic Evolution and function of proteins

Activity 2-0 - Introduction to Plasmid and Enzyme p450

Activity 2-1 - Plasmids Isolation

Activity 2-2 - DNA Analysis via Restriction Enzymes

Project 2-3 - Identifying Plasmid Restriction Mapping Using Agarose Gel Electrophoresis

Activity 2-4 - Bacterial Transformation with Engineered Plasmid

Activity 3-1 - Identifying Genetically Modified Organisms Through DNA Analysis

Activity 3-2 - Primer Design and Barcoding via Bioinformatics

Activity 4-1 - RNA Extraction via Guanidine Thiocyanate and Phenol

Activity 4-2 - Reverse Transcription Polymerase Chain Reaction

Activity 4-3 - Real-Time Quantitative Polymerase Chain Reaction

Activity 4-4 - DNA Sequencing with SeqStudio Genetic Analyzer

Index

Glossary

Detailed Licensing

Detailed Licensing

Licensing

A detailed breakdown of this resource's licensing can be found in [Back Matter/Detailed Licensing](#).

Activity 1-1 – Data on Human Traits

Learning Objectives

- Define and distinguish between **genotype** and **phenotype**.
- Explain how **inheritance patterns** work (autosomal dominant, autosomal recessive, sex-linked).
- Identify common **Mendelian traits** in yourself and others.
- Describe the relationship between **genetic traits**, **environment**, and **phenotypic expression**.
- Use the **OMIM** and **NCBI** databases to explore genetic traits and their associated genes.
- Reflect on the influence of **sexual selection** and **evolutionary preference** on trait frequency.

Definition: Term

- **Genotype:** The genetic makeup (alleles) of an organism, inherited from both parents.
- **Phenotype:** The observable traits or characteristics resulting from the genotype and environment.
- **Allele:** One of the different forms of a gene (e.g., dominant or recessive).
- **Autosomal Dominant:** A trait that requires only one copy of the dominant allele to be expressed.
- **Autosomal Recessive:** A trait that is only expressed if both alleles are recessive.
- **Sex-linked Trait:** A trait associated with genes located on the X or Y chromosome.
- **Carrier:** A person who has one recessive allele but does not show the trait.
- **OMIM:** A genetic disorder database (Online Mendelian Inheritance in Man).
- **NCBI:** National Center for Biotechnology Information – a tool for exploring genes and proteins.

Note

- Review the difference between **dominant and recessive inheritance**.
- Think about common traits you observe in yourself or your family.
- Skim through the **OMIM.org** and **NCBI Gene** websites.
- Come prepared to observe, record, and discuss physical traits respectfully and scientifically.

Introduction to Genotype, Phenotype, and Inheritance

Phenotype refers to the observable traits or characteristics displayed by an organism—such as eye color, blood type, or the ability to roll one's tongue. **Genotype**, on the other hand, is the underlying genetic makeup that determines these traits. In essence, the phenotype is the product of both genotype and environmental interactions. For example, a person may have the genotype for tall stature, but poor nutrition during growth can result in shorter height—a clear demonstration of gene-environment interplay. In this lab, students will focus on traits that are primarily influenced by genetics to observe inheritance patterns.

Phenotype can be thought of as the "**performance on stage**", while **genotype** is the "**script written backstage**." Just like actors follow a script (with room for improvisation), organisms express physical and behavioral traits based on their genetic instructions—but those traits may also be influenced by the environment. For instance, let's imagine two genetically identical seeds (same genotype). If one is planted in nutrient-rich soil with ample sunlight and water, while the other is planted in poor soil and receives little light, they will grow into very different plants—**demonstrating different phenotypes** despite sharing the same genotype. This is similar to how a child genetically predisposed to be tall may not reach their full height potential if malnourished during critical developmental years.

This interaction between genes and the environment is central to fields such as **personalized medicine**, where doctors tailor treatments based on a patient's genotype while also considering lifestyle and environmental factors. For example, some individuals metabolize drugs like warfarin or codeine differently due to genetic variants in liver enzymes—an application of genotype-phenotype analysis in clinical settings.

The primary goal of this lab is to help students understand how certain human traits are inherited and to introduce them to bioinformatics tools used to study human genetic disorders. The activities are designed to build practical skills in identifying patterns of inheritance—such as autosomal dominant, autosomal recessive, and sex-linked traits—and to explore the real-world utility of online genetic databases like **NCBI** and **OMIM**. Ultimately, students will not only analyze their own traits but also research genetic diseases, obtain sequence data, and use **BLAST** to identify homologous sequences, deepening their understanding of genetics and molecular biology.

Modes of Inheritance: Autosomal and Sex-Linked Traits

Traits in humans are inherited according to well-established patterns, depending on where the gene responsible for a trait is located and whether the allele is dominant or recessive. These patterns help us predict how specific characteristics or genetic disorders may be passed from one generation to the next. The two major categories of inheritance that we will focus on are **autosomal inheritance** and **sex-linked inheritance**.

- **Autosomal dominant traits** are caused by genes located on the autosomes, which are the 22 pairs of chromosomes that do not determine sex. A trait is considered dominant if just one copy of the dominant allele—received from either parent—is enough to express the trait. This means that a person only needs one "strong" version of the gene for the characteristic to be visible. For example, having a **widow's peak** (a V-shaped point in the hairline) is often cited as an autosomal dominant trait. If a child inherits the dominant allele from just one parent, they are likely to show the trait, regardless of the allele they inherit from the other parent.
- **Autosomal recessive traits** also involve genes on the autosomes, but in this case, both alleles must be the recessive form for the trait to be expressed. If an individual has only one copy of the recessive allele, they are considered a **carrier** and typically do not show the trait themselves. The trait is only observed when the person inherits the recessive allele from **both** parents. A classic example is **attached earlobes**, which are commonly used in educational settings to demonstrate autosomal recessive inheritance. If both parents are carriers, there is a 25% chance that their child will inherit two copies of the recessive allele and display the trait.
- **Sex-linked traits** are associated with genes located on the sex chromosomes, particularly the X chromosome. Since females have two X chromosomes (XX) and males have one X and one Y (XY), inheritance of sex-linked traits often differs between the sexes. **X-linked recessive traits**, such as **red-green color blindness**, are more commonly expressed in males. This is because males only have one X chromosome, so a single recessive allele on that X is enough for the trait to appear—there is no second X to potentially mask the effect. In contrast, females must inherit two copies of the recessive allele (one from each parent) to express the trait, which is less likely. **X-linked dominant traits** can affect both males and females but are often more easily transmitted from mothers to their children since mothers contribute one of their two X chromosomes to each child. The Y chromosome carries very few genes, so **Y-linked inheritance** is rare and usually involves traits passed strictly from father to son.

Understanding these different modes of inheritance—autosomal dominant, autosomal recessive, and sex-linked—is essential for predicting how traits may appear in families and populations. More importantly, these principles form the foundation for analyzing and diagnosing **genetic disorders**, many of which follow these same classic **Mendelian patterns**. Mastering this knowledge will help you not only understand your own inherited traits, but also begin to explore the genetic basis of human health and disease.

Activity 1-1: Collecting Data on Human Traits

This activity introduces students to foundational concepts in genetics through direct, observational data collection. Students will investigate common **human phenotypic traits**—physical characteristics that result from their underlying **genotype** (genetic makeup). These traits are examples of **Mendelian inheritance**, where a single gene controls the expression of a trait and follows predictable patterns (dominant vs. recessive).

By gathering data on themselves and their peers, students will begin to understand how often specific alleles appear in a small "population" (their class). For example, **tongue rolling** is often cited as a dominant trait. If 70% of the class can roll their tongue, this suggests a higher frequency of the dominant allele in this mini-population. However, students will also observe traits like **attached vs. free earlobes**, **dimples**, or **widow's peak**—traits typically used in genetics education to model **autosomal**

inheritance patterns. This activity is not only about counting traits; it's about **visualizing gene frequency**, **inferring genotypes**, and exploring what our genes say about us—even in a small population.

More importantly, it opens a fascinating question: *Do the most common traits in a population match the traits we find most attractive or desirable?* From an evolutionary and anthropological perspective, physical attractiveness is shaped by what traits were perceived as favorable for mating. Yet, this doesn't always align with dominant traits. For example, **blue eyes** are recessive but are often considered attractive in some cultures. Similarly, **dimples** (dominant) and **freckles** (sometimes recessive) are sometimes romanticized features. Students will reflect on whether the traits most commonly found in the class are also those considered most "attractive" or "desirable," and what this might say about **sexual selection** vs. **natural selection**. Imagine we're studying a population to identify traits associated with perceived **attractiveness**. Evolution isn't just about survival—it's also about reproduction. In this light, the traits that become more common over generations may not only be "dominant" genetically, but also **selected** because of attractiveness or social preference.

Ask yourself:

- Are traits like **dimples**, **widow's peak**, or **eye color** more common because they're dominant or because they're seen as attractive?
- Are there **recessive traits** that are still prevalent because they're desirable (e.g., blue eyes)?
- If we view our class as a sample population, what traits do we find attractive—even if they're not the most genetically dominant?

This exercise encourages you to reflect on how **sexual selection** (mate preference) may shape populations over time, sometimes in ways that contrast with pure Mendelian expectations. Dominance does not always equal desirability.

Exploring Human Traits and Genetic Inheritance in Our Classroom Population

Objective:

To observe and analyze the distribution of common inherited traits within our class. We will also explore how dominant and recessive traits appear in a population and what these patterns may reveal about attractiveness and evolutionary selection.

Instructions:

1. **Trait Type Key:**
 - **BOLD** Traits are **dominant**.
 - *Italicized* Traits are **recessive**.
2. **Observe Yourself First:** Examine your own physical traits listed in **Table 1-1**. Use a mirror or ask a partner to help. Record your **phenotype** (what you observe) and try to **infer your genotype** (your possible allele combinations).
3. **Survey Your Classmates:** In small groups, observe and record the same traits in every member of your class. You can ask classmates or visually confirm traits when appropriate and respectful. Tally how many classmates express each trait.
4. **Use Table 1-1 to Record Your Data:** For each trait, write:
 - A **checkmark** in the column applies to your phenotype.
 - Your inferred genotype (based on your phenotype and what you know about inheritance).
 - The number of classmates displaying that same trait.

Explore Further with OMIM: For each trait, visit [OMIM.org](https://omim.org) and enter the OMIM number to learn more about the associated gene, its chromosomal location, and any relevant genetic disorders.

Table 1-1: Genetic Trait Survey

Name of Trait	Description	OMIM Number	Your Phenotype	Your Genotype	# of Classmates With Trait
Earlobes - Free	Earlobes hang free	128900			
<i>Earlobes - Attached</i>	Earlobes attached to the side of the head				
Widow's Peak	Pointed frontal hairline	194000			
<i>No Widow's Peak</i>	No point in frontal hairline				
Can Roll Tongue	Can curl tongue upwards	189300			
<i>Cannot Roll Tongue</i>	Cannot curl tongue				
Dimples, Facial - Present	Dimple on cheeks when smiling	126100			
<i>Dimples - Absent</i>	No cheek dimple				
Pigmented Iris - Present	Eye color is brown or green				
<i>No Pigmented Iris</i>	Blue or gray eyes				
Mid-digital Hair - Present	Hair on middle segment of fingers	157200			
<i>No Mid-digital Hair</i>	No hair on mid-segment of fingers				

Post-Lecture Objectives

- **Genotype is the blueprint; phenotype is the result.**
- Traits follow **predictable inheritance patterns**—but **environment** and **natural/sexual selection** can modify what we see.
- Not all dominant traits are common, and not all recessive traits are rare—**evolutionary preference** and **cultural desirability** play a role.
- Bioinformatics tools like **OMIM** and **NCBI** allow scientists (and you!) to trace genes to chromosomes, diseases, and proteins.
- Trait observation and phenotype/genotype inference
- Data collection and population analysis
- Application of inheritance patterns (dominant, recessive, sex-linked)
- Use of genetic databases (OMIM, NCBI)
- Critical thinking about **evolutionary biology** and **mate preference**

This page titled [Activity 1-1 – Data on Human Traits](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Victor Pham](#).

Activity 1-2 - Human Genetic Inheritance and Diseases

Learning Objectives

- Define a **gene mutation** and distinguish it from deletions and duplications.
- Identify a human **monogenic disease** caused by a point or repeat mutation.
- Use **OMIM** and **NCBI** to locate a gene, identify the associated mutation, and determine its effects on protein function.
- Describe how a gene mutation translates into **phenotypic symptoms** and impacts **life expectancy**.
- Determine the **inheritance pattern** (autosomal dominant/recessive or X-linked) of a genetic disease.
- Navigate DNA, mRNA, and protein sequences using **NCBI Gene** and **NCBI Protein** databases.
- Analyze how genetic mutations disrupt protein function, contributing to disease.

Definition: Term

- **Mutation:** A change in the DNA sequence (e.g., substitution, insertion, or expansion).
- **Monogenic Disorder:** A disease caused by mutation in a single gene.
- **Point Mutation:** A single base substitution in DNA that can change a protein's function.
- **Repeat Expansion:** A mutation where a short DNA sequence is repeated more times than normal.
- **Genotype:** An individual's combination of alleles for a gene.
- **Phenotype:** Observable traits or symptoms that result from genotype.
- **Inheritance Pattern:** The way a trait or disease is passed down (dominant, recessive, X-linked).
- **NCBI Gene:** A database used to find genes, sequences, and genomic context.
- **OMIM:** Online Mendelian Inheritance in Man—a database of human genetic disorders.
- **NP Accession Number:** A unique identifier for a specific protein sequence in NCBI.

Note

- Review how **DNA → RNA → Protein** works (Central Dogma).
- Familiarize yourself with basic **gene structure**: exons, introns, coding vs. non-coding regions.
- Visit [OMIM](#) and [NCBI Gene](#) and explore example genes like **CFTR**, **HBB**, or **HTT**.
- Think about the relationship between **genetic mutation** and **disease phenotype**.
- Bring your laptop and a willingness to explore!

Activity 1-2A: Researching Human Genetic Diseases

Background

Genetic diseases are caused by changes—or **mutations**—in an organism's DNA sequence. These **mutations** can occur in a single gene or across multiple genes, and their effects range from mild to severe. In this activity, you will investigate a **monogenic (single-gene) disorder** caused by a specific **mutation**, not a deletion or duplication. You'll explore where in the genome the gene is located, what goes wrong in the gene, how it affects people, and how it's inherited. A **mutation** is a change in the DNA sequence of a gene. This might be as small as a single-letter substitution (called a point mutation), or as big as inserting or removing segments. For this activity, focus on **mutation-based diseases**, like:

- **Cystic Fibrosis (CFTR gene, Chromosome 7)** – Caused by a point mutation (most commonly $\Delta F508$) that results in thick, sticky mucus buildup.
- **Sickle Cell Disease (HBB gene, Chromosome 11)** – A single base substitution leads to the production of abnormal hemoglobin, affecting red blood cells.
- **Huntington's Disease (HTT gene, Chromosome 4)** – Caused by a repeat expansion mutation leading to progressive brain degeneration.

Do **not** choose disorders that are due to large **deletions or duplications** of genes or entire chromosome segments (e.g., Down syndrome caused by trisomy 21 is not allowed).

Your DNA is packaged into 23 pairs of **chromosomes**, each carrying hundreds to thousands of genes. Every gene has a specific “address” on a chromosome. For example, the CFTR gene is found on **chromosome 7**. In this activity, you’ll pick **one chromosome**, find a mutated gene located there, and research the genetic disease it causes. Each gene has two copies (alleles), one from each parent. Depending on the mutation and the gene, diseases can follow different **modes of inheritance**:

- **Autosomal Dominant (AD)**: Only one mutated copy is needed to cause the disease (e.g., Huntington’s).
- **Autosomal Recessive (AR)**: Both gene copies must be mutated for the disease to appear (e.g., Cystic fibrosis).
- **X-linked**: The mutated gene is on the X chromosome. Since males have one X and one Y, they are more affected by X-linked recessive diseases (e.g., Duchenne muscular dystrophy).

This investigation helps you link **genotype (DNA mutation)** to **phenotype (observable traits)**—a key theme in genetics. You’ll use **OMIM** and **NCBI** to find:

- The **gene** responsible for the disease.
- The **DNA sequence** (in base pairs).
- How the **mutation** alters gene function.
- What **phenotypic traits** (symptoms) appear.
- **Inheritance pattern**, and how long individuals with the disease typically live.

Instructions:

1. Choose any chromosome (1–22, or X/Y). Use the **NCBI**, **Google**, or **ChatGPT** to explore it.
2. Scroll through the gene list and click to learn more about each one. Choose a gene that is known to cause a disease due to **mutation**, not deletion or duplication. You can also search for diseases in **OMIM** and trace them back to their gene.
3. Once you’ve chosen a gene and its associated disease, fill out **Table 1-2A** with the following:
 - Names of your group members (or individual)
 - Chromosome number
 - Disease name
 - Gene name (must be on selected chromosome)
 - Mode of inheritance (**Autosomal Dominant** or **Autosomal Recessive**)
 - Number of base pairs in the gene
 - Life expectancy of individuals with the disease
 - Most common symptoms
4. One student per group will submit the completed table before the end of class.

Table 1-2A: Genetic Disease Research Summary

Group Members	Chromosome Number	Disease Name	Gene Name	Inheritance (AD/AR)	Gene Size (Base Pairs)	Life Expectancy	Common Symptoms

Example: Table 1-2A – Genetic Disease Research Summary

Group Members	Chromosome Number	Disease Name	Gene Name	Inheritance (AD/AR)	Gene Size (Base Pairs)	Life Expectancy	Common Symptoms
Alex, Jordan, Sam	7	Cystic Fibrosis	CFTR	Autosomal Recessive	~189,000 bp	Median: ~40–50 years (with treatment)	Thick mucus in lungs, chronic lung infections, poor growth, salty sweat, infertility

Activity 1-2B: Exploring Gene and Protein Sequences

This activity helps you visualize how genetic information flows from **DNA to RNA to protein**, a process known as the **Central Dogma of Molecular Biology**. When a mutation occurs in a gene, it can affect the structure or function of the protein that the gene codes for. This is often the root cause of many genetic diseases. For example, the **BRCA1 gene** (linked to breast and ovarian cancer) is located on chromosome 17. When functioning properly, the BRCA1 protein helps repair damaged DNA. A mutation in BRCA1 can lead to a malfunctioning protein, increasing cancer risk because damaged DNA isn't properly repaired. In this activity, you will retrieve and analyze different forms of your gene's sequence:

- **Genomic DNA (NC_ or NG_)**: The full stretch of DNA that includes the gene, including introns (non-coding regions) and exons (coding regions).
- **mRNA (NM_ or XM_)**: This is the "cleaned-up" transcript with introns removed—what the ribosome reads to build a protein.
- **Protein (NP_ or XP_)**: This is the final, functional product. You'll explore its name, function, and sequence.

Protocol

1. Using either **NCBI**, **Google**, or **ChatGPT** to determine the name of the **protein** that your gene expresses.
2. Visit NCBI protein (i.e., <https://www.ncbi.nlm.nih.gov/protein>)
3. Input your protein's name in the search bar. (Make sure the proteins belongs to *Homo sapeins*)
4. Click on the correct protein entry and scroll down the page on the website, and **copy the full protein sequence**
 - On the same webpage, you can find and copy the **NP accession number** or **GeneBank code**
5. Save the sequence of NP code for the next lab.

Data Collection Table for Activity 1-2B

Group Members	Gene Name	Disease Name	Protein Name	Protein Sequence	NP Accession #	Notes on Protein Function

Example Entry: BRCA1

Group Members	Gene Name	Chromosome Number	Disease Name	Protein Name	NP Accession #	Notes on Protein Function
Sarah & Liam	BRCA1	17	Breast cancer	Breast cancer type 1 susceptibility protein	NP_009225.1	Helps repair DNA breaks; tumor suppressor

What Should You Now Understand?

- How to **trace a disease back to a specific gene** and mutation.
- How a **single mutation** can lead to **major disruptions in protein function**.
- That **genotype (mutation)** leads to **phenotype (symptoms)** through faulty proteins.
- How **modes of inheritance** impact who gets the disease and how it runs in families.
- How **bioinformatics tools** (OMIM, NCBI) empower researchers to investigate diseases in silico.
- Navigating **genomic and protein databases** (NCBI, OMIM).
- Understanding and classifying **mutation types**.
- Connecting **gene mutations** → **protein dysfunction** → **symptoms**.
- Organizing data about gene size, symptoms, life expectancy, and inheritance.

- Thinking critically about **genetic research and human health**.

This page titled [Activity 1-2 - Human Genetic Inheritance and Diseases](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Victor Pham](#).

Activity 1-3 - Genetic evolution and Identifying Homologs

Learning Objectives

- Define what BLAST and BLASTp are and why they are used in evolutionary biology.
- Explain the difference between homologous and analogous sequences.
- Describe how protein conservation can reveal functional and evolutionary insights.
- Recognize the strengths of BLASTp in detecting distant evolutionary relationships.

Note

Students should read:

- The introductory text on BLAST and BLASTp provided in the chapter.
- The example comparing jaw bones and ear bones in vertebrates.
- Definitions of homologs vs. analogs.
- Overview of how BLASTp works (breaking sequences into words, E-values, identity %).

Definition: Term

- **BLAST (Basic Local Alignment Search Tool):** A bioinformatics tool that finds regions of similarity between biological sequences.
- **BLASTp:** A version of BLAST that compares protein sequences, useful for finding distant evolutionary relationships.
- **Homologs:** Genes or proteins that share a common ancestor and often a similar function.
- **Analogous:** Genes or proteins with similar function but different evolutionary origins (convergent evolution).
- **E-value (Expect Value):** A statistical measure that estimates the likelihood a match occurred by chance. Smaller = more significant.
- **Identity %:** The percentage of amino acids that exactly match between two protein sequences.
- **Query Coverage:** The percentage of your protein that aligns with a match in the database.
- **Functional Domain:** A conserved region in a protein that is critical for its biological function.

Exploring Protein Homology Using BLASTp

What is BLAST?

BLAST (Basic Local Alignment Search Tool) is a powerful **Bioinformatics** tool used to compare your sequence (DNA or protein) against a vast database of sequences stored in NCBI. It helps identify similar sequences, called **homologs**, which likely share a common evolutionary origin. There are different types of BLAST:

- **BLASTn** compares **nucleotide sequences** (DNA or RNA).
- **BLASTp** compares **protein sequences** (amino acids).

In this lab, we use **BLASTp**. Why? Because proteins evolve more slowly than DNA, and BLASTp is better at recognizing distant evolutionary relationships. Also, amino acids have different chemical properties, which helps the software make better alignments—even when the DNA changes, the protein may still “look” similar functionally. Imagine a mutation changes a DNA codon, but the new codon still codes for an identical amino acid (like glutamic acid → aspartic acid). BLASTp recognizes the similarity based on function, even when the nucleotide sequence looks different. BLASTn would miss this.

Here’s an interesting example of protein homology and evolution: **In mammals, the tiny bones in your middle ear (malleus, incus, and stapes) evolved from jaw bones in reptilian ancestors.** The **malleus** and **incus** were once part of the jaw joint in early vertebrates. Over time, they shifted function and became part of the hearing apparatus in mammals. So if you compare the proteins involved in forming middle ear bones in humans and jawbones in reptiles, you might find **homologous proteins**—same

evolutionary origin, different anatomical roles. This is a great example of how protein conservation gives us clues about evolutionary history.

- **Homologs:** Sequences that are similar because they come from a **common ancestor**. (e.g., human and chimp hemoglobin).
- **Analogs:** Sequences or structures that serve **similar functions but evolved independently**—a process called **convergent evolution**. (e.g., bat wings and insect wings).

How BLASTp Works

When you paste your protein into BLASTp, **BLAST breaks the sequence into short segments (called "words")**. It scans the database looking for matching words. Once a match is found, it extends the alignment on both sides. It scores the quality of alignment using: **Identity** (% of amino acids that match exactly) and **E-value** (probability the match is just by chance—the smaller, the better). BLAST ranks and displays the best matches (homologs) in a user-friendly way: A **colorful graphic** (where red = best hit), a **description table** with identity %, E-value, and coverage, and the actual **sequence alignments**.

Lab Protocol

1. Find Your Protein

- Go to NCBI **Protein** Database (<https://www.ncbi.nlm.nih.gov/protein>)
- Search for your protein of interest (e.g., "sonic hedgehog").
- Click on the result and look for the **GenBank accession number** (e.g., NP_000257.2, BAA33523.2, etc).



2. Launch BLASTp

- On the right side of the page, under **"Analyze this sequence,"** click **"Run BLAST"**.
- This sends your protein into the **BLASTp** query form automatically.

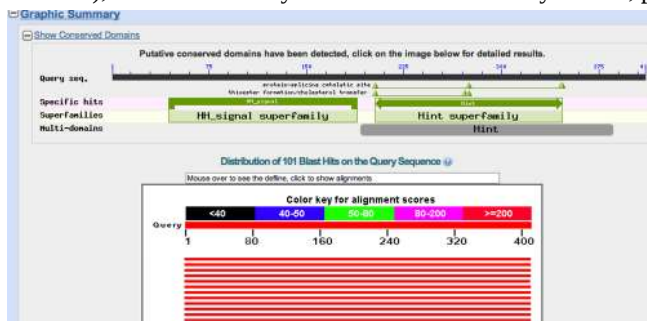


3. Set Up Your Search

- Confirm that **BLASTp** is selected (top left).
- Under **Organism**, search for **one species at a time** (e.g., *Arabidopsis thaliana*, *E. coli*, etc.).
- Click **"BLAST"** at the bottom to begin.

4. Review Your Results. You'll see three main sections:

- Graphic Summary:** Shows where the best hits align. Red bars mean highly similar regions.
 - The **Graphic Summary** is a visual snapshot showing how well other proteins in the database match your query protein. The position of each bar shows *where* in your protein the match occurred, and the color of the bar indicates *how strong* that match is. Colors are used to represent similarity scores: **red bars** indicate very strong matches (high similarity), **orange or pink bars** show moderate matches, **green or blue bars** reflect lower similarity, and **black or gray bars** represent weak or non-significant matches. For example, if you see a red bar for a chimpanzee protein and a green bar for a fruit fly protein, it suggests that the chimpanzee version is highly similar to your query protein (likely the same function), while the fruit fly version is more distantly related, possibly sharing only one conserved domain.



- Descriptions Table:** Includes Identity %, E-value (smaller is better), and Query Coverage.
 - The **Descriptions Table** gives you a detailed breakdown of each top “hit”—each protein that had some level of similarity to your query. The **Max Score** tells you how well the sequences align overall, with higher scores indicating better matches. **Query Coverage** shows the percentage of your protein that aligned with the match; a higher coverage means a longer stretch of similarity. The **E-value** (or Expect value) is a critical measure: it tells you the probability that this match occurred by random chance. The lower the E-value, the more significant the match is—values close to 0.0 suggest that the sequences are almost certainly homologous. Finally, the **Percent Identity** tells you how similar the actual amino acid sequences are; values above 90% usually indicate the protein is nearly identical and probably has the same function in different organisms. For instance, if your human protein aligns with a mouse protein with 98% identity, 100% query coverage, and an E-value of 0.0, that means the two proteins are almost identical—likely performing the same function in both species.

Descriptions

Sequences producing significant alignments:

Select: **All** None Selected:0

Alignments Download GenPept Graphics Distance tree of results Multiple alignment

Description	Max score	Total score	Query cover	E value	Ident	Accession
☐ hedgehog [Homo sapiens]	840	840	100%	0.0	100%	BAA33523.2
☐ indian hedgehog protein preproprotein [Homo sapiens]	837	837	100%	0.0	99%	NP_002172.2
☐ PREDICTED: indian hedgehog protein [Macaca mulatta]	833	833	100%	0.0	99%	XP_001092385.2
☐ PREDICTED: indian hedgehog protein [Rhinopithecus roxellana]	832	832	100%	0.0	99%	XP_010358565.1
☐ PREDICTED: indian hedgehog protein [Macaca nemestrina]	832	832	99%	0.0	99%	XP_011738989.1
☐ PREDICTED: indian hedgehog protein [Pongo abelii]	832	832	100%	0.0	99%	XP_002812916.1
☐ PREDICTED: indian hedgehog protein [Pan troglodytes]	826	826	100%	0.0	98%	XP_526034.2
☐ PREDICTED: indian hedgehog protein [Saimiri boliviensis boliviensis]	822	822	100%	0.0	98%	XP_003925566.1
☐ PREDICTED: indian hedgehog protein [Aotus nancymaae]	819	819	100%	0.0	97%	XP_012301804.1

o **Alignments:** Detailed comparison of your protein to others.

- The **Alignments** section lets you dig into the exact amino acid-by-amino acid comparison between your protein and the matching protein. It shows whether the amino acids are exactly the same (**identical residues**) or if they are biochemically similar (**conservative substitutions**), and highlights regions that match. This is where you can see not only that proteins are similar but also which *parts* are similar—which is key to understanding whether functional domains are conserved. This detailed view helps you ask deeper biological questions: for example, if only the middle part of the protein aligns across species, maybe that region is a conserved domain important for binding DNA or ATP. If the alignment only occurs in one part of the protein, it might not be the entire protein that's conserved—just a functional part. This is common with proteins that share **functional domains** but are otherwise unrelated.

Alignments

Download GenPept Graphics

hedgehog [Homo sapiens]
Sequence ID: [gb|BAA33523.2](#) Length: 411 Number of Matches: 1

Range 1: 1 to 411 GenPept Graphics

Score: 840 bits(2170) Expect: 0.0 Method: Compositional matrix adjust. Identities: 411/411(100%) Positives: 411/411(100%) Gaps: 0/411(0%)

Query 1 MSFARLRFRLIFCLVLLLLVVFAMGCCGRVVGSRPRPRKLVLATKQFSFVPEET 60
Sbjct 1 MSFARLRFRLIFCLVLLLLVVFAMGCCGRVVGSRPRPRKLVLATKQFSFVPEET 60

Query 61 LGASRYEGKIANSERFKELTFTYTPDIFKDEENTGADRLATQCKRDLNSLAEVSN 120
Sbjct 61 LGASRYEGKIANSERFKELTFTYTPDIFKDEENTGADRLATQCKRDLNSLAEVSN 120

Query 121 QWPKLVRTGNDGDHSESLYEGRAVDITSGRKNKYGLLRLAVEAGFDMVY 180
Sbjct 121 QWPKLVRTGNDGDHSESLYEGRAVDITSGRKNKYGLLRLAVEAGFDMVY 180

Query 181 ESKAHVHCIVKSESAARTGOCFFAGQVRLSGARVALSAVRPGDVLAMGDSPT 240
Sbjct 181 ESKAHVHCIVKSESAARTGOCFFAGQVRLSGARVALSAVRPGDVLAMGDSPT 240

Query 241 SDVLLDLREPHRLRAVQVITQDFRRLALCPAHLFTADHETPAARFATFASVQ 300
Sbjct 241 SDVLLDLREPHRLRAVQVITQDFRRLALCPAHLFTADHETPAARFATFASVQ 300

Query 301 GQTVLVAGAPGLQFARVAAVSTVVALGATAFLTKRGLVVEDVVASCFANVADHLAG 360
Sbjct 301 GQTVLVAGAPGLQFARVAAVSTVVALGATAFLTKRGLVVEDVVASCFANVADHLAG 360

Query 361 FWPLKLFHSLAWGNTFCGQVNTYFQLLRGLRLGLLGRSFLPLKSGAGS 411
Sbjct 361 FWPLKLFHSLAWGNTFCGQVNTYFQLLRGLRLGLLGRSFLPLKSGAGS 411

Related Information
[Gene - associated gene details](#)

Record Your Observations

Use the table below to document your results from different clades (groups of organisms). For each species, note the **color bar**, **% identity**, and **E-value**. Then, hypothesize the function of the protein in that organism. (You should look up the function of your proteins in these organisms. If the functions are unknown, you can come up with your own hypothesis of its functions)

Clade / Organism	Bar Color	Identity %	E-value	Hypothesized Function
Primates (9443)				
Marsupials (9263)				
Monotremes (9255)				
Birds (8782)				
Lizards (8504)				
Amphibians (8292)				
Fishes (117569)				
Fruit Flies (7211)				
Sea Urchins (7625)				
Sponges (6040)				
Arabidopsis (3702)				
Yeast (4932)				
E. coli (562)				

Reflection Questions

1. What organisms had the most conserved (similar) version of your protein?
2. How do the protein's functions vary between simpler organisms (like bacteria) and more complex ones (like mammals)?
3. Can you identify a trend in protein conservation across evolution?
4. What does this say about the importance of your protein?

Example BLASTp Results Table

Query protein: Human MYH7 (Myosin heavy chain 7)

Function in humans: Part of the motor protein complex for cardiac and skeletal muscle contraction.

Clade / Organism	Bar Color	Identity %	E-value	Hypothesized Function
Primates (9443)	Red	99%	0.0	Muscle contraction in heart and limbs
Marsupials (9263)	Red	96%	0.0	Skeletal and cardiac muscle contraction
Monotremes (9255)	Red	94%	0.0	Same – contractile protein in heart and skeletal muscle
Birds (8782)	Red	85%	2e-100	Flight and leg muscle contraction
Lizards (8504)	Red	83%	5e-90	Muscle movement (locomotion, tail movement)
Amphibians (8292)	Orange	78%	3e-80	Swimming muscle function; limb movement
Fishes (117569)	Orange	70%	2e-60	Swimming muscle movement; tail fin muscle control
Fruit Flies (7211)	Green	35%	0.003	Muscle contraction in wings and legs; partial homolog
Sea Urchins (7625)	Green	32%	0.4	Tentacle or tube foot movement (using actomyosin system)
Sponges (6040)	Gray	18%	10	No true muscles; possible ancient cytoskeletal role
Arabidopsis (3702) (Plant)	Gray	20%	8	Cytoplasmic streaming (actin-myosin-like transport system)
Yeast (4932)	Gray	22%	6	Organelle transport, cell division via actin-myosin system
E. coli (562)	Black	5%	>100	No homolog found; prokaryotes lack myosin-like proteins

Example Observations:

- **High similarity in vertebrates**, suggesting MYH7's role is critical in muscle function and has been conserved.
- **Fruit flies and sea urchins** have partial matches—they may use similar motor proteins for movement.
- **Plants and fungi** don't have muscles, but they still use actin-myosin-like systems for internal transport.
- **Bacteria (like E. coli)** don't have homologs—makes sense since they don't use motor proteins in the same way.

Post-Lecture Objectives

- Use BLASTp to analyze a protein's conservation across different species.
- Interpret E-values, bar colors, and identity % to assess evolutionary relatedness.
- Hypothesize how conserved proteins may function differently in different organisms.
- Construct a hypothetical evolutionary timeline for a protein's role across clades.

This page titled [Activity 1-3 - Genetic evolution and Identifying Homologs](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Victor Pham](#).

Activity 1-4 - PowerPoint Presentation on the Bioinformatic Evolution and function of proteins

Bioinformatics PowerPoint Presentation

This is not just about disease. It's about uncovering hidden stories written in our DNA and asking, "How did this gene evolve? What *else* might it be doing in other organisms?" You'll use bioinformatics tools like BLAST to compare gene sequences across different species and form a hypothesis about the evolutionary *function* of your disease-associated gene.

Presentation Objective

- Use **NCBI BLASTp** and sequence data to identify when your disease-associated gene may have emerged across the evolutionary tree.
- Investigate how this gene (or protein) might be **beneficial or functional in other species**. What role might it serve in animals that don't get the disease? Has evolution repurposed it?

This project is about connecting dots between **genetics**, **function**, and **evolution**. Think like evolutionary detectives!

How to Approach Your Hypothesis

Ask questions like:

1. Why does this gene exist in birds or reptiles?
2. What biological process might it be involved in?
3. What happens if this gene is mutated in humans — and how could that give clues to its *original* role in evolution?
4. Could this gene's original function be **reused** or **repurposed** in different animals?

Grading Rubric

Category	Points	Description
Disease Overview	5 pts	Brief, clear, relevant. What's the disease? Which gene is involved? Why do we care?
Evolutionary Timeline	5 pts	Where do you see the gene in the tree of life? What major clades have it? Where does it seem to "appear"?
Evolutionary Function Hypothesis	20 pts	This is the <i>most important</i> part. Propose how the gene/protein might help other species. How might it be repurposed for a different function? Could it be part of hearing in one species, and jaw formation in another? Be bold, creative, and science-based.
Slide & Time Limits	-10 pts	Max 5 content slides (plus title/reference). Max 8 minutes . Practice ahead!
Quiz Question	N/A	Write one multiple-choice question for your peers. (This may appear on the quiz!)

- Your goal is to **tell a story**, not just dump data.
- We're not asking for certainty—this is **hypothesis-driven science**.
- Be *curious*, be *creative*, and support your ideas with evidence from BLAST and published data (if available).
- Make it fascinating enough that your classmates remember your presentation when the quiz comes around!

Example

Let's say your gene is **EYA1**, associated with **Branchio-Oto-Renal syndrome** (a disorder affecting ear and kidney development in humans).

- You BLAST the **EYA1 protein**, and you find it exists in **amphibians, birds, fish**, even in **invertebrates**.
- But in non-humans, it's not linked to disease — instead, it's part of **ear development**, or even **jaw and gill structure** in fish.
- **Your hypothesis:** This gene originally evolved to shape **pharyngeal arches** (structures in embryos that become jaws, gills, and ears), and in humans, mutations disrupt that process.

You could explain how a gene critical for ancient **fish gill arches** eventually became part of **human ear structure**, and when mutated, causes hearing problems. That's a *wild* evolutionary journey!

Example Student Presentation:

Title: "The Evolutionary Journey of the *EYA1* Gene"

"Did you know that the tiny bones inside your ear once helped ancient fish breathe underwater? In our project, we discovered how a gene associated with human hearing loss actually helped build gills in our distant aquatic ancestors..."

Slide 1: Disease Overview – BOR Syndrome & the EYA1 Gene (5 pts)

- **Branchio-Oto-Renal (BOR) Syndrome**
 - Rare genetic disorder affecting **ears, neck (branchial arches)**, and **kidneys**
 - Symptoms: hearing loss, malformed ears, neck cysts, kidney issues
 - Caused by mutations in the **EYA1** gene, a transcriptional activator during early development
- **EYA1 Gene Function in Humans**
 - Essential for the development of **ear structures** and **kidneys**
 - Expressed during embryonic formation of the **pharyngeal arches**

Slide 2: Evolutionary Timeline (5 pts)

Clade (Taxid)	Identity %	Hypothesized Function	Evolutionary Advantage
Primates (9443)	~99%	Coordination of ear , jaw , and kidney development	Advanced sensory perception, vocalization, and filtration
Birds (8782)	~88%	Development of balance organs and lightweight skull structures	Enhanced flight coordination and environmental sensing
Amphibians (8292)	~85%	Formation of inner ear structures and jaw bones	Transition to land: improved balance and hearing in air
Fish (117569)	~82%	Development of gill arches and jaw cartilage	Supports respiration and feeding structures in aquatic environments
Invertebrates (e.g., Sea Urchins)	~50%	Possibly involved in sensory organ precursor cells	Early role in organizing body plan or sensory input
Bacteria / Fungi	<35% (if any)	Unclear or non-existent	Potential origin of ancestral gene segments or regulators

Hypothesis: The EYA1 gene originated in early vertebrates or before, and was repurposed across species for different developmental structures.

Slide 3: Evolutionary Hypothesis – EYA1 in Other Species (20 pts)

 What does EYA1 do in non-humans?

- **Fish:**
 - EYA1 helps form **gill arches** and cranial cartilage
 - Similar structures to mammalian **middle ear bones**
- **Birds and Amphibians:**
 - Involved in **inner ear** formation
 - Required for sensory cell development and balance
- **Hypothesis:**
 - EYA1 likely evolved to control **branchial arch development** in early fish
 - These arches evolved into **jaws** and **ears** in mammals
 - In humans, EYA1 mutations cause malformations in these ancient structures
 - **Fun Fact:** The **stapes** (middle ear bone in humans) evolved from the **hyomandibula** — a gill support bone in ancient fish. EYA1 helped build both!

Slide 4: Conclusions & Quiz Question

- EYA1 is a deeply conserved gene repurposed across evolution
- Originally involved in **gill arch** formation, later adapted for **ears** and **kidneys**
- Mutations in humans reveal its vital role in embryonic development

Quiz Question (Multiple Choice):

1. What structure did the mammalian middle ear bones evolve from in ancient fish?

- A. Vertebral column
- B. Fin bones
- C. Gill arches
- D. Tail spine

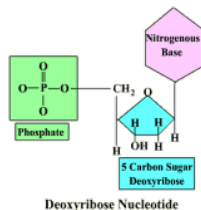
✓ Correct answer: **C. Gill arches**

This page titled [Activity 1-4 - PowerPoint Presentation on the Bioinformatic Evolution and function of proteins](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Victor Pham](#).

Activity 2-0 - Introduction to Plasmid and Enzyme p450

Nucleotides – The Building Blocks of DNA

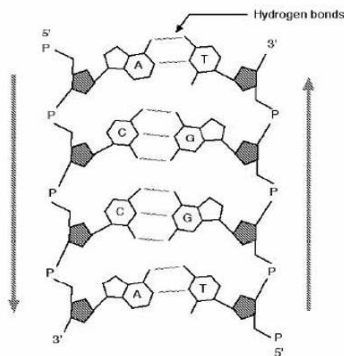
DNA, or **Deoxyribonucleic Acid**, is the molecular blueprint of life, storing genetic information that dictates everything from eye color to disease susceptibility. At its core, DNA is made up of **nucleotides**, which are the basic building blocks. Each nucleotide consists of three components: a **deoxyribose sugar**, a **phosphate group**, and a **nitrogenous base**. The sugar is a five-carbon molecule, and the carbons are numbered from **1' (one prime)** to **5' (five prime)**. This numbering becomes crucial in understanding DNA's directionality and how the molecule is assembled. The four nitrogenous bases are **Adenine (A)**, **Thymine (T)**, **Guanine (G)**, and **Cytosine (C)**—and they are what make up the unique code of genetic information. For example, a gene that codes for the insulin protein will have a specific sequence of these bases that is different from the gene that codes for hemoglobin.



Phosphodiester Bonds and Directionality (5' to 3')

Nucleotides connect to each other via **phosphodiester bonds**, which form between the **3' hydroxyl (–OH)** group of one sugar and the **5' phosphate group** of the next nucleotide. This linkage forms a sugar-phosphate backbone that is both strong and flexible, providing structural support to the DNA molecule. Because of this bonding, each DNA strand has directionality—one end has a **free 5' phosphate group**, and the other end has a **free 3' hydroxyl group**. This is critical during **DNA replication and transcription**, as enzymes like DNA polymerase and RNA polymerase can only add new nucleotides to the **3' end**, so DNA is always synthesized in the **5' to 3' direction**. Think of it like writing a sentence in English: you must start from the left (5') and move to the right (3')—writing in the reverse would result in gibberish or unusable code.

Although a single DNA strand can carry information, **functional DNA in living cells exists as a double-stranded helix**. Two DNA strands align in an **antiparallel** fashion—one runs in a **5' to 3'** direction, and the other runs in the **3' to 5'** direction. This orientation is necessary for the nitrogenous bases to **hydrogen bond correctly** in the center of the helix. If both strands ran in the same direction, the base pairs would not align properly for hydrogen bonding, and the helix would not form. Imagine a zipper: both sides must align oppositely for the teeth (bases) to interlock. The classic image of DNA's **double helix**, first revealed by Rosalind Franklin's X-ray crystallography and modeled by Watson and Crick, is built upon this principle of antiparallel strand alignment.



Complementary Base Pairing and Stability

The two DNA strands are held together through **specific base pairing**—Adenine (A) always pairs with Thymine (T), and Guanine (G) always pairs with Cytosine (C). These pairings are stabilized by **hydrogen bonds**: **A-T pairs form two hydrogen bonds**, while **G-C pairs form three**. The extra bond in G-C pairs makes them **more stable** and harder to break. This difference has practical implications. For instance, **genes or DNA regions rich in G-C content are more stable and harder to denature**, which is an important consideration in **PCR (polymerase chain reaction)** when designing primers or setting annealing temperatures.

Similarly, **promoter regions in some species tend to have more A-T content**, making them easier for the transcription machinery to unwind and access.

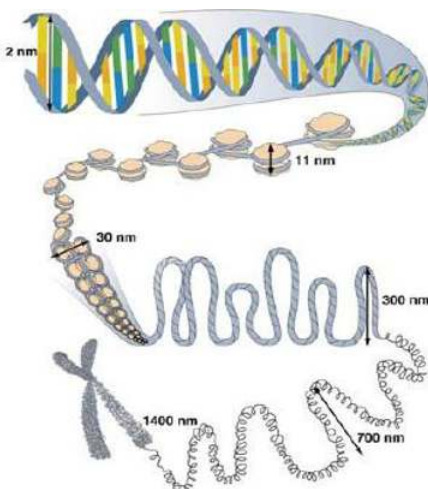
One fascinating aspect of DNA is that any two single strands with **complementary sequences** can find each other and **anneal**—or stick together—via base pairing. This principle is widely used in biotechnology and molecular biology. For example, during **PCR**, synthetic DNA primers bind (anneal) to specific complementary sequences on target DNA to initiate replication. In **gene cloning**, scientists design complementary overhangs to ligate DNA fragments. In **Southern blotting**, labeled DNA probes bind to complementary DNA in a sample, allowing for gene detection. This ability of DNA to self-recognize and pair with its complement underpins much of modern genetic technology.

Chromosomes vs. Plasmids:

In biology, the **chromosome** is the main genetic blueprint of a cell. In **bacteria**, this typically takes the form of a **single circular chromosome**, which carries all the essential genes required for survival, such as those involved in DNA replication, cell division, metabolism, and stress response. In contrast, **eukaryotic cells**—like those in humans, animals, or plants—have **multiple linear chromosomes** housed in a membrane-bound nucleus. For example, human cells contain **46 chromosomes** (23 pairs), which are tightly packed using **histone proteins** to form **chromatin**. During cell division, this chromatin condenses and becomes visible under a light microscope, which is a hallmark of mitosis and meiosis.

In addition to chromosomes, many cells—especially bacteria—also contain **plasmids**. Plasmids are **small, circular, double-stranded DNA molecules** that replicate **independently of the chromosome**. While not essential for basic survival, plasmids can carry advantageous genes, such as those for **antibiotic resistance**, **virulence**, or **metabolic specialization**. For instance, some plasmids enable bacteria to break down unusual sugars or resist antibiotics like tetracycline. Unlike chromosomes, which exist in one or two copies per cell, **plasmid copy number varies**—**natural plasmids** usually have fewer than 30 copies, while **engineered plasmids** (used in labs) can reach **hundreds of copies per cell**, allowing high levels of gene expression.

For a plasmid to be maintained in a cell, it must be able to replicate. This begins at a specific sequence known as the **origin of replication (ori)**. The ori is where proteins like **DNA helicase** (which unwinds the double helix) and **DNA polymerase** (which synthesizes new DNA) bind and initiate the replication process. The efficiency of plasmid replication depends on the type of ori and its compatibility with the host. For instance, **ColE1** is a common ori used in *E. coli* plasmids, enabling multiple rounds of replication and high plasmid yield. This feature is incredibly useful in genetic engineering and biotechnology, where researchers introduce foreign genes into plasmids and use bacterial hosts to replicate and express these genes. But getting a plasmid into a bacterium is not trivial—this is where the process of **transformation** comes in.



Transformation is the process of introducing **foreign DNA (like a plasmid)** into a bacterial cell. However, this process is **naturally inefficient**—only about **1 in every 10,000 bacterial cells** will successfully take up a plasmid during a standard chemical transformation (e.g., using calcium chloride and heat shock). To identify which cells have been successfully transformed, researchers use **selectable markers**, most commonly **antibiotic resistance genes**. For example, the **ampicillin resistance gene (bla)** encodes the **β -lactamase enzyme**, which breaks down ampicillin. When bacteria are plated on agar containing ampicillin,

only the transformed cells carrying the plasmid survive and grow, forming colonies. Untransformed cells are killed by the antibiotic, making selection straightforward and efficient.

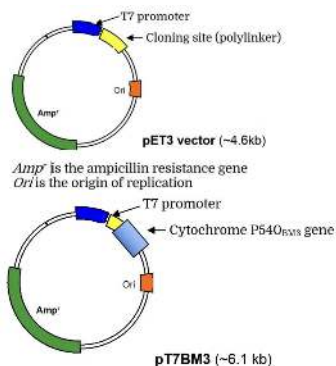
Plasmids used in molecular biology are designed for **flexibility and ease of manipulation**. One key feature is the **polylinker**, also called the **multiple cloning site (MCS)**. This is a **short synthetic DNA segment** that contains several unique **restriction enzyme recognition sites**, allowing scientists to insert genes of interest at precise locations. These sites are typically **palindromic sequences**—meaning they read the same from 5' to 3' on both strands. For example, the sequence **GAATTC** (cut by **EcoRI**) reads the same on the complementary strand: 5'-GAATTC-3' and 3'-CTTAAG-5'. Restriction enzymes recognize these palindromes and cut them, creating sticky or blunt ends that facilitate **gene insertion during cloning**.

Expression of Cytochrome P450 in Transformed *E. coli*

In the lab described, two strains of *E. coli* are transformed with plasmids to study the expression of **Cytochrome P450 BM3**, a protein that is not naturally found in *E. coli* but comes from ***Bacillus megaterium***. One strain receives an **empty vector (pET3a)**, while the other receives **pET3a with the P450BM3 gene inserted (pT7BM3)**. These plasmids carry the **ampicillin resistance gene** for selection and are designed to express the P450 gene **only under certain conditions**.

The **pET3a vector** contains a **T7 promoter**, which is recognized only by **T7 RNA polymerase**, not by *E. coli*'s natural polymerase. The *E. coli* strain used here has been genetically modified to contain **T7 RNA polymerase integrated into its chromosome**, under control of an **inducible promoter** (like the lac operon). When researchers add **IPTG**, a molecule that mimics lactose, it induces T7 polymerase production. This polymerase then binds to the T7 promoter on the plasmid and transcribes the P450 gene **at very high rates**—up to **five times faster** than native *E. coli* RNA polymerase.

This regulated, high-efficiency system allows researchers to **"turn on" protein expression only when needed**, avoiding the potential toxicity or resource drain of constant overexpression. Once expressed, the **Cytochrome P450 BM3 enzyme** can be isolated and studied for its role in **drug metabolism, biotransformation**, or other biochemical processes.



Introduction to Cytochrome P450_{BM3}:

Cytochrome P450 enzymes (often abbreviated as CYPs) form a vast and highly conserved superfamily of heme-thiolate proteins found across all domains of life—from bacteria to humans. These enzymes catalyze a wide array of oxidation reactions, making them crucial for various biological functions. In humans, for instance, P450s are vital in drug metabolism, cholesterol synthesis, and detoxification of xenobiotics (foreign substances). For example, human CYP3A4 metabolizes more than 50% of pharmaceuticals, impacting drug dosing and drug-drug interactions. In this course, however, our focus will be on the **prokaryotic enzyme cytochrome P450BM3**, originally isolated from the bacterium *Bacillus megaterium*. This bacterial enzyme is among the best-characterized P450s and serves as an excellent model system due to its high activity and simplicity of expression in *E. coli*.

All cytochrome P450 enzymes contain a **heme prosthetic group**, which is responsible for their catalytic activity and gives them their characteristic reddish-brown color. This heme contains a central **iron atom (Fe)** coordinated to four nitrogen atoms of a porphyrin ring and covalently bound to a cysteine residue from the protein. The heme iron cycles between Fe^{3+} (ferric) and Fe^{2+} (ferrous) states during catalysis. Importantly, it is this iron atom that binds oxygen and facilitates its activation for incorporation into substrates. For example, when *lauric acid* (a 12-carbon saturated fatty acid) is used as a substrate in the lab, one of the oxygen atoms from molecular O_2 is inserted into the fatty acid, forming **hydroxy-lauric acid**, while the second oxygen atom is reduced to water.

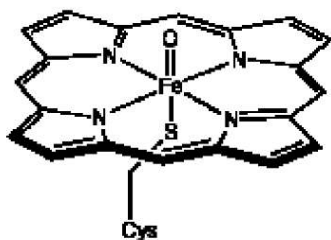


Figure 1. Heme group of a cytochrome P450 protein.
(From the coordination chemistry page of Saint John's University.)

The primary reaction catalyzed by P450 enzymes is **hydroxylation**, where a hydrogen atom in a substrate is replaced with a hydroxyl group (-OH). P450BM3, specifically, is a **fatty acid hydroxylase**—meaning it hydroxylates medium- to long-chain fatty acids like lauric acid. This type of enzyme is also called a **monooxygenase**, because it incorporates one atom of molecular oxygen into the substrate and reduces the other to water. This distinguishes it from **dioxygenases**, which insert both oxygen atoms into the product. This kind of reaction is highly valuable in biotechnology and pharmaceutical industries, where selective oxidation is difficult to achieve with synthetic catalysts. For instance, engineered P450s can be used to produce drug metabolites or to hydroxylate complex molecules like steroids and vitamins.

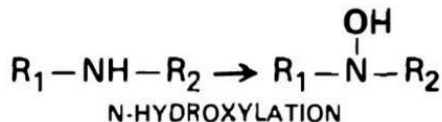


Figure 2. A standard hydroxylation reaction. From Nebert *et al. Ann Rev Biochem.* 1987.

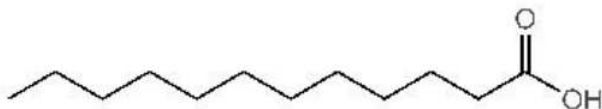


Figure 3. The structure of lauric acid.
Modified from www.pherobase.com.

For P450 enzymes to function, electrons must be transferred to the heme iron to enable oxygen activation. These electrons are typically supplied by **NADPH**, which gets oxidized to NADP^+ in the process. In class I P450 systems (mostly bacterial and mitochondrial), electrons flow from NADPH to **ferredoxin reductase** (which contains FAD), then to **ferredoxin** (an iron-sulfur protein), and finally to the P450 heme. Class II systems (more common in eukaryotes), on the other hand, transfer electrons via **cytochrome P450 reductase**, a protein containing both **FAD** and **FMN** domains. What's unique about P450BM3 is that it is a **rare bacterial class II enzyme** and is **fused into a single polypeptide chain**—combining the heme, FAD, and FMN domains. This structural integration allows for rapid intramolecular electron transfer, making P450BM3 **one of the fastest P450 enzymes known**. In contrast, most P450 systems are multipart, requiring complex protein-protein interactions to complete the electron transfer chain.

The catalytic cycle of P450 enzymes is a multi-step process beginning with substrate binding. Initially, the heme iron is in the ferric state (Fe^{3+}) and coordinated to a water molecule. When a substrate (like lauric acid) binds to the enzyme's active site, it **displaces this water**, which triggers a conformational change and allows the heme to accept an electron from NADPH via the FAD/FMN cofactors. This reduces Fe^{3+} to Fe^{2+} , enabling it to bind molecular oxygen (O_2). A second electron is transferred, reducing O_2 to a **superoxide**, which undergoes **protonation** and **O-O bond cleavage**. One oxygen atom is released as water, and the other is converted into a highly reactive iron-oxo species capable of abstracting a hydrogen from the substrate. This results in the **insertion of an OH group into the substrate**—a process central to drug metabolism and biosynthesis. The cycle then resets as the hydroxylated product leaves the active site.

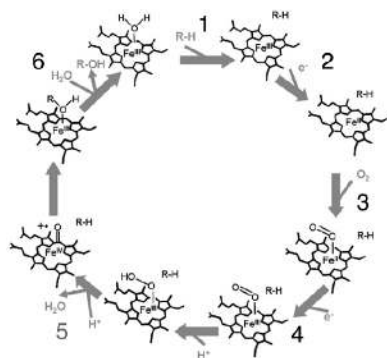


Figure 4. Model of P450 hydroxylation.

P450_{BM3} in the Lab

In this lab course, you will work with *E. coli* cells transformed with either an **empty plasmid (pET3)** or a **recombinant plasmid (pT7BM3)** carrying the gene encoding for P450_{BM3} under control of a **T7 promoter**. This system utilizes *E. coli* strains with a **chromosomally integrated T7 RNA polymerase gene**, which allows for **inducible, high-level expression** of the P450 gene. This is because the T7 RNA polymerase, unlike native *E. coli* RNA polymerase, only transcribes genes downstream of a T7 promoter and does so at a significantly faster rate (up to five times faster). The pET3 plasmid also contains an **ampicillin resistance gene**, which facilitates selection of successfully transformed *E. coli* using antibiotic-containing media.

Each group in the lab receives two cultures: one with the empty vector (control) and one with the vector containing the P450_{BM3} gene. Your objective is to **induce expression of the P450 protein, purify it, and study its enzymatic kinetics** using lauric acid as a substrate. These experiments allow you to explore fundamental concepts in molecular cloning, protein expression, and enzymology—all through the lens of a real, biotechnologically relevant enzyme.

Cytochrome P450_{BM3} is not just another enzyme—it is a *biocatalytic powerhouse*. Its efficiency, simplicity, and amenability to engineering make it a model enzyme for studying P450 catalysis and for developing new biotechnological applications. Because all the electron transfer components are within a **single polypeptide**, there are fewer bottlenecks in the catalytic cycle. This makes it ideal for research into **enzyme kinetics, biocatalyst engineering, and drug metabolism modeling**. Its native substrate, lauric acid, provides a straightforward system to monitor activity, and mutations in the active site have allowed scientists to alter substrate specificity—leading to engineered variants that can hydroxylate drugs, pollutants, or complex natural products.

Big Picture

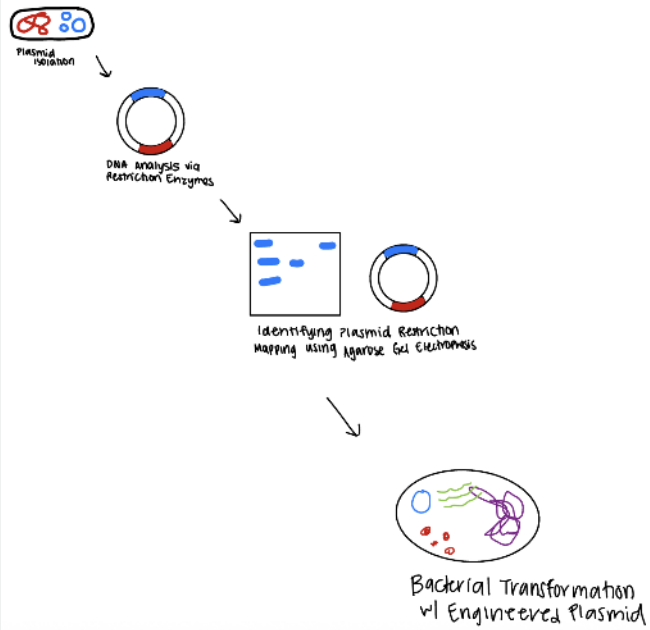


Image of a flow chart summarizing the big picture of genetic engineering via plasmid. Image created by Dr. Victor Pham's student, Geneva Anh Thy Doan.

This page titled [Activity 2-0 - Introduction to Plasmid and Enzyme p450](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Victor Pham](#).

Activity 2-1 - Plasmids Isolation

Learning Objectives

- Define what a plasmid is and explain why it is biologically and biotechnologically important.
- Describe how plasmids can carry genes like antibiotic resistance or metabolic enzymes (e.g., P450).
- Explain how plasmids are introduced into bacteria and how they replicate.
- Outline the key steps and principles behind the plasmid mini prep procedure.
- Understand the purpose and function of buffers P1, P2, P3, binding buffers, and wash/elution steps.

Definition: Term

- **Plasmid:** A small, circular, double-stranded DNA molecule found in bacteria, independent of chromosomal DNA.
- **Supercoiling:** A compact DNA structure resulting from the twisting of the DNA helix, aiding plasmid stability and extraction.
- **Vector:** A vehicle used to deliver genetic material into cells (e.g., plasmids).
- **Transformation:** The process of introducing foreign DNA (like a plasmid) into a bacterial cell.
- **Miniprep:** A laboratory technique used to isolate plasmid DNA from bacterial cells.
- **Alkaline lysis:** A method using high pH to break open cells and selectively isolate plasmid DNA.
- **EDTA:** A molecule that chelates metal ions, protecting DNA from degradation by nucleases.
- **SDS:** A detergent that lyses cells and denatures proteins.
- **Potassium acetate:** Neutralizes the lysis solution and causes precipitation of chromosomal DNA and proteins.
- **Spin column:** A silica-based tool that binds DNA in high-salt conditions for purification.

Note

- Plasmids are powerful tools in genetic engineering, especially for cloning and expression studies.
- Each buffer in the miniprep has a distinct role in protecting, releasing, purifying, or binding DNA.
- Understanding the chemical basis of lysis and DNA precipitation helps troubleshoot common lab issues (e.g., low DNA yield or contamination).
- Hands-on practice with miniprep connects molecular theory with real-world biotechnology workflows.

What is a Plasmid and Why It Matters

A **plasmid** is a small, circular, double-stranded DNA molecule distinct from the chromosomal DNA found in bacteria. These plasmids often carry genes that provide a survival advantage, such as **antibiotic resistance** or **metal ion resistance**. For example, a plasmid carrying a gene for ampicillin resistance allows bacteria to survive in media containing this antibiotic. Plasmids are supercoiled, making them more compact, more stable, and easier to isolate than chromosomal DNA. In molecular biology, plasmids are used as **vectors**—vehicles to transfer genetic material. If a researcher wants to study the function of the **cytochrome P450 gene**, they can clone this gene into a plasmid, transform it into *E. coli*, and then harvest that plasmid from the bacteria after amplification.

After inserting the P450 gene into a plasmid, the next step is **transformation**—introducing the recombinant plasmid into **competent** *E. coli* cells. These bacteria, once transformed, replicate the plasmid along with their own DNA during cell division. This natural amplification ensures that from a single transformed colony, you can culture millions of bacteria, each carrying **multiple copies of the plasmid**. This is essential because it provides a **mass production system** for the DNA you want to study or manipulate.

Principle of Plasmid Mini Prep

To analyze the plasmid DNA (e.g., verify the presence of the P450 gene by restriction digestion), we need to extract it from the bacterial cells. The **Plasmid Mini Prep** is a **quick and efficient method** to purify plasmid DNA, separating it from chromosomal

DNA, proteins, RNA, and other cellular components. This method relies on **alkaline lysis**, developed by Birnboim and Doly in 1979, which takes advantage of structural differences between plasmid and genomic DNA to selectively isolate plasmids.

1. Resuspension in Buffer P1

Bacterial cells from a liquid culture are first harvested by centrifugation. The pellet is then **resuspended in Buffer P1**, which contains:

- **EDTA:** A chelating agent that binds divalent metal ions (like Mg^{2+} and Ca^{2+}). These ions are essential cofactors for nucleases like **DNase**, which could degrade DNA. By removing metal ions, EDTA **inhibits DNases**, protecting the plasmid DNA. If EDTA were omitted, any endogenous DNases could become active during cell lysis and degrade your plasmid DNA, leading to poor yields or fragmented plasmids.
- **RNase A:** An enzyme that degrades RNA. This ensures that the final preparation is free from RNA contamination, allowing you to work exclusively with DNA.

2. Lysis with Buffer P2

After resuspension, **Buffer P2** is added. This step causes the cell to **burst open**, releasing its contents. Both DNA types become single-stranded, and cellular proteins and lipids start to precipitate or denature. This solution contains:

- **Sodium dodecyl sulfate (SDS):** An anionic detergent that **disrupts the cell membrane** and **denatures proteins**. SDS disrupts the lipid bilayer of the bacterial membrane, much like soap breaks down grease.
- **NaOH (alkaline):** Raises the pH drastically, causing denaturation of chromosomal and plasmid DNA. The high pH simultaneously denatures the bacterial proteins and unwinds all DNA strands.

3. Neutralization with Buffer P3

To stop the harsh alkaline reaction, **Buffer P3** is added. At this point, a critical physical property difference is exploited: **plasmid DNA is small and supercoiled**, so it can quickly reanneal and stay in solution. In contrast, **chromosomal DNA is large** and entangled, so it **aggregates and precipitates**. It contains:

- **Potassium acetate:** It neutralizes the pH, allowing DNA strands to renature. The potassium ions also form an insoluble complex with SDS and denatured proteins, which precipitate out of solution. Think of plasmid DNA like a rubber band—it snaps back into shape when tension is released. Genomic DNA is like a long rope—it tangles up and settles out.

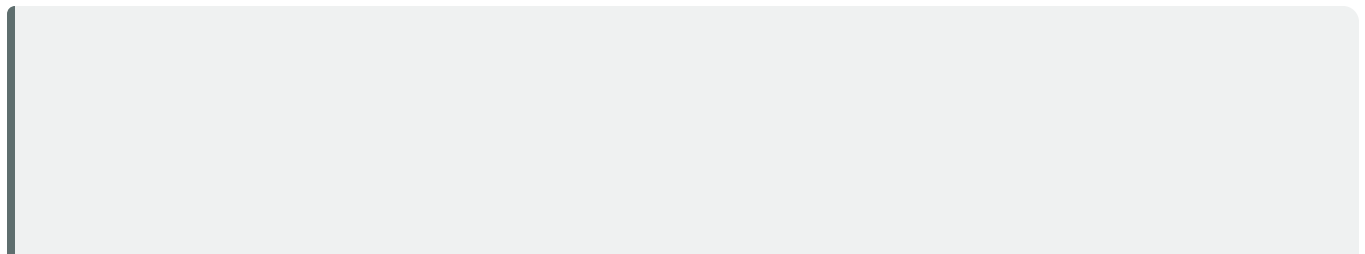
4. Centrifugation and DNA Binding

The resulting mixture is centrifuged at high speed. The supernatant is then passed through a **silica spin column**. Under high-salt conditions, DNA binds to silica. Other contaminants pass through during the wash steps.

- The **pellet** contains cell debris, precipitated proteins, and genomic DNA.
- The **supernatant** contains your clean plasmid DNA. DNA sticks to silica under certain salt conditions much like Velcro attaches when pressed together. Later, it can be eluted with a low-salt or water-based buffer.

5. Elution

Finally, the bound plasmid DNA is eluted with a low-salt buffer or sterile water. This DNA can now be used for downstream applications like **restriction digestion**, **sequencing**, **PCR**, or **transformation into another organism**. For example, if your plasmid carries the P450 gene, a restriction digest can be used to confirm its presence. Specific enzymes will cut the DNA at known sites, generating fragments of predictable sizes when run on an agarose gel. The pattern will tell you whether your plasmid carries the gene and whether it's in the right orientation.



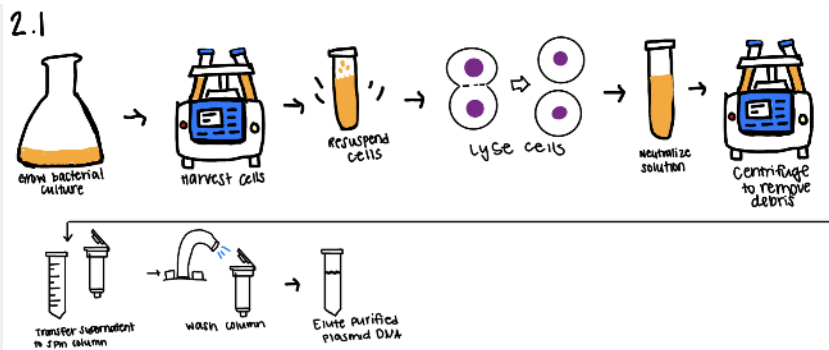


Image of a flow chart summarizing Plasmid Mini Prep. Image created by Dr. Victor Pham's student, Geneva Anh Thy Doan.

Plasmid Mini Prep Protocol (ZymoPURE Kit)

Goal: Isolate plasmid DNA from *E. coli* cultures containing either **pET3** or **pT7BM3** plasmids.

Materials Needed:

- Disinfectant spray
- 70% Ethanol (Flinn Scientific #E0013)
- Biohazard tip waste and liquid waste containers
- Microcentrifuge (Carolina #214061)
- 1.5 mL microcentrifuge tubes and tube rack
- *E. coli* culture with **pET3** plasmid
- *E. coli* culture with **pT7BM3** plasmid
- ZymoPURE Plasmid Miniprep Kit (Zymo #D4211)
- **Buffer P1** (keep refrigerated at 4°C)
- Micropipettes: P10, P20, P200, P1000 + sterile tips

Safety note: Disinfect your workspace and wash your hands before and after the procedure. Use proper pipetting technique and wear gloves throughout.

Procedure

1. Each team member should handle one culture type:
 - Group A: pET3 plasmid
 - Group B: pT7BM3 plasmid
2. Use a **P1000 pipette** to transfer **1.5 mL of bacterial culture** into a labeled 1.5 mL microcentrifuge tube.
 - Label clearly with group name and plasmid type.
3. Centrifuge the tube at **maximum speed ($\geq 10,000$ rpm)** for **1 minute**.
4. Carefully pour off or pipette out the supernatant without disturbing the white pellet at the bottom.
5. Repeat steps 2–4 **three more times**, using the **same microcentrifuge tube** each time.
 - This concentrates your bacteria into one pellet for better plasmid yield.
6. Add **250 μ L of Buffer P1** using a new tip.
7. Pipette up and down or vortex gently until the pellet is fully dissolved—**no clumps should remain**.
8. Add **250 μ L of Buffer P2** to the tube.
9. Gently **invert the tube 6 times** (do NOT vortex). The solution should turn **viscous and purple**.
10. Let the tube **sit for 3 minutes** (no more than 5 minutes).
11. Add **250 μ L of Buffer P3** and immediately **invert 6 times**. The solution will become **cloudy and yellow** as cell debris and proteins precipitate.
12. Spin the tube at **maximum speed for 5 minutes**.

13. You will see a **white pellet** at the bottom—this contains cellular debris and chromosomal DNA.
14. Using a **P1000 pipette**, carefully transfer all of the **clear supernatant** into a **new labeled 1.5 mL tube**, avoiding the white pellet.
 - Discard the tube with the pellet.
15. Add **260 µL of Binding Buffer** to the new tube.
16. Invert gently for 15 seconds.
17. Pour the mixture into the **Zymo spin column** (attached to a collection tube).
18. Let it sit for 1 minute, then **centrifuge for 1 minute**.
 - Discard the flow-through from the collection tube.
 - Save the Spin Column, which contains your plasmids in the silica beads.
 - Reattach the collection tube to the spin column
19. Add **800 µL of Wash 1**, centrifuge for 1 minute, and discard flow-through.
20. Add **800 µL of Wash 2**, centrifuge for 1 minute, and discard flow-through.
21. Add **200 µL of Wash 2** again, centrifuge for 1 minute, and discard flow-through.
22. Centrifuge the spin column **one more time for 1 minute** to remove any leftover ethanol from the wash buffers.
 - Residual ethanol can inhibit downstream reactions like PCR or restriction digest.
 - (If there's still liquid in the spin column, continue spinning it down until there's no more liquid remaining)
23. Move the spin column into a **new, labeled 1.5 mL microcentrifuge tube**.
24. Snap off the cap of the tube.
25. Add **25 µL of Elution Buffer** directly to the white silica membrane of the spin column.
26. Let it sit for **1 minute**, then **centrifuge for 2 minutes** at maximum speed.
27. Cap and label the tube as either **“Plasmid – pT7BM3”** or **“Plasmid – pET3a”**.
28. Place the tube in the freezer for storage.
 - Discard the used Silica spin column.
29. Measure your DNA via nanodrop

Parameter	Ideal Range	Notes
Concentration	>0.1 µg/µL	This also depends on how much bacteria was used from step 1
A260/A280	>1.8	Protein/phenol contamination indicator
A260/A230	>2.0	Salt/guanidine/alcohol contamination indicator

Summary

- Plasmids are essential tools for gene cloning, protein production, and synthetic biology.
- Transformation allows bacteria to amplify plasmids, creating a renewable source of genetic material.
- The mini prep process leverages molecular biology principles (like DNA binding to silica in salt) to purify plasmid DNA.
- Proper buffer usage and careful pipetting are critical for successful DNA extraction.

This page titled [Activity 2-1 - Plasmids Isolation](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Victor Pham](#).

Activity 2-2 - DNA Analysis via Restriction Enzymes

Learning Objectives

- Explain the purpose of using restriction enzymes to analyze plasmid DNA.
- Predict the outcome of restriction digests based on plasmid maps and insert size.
- Describe how bacterial transformation reveals phenotypic traits tied to gene expression.
- Interpret gel electrophoresis banding patterns to identify which plasmid contains the P450 gene.
- Understand why combining genotypic and phenotypic approaches increases experimental accuracy.

Definition: Term

- **Plasmid:** A small, circular DNA molecule that replicates independently of the bacterial chromosome. Used as a vector to carry genes of interest.
- **Restriction Enzyme:** A protein that cuts DNA at specific recognition sequences. Examples: EcoRI and BamHI.
- **Sticky Ends:** Single-stranded overhangs produced when restriction enzymes cut DNA asymmetrically. Useful for ligating inserts.
- **Blunt Ends:** DNA ends with no overhangs; both strands are cut at the same position. Harder to ligate.
- **Agarose Gel Electrophoresis:** A technique that separates DNA fragments by size using an electric current.
- **Transformation:** The process of introducing plasmid DNA into bacteria (e.g., *E. coli*).
- **Selectable Marker:** A gene (often conferring antibiotic resistance) used to identify cells that have taken up a plasmid.
- **Phenotypic Screening:** Observing traits (e.g., color change, growth on antibiotics) in transformed bacteria to infer plasmid content.

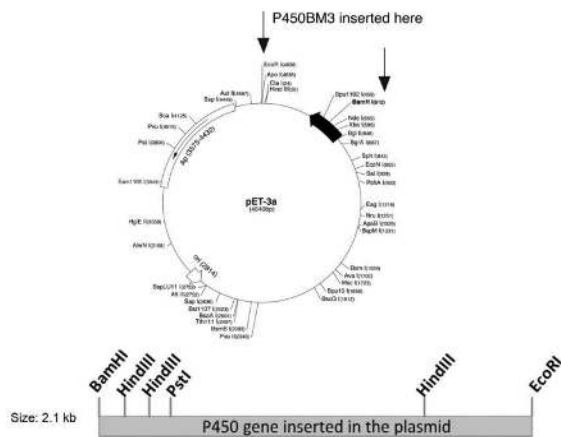
Identifying the P450 Plasmid

In this part of your molecular biology project, your goal is to determine which of the two plasmids contains the gene encoding **cytochrome P450**, an important enzyme involved in drug metabolism. To make this identification, you will use **two complementary molecular techniques**: (1) **restriction enzyme digestion followed by gel electrophoresis**, and (2) **bacterial transformation followed by phenotypic observation**.

Why use two approaches? In science, especially molecular biology, it's essential to confirm findings through **multiple lines of evidence**. Each method provides a different type of information: one gives a **genotypic view** (DNA content and size), while the other shows the **phenotypic outcome** (what traits the plasmid expresses in a living organism).

Direct DNA Analysis via Restriction Enzymes and Gel Electrophoresis

One of the most reliable ways to determine the identity of a plasmid is to use **restriction enzymes** to cut the plasmid DNA at specific sites. These enzymes act like molecular scissors, cutting at unique, short **palindromic sequences** in the DNA. Suppose you have **Plasmid A** and **Plasmid B**, and only one of them has the P450 gene inserted using **EcoRI** and **BamHI**. If you digest both plasmids with these enzymes and run the digested DNA on an **agarose gel**, the plasmid with the insert (let's say Plasmid A) will show two bands: one corresponding to the vector backbone and one to the inserted P450 gene. Plasmid B, without the insert, would show only one band—the empty backbone. By comparing the **pattern of DNA bands** (called a *restriction digest pattern*) on the gel, you can infer which plasmid contains the insert. This process is known as **restriction mapping**.



Confirming by Phenotypic Analysis After Transformation

After digesting and analyzing the DNA, you'll use another technique: **bacterial transformation**. This involves introducing each plasmid into *E. coli* cells and observing the **resulting traits**. Some phenotypic traits may be directly or indirectly linked to the expression of the inserted gene. If the P450 gene is expressed under a strong promoter (like T7), and this protein causes a noticeable color change in the media or confers resistance to a certain chemical, transformed bacteria carrying the correct plasmid (pT7BM3) might show this trait. Those carrying the empty plasmid would not. Even if there's no clear visible phenotype from the P450 itself, the plasmid may carry a **selectable marker** (e.g., ampicillin resistance), which ensures that only transformed bacteria grow on ampicillin-containing plates.

Plasmids

Plasmids are small circular DNA molecules that replicate independently of the bacterial chromosome. Because of their ease of use and modifiability, plasmids are fundamental tools in molecular biology labs.

1. Origin of Replication (Ori):

- The site where replication begins.
- Ensures that the plasmid is copied inside *E. coli*.
- Without it, the plasmid would be lost during cell division.

2. Selectable Marker Gene:

- Allows you to identify which *E. coli* cells have successfully taken up the plasmid.
- Commonly used markers include **antibiotic resistance genes**.
- For example, **pET3a** plasmid encodes **β -lactamase**, an enzyme that breaks down **ampicillin**. Only bacteria with this plasmid will grow on ampicillin plates.

3. Multiple Cloning Site (MCS):

- A short DNA region containing many restriction sites.
- Facilitates the insertion of foreign DNA (like your P450 gene).
- Example: MCS with **EcoRI** and **BamHI** allows the gene to be inserted at these specific sites.

Restriction enzymes (also called restriction endonucleases) are proteins that cut DNA at specific sequences. These enzymes are isolated from bacteria, where they originally served as a defense mechanism against viruses. Each enzyme cuts at a precise spot within or near this site. Many restriction enzymes recognize **palindromes**—a sequence that reads the same forwards and backwards. For instance:

- **EcoRI** recognizes: 5' GAATTC 3'
- **BamHI** recognizes: 5' GGATCC 3'



There are also two different types of restriction cuts (**Figure 2.2.1**). One is a **Sticky ends**, where there is an overhanging single-stranded sequence left after cutting. These can base pair with complementary sequences. For example, EcoRI leaves sticky ends, which is useful for ligating (joining) DNA fragment. The other type is the **Blunt ends**, where there is no overhangs, both DNA strands end at the same position. This makes it harder to ligate, but still usable with blunt-end cloning techniques.

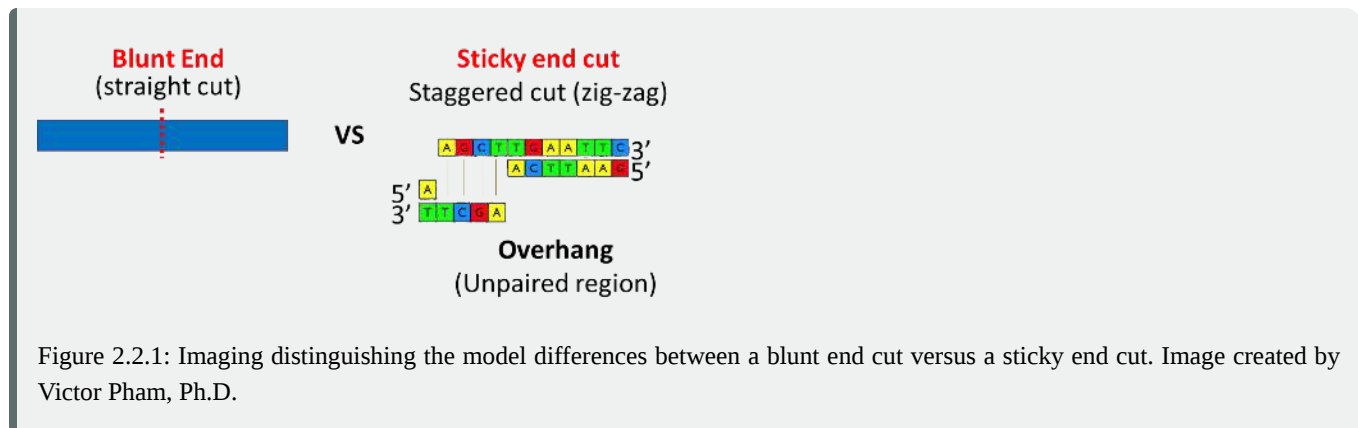


Figure 2.2.1: Imaging distinguishing the model differences between a blunt end cut versus a sticky end cut. Image created by Victor Pham, Ph.D.

When you cut a plasmid with restriction enzymes, you generate **predictable DNA fragment sizes**. These can be visualized on an agarose gel stained with a DNA dye like ethidium bromide or SYBR Safe. The pattern of bands acts like a fingerprint for that plasmid.

Plasmid A cut with EcoRI and BamHI:

- Vector backbone: 5,000 bp
- P450 insert: 1,500 bp
 - → On the gel, you'll see **two bands**, one at ~5 kb and another at ~1.5 kb.

Plasmid B (empty vector):

- Only 5,000 bp
 - → One band at ~5 kb.

This difference confirms which plasmid contains the insert.

Components of a Restriction Digest Reaction

A standard restriction digest reaction is carefully formulated using several critical components, each serving a distinct role in ensuring successful cleavage of plasmid DNA. The primary component is the **plasmid DNA**, which contains the sequence of interest and serves as the target for enzymatic cleavage. To cut the DNA at specific sites, **restriction enzymes** are used. These enzymes recognize unique short DNA sequences—called recognition sites—and cleave the DNA precisely at those locations, allowing for structural analysis or modification of the plasmid.

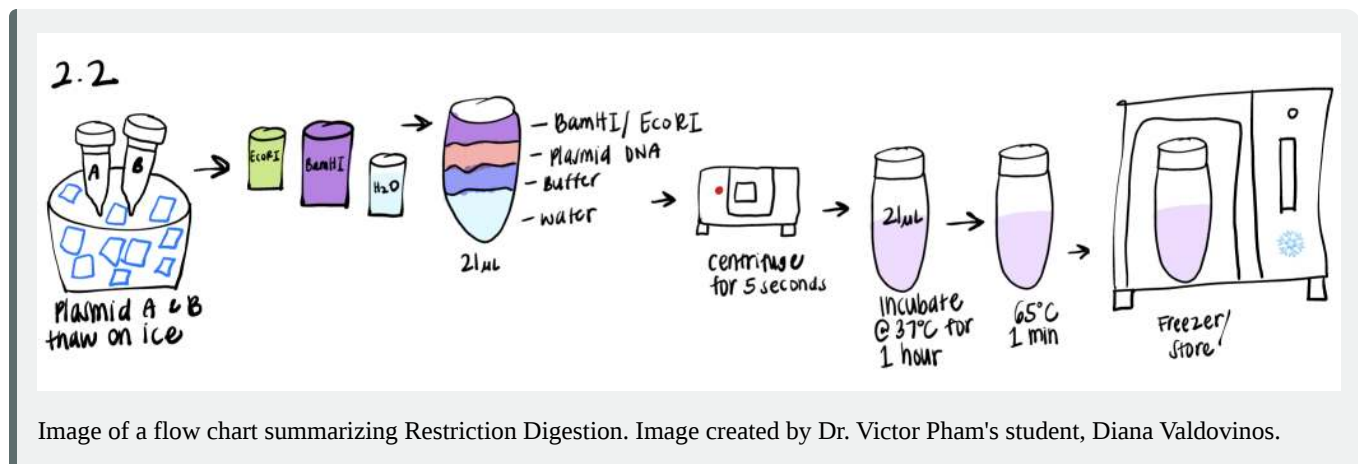
The efficiency of these enzymes depends on the chemical environment, which is maintained by a specially formulated **buffer**. Each restriction enzyme requires a particular buffer composition that includes the optimal pH and ion concentrations for catalytic activity. To further support enzyme stability, **Bovine Serum Albumin (BSA)** is often included. BSA prevents the enzyme from adhering to the walls of the microcentrifuge tube and helps maintain its functional three-dimensional structure, especially in dilute reactions. Additionally, **RNase** is frequently added to eliminate RNA contaminants that may be present from the plasmid preparation process, ensuring that the final gel analysis is not confounded by unwanted nucleic acid bands.

Once the digestion is complete, the enzyme must be **heat-inactivated**—typically by raising the temperature to denature its structure—thereby stopping further enzymatic activity. Before loading the digested DNA onto an agarose gel for electrophoresis, a

loading dye mixture is added. This dye contains multiple important additives. **Glycerol** increases the density of the sample so that it settles into the gel wells. **Bromophenol blue** serves as a tracking dye, migrating through the gel and allowing researchers to visually monitor DNA progression. The dye mixture also contains **SDS** (a detergent that denatures proteins) and **EDTA** (a chelating agent that binds divalent metal ions like Mg^{2+}), both of which help stop any residual enzyme activity by interfering with the enzyme's structure or required cofactors.

During gel electrophoresis, different DNA conformations can behave uniquely. **Supercoiled plasmids**, which are tightly compacted circular forms of DNA, migrate more rapidly through the gel and appear lower than expected. This compact shape can sometimes produce **smears or streaks**, especially if the digestion was incomplete. In contrast, **linear DNA**, generated by complete restriction digestion, migrates at a consistent rate according to its size, appearing as clear bands. **Nicked or relaxed DNA**, which results from single-strand breaks in the plasmid, migrates more slowly and often appears as a separate band above the linear form.

To conclusively identify whether a plasmid contains your gene of interest—such as the P450 gene—it is best to use a **dual-approach** strategy. First, **restriction digestion** provides a molecular fingerprint of the plasmid by revealing the size and number of DNA fragments. This allows you to infer the presence of inserted sequences. Second, **transformation and phenotypic screening** in *E. coli* cells offer functional confirmation by showing whether the gene is expressed or confers a specific trait (e.g., antibiotic resistance or color change). By integrating both molecular and phenotypic evidence, researchers significantly increase the **accuracy and reliability** of their findings. This is a standard practice in molecular biology and ensures the correct identification of the P450-containing plasmid before proceeding to more advanced applications like expression studies, enzymatic assays, or protein purification.



Restriction Digest Protocol

Objective: To perform a restriction enzyme digestion of two plasmids (A and B) using EcoRI and BamHI, allowing you to determine which plasmid contains the P450 gene.

Materials Needed:

- Micropipettes and sterile tips (filter tips preferred)
- 1.5 mL microcentrifuge tubes
- Nuclease-free water (e.g., Invitrogen, #AM9906)
- Plasmid DNA samples (from previous lab – plasmid A and plasmid B)
- EcoRI restriction enzyme kit (NEB, #R0101T)
- BamHI restriction enzyme kit (NEB, #R0136T)
- 10X Restriction buffer (comes with enzyme kits)
- Water bath set to 37°C
- Water bath set to 65°C

Procedure:

1. Remove the plasmid A and plasmid B tubes from the freezer and allow them to thaw completely on ice.
2. Using a marker, clearly label two new 1.5 mL microcentrifuge tubes:
3. One as "**Plasmid A Digest**"
4. The other is "**Plasmid B Digest**"
5. Into each labeled tube, add the following components in the given order. Use fresh pipette tips for each transfer to avoid contamination:
 - This protocol will be for a **double digestion** using both *EcoRI* and *BamHI*. If you're doing a **single digest**, just add 2 μL of either *EcoRI* or *BamHI*, and increase the water volume to 12 μL instead of 10 μL .
 - **Final reaction volume: 21 μL .** (Figure 2.2.2) (The buffers and reagents are added in no particular order.)
 - 5 μL of plasmid DNA (either A or B)
 - 2 μL of *EcoRI* enzyme
 - 2 μL of *BamHI* enzyme
 - 2 μL of 10X restriction buffer
 - 10 μL of nuclease-free water
6. Spin down your samples using a microcentrifuge machine for under 5 seconds. (This will bring all condensation and side liquid down to the bottom)
7. Place your tubes into a 37°C water bath and incubate for **1 hour**. This allows the enzymes to cut the DNA at their recognition sites.
8. After digestion is complete, place the tubes at 65°C for 1 minute, spin down for under 5 seconds, and then store in the freezer until you're ready to load them on an agarose gel for analysis.

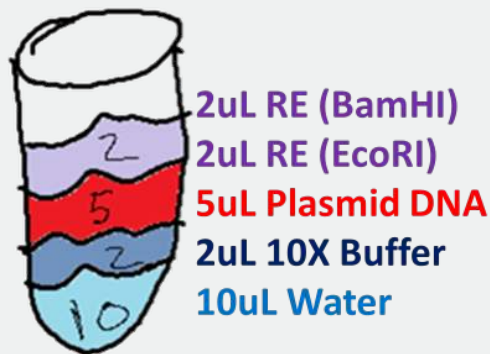


Figure 2.2.2: Simplified image depicting how much volume of each reagent. Image created by Victor Pham, Ph.D.

Post-Lecture Objectives

- Setting up restriction digest reactions with proper enzyme and buffer combinations.
- Using micropipettes with sterile technique.
- Running and interpreting an agarose gel.
- Transforming bacteria and analyzing colony growth.
- Synthesizing multiple lines of evidence to draw conclusions in a molecular biology experiment.
- **DNA Banding Patterns:** You should be able to match your gel results to a predicted pattern based on the expected insert size (e.g., 5,000 bp for the backbone and 1,500 bp for the P450 insert).
- **Genotypic Confirmation:** If you see two bands on the gel, you likely have the plasmid containing the insert (P450 gene). One band likely means it's the empty vector.
- **Phenotypic Confirmation:** Did the transformed colonies survive on ampicillin? Did you observe any color changes or growth differences that indicate gene expression?

- **Data Cross-Validation:** You've learned that combining both approaches—genotypic (gel) and phenotypic (growth/trait)—increases your confidence in identifying the correct plasmid.

This page titled [Activity 2-2 - DNA Analysis via Restriction Enzymes](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Victor Pham](#).

Project 2-3 - Identifying Plasmid Restriction Mapping Using Agarose Gel Electrophoresis

Learning Objectives

- Explain the basic principles of agarose gel electrophoresis.
- Describe how DNA size and shape affect migration through a gel.
- Identify the roles of loading dye, DNA ladder, and SYBR/Gel Green.
- Predict how supercoiled, linear, and nicked plasmid DNA will migrate differently.
- Outline how agarose gel electrophoresis can be used to distinguish between plasmid A and plasmid B.

Definition: Term

- **Agarose Gel Electrophoresis:** A technique that separates DNA fragments by size and shape using an electric field through a gel matrix.
- **Agarose:** A polysaccharide derived from seaweed that forms a porous gel used to separate DNA fragments.
- **DNA Ladder:** A molecular ruler—a standard set of DNA fragments of known sizes used to estimate the length of unknown DNA samples.
- **Loading Dye:** A dye that contains glycerol and tracking dyes to help load DNA into wells and visualize the progress of electrophoresis.
- **SYBR Green / Gel Green:** Fluorescent dyes that intercalate into DNA and glow under blue light, allowing visualization of DNA bands.
- **Supercoiled DNA:** A tightly coiled plasmid form that migrates faster than linear or relaxed circular forms.
- **Linear DNA:** DNA that has been cut with restriction enzymes, resulting in a straight, uncoiled molecule.
- **Relaxed Circular DNA:** An open-circle form of plasmid DNA that migrates more slowly than supercoiled or linear forms.
- **Mutagen:** A chemical agent that can cause genetic mutations (e.g., SYBR Green).
- **Electrophoresis Buffer (TBE):** A solution that conducts electricity and maintains pH during gel electrophoresis.

Identifying Your P450 Plasmid Using Agarose Gel Electrophoresis

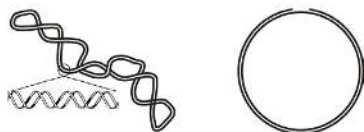
In this lab, you will identify which plasmid (A or B) contains the gene for the P450 enzyme. You'll do this using **agarose gel electrophoresis**, a technique that separates DNA fragments by size and shape so we can visualize and analyze them.

Agarose Gel Electrophoresis

Agarose gel electrophoresis is a common method used in molecular biology to separate DNA fragments. The gel is made from **agarose**, a sugar polymer derived from seaweed, that forms a mesh-like structure when cooled. This structure creates a molecular “sieve” through which DNA molecules can move. DNA is **negatively charged** due to its **phosphate backbone**. When an electric field is applied across the gel, DNA fragments will migrate from the **negative end (black)** toward the **positive end (red)**. Smaller DNA fragments can weave through the pores in the gel more easily and therefore move **faster and farther** than larger fragments.

DNA fragments don't always migrate based on size alone—**shape matters too**. When analyzing your gel, you may see bands that represent different shapes of the same DNA molecule. Plasmid DNA can come in different forms:

- **Supercoiled DNA** is tightly wound and very compact. It migrates the **fastest** through the gel.
- **Linear DNA**, which results from cutting circular DNA with restriction enzymes, moves at a **moderate** speed.
- **Relaxed circular DNA**, such as nicked plasmids, is more open and bulky, so it moves **slowly**.

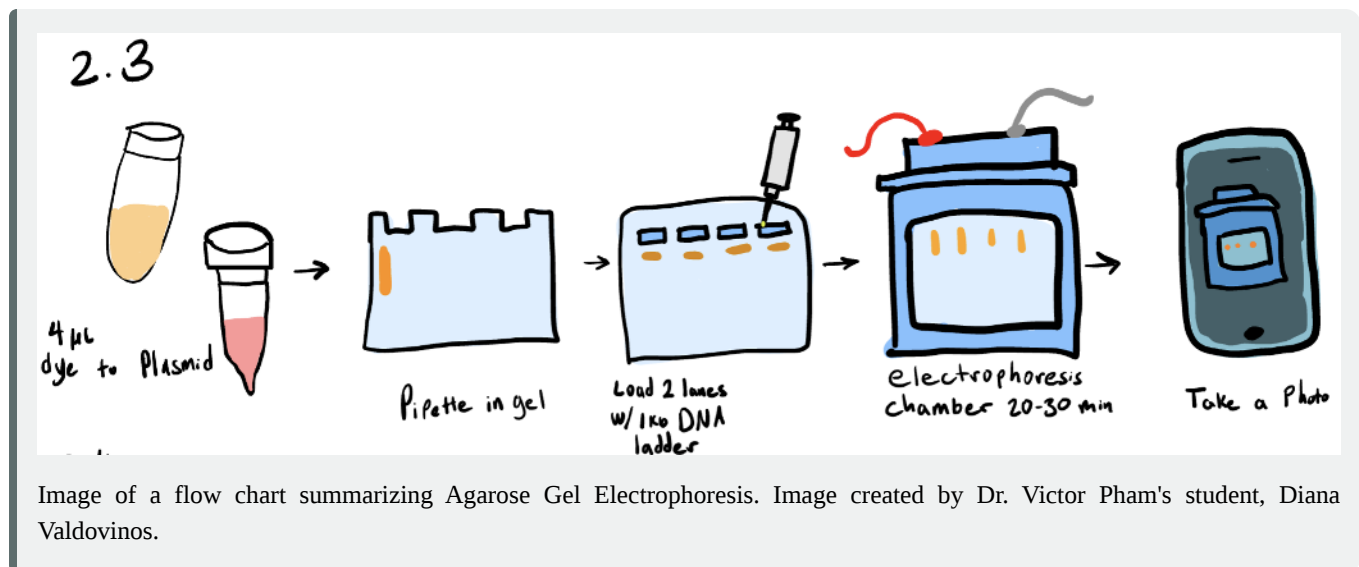
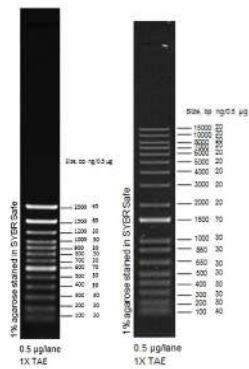


DNA is colorless, so we need a **stain** to see it in the gel. **SYBR Green** (or **Gel Green**) is a **fluorescent dye** that slips in between DNA base pairs and **glows green** under blue light. This allows you to visualize the DNA bands safely without using UV light (unlike the older dye, ethidium bromide).

⚠ SYBR Green is a **mutagen**, which means it can alter DNA and is potentially harmful. Always wear gloves and avoid skin contact.

To estimate the size of your DNA fragments, you'll use a **DNA ladder**. This is a commercially prepared sample with **known DNA fragment sizes**. You can compare your unknown bands to the ladder like a ruler to determine how long each fragment is (in base pairs or kilobases).

Before loading DNA samples into the gel, you'll mix them with a **loading dye**, which contains **Glycerol** that makes the sample sink into the well, and contains **tracking dyes**, like bromophenol blue, that migrate ahead of the DNA and help you monitor how far the sample has run.



Materials

- 0.8% agarose gel
- SYBR Green (Invitrogen #S33102) or Gel Green (EmbiTec #EC-1995)
- MiniOne Gel Electrophoresis System
- 1x TBE Buffer (Fisher Scientific #BP13334)
- Digested plasmids A and B (from your previous lab)
- 6X Loading dye
- 1 kb DNA Ladder (Invitrogen #10488-085)
- Micropipettes and filter tips

Procedure

1. Add **4 μL** of 6X loading dye to your **20 μL** digested plasmid sample. Mix gently by pipetting up and down.
 2. Carefully pipette the full **24 μL** of your prepared sample into the assigned well in the gel.
 3. The instructor will load **two lanes** with **10 μL of 1 kb DNA ladder** each. This gives you a reference to compare the sizes of your fragments.
 4. Close the lid on the electrophoresis chamber and run the gel using the MiniOne system for approximately **20–30 minutes**, or until the tracking dye has migrated halfway.
 5. Use the MiniOne viewer to turn on the blue light and observe the glowing bands. **Take a picture** of your gel using your phone camera for later analysis.
- ⚠ Always **wear gloves** when handling gels, stains, and buffers.
 - Never look **directly into the blue light**—it can damage your eyes.
 - ⚠ SYBR Green and Gel Green are **potential carcinogens**. Avoid skin and eye contact.
-

Post-Lecture:

Objectives

- Analyze an agarose gel to determine which plasmid contains the P450 gene.
 - Interpret differences in band patterns based on DNA size and structure.
 - Explain the role of each reagent and material used during gel electrophoresis.
 - Apply electrophoresis results to support conclusions about molecular identity.
 - Reflect on common mistakes and best practices when loading and running gels.
-

This page titled [Project 2-3 - Identifying Plasmid Restriction Mapping Using Agarose Gel Electrophoresis](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Victor Pham](#).

Activity 2-4 - Bacterial Transformation with Engineered Plasmid

Learning Objectives

- Define *bacterial transformation* and explain its historical significance.
- Identify the role of plasmids in transformation experiments.
- Describe how calcium chloride treatment and heat shock facilitate DNA uptake.
- Understand the purpose of selective vs. non-selective plates.
- Predict what transformation results might look like and what they mean.

Definition: Term

- **Bacterial transformation:** A process by which bacteria take up foreign DNA from their surroundings.
- **Plasmid:** A small, circular DNA molecule separate from chromosomal DNA that can carry genes of interest.
- **Competent cells:** Bacteria that are capable of taking up external DNA due to increased membrane permeability.
- **Calcium chloride (CaCl₂):** A chemical used to prepare competent cells by neutralizing membrane charges.
- **Heat shock:** A brief, controlled temperature increase that helps plasmid DNA enter the cell.
- **Selective medium:** Growth medium that includes an antibiotic (e.g., ampicillin) to allow only transformed cells to grow.
- **β-lactamase:** An enzyme encoded by a resistance gene on the plasmid that deactivates ampicillin.
- **LB broth/agar:** Nutrient-rich medium that supports bacterial growth (stands for Luria-Bertani)

What is Bacterial Transformation?

Bacterial transformation is a foundational technique in molecular biology in which foreign DNA, typically in the form of plasmids, is introduced into bacterial cells. This concept has its roots in the classical experiment by **Frederick Griffith in 1928**, who observed that non-virulent *Streptococcus pneumoniae* could become virulent when exposed to heat-killed virulent strains. While Griffith didn't know it at the time, the key takeaway was that **some "transforming principle" was being transferred**—a mystery later solved in 1944 by **Avery, MacLeod, and McCarty**, who demonstrated that **DNA is the hereditary material** responsible for transformation. These discoveries laid the groundwork for modern genetic engineering.

In the lab, this process is harnessed to introduce **engineered plasmids**—small, circular DNA molecules—into bacterial cells like *E. coli*. These plasmids often carry genes of interest, such as those encoding proteins (e.g., cytochrome P450 enzymes), or selectable markers like antibiotic resistance.

Competency

For transformation to occur, bacterial cells must be in a special state known as **competence**, where their cell membranes are temporarily more permeable to DNA. While some bacteria like *Bacillus subtilis* can naturally take up DNA from the environment, **most laboratory strains of *E. coli* are not naturally competent**. Therefore, they must be **artificially induced into a competent state**, usually by chemical or electrical methods.

In this experiment, you are using the **calcium chloride (CaCl₂) method** to induce competency. Treating *E. coli* cells with ice-cold calcium chloride makes the cell wall more permeable, possibly by neutralizing the negative charges on both the DNA and the cell membrane, reducing electrostatic repulsion. After this treatment, cells are kept on **ice** to stabilize them in a competent state until they are ready for transformation.

The actual transformation process begins by mixing the plasmid DNA (in this case, your **P450 plasmid**) with the competent *E. coli* cells. This is followed by a brief but critical step known as **heat shock**, typically at **42°C for 30–45 seconds**. The sudden temperature change is thought to create a thermal imbalance across the membrane, forming transient pores through which the plasmid DNA can enter the bacterial cytoplasm. **Not all cells will take up the plasmid**, but those that do have the potential to express the genes it carries. Without the heat shock, transformation efficiency is drastically reduced, although some DNA may still be taken up passively during the incubation on ice. This highlights the **importance of the heat shock step** in increasing transformation success.

Immediately after heat shock, the transformed cells are incubated with a **nutrient-rich medium (like LB broth)** at **37°C for 30–60 minutes**. This recovery phase serves several critical purposes. It allows the cells to repair their membranes. It gives time for **gene expression** to begin—especially the expression of the **β -lactamase enzyme**, encoded by the **Ampicillin resistance gene** on your plasmid. This enzyme hydrolyzes ampicillin, neutralizing its bactericidal effect. Without this recovery time, even successfully transformed cells may not survive antibiotic selection because they haven't had a chance to produce enough β -lactamase. This is a classic example of how **gene expression timing impacts cell survival in selective environments**.

After recovery, the transformed cells are **plated on two types of agar plates**:

- **LB + Ampicillin Plate:** This is the **selective plate**, designed to allow only the bacteria that have successfully taken up the plasmid to grow. If a colony forms here, it suggests the cell has incorporated the plasmid and is producing β -lactamase.
- **LB Only Plate:** This is a **control plate**. All bacteria, transformed or not, can grow here because there is no antibiotic to inhibit them. This plate helps verify that the cells were viable post-transformation and that any lack of growth on the Ampicillin plate is not due to cell death or poor handling.

This comparison is key to understanding the role of **selectable markers** in molecular cloning. Only transformed cells that express the **Ampicillin Resistance gene** will survive on the Ampicillin plate, allowing for **easy screening** of successful transformations.

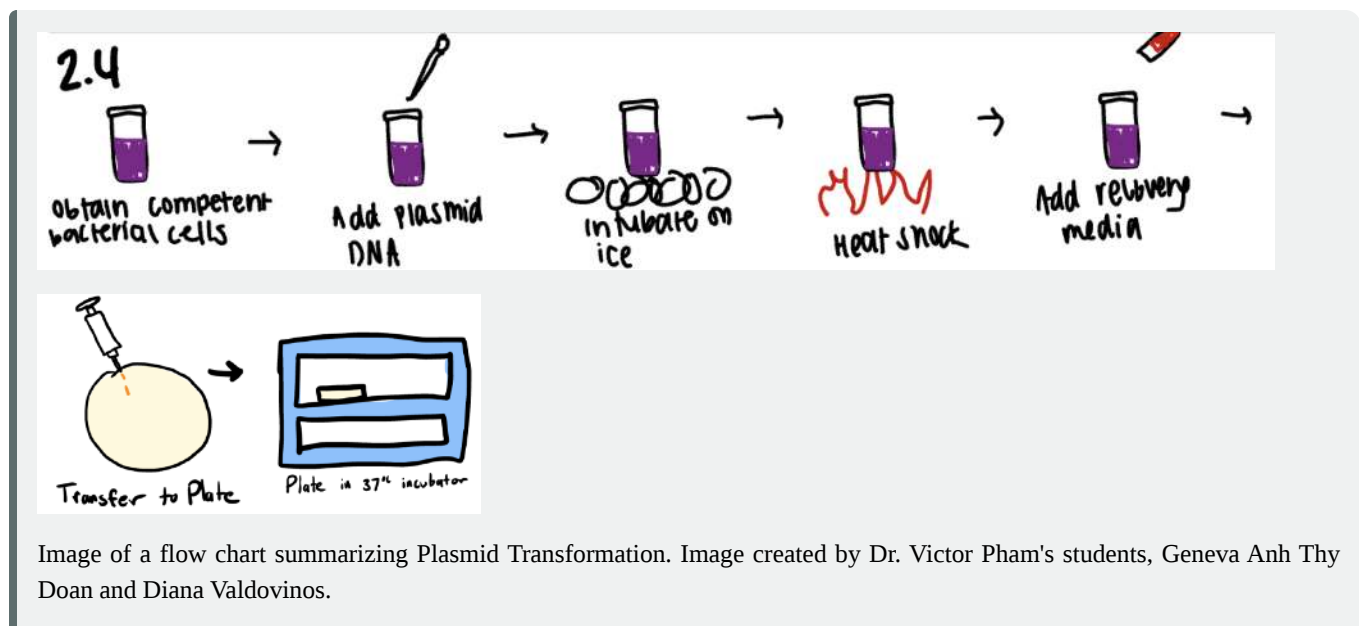
When you observe the plates in the next class, you should expect:

- **Numerous colonies on the LB-only plate**, representing all viable cells (both transformed and untransformed).
- **Fewer colonies on the LB + Ampicillin plate**, representing only those cells that were successfully transformed with the P450 plasmid.

The presence of colonies on the selective plate **confirms plasmid uptake**. If no colonies are seen on the Ampicillin plate but are present on the LB-only plate, this indicates that **transformation was unsuccessful** or **recovery time was insufficient** for the expression of the resistance gene.

In this experiment, your plasmid encodes a **cytochrome P450 enzyme**, which is often used in drug metabolism studies or biocatalysis research. By transforming this plasmid into *E. coli*, you're essentially programming the bacteria to produce the P450 enzyme. If you later induce the expression of this gene (e.g., using IPTG if under a lac promoter), you can purify the enzyme for experiments like:

- Substrate metabolism assays
- Spectrophotometric P450 activity measurements
- Structural or kinetic studies



Materials:

- **Plasmid DNA** (isolated in the previous lab)
- **Competent *E. coli*** cells (prepared using 0.1 M CaCl₂)
- **Micropipettes** and sterile **tips**
- **LB broth** (Luria-Bertani media) – 1 tube per group
- **LB agar plates** – 2 per group (without antibiotics)
- **LB agar + Ampicillin (100 µg/mL)** plates – 2 per group
- **42°C water bath** with a thermometer (temperature must be exact)
- **37°C shaking incubator** and **37°C still incubator**
- **Sterile inoculation loop** or **bacterial spreader**
- **Bunsen burner** with sparker (for sterilization)
- **70% ethanol** and **bleach** (for surface disinfection)
- **0.1 M Calcium Chloride solution** (Carolina #851840)

Transformation Protocol

1. Transfer **50 µL of your competent bacteria into a 1.5mL centrifuge tube**.
 - **(Optional)** Mix in 50µL of 0.1M CaCl₂ to your bacteria
2. Place the tube on ice and let it sit for **at least 2 minutes**.
3. Add **10 µL of your P450 plasmid DNA (Not Digested)** into the tube with the competent cells. **Gently** mix via pipetting—**do not cause bubbles**. Make sure all liquid is at the bottom of the tube. Place the tube back on ice for **another 2 minutes**.
4. Carefully place the tube in a **42°C water bath** for **exactly 45 seconds**. This creates a thermal imbalance that helps the plasmid enter the cells. Remove the tube promptly and return it to ice for 1 minutes.
5. **(Optional)** Add **500 µL of LB broth** (room temperature) to the tube. This provides nutrients for the cells to begin recovering and expressing the plasmid genes.
 - Place the tube in a **37°C shaking incubator** (or water bath shaker) and incubate for **30 minutes**. This allows the cells to begin producing the **β-lactamase enzyme**, which will protect them from ampicillin.
6. Using a pipette, transfer the half of the mixture onto one **LB agar + Ampicillin plate** and the other half to an **LB-only plate (Optional)**. Use a sterile spreader or gently shake the plate side to side to spread the liquid evenly.
7. Let the liquid absorb into the plate for about **1 minutes**. Then **invert** the plate (agar side up) and place it in a **37°C incubator overnight**.
8. During the next lab session, examine both the **LB-only plate** and the **LB + Ampicillin plate**:
 - Count the number of bacterial colonies on each.
 - Note differences in growth between the plates.
 - Observe and record any differences in **colony color or appearance** on the Ampicillin plate, and compare your results with your classmates.
 - Record all your findings in your lab notebook.

Post-Lecture Objectives

- Explain each step of the transformation process and its purpose.
- Interpret your results from the LB and LB + Ampicillin plates.
- Determine whether transformation was successful and why.
- Connect transformation results to the concept of gene expression (e.g., production of β-lactamase or P450).
- Explain how selectable markers help us identify transformed cells.

Key Takeaways

- Transformation is a controlled process to introduce new genes into bacteria.
- Success depends on precise steps: making cells competent, adding DNA, heat shock, and recovery.
- Selective media (like Ampicillin plates) ensure that only transformed cells grow.
- The expression of new genes (like P450) can now be studied or used for further applications.
- This technique is foundational in everything from cloning to synthetic biology to drug development.

This page titled [Activity 2-4 - Bacterial Transformation with Engineered Plasmid](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Victor Pham](#).

Activity 3-1 - Identifying Genetically Modified Organisms Through DNA Analysis

Learning Objectives

- Define what genetically modified organisms (GMOs) are and explain how they are created.
- Describe the purpose of common genetic elements found in GM crops (e.g., 35S promoter, NOS terminator).
- Understand the societal context and global regulation of GMO labeling.
- Outline the steps required to detect GMOs in food using DNA extraction, PCR, and gel electrophoresis.
- Interpret why specific primers are used in the detection of GM markers.
- Prepare PCR-ready primer dilutions and determine how to calculate and normalize DNA concentration for reactions.

Definition: Term

- **Genetically Modified Organism (GMO):** An organism whose genome has been altered using recombinant DNA technology to express traits like pest resistance or improved nutrition.
- **35S Promoter:** A regulatory sequence from the Cauliflower Mosaic Virus that drives strong gene expression in plant cells. Common in GM crops.
- **NOS Terminator:** A sequence from *Agrobacterium tumefaciens* that signals the end of transcription; used to terminate transgene expression.
- **PCR (Polymerase Chain Reaction):** A technique to amplify specific DNA sequences, making millions of copies of a target region.
- **Primer:** A short strand of nucleotides that serves as a starting point for DNA synthesis during PCR.
- **SYBR Safe Dye:** A fluorescent dye that binds DNA and allows visualization under UV or blue light during gel electrophoresis.
- **DNA Extraction:** The process of isolating DNA from a biological sample.
- **Nanodrop Spectrophotometry:** A method to measure DNA concentration and purity using absorbance at 260 nm.
- **GMO Marker:** A DNA sequence specifically inserted into GM organisms, often targeted for detection (e.g., 35S, NOS, selectable markers).
- **Control Reaction:** A reaction used to confirm that your PCR or extraction worked (positive) or to check for contamination (negative).

Are You Eating Genetically Modified Foods?

Genetically modified (GM) foods are derived from organisms whose DNA has been altered using genetic engineering techniques. These modifications are often made to improve crop yield, pest resistance, or nutritional content. In the United States, food products must be labeled as containing GMOs only if **more than 5%** of their content is genetically modified. For instance, a breakfast cereal made primarily from genetically engineered corn would be labeled only if the GM content exceeds that threshold. Interestingly, the term “**GMO-free**” can still be used if the food contains **less than 5% GM ingredients**. By contrast, many countries in Europe and Asia have stricter policies, requiring labeling for GM content exceeding just 1%. This variation in global standards highlights the ongoing debate around transparency and consumer rights when it comes to food labeling.

The central goal of this lab is to engage students in a real-world application of molecular biology: detecting GMOs in common foods. Students will collect plant-based items from a grocery store—like corn chips, soy milk, tofu, or veggie burgers—and analyze them in the lab to see if they contain genetically modified DNA. The investigation unfolds through five key stages:

1. **Food Collection** – Students select processed or raw plant-based food items suspected of containing GM ingredients.
2. **DNA Extraction** – Genetic material is extracted from the food samples using a simple lysis buffer and ethanol precipitation method.
3. **PCR Amplification** – Specific DNA sequences associated with GMOs are amplified using the polymerase chain reaction (PCR).

4. **Gel Electrophoresis** – Amplified DNA fragments are separated and visualized to confirm the presence or absence of GMO-specific sequences.
5. **RT-qPCR (Reverse Transcription & Quantitative PCR)** – An advanced analysis is optionally performed to quantify gene expression levels or detect RNA transcripts, particularly useful if the GM trait involves a transgene under active expression in the plant.

This activity introduces students to a variety of sophisticated molecular biology techniques, reinforcing both theoretical and practical concepts:

- **Genetic Engineering and Transformation:** GMOs are created by inserting new genes into an organism's genome. For example, Bt corn contains a gene from the bacterium *Bacillus thuringiensis* that produces a protein toxic to insect pests. The lab helps students understand how these genes are introduced and expressed.
- **DNA Transcription and Translation:** If a plant is genetically modified to produce a protein (e.g., a pest-resistant enzyme), its DNA is first transcribed into mRNA, then translated into the protein. RT-qPCR can help detect these transcripts, showing active expression of the transgene.
- **Gene Regulation:** Not all genes are turned on all the time. GM traits often include regulatory sequences like promoters to control when and where a gene is active. The presence of a promoter like the cauliflower mosaic virus 35S promoter (CaMV 35S) is a common target in GMO detection.
- **DNA Replication and PCR:** PCR is a technique that mimics natural DNA replication but in a controlled, amplified way. Using primers specific to known GM sequences, students can selectively copy parts of the genome to see if their food contains those sequences.
- **Plant Development and Physiology:** Understanding which plant tissues are best for DNA extraction, and how gene expression varies in different plant parts, gives students insight into basic plant biology.
- **Agricultural and Environmental Science:** The project connects lab science to agricultural practices—why certain GM crops are developed, how they're used in farming, and what the broader ecological consequences might be.

GMO Detection Guided Approach

The GMO detection kit used in this activity encourages a **guided inquiry-based learning** method. This means students follow structured protocols but also make critical decisions along the way—like choosing appropriate controls, troubleshooting PCR issues, and interpreting gel results. Students will ask questions like:

- **Was DNA extraction successful?** A control PCR reaction targeting a plant gene (e.g., tubulin or rbcL) ensures the sample contains amplifiable plant DNA.
- **Did PCR work as expected?** Positive and negative controls help interpret results. A successful PCR will show clear bands on a gel, while no bands could suggest failed amplification.
- **Is there contamination?** Contamination can lead to false positives. Students will use sterile technique and proper controls to validate their results.

GM Crops

The experiment doesn't end at the gel. Students are encouraged to explore the broader implications of GMO technology by engaging in a structured debate on its risks and benefits.

Concerns About GM Crops:

- **Super-Weeds:** When herbicide-resistant crops like "Roundup Ready" soybeans cross-pollinate with wild relatives, the result can be herbicide-resistant weeds that are harder to control.
- **Super-Bugs:** Insects exposed to Bt crops over many generations may evolve resistance, rendering the Bt trait ineffective.
- **Allergic Reactions:** Introducing new genes can lead to the expression of novel proteins. If these proteins resemble allergens, they could pose a health risk to sensitive individuals.
- **Antibiotic Resistance:** Some GM crops use antibiotic resistance genes as markers during development. There is concern these genes could be transferred to gut bacteria, potentially contributing to antibiotic resistance.
- **Public Health Risks:** While most GMOs are considered safe, critics argue that long-term health effects have not been fully studied, particularly for traits affecting metabolism or protein expression.

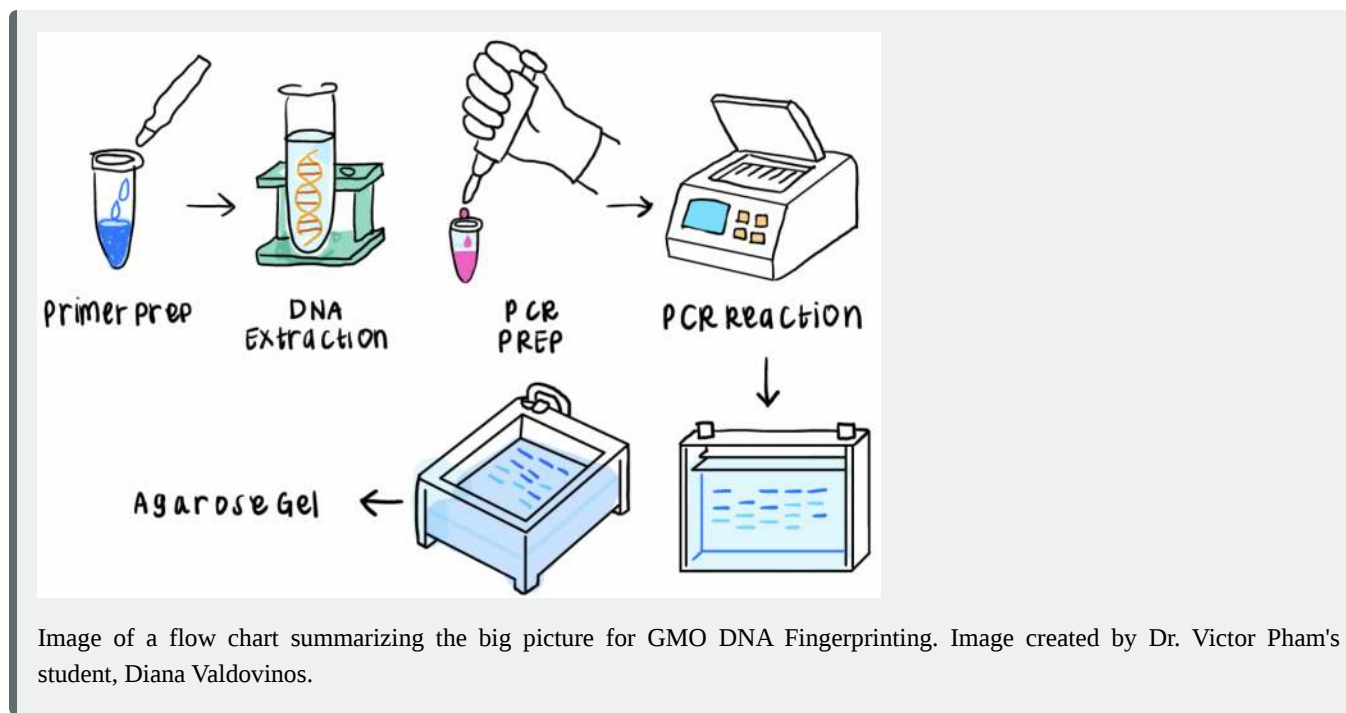
Benefits of GM Crops:

- **Environmental Impact:** GM crops that require fewer chemical pesticides reduce environmental pollution and protect beneficial insect populations.
- **Soil Conservation:** Crops engineered for drought resistance or pest resistance can grow in less-than-ideal soil, reducing the need to clear additional farmland.
- **Nutritional Benefits:** “Golden rice,” genetically engineered to produce beta-carotene, helps combat vitamin A deficiency in regions where rice is a staple food.
- **Agricultural Expansion:** GM technology enables farming in areas previously unsuitable for agriculture due to poor soil or harsh climates, improving food security.

Through this lab, students develop a comprehensive understanding of how molecular biology can be applied to everyday problems—like verifying the presence of GMOs in food. They not only learn lab techniques like DNA extraction, PCR, and gel electrophoresis, but also explore the real-world implications of genetic engineering. By combining technical skills with ethical and environmental discussions, this activity fosters critical thinking and gives students the tools to make informed decisions about biotechnology and its role in our global food system.

Objective

Test store-bought plant-based foods (like corn chips or veggie burgers) to check for the presence of genetically modified organisms (GMOs) using DNA extraction, PCR, and gel electrophoresis.



Day 1 - Part 1 – Primer Prep

Diluting Primers (Stock Prep - Instructor will do this)

- For every 1 nmol of dried primer, add **10 μ L** of distilled water (dH_2O).
 - → This creates a **100 μ M stock solution** (also called 100X).

Make a Primer Master Mix Stock (for each primer set - Students will do)

- Obtain 6 centrifuge tubes and label each of them with: **10X, Gene Name, "Master", and Group Name**
 - For example: **10X Zss Master, Vp**
- **List of gene names:**

- Zss (Control)
- Lectin
- 35S
- NOS
- Invertase
- SPS
- In each tube, add the following to obtain a 2 μ M solution (10X primer mix)
 - Mix:
 - 240 μ L diH₂O
 - 5 μ L Forward primer (of that specific gene)
 - 5 μ L Reverse primer (of that specific gene)

Primer Selection

There are **five different GMO-related primer sets** available (plus one control).

Important Note:

- If your sample tests positive for any of these GMO markers, that means the food contains a GMO sequence.
- However, if your result is negative, that does not mean the food is organic — it may still contain GMOs that we aren't testing for.
 - We are only using a limited number of primers, so a negative result simply means that none of the specific sequences tested were found.

Primer Options

1. (REQUIRED) Plant Master Mix (Control)

- **Purpose:** This is not a GMO marker. It's a positive control to confirm your **DNA extraction worked** and that plant DNA is present in your sample.
- **Why it's used:** If this doesn't show a band in PCR, it means your extraction or PCR failed — not that the food is GMO-free.

2. Invertase

- **Purpose:** Detects the **invertase gene**, which helps plants **break down sucrose into glucose and fructose**.

3. Sucrose Phosphate Synthase (SPS)

- **Purpose:** SPS is involved in the **biosynthesis of sucrose**, an important sugar in plants.

4. Lectin

- **Purpose:** Detects the **lectin gene**, a natural antimicrobial gene.

5. 35S Promoter

- **Purpose:** This sequence comes from the **Cauliflower Mosaic Virus** and is used to **turn on genes** in GMO crops.
- **Why it's engineered:** It acts like an "on-switch" to make sure inserted GMO genes are **actively expressed** in the plant. It's one of the **most common** elements in GMO crops.

6. NOS Terminator

- **Purpose:** This sequence comes from the **nopaline synthase gene** in *Agrobacterium tumefaciens*.
- **Why it's engineered:** It acts like a "stop sign" for transcription, making sure the inserted GMO gene **turns off properly** after it's expressed. It's often paired with the 35S promoter in GMO constructs.

Day 1 - Part 1 - Primer Prep



Primer Prep

for every 1 nm of dried primer, add 10 μL of distilled water (dH_2O).

→ creates 100 μM stock solution (also called 100x)



Make a primer mix

- 6 centrifuge tubes & label each with: 10x, Gene name, "Master", & Group name
- In each tube, add the following to obtain a 2 μM solution (10x primer mix)
 - mix:
 - 240 μL dH_2O
 - 5 μL Forward Primer
 - 5 μL Reverse Primer

Image of a flow chart summarizing Primer Prep. Image created by Dr. Victor Pham's student, Diana Valdovinos.

Day 1 - Part 2 – DNA Extraction

Materials:

- Food sample (individually done), distilled water, grinder, microcentrifuge tubes

Procedures

1. Weigh **1 gram** of your plant-based food sample.
 - Freeze down another 1 gram of your plant-based food for RT-PCR (Next lab).
2. Add **5 mL** of distilled water.
3. Grind or mash the sample for **2 minutes** until well mixed.
4. Transfer **500 μL** into a tube labeled **"For PCR"**
5. Add **1 mL of InstaGene solution** to your **"For PCR"** tube, and then mix by inverting at least 6 times.
6. **Boil** this sample for **15 minutes** via heat block at 95°C (or hot boiling water on a heatblock)
7. **Spin** the tube in a microcentrifuge machine for **5 minutes**.
8. Carefully transfer the **clear supernatant** to a **new tube** and label it **"For PCR, plant name, initial, date"**
 - You can discard the original tube with the pellet.
9. Use the **Nanodrop** to measure your DNA concentration.
10. Record the purity: 260/280 _____ and 260/230 _____
11. Record the concentration: _____ $\text{ng}/\mu\text{L}$. (Write this concentration on your tube)
 - If your sample is more concentrated than **100 $\text{ng}/\mu\text{L}$ (or 0.1 $\mu\text{g}/\mu\text{L}$)**, dilute it:
 - Add dH_2O to bring the concentration down to 100 $\text{ng}/\mu\text{L}$.
 - If the concentration is already less than 100 $\text{ng}/\mu\text{L}$, no dilution is needed

12. Freeze down your sample

Day 1 - Part 2 - DNA Primer



Materials :

- Food sample (individually done), distilled water, grinder, microcentrifuge tubes

Procedures

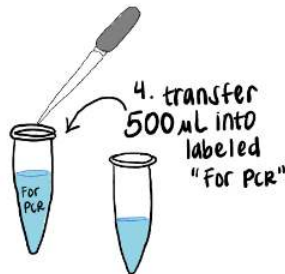
1. Weigh 1 gram of your plant-based food sample



2. Freeze down another 1 gram of your plant sample for PT-PCR (Next lab)



3. Grind/smash sample for 2 mins until well mixed



4. transfer 500 μ L into labeled "For PCR"



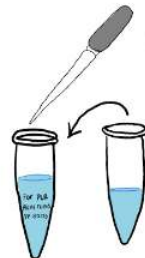
5. add 1 mL Instagene Solution to your "For PCR", invert at least 6 times



6. Boil Sample for 15 mins via heat block at 45°C



7. Spin the tube in microcentrifuge for 5 mins.



8. transfer Supernatant to new tube & label: "For PCR, Plant name, Initial, date"

- discard original tube w/ pellet.



9. Use NanoDrop to measure DNA drop

Image of a flow chart summarizing DNA Extraction. Image created by Dr. Victor Pham's student, Diana Valdovinos.

Day 2 - Part 1 – PCR Prep (Individual Work)

1. Obtain six PCR tubes with cap.

- Label each tube clearly:

- Primer gene names on the **side**
- Initial or name on the **top**

2. Add ____ **μL of DNA (you need 200ng)** into each PCR tube.
 - Volume calculation for DNA Example: If you have DNA is 50 ng/μL → $200\text{ng} \div 50\text{ng}/\mu\text{L} = 4 \mu\text{L}$ needed
3. Add **enough water** to bring the total to **12 μL** (i.e., If 4 μL DNA, add 8 μL water)
4. Into each PCR tube, use **DIFFERENT** pipet tips to add:
 - **10 μL of 2X Master Mix**
 - **8 μL of your 2 μM primer Master (from Part 1)**
 - You should have 6 tubes in total, each having different primers

Note: The total volume should be 30 uL. Run PCR immediately.

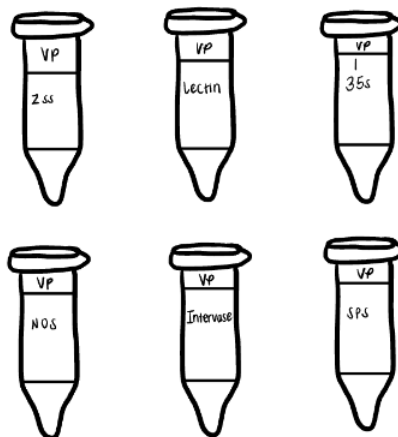
Day 2 - Part 2 – PCR Reaction

1. Please place them in the thermal cycler (PCR machine).
2. Set the machine to the following cycle:

Step	Temperature	Time
Initial denaturation	95°C	2 min
Denaturation	95°C	1 min
Annealing	59°C	1 min
Extension	72°C	2 min
Repeat steps 2–4 for 40 cycles		
Final extension	72°C	10 min
Hold	4°C	∞

Day-2-part 1-PCR prep

1. Obtain 6 tubes
With cap
2. add _____ μ L
of DNA
(you need 200 ng)
3. Add enough
water to
bring total
to 12 μ L



Day-2-part 1-PCR prep

* Place in thermal cycler
(PCR Machine)



PCR reaction

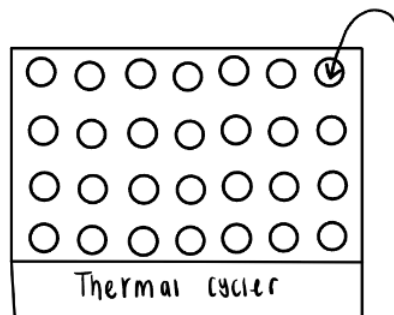


Image of a flow chart summarizing PCR Prep and PCR Reaction. Image created by Dr. Victor Pham's student, Diana Valdovinos.

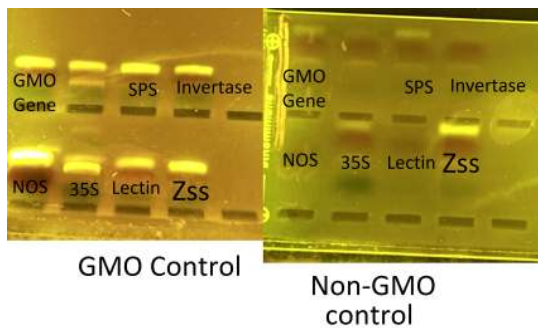
Day 3 - Part 1 – Agarose Gel (To Visualize PCR Results)

Materials:

- Agarose, TAE buffer, SYBR dye, gel box, loading dye, UV transilluminator

Procedures:

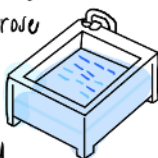
1. Prepare a **0.8% agarose gel**:
 - Mix **0.24 g of agarose** into **30 mL of 1X TAE buffer**.
 - Microwave until fully dissolved.
 - Let it cool (warm to the touch), then add:
 - **3 μL of 10,000X SYBR Safe dye** (for DNA visualization)
 - This should be added before the gel solidifies!
 - Pour the gel into a casting tray and let it set **in the dark**.
2. Mix your samples before loading:
 - **20 μL of your PCR sample**
 - **4 μL of 6X DNA loading dye**
3. Carefully load your samples into the gel wells.
4. Run the gel in **1X TAE buffer** for **30–45 minutes**, in the **dark**.
5. Visualize your results under UV light or a blue light transilluminator.
6. Compare your bands to the **GMO-positive** and **GMO-negative controls**.



Day-3-Part 1- Agarose Gel

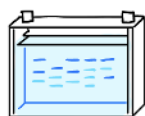
1. prepare 0.8% agarose gel

- mix 0.24g of agarose into 30mL of 1x TAE buffer
- microwave ^{sss} until fully dissolved



2. mix your sample before loading

- 20µL of your PCR sample
- 4µL of 6x DNA loading dye



3. load samples into gel wells & run in gel in 1x TAE buffer for 30-40 mins, in dark.

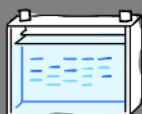


Image of a flow chart summarizing Agarose Gel Electrophoresis. Image created by Dr. Victor Pham's student, Diana Valdovinos.

Post-Lecture Objectives

- Analyze gel electrophoresis results to determine whether their food sample contains GMO DNA.
- Explain the function of the genetic elements detected by their chosen primers (e.g., what role the 35S promoter plays).
- Evaluate the reliability of their PCR data using positive and negative controls.
- Identify potential sources of error in DNA extraction or PCR and propose troubleshooting steps.
- Interpret why a negative result doesn't necessarily indicate a GMO-free food sample.
- Discuss the implications of GMO use from agricultural, environmental, and ethical perspectives.

This page titled [Activity 3-1 - Identifying Genetically Modified Organisms Through DNA Analysis](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Victor Pham](#).

Activity 3-2 - Primer Design and Barcoding via Bioinformatics

Learning Objectives

- Understand what a primer is and how it is used in PCR.
- Use NCBI BLAST to identify the organism and gene associated with a primer sequence.
- Determine the function of a protein using gene/protein names.
- Analyze primer properties (length, GC content, T_m).
- Design your own primers from a gene sequence using basic design rules.

Definition: Term

- **Primer:** A short DNA sequence (usually 18–25 nucleotides) used to initiate DNA synthesis in PCR.
- **PCR (Polymerase Chain Reaction):** A technique to amplify a specific DNA sequence using primers and DNA polymerase.
- **BLAST (Basic Local Alignment Search Tool):** A bioinformatics tool that compares DNA or protein sequences to known sequences in databases.
- **T_m (Melting Temperature):** The temperature at which 50% of the DNA duplex dissociates; important for primer binding stability.
- **GC Content:** The percentage of guanine (G) and cytosine (C) in a DNA sequence; affects melting temperature and primer stability.
- **Reverse Complement:** A sequence written in the reverse 5' → 3' order with complementary bases (A ↔ T, G ↔ C); used for reverse primers.

Introduction

In this lab, you'll learn how to use a **primer sequence** (a short DNA segment) to identify which **organism** and **protein** the sequence comes from. This kind of analysis is commonly used in **genetically modified (GM) food testing**, forensic science, and biotech research.

Primers are short DNA sequences used in **PCR** (Polymerase Chain Reaction) to amplify specific genes. Scientists design primers to target a specific gene in a specific organism. But if you're given a **mystery primer**—like in this lab—you can run a **BLAST search** to reveal its origin!

Part 1: Using NCBI BLAST to Identify the Organism

1. Go to the NCBI BLAST website
2. Choose “**Nucleotide BLAST**”
3. Paste one of the GM sequences (e.g., TTG CGC CTG AAC GGA TAT) into the “**Enter Query Sequence**” box.
4. Scroll down and click “**BLAST**”
5. After a few seconds, you'll see a list of **matches** from various organisms.
 - Let's say the BLAST result shows a strong match with *Oryza sativa Japonica Group* (a type of rice). This means the DNA primer likely comes from rice.

Standard Nucleotide BLAST

blastn blastp blastx tblastn **tblastx**

BLASTN programs search nucleotide databases using a nucleotide query. mo

Enter Query Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) [?](#) Clear Query subrange [?](#)

TTG CGC CTG AAC GGA TAT From To

Or, upload file No file chosen [?](#)

Job Title

Enter a descriptive title for your BLAST search [?](#)

☐ Align two or more sequences [?](#)

Descriptions Graphic Summary Alignments **Taxonomy**

Reports Lineage Organism **Taxonomy**

100 sequences selected [?](#)

Taxonomy	Number of hits	Number of Organisms
☐ cellular organisms	100	58
☐ Bacteria	77	36
☐ Pseudomonadota	38	24
☐ Gammaproteobacteria	20	9
☐ Enterobacterales	16	6

Part 2: Determining the Protein and Function

- In the BLAST results, look for the **gene or protein name** associated with your primer. It might say something like:
 - Protein: **Sucrose phosphate synthase (SPS)**
 - Organism: *Oryza sativa Japonica Group*
- Then, search **Google or ChatGPT** for: "Sucrose phosphate synthase function"
 - You'll find it's involved in **sugar biosynthesis**, which is important in plants to store energy.

Descriptions Graphic Summary **Alignments** Taxonomy

Alignment view: Pairwise ☐ CDS feature [? Restore defaults](#)

100 sequences selected [?](#)

[Download](#) [GenBank](#) [Graphics](#)

Yersinia mollaretii ATCC 43969 chromosome, complete genome
Sequence ID: [CP054043.1](#) Length: 4603534 Number of Matches: 1

Range 1: 511085 to 511102 [GenBank](#) [Graphics](#) [? View Match](#) [? Previous Match](#)

Score	Expect	Identities	Gaps	Strand
36.2 bits(18)	30	18/18(100%)	0/18(0%)	Plus/Minus

Query 1 TTGCGCTTACCGATAT 18
Sbjct 511102 TTGCGCTTACCGATAT 511085

[Download](#) [GenBank](#) [Graphics](#)

Oryza sativa Japonica Group sucrose phosphate synthase 3 gene, partial cds
Sequence ID: [KT225497.1](#) Length: 1092 Number of Matches: 1

Range 1: 63 to 80 [GenBank](#) [Graphics](#) [? View Match](#) [? Previous Match](#)

Score	Expect	Identities	Gaps	Strand
36.2 bits(18)	30	18/18(100%)	0/18(0%)	Plus/Plus

Query 1 TTGCGCTTACCGATAT 18
Sbjct 63 TTGCGCTTACCGATAT 80

Part 3: Complete the Bioinformatics Primer Identification Table

Instructions: Use each primer sequence below and perform a **NCBI Nucleotide BLAST** search (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) to identify:

- The **organism** from which the sequence is derived (check the **Taxonomy** section).
- The **protein** encoded by the sequence (check **alignment**).
- The **function** of the protein (check Google or ChatGPT).

Primer Sequence	Organism (from BLAST)	Protein Name	Protein Function
CCA CGT CTT CAA AGC AAG TGG			
TCC TCT CCA AAT GAA ATG AAC TTC C			
GCA TGA CGT TAT TTA TGA GAT GGG			
GAC ACC GCG CGC GAT AAT TTA TCC			
GCC CTC TAC TCC ACC CCC ATC C			
GCC CAT CTG CAA GCC TTT TTG TG			
CCG CTG TAT CAC AAG GGC TGG TAC C			
GGA GCC CGT GTA GAG CAT GAC GAT C			

Example (Filled)

Primer Sequence	Organism	Protein Name	Protein Function
SPS-F: TTG CGC CTG AAC GGA TAT SPS-R: GGA GAA GCA CTG GAC GAG G	<i>Oryza sativa</i> (Japonica Group)	Sucrose Phosphate Synthase	Catalyzes synthesis of sucrose-6-phosphate from UDP-glucose and fructose-6-phosphate in plants

Primer Design Rules

When scientists design primers, they use some golden rules to make sure they work well in PCR:

Rule	Why it matters
Primer Length (18–25 nucleotides)	Long enough to bind specifically , short enough to bind easily .
GC Content (40–60%)	GC pairs are more stable than AT. Too low = weak binding. Too high = non-specific binding.
Melting Temperature (T_m)	Forward and reverse primers should have similar T _m (within 3–5°C) for efficient PCR.

Part 4: Check Your Primer Properties Using Oligo/primer Calculator

Instructions:

- Go to a primer analysis tool, such as:
 - <https://www.biosyn.com/gizmo/tools/o...calculator.ht>

2. Paste each primer sequence into the tool.
3. Record:
 - **Primer Length** (in base pairs)
 - **% GC Content**
 - **Melting Temperature (T_m)** in °C

Primer Properties Analysis Table

Primer Sequence	Length (bp)	% GC Content	Melting Temp (°C)
CCA CGT CTT CAA AGC AAG TGG			
TCC TCT CCA AAT GAA ATG AAC TTC C			
GCA TGA CGT TAT TTA TGA GAT GGG			
GAC ACC GCG CGC GAT AAT TTA TCC			
GCC CTC TAC TCC ACC CCC ATC C			
GCC CAT CTG CAA GCC TTT TTG TG			
CCG CTG TAT CAC AAG GGC TGG TAC C			
GGA GCC CGT GTA GAG CAT GAC GAT C			

Example

Primer Sequence	Length (bp)	% GC Content	Melting Temp (°C)
TTG CGC CTG AAC GGA TAT	18 bp	50.0%	52.3°C

Part 5: Designing Primers from a Gene Sequence

Imagine you're given a long DNA sequence from the **Insulin gene (Homo sapiens)**. You need to design two primers: one forward and one reverse. Your goal is to:

- Start from the **beginning (5' end)** and **end at the 3' end** of the region of interest.
- Have **similar melting temperatures (T_m)** — ideally within 3–5°C of each other.
- Be **18–25 nucleotides** long.
- Have a **GC content between 40–60%**.
- Avoid secondary structures like hairpins and dimers.

Look for your nucleotide sequence from <https://www.ncbi.nlm.nih.gov/nucleotide/>

Example sequence (for insulin)

5' ctggggacct gaccagccg cagcctttgt gaaccaacac ctgtgcggct cacacctggt
ggaagctctc tacctagtgt gcggggaacg aggcttctc tacacacca agaccgccg

ggaggcagag gacctgcagg tggggcaggt ggagctgggc gggggccctg gtgcaggcag
cctgcagccc ttggccctgg aggggtccct gcagaagcgt ggcatgtgg aacaatgtg
taccagcatc tgctccctct accagctgga gaactactgc aacta 3'

Procedure:

1. Start with the **5' end** for the forward primer
2. Highlight the **first and last 18–25 bases** of the coding strand (5' to 3').
 - o Try a 20 bp segment and **calculate its Tm** using a tool like [OligoCalc](#).
 - 5' ctggggacct gaccagccg cagcctttgt gaaccaacac ctgtcggct cacacctggt
 - Temperature: 62°C
 - o Continue shifting or adjusting the bases until you reach an ideal temperature of interest
 - 5' ctggggacct gaccagccg cagcctttgt gaaccaacac ctgtcggct cacacctggt
 - Temperature: 59°C
 - If the Tm is too low or too high, you can shift forward or backward a base or change the primer length slightly to optimize Tm.
 - o Now go to the **end of the gene** (3' end), and pick a sequence that runs in the **reverse complement direction**.
 - taccagcatc tgctccctct accagctgga gaactactgc aacta 3'
 - Temperature: 50°C
 - o Continue shifting or adjusting the bases until you reach an ideal temperature of interest
 - taccagcatc tgctccctct accagctgga gaactactgc aacta 3'
 - Temperature: 58°C
 - Take the **reverse complement** for the reverse primer only
 - For example, 5' AAAGGG 3'
 - 5 CCCTTT 3'

Primer	Sequence	Tm (°C)
Forward Primer	acc tga ccc agc cgc agc ct	~59°C
Reverse Primer	CTC CAG CTG GTA GAG GGA GC	~58°C

This is what you submit to me. These are well-matched primers, ready for use in PCR.

Extra Credit Assignment

Choose **any gene of your choice** (can be human or from another species) and follow the same procedure to design one **forward** and one **reverse primer** for that gene. Use NCBI or Ensembl to retrieve the gene sequence.

What to Submit:

Primer Type	Sequence (5' → 3')	Length (bp)	GC Content (%)	Tm (°C)
Forward Primer				
Reverse Primer				

Gene Name:

Organism:

Link to Sequence (NCBI or Ensembl):

Brief reflection (1–2 sentences): What was the most interesting or difficult part of designing your own primers?

Note: Full credit only if the primers meet length, GC%, and Tm requirements.

Summary

In this lab, you explored how a **short DNA sequence (primer)** can be traced back to its **source organism and protein function** using the **NCBI BLAST tool**. You also learned how to check a primer's suitability for PCR by analyzing **length, GC content, and melting temperature**. Finally, you practiced **designing your own primers** from a gene sequence, learning how to shift and adjust until you hit the ideal temperature zone for reliable amplification.

This page titled [Activity 3-2 - Primer Design and Barcoding via Bioinformatics](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Victor Pham](#).

Activity 4-1 - RNA Extraction via Guanidine Thiocyanate and Phenol

Learning Objectives

- Explain the purpose and importance of RNA extraction in molecular biology experiments.
- Identify the role of each chemical used during RNA extraction (e.g., RNA-Bee, phenol, guanidine thiocyanate, chloroform, isopropanol, ethanol).
- Describe the molecular basis for phase separation and RNA precipitation.
- Perform a step-by-step RNA extraction protocol with attention to purity and yield.
- Evaluate RNA quality and concentration using a spectrophotometer (e.g., Nanodrop).

Definition: Term

- **RNA Extraction:** A laboratory process to isolate RNA from cells or tissues for use in downstream applications.
- **RNase:** An enzyme that degrades RNA; often a source of contamination in RNA work.
- **RNA-Bee:** A reagent that contains phenol and guanidine thiocyanate to lyse cells and preserve RNA.
- **Phenol:** A nonpolar organic compound that denatures proteins and helps separate them from RNA.
- **Guanidine Thiocyanate:** A chaotropic salt that disrupts proteins and inactivates RNases.
- **Chloroform:** A nonpolar solvent used with phenol to enhance phase separation during RNA extraction.
- **Aqueous Phase:** The upper layer after centrifugation, which contains RNA and other water-soluble molecules.
- **Isopropanol:** A polar organic solvent that precipitates RNA out of solution.
- **Ethanol Wash:** A step to remove residual salts and impurities from the RNA pellet.
- **A260/A280 Ratio:** A purity check using a spectrophotometer; ideal value for RNA is ~2.0.
- **Nanodrop:** A microvolume spectrophotometer used to assess nucleic acid concentration and purity.

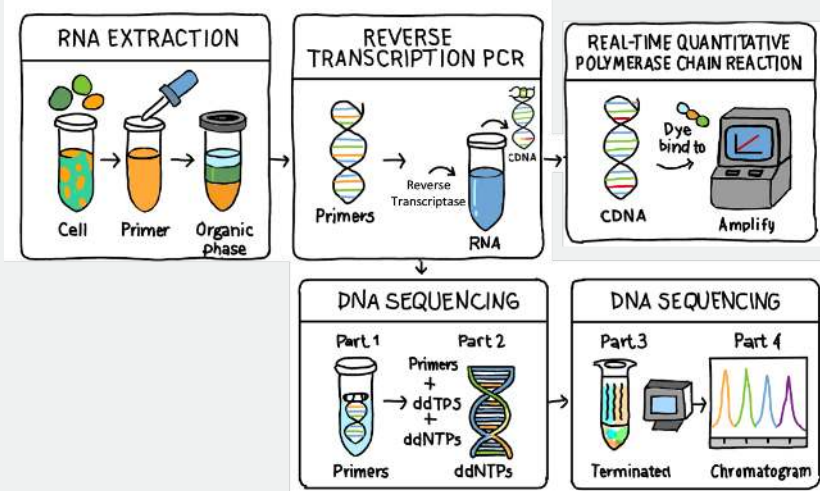


Image of a flow chart summarizing big picture flow chart for DNA Barcoding. Image created by Dr. Victor Pham's student, Geneva Anh Thy Doan.

RNA Extraction

RNA extraction is a fundamental technique in molecular biology that allows scientists to isolate RNA from cells for downstream applications like RT-qPCR, RNA sequencing, or gene expression analysis. RNA must be extracted carefully because it is prone to degradation by RNases—enzymes present both inside and outside cells. There's a misconception that one may use **RIPA buffer**, which is an excellent reagent for protein extraction. However, it lacks the strong chaotropic and phase-separation properties needed

for clean RNA isolation. Also, it doesn't inactivate RNases as effectively as RNA-Bee. This is why RIPA is avoided in RNA work—it may allow RNA degradation.

To extract RNA efficiently, cells are disrupted using a powerful reagent like **RNA-Bee**, which contains **phenol** and **guanidine thiocyanate**. **Guanidine thiocyanate** is a **chaotropic salt**, which breaks down protein structures and denatures enzymes (including RNases), preserving the integrity of the RNA. **Phenol** plays a different but equally important role. It's **nonpolar**, meaning it does not mix well with polar substances like RNA. When mixed with cell lysates, phenol helps **separate the polar nucleic acids (RNA and DNA)** from nonpolar molecules such as lipids and proteins. As a result, nucleic acids stay in the **aqueous (top) layer**, while proteins and lipids settle in the **organic (bottom) layer**. Think of phenol as oil in a salad dressing—it helps separate water-loving (hydrophilic) ingredients from fat-loving (hydrophobic) ones.

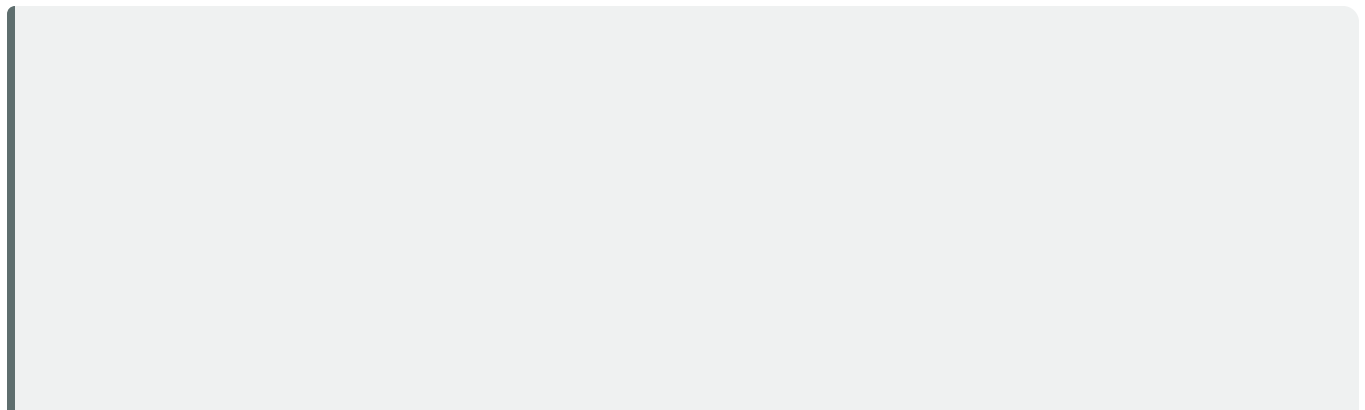
Once cells are lysed in RNA-Bee, **chloroform** is added. Chloroform is another **nonpolar solvent**, and it mixes with phenol to form a denser organic layer. When centrifuged, the mixture separates into three layers. Imagine shaking a bottle of oil, vinegar, and honey. Eventually, the heaviest (honey) settles at the bottom, and the lightest (vinegar) floats to the top. This clean separation is crucial so you can carefully collect the top layer (aqueous phase) that contains pure RNA.

- **Aqueous (top)** – contains **RNA**.
- **Interphase (middle)** – DNA and some proteins.
- **Organic (bottom)** – phenol, chloroform, denatured proteins, and lipids.

The next step is **RNA precipitation**, where **isopropanol** (a type of alcohol) is added to the aqueous phase. RNA is **less soluble in alcohol**, especially at cold temperatures, so it comes out of solution and forms a visible pellet after centrifugation. To maximize precipitation, the sample is usually **incubated overnight at 4°C**, which helps RNA clump together into a pellet you can spin down easily.

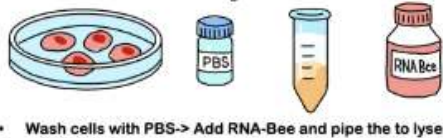
Even after RNA is precipitated with **isopropanol**, some salts and impurities might remain. To clean the pellet further, **100% ethanol** is added. Ethanol also helps **remove any residual isopropanol**, which can interfere with downstream reactions such as **reverse transcription or PCR**. However, if 100% ethanol is not available, 95% ethanol containing methanol can be used as an alternative for the wash step. While slightly less efficient due to its water content, 95% ethanol is still effective at removing salts and impurities. The small amount of methanol present should not pose a problem as long as the RNA pellet is thoroughly **air-dried** before resuspension. For best results, an **additional ethanol wash** can be included to enhance RNA purity when using 95% ethanol, and care should be taken to allow complete evaporation of any residual solvent.

Once dry, the purified RNA is resuspended in **RNA-free water**. You can measure its concentration using a **Nanodrop spectrophotometer**, which detects RNA at 260 nm. The **purity** of the RNA can be assessed by the **A260/A280 ratio**. An ideal A260/A280 ratio for pure RNA is typically around **2.0**. This indicates minimal contamination from proteins, phenol, or other organic compounds, which tend to absorb more strongly at 280 nm. If the ratio is significantly lower than 2.0 (e.g., 1.6–1.8), it suggests protein contamination or incomplete removal of phenol. In such cases, you may want to repeat the ethanol wash or perform a column-based cleanup for higher purity. Additionally, you can check the **A260/A230 ratio**, which ideally should be **2.0–2.2**. Lower values here may indicate contamination by salts, carbohydrates, or residual guanidine from the RNA-Bee reagent. Ensuring both ratios are within ideal ranges will give you the best results in sensitive downstream applications like RT-PCR, qPCR, or RNA-seq. Once quantified and verified, your RNA is ready to use or can be stored long-term at **–80°C** to preserve its integrity.

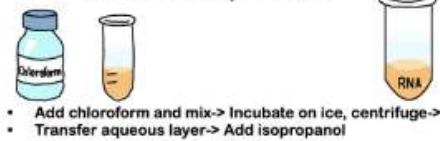


RNA EXTRACTION

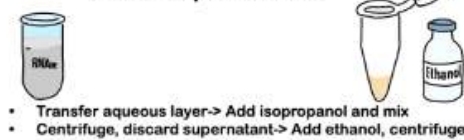
1. Harvest & Lyse Cells



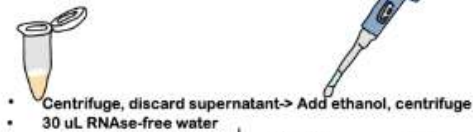
2. Phase Separation



3. Precipitate RNA



4. Wash RNA Pellet



5. Resuspend RNA



Image of a flow chart summarizing RNA Extraction. Image created by Dr. Victor Pham's student, Geneva Anh Thy Doan.

Protocol for RNA Extraction

Materials

- Treated or control cells (in 10 cm dish)
- RNA-Bee reagent
- Chloroform
- Isopropanol (100%)
- Ethanol (100%)
- RNase-free water
- 1X PBS
- 2 mL centrifuge tubes
- Pipettes and tips
- Ice bucket
- Centrifuge (4°C)
- Nanodrop or other RNA quantification tool

Protocol Steps

1. Obtain your samples, tissues, or cells (Ideally under 0.2g).
2. Wash the cells **twice** with 1mL of **1X PBS** or with **Nuclease-free water** to remove any contaminants.
 - To get rid of the PBS, you will briefly centrifuge for 1-5 seconds and discard the liquid down the sink.

3. Add **1 mL of RNA-Bee** directly to the cells.
4. Pipette the cells gently, and then transfer the entire lysate to a **2-mL centrifuge tube**.
5. Incubate the tubes on **ice for 5 minutes** to let the lysis complete.
6. Add **600 µL of cold chloroform** to the tube.
7. **Vortex or invert the tube vigorously** for about 10–15 seconds to mix well.
8. Let the tube sit on **ice for 5 minutes** to allow phase separation.
9. Spin the tube at **12,000 rpm for 5 minutes at 4°C (Recommended, but 4°C is not required)**.
 - OR spin down via mini-centrifuge machine (6000 rpm) for 10 minutes if you do not have a fast centrifuge machine.
10. After spinning, you'll see three layers. Carefully remove only the **top aqueous layer** (which contains the RNA) and transfer it to a new tube.
 - It's okay not to capture all of the upper solution, so long as you do not transfer the other two layers.
 - Discard the bottom two layers
11. Add **600 µL of cold 100% isopropanol** to the new tube.
12. Mix by gently inverting the tube 6-8 times.
13. Incubate the tube at **-80°C** for 20 minutes to allow RNA to precipitate.
 - Alternatively, if you do not have a -80°C, then incubate at -20°C for 1 hour.
14. Spin the tube at **12,000 rpm for 5 minutes at 4°C. (Recommended, but 4°C is not required)**
 - OR spin down via mini-centrifuge machine (6000 rpm) for 10 minutes if you do not have a fast centrifuge machine.
15. **Carefully discard the supernatant by micropipetting.** You should see a white pellet (RNA) at the bottom.
16. Add **1 mL of 100% ethanol** to the pellet to wash away impurities.
 - You can use 70% Ethanol as an alternative.
17. Centrifuge at **6000 rpm for 5 minutes at 4°C. (Recommended, but 4°C is not required)**
18. **Discard the ethanol first**, and then let the pellet **air-dry** under a BSL-2 hood vent for 5 minutes.
19. Add **30 µL of RNase-free water** to the RNA pellet.
20. Gently pipette up and down to dissolve the RNA.
21. Use a **Nanodrop** or similar device to measure RNA **concentration** and **purity**.
22. Store the RNA at **-80°C** until ready to use.

Category	Concentration (ng/µL)	A260/A280 Ratio	A260/A230 Ratio	Interpretation
Very Good	≥ 100	2.0 – 2.1	≥ 2.0	Highly pure RNA; no protein (280nm) or small molecule (230nm) contamination. Ideal for RT-qPCR, RNA-seq.
Good	50 – 100	1.9 – 2.0	1.8 – 2.0	Clean RNA with minimal contaminants. Suitable for most downstream applications.
Average	20 – 50	1.7 – 1.9	1.5 – 1.8	Some protein, salt, or reagent contamination. May affect sensitive applications.
Poor	< 20	< 1.7 or > 2.2	< 1.5	Likely degraded or contaminated RNA. Not ideal for most downstream uses. Re-extraction may be needed.

Post-Lecture Objectives

- Confidently explain why RNA is more delicate to isolate than DNA or protein.
- Identify each layer after centrifugation and explain what is found in each.
- Perform an RNA extraction protocol and troubleshoot problems like low yield or poor purity.
- Properly interpret Nanodrop data, including A260/A280 and A260/A230 ratios.

- Understand how solvent polarity, temperature, and salt concentration affect RNA yield and purity.

This page titled [Activity 4-1 - RNA Extraction via Guanidine Thiocyanate and Phenol](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Victor Pham](#).

Activity 4-2 - Reverse Transcription Polymerase Chain Reaction

Learning Objectives

- Define RT-PCR and explain its importance in studying gene expression.
- Differentiate between DNA, RNA, and cDNA in the context of RT-PCR.
- Describe the purpose and function of each key reagent used in the RT-PCR process.
- Outline and explain each step in the RT-PCR protocol.
- Interpret how RT-PCR results can confirm gene expression in genetically modified organisms (GMOs).

Definition: Term

RT-PCR: A technique that first converts RNA into complementary DNA (cDNA) and then amplifies it using PCR to detect gene expression.

Reverse Transcriptase: An enzyme that synthesizes DNA from an RNA template.

cDNA: Complementary DNA synthesized from an RNA template via reverse transcription.

mRNA: Messenger RNA; a type of RNA that reflects active gene expression by serving as a template for protein synthesis.

Oligo dT Primer: A primer made of deoxythymidine that binds to the poly-A tail of eukaryotic mRNA.

Random Hexamer Primer: A mix of random six-nucleotide sequences that bind various RNA types, including degraded or non-polyadenylated RNA.

dNTPs: Deoxynucleotide triphosphates – the building blocks for DNA synthesis.

RNase Inhibitor: A molecule that prevents RNA degradation by inhibiting RNase enzymes.

MgCl₂: Magnesium chloride; provides Mg²⁺ ions, which are essential cofactors for enzyme activity in RT-PCR.

Master Mix: A premade solution containing most reagents required for a reaction, improving consistency and reducing pipetting errors.

Reverse Transcription Polymerase Chain Reaction (RT-PCR)

RT-PCR is a technique used when you want to analyze gene expression — in other words, what genes are actively being used by a cell. When scientists want to understand which genes are being actively used by an organism — in other words, which genes are "**turned on**" and making proteins — they use a technique called **Reverse Transcription Polymerase Chain Reaction (RT-PCR)**. This method is especially helpful when studying **genetically modified (GM) organisms**, which are often engineered to express specific genes that give them special traits, such as resistance to pests, herbicides, or drought. To figure out which genes are active, we need to study **RNA**, not DNA. This is because **RNA is the intermediate step between DNA and proteins**. If DNA is the instruction manual, RNA is the working copy that cells use to build proteins. So, detecting a specific RNA transcript means the gene is **actively being expressed** in that moment. However, traditional PCR techniques only work on DNA, not RNA. That's why RT-PCR includes a critical first step: **reverse transcription**, which converts RNA into complementary DNA (**cDNA**). Once we have the cDNA, we can use PCR to amplify specific gene regions and study them.

Let's take a practical example, imagine we are testing a sample of **GM corn** engineered to express the **Bt toxin gene**, which allows it to kill insect pests. If we extract RNA from the corn tissue and detect the **Bt mRNA**, it means the gene is not only present in the genome, but also **actively being used**. By reverse transcribing this RNA into cDNA and amplifying it with PCR, we can confirm that the gene is expressed.

The RT-PCR process involves several key steps. First, the extracted RNA is mixed with two types of primers: **oligo dT primers**, which bind to the poly-A tail of messenger RNA (mRNA), and **random primers**, which help capture a broad range of RNA transcripts. This mix is then **heated to 70°C** to denature secondary structures and allow primers to anneal, and then **cooled to 4°C** to stabilize the primer-RNA binding. After this, a **master mix** is added. This contains the enzyme **reverse transcriptase**, **nucleotides (dNTPs)**, **magnesium ions (MgCl₂)**, **reaction buffer**, and an **RNase inhibitor** to protect the RNA. The mixture is incubated at **42°C for one hour**, which is the ideal temperature for the enzyme to build cDNA. Finally, the reaction is heated to **70°C for 15 minutes** to inactivate the enzyme, and the cDNA is ready for use.

RT-PCR Reagents and Their Roles

Primers: Oligo dT and Random Hexamers

In reverse transcription, **primers are crucial** because the reverse transcriptase enzyme cannot start making DNA on its own—it needs a starting point. That's what primers provide: a short stretch of nucleotides that bind (anneal) to the RNA and allow the enzyme to begin synthesizing cDNA. Combining **oligo dT and random primers** gives a more **comprehensive snapshot** of gene expression — you target mRNA specifically with oligo dT, while also covering less conventional or partially degraded transcripts with random primers.

- **Oligo dT Primers:** These are short sequences of deoxythymidine (dT) bases — for example, (dT)_{12–18} — that specifically bind to the **poly-A tail** found at the 3' end of **eukaryotic mRNA**. Most mature mRNAs in eukaryotic cells have a long tail of adenine bases (poly-A), which helps protect the RNA and aids in translation. Oligo dT primers are **selective**: they ensure that you're mainly reverse-transcribing **mRNA**, which is what you want when analyzing **gene expression**.
 - *Advantage:* Enriches for mature mRNA (not rRNA or tRNA).
 - *Limitation:* Won't capture RNAs without poly-A tails (e.g., bacterial RNA, histone mRNAs, or degraded mRNAs).
- **Random Primers (Random Hexamers):** These are **short sequences of 6 nucleotides** with random base combinations (e.g., NNNNNN). Because they are so short and randomly designed, they can bind to **multiple points along all types of RNA**—not just mRNA. They are useful for **generating more complete cDNA coverage** of all RNA types, including partially degraded RNA, or for detecting non-polyadenylated transcripts like **rRNA or viral RNA**.
 - *Advantage:* Increases coverage and works even with degraded RNA.
 - *Limitation:* May include rRNA or other non-mRNA transcripts if not careful.

Reverse Transcriptase Enzyme

This is the star of the show. **Reverse transcriptase** is an RNA-dependent DNA polymerase — meaning it reads RNA and makes a **complementary DNA strand (cDNA)**. The enzyme works best around 42°C, where it can process RNA efficiently without being too sensitive to temperature fluctuations. Most commonly used reverse transcriptases come from retroviruses like **Moloney Murine Leukemia Virus (M-MLV)** or **Avian Myeloblastosis Virus (AMV)**.

dNTP Mix (Deoxynucleotide Triphosphates)

These are the building blocks of DNA: **dATP, dTTP, dCTP, and dGTP**. The reverse transcriptase needs these nucleotides to synthesize the new DNA strand. Without them, the reaction would stop immediately. Think of them as the **bricks** the enzyme uses to build the cDNA.

Magnesium Chloride (MgCl₂)

Magnesium ions (Mg²⁺) are a **critical cofactor** for the activity of reverse transcriptase. They help stabilize the interaction between the enzyme and the nucleotides, and ensure proper catalytic function. If magnesium levels are too low, the enzyme won't work efficiently. If too high, it can promote nonspecific binding or errors.

Reaction Buffer

This provides the **optimal pH, salt concentration, and other conditions** for the enzyme to function. Enzymes are very sensitive to their environment, and the buffer ensures the reverse transcription occurs smoothly. It usually contains **Tris-HCl, KCl, and sometimes DTT** (dithiothreitol, a reducing agent that stabilizes the enzyme).

RNase Inhibitor

RNA is **very fragile** and prone to degradation — especially by **RNases**, which are enzymes that break down RNA and are found everywhere (on skin, dust, etc.). Even a small contamination can ruin the sample. RNase inhibitors bind to any RNases present and **prevent them from degrading your RNA** during the reverse transcription step.

Steps in RT-PCR

- **70°C (5 min):** This initial heating step unfolds secondary RNA structures and allows primers to bind properly. Many RNAs form complex loops or hairpins, and heating helps **relax** them for better primer annealing. Note that this is prior to adding the

enzyme **reverse transcriptase**

- **4°C (10 min):** Cooling stabilizes the primer-RNA hybrids, making sure the primers stay attached.
- **42°C (60 min):** This is the ideal temperature for **reverse transcriptase** to synthesize the DNA.
- **70°C (15 min):** This final heating step **inactivates the enzyme**, stopping the reaction and making your cDNA stable for storage or downstream use.
- **4°C (Hold):** Keeps the reaction cool and stable until you're ready to move forward.

REVERSE TRANSCRIPTION PCR (RT-PCR)

Objective: Convert RNA into cDNA for analysis of GM gene expression

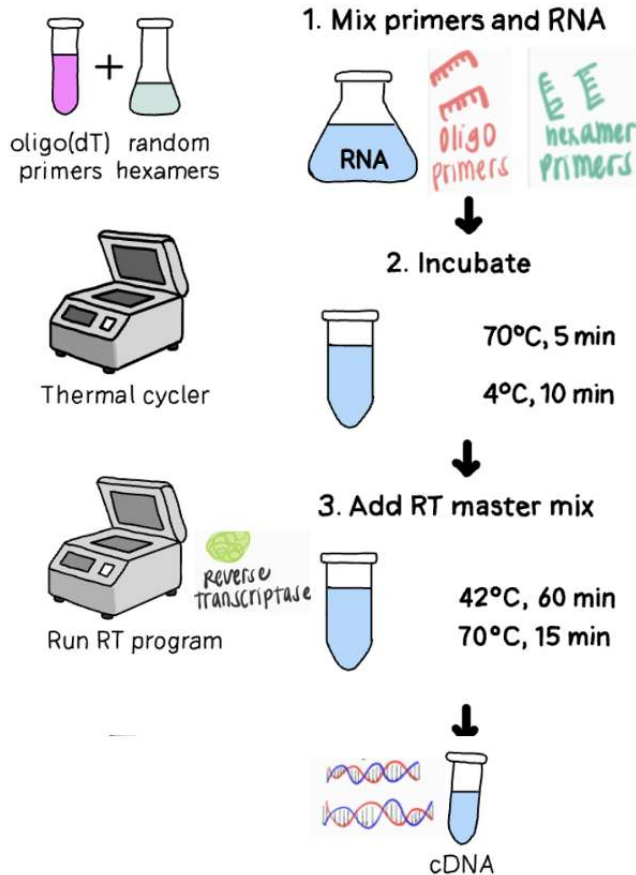


Image of a flow chart summarizing Reverse Transcriptase. Image created by Dr. Victor Pham's student, Geneva Anh Thy Doan.

Reverse Transcription PCR (RT-PCR)

Objective: Convert RNA into cDNA for later analysis of GM gene expression.

Materials Needed (Per RNA Sample):

- **Extracted RNA sample**
- **Nuclease-free water**
- **GoScript Master Mix Kit:**
 - **Oligo(dT) primers** (500 µg/mL)
 - **Random hexamer primers** (500 µg/mL)

- **5X RT Buffer**
- **25 mM MgCl₂**
- **10 mM dNTP mix**
- **RNase inhibitor** (40 U/μL)
- **GoScript Reverse Transcriptase** (200 U/μL)

 Note

 Create a mastermix, and then multiply these volumes by the number of samples plus 1 extra to avoid running short when pipetting.

Protocol

1. In a PCR tube, combine:
 - **X μL (up to 1 μg) of RNA**
 - **1 μL oligo(dT) primers**
 - **1 μL random hexamer primers**
 - Add **nuclease-free water** to reach a final volume of **10 μL**.
2. Mix gently by pipetting up and down.
3. Place the tubes in a thermal cycler (or water bath):
 - Heat at **70°C for 5 minutes** (to remove RNA secondary structure).
 - Immediately cool to **4°C for at least 10 minutes** (to help primers anneal or bind).
4. In a fresh tube, prepare a **RT Master Mix** with your group by adding the following solution below: (Note: Create a Mastermix with your group, and then transfer 10uL to each member's samples)
 - **2.0 μL 5X RT buffer**
 - **1.2 μL MgCl₂**
 - **2.0 μL dNTP mix**
 - **0.5 μL RNase inhibitor**
 - **4.3 μL Reverse transcriptase enzyme**
5. After primer annealing is done, only add **10 μL of the RT Master Mix** to your PCR tube from step 1 (after your PCR is completed).
6. Place tubes in the thermal cycler and run the following program:
 - **42°C for 60 minutes** (Reverse transcription: RNA → cDNA)
 - **70°C for 15 minutes** (Enzyme inactivation)
 - **Hold at 4°C** (this is optional, for temporary storage)
7. You can either store it at **-20°C** for future use, or use it immediately for **PCR/qPCR** to detect gene expression.

 Note

- Always wear gloves to avoid RNase contamination.
- Use clean, RNase-free tips and tubes.
- Keep RNA and enzymes on ice as much as possible.
- Label tubes clearly

Post-Lecture Objectives

- RT-PCR bridges the gap between transcriptomics and molecular detection by turning RNA into cDNA for amplification.
- The use of both oligo dT and random primers ensures better coverage of RNA transcripts, especially when RNA is partially degraded.
- The reagents in the RT master mix play very specific roles that must be optimized for a successful reaction.

- Proper temperature control at each stage ensures enzyme activity, primer annealing, and RNA integrity.
- RT-PCR provides **functional evidence** of gene activity, especially useful in GMO validation (e.g., detecting Bt gene expression in GM corn).

Concepts

RNA as Evidence of Gene Expression: Presence of mRNA = gene is actively being used

Why Convert to cDNA? DNA is more stable and compatible with PCR; RNA is not

Enzyme Sensitivity: Reverse transcriptase and RNA are both fragile – temperature and contamination control are key

Specificity from Primers: Primer choice affects what kind of RNA gets reverse-transcribed (mRNA vs. total RNA)

Application in GMOs: Confirms if inserted genes like Bt are not only present but actively transcribed

This page titled [Activity 4-2 - Reverse Transcription Polymerase Chain Reaction](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Victor Pham](#).

Activity 4-3 - Real-Time Quantitative Polymerase Chain Reaction

Learning Objectives

- Distinguish between **qPCR** and **RT-PCR**, and understand how they are connected.
- Explain how **SYBR Green** is used to detect and quantify DNA amplification.
- Describe the **thermal cycling steps** in a qPCR reaction.
- Set up a qPCR reaction using primers, cDNA, and master mix.
- Interpret **Ct values** and **melting curves** to assess gene expression and amplification quality.

Definition: Term

qPCR (Quantitative PCR): A method to detect and measure the amount of a specific DNA sequence in real time using fluorescence.

RT-PCR (Reverse Transcription PCR): A method that converts RNA into cDNA, often used before qPCR to study gene expression.

SYBR Green: A fluorescent dye that binds to double-stranded DNA, used to monitor amplification during qPCR.

Ct Value (Cycle Threshold): The cycle number at which the fluorescence signal exceeds background; inversely related to the amount of target DNA.

Melting Curve Analysis: A post-PCR step that slowly heats the sample to identify the specificity of the amplified product.

Primer-Dimers: Unintended short double-stranded DNA products formed by primers binding to each other instead of the target sequence.

Real-Time Quantitative Polymerase Chain Reaction (qPCR)

Quantitative Polymerase Chain Reaction (qPCR), sometimes called Real-time PCR, should not be confused with Reverse Transcription PCR (RT-PCR). Once the cDNA has been synthesized from RT-PCR, we can perform **quantitative PCR (qPCR)** to determine **how much of a specific gene is present**. Unlike regular PCR, which only shows if a gene is present or absent, qPCR gives us **real-time, quantitative information** about gene expression. This is done using a special dye called **SYBR Green**, which binds to **double-stranded DNA**. As PCR amplifies the cDNA, more double-stranded DNA is created, and the fluorescence from SYBR Green increases. A **qPCR machine** detects this fluorescence during each cycle, giving a continuous readout of DNA amplification. Let's return to our Bt corn example. If we run qPCR with primers specific to the Bt gene and see a **strong fluorescent signal**, it tells us the gene is not just present — it is being **highly expressed** in the tissue we tested. This gives us quantitative evidence of gene activity.

The thermal cycling conditions for qPCR are carefully controlled. Each cycle includes a **denaturation step at 95°C**, where double-stranded DNA separates into single strands; an **annealing step at 55°C**, where primers bind to their target sequences; and an **extension step at 72°C**, where the DNA polymerase builds new strands. These steps are repeated for about **40 cycles**. After the amplification is complete, a **melting curve analysis** is performed. During this step, the sample is slowly heated, and the qPCR machine tracks the drop in SYBR Green fluorescence as the DNA melts (denatures). A **sharp, single melting peak** indicates a specific product, while multiple peaks or irregular curves may suggest nonspecific amplification or primer-dimers.

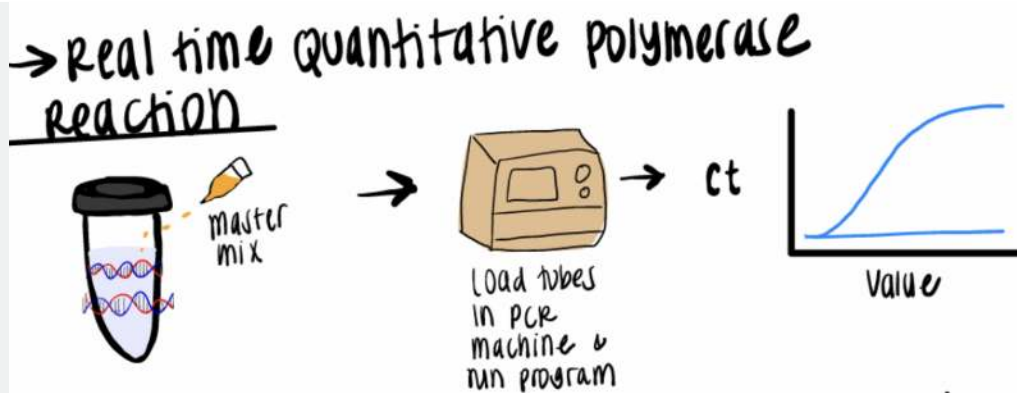


Image of a flow chart summarizing Real-time quantitative PCR. Image created by Dr. Victor Pham's student, Diana Valdovinos.

qPCR Protocol with SYBR Green

Objective: Use qPCR to detect and measure how much of a target gene is being expressed in a genetically modified sample by using cDNA and SYBR Green dye.

Materials:

- SYBR Green Master Mix (2X)
- Forward primer (200 μ M stock)
- Reverse primer (200 μ M stock)
- Synthesized cDNA sample (from RT step)
- Nuclease-free water
- qPCR tubes or 96-well plate
- qPCR machine

Procedure:

1. In a clean tube, create a **qPCR MasterMix-Primer name** with your group by mixing the following solution: *(Note: You need to create another qPCR MasterMix for different primers)*
 - 10 μ L of SYBR Green Master Mix (1X final concentration)
 - 5 μ L of your **primer mix** (from Activity 3.1, or 20 μ M Forward and 20 μ M Reverse Primers)
2. Add 15 μ L of the **qPCR Master Mix** to your qPCR tube (Note: qPCR tube contains a **flat cap**)
3. Into your qPCR tube:
 - Add up to 5 μ L of **cDNA** (50 ng)
 - Adjust the remaining volume to 5 μ L with Nuclease-free water
4. Load the tubes/plate into the qPCR machine and run this program:

Step	Temp (°C)	Time	Purpose
Initial Denaturation	95°C	10 sec	Unwind DNA strands
Annealing	55°C	30 sec	Primers bind to DNA
Extension	72°C	30 sec	DNA is copied
Repeat above steps for 40 cycles			
Increase the temperature slowly from 55°C to 95°C	Go up 0.5°C every 5 seconds		
Final SYBR Imaging			

- The qPCR machine will detect **fluorescence from SYBR Green**, which binds to double-stranded DNA.
- The more gene expression, the earlier the signal appears (lower Ct value).
- Use the software to create a **bar graph** of expression levels across samples.

Post-Lecture Objectives

- You prepared a qPCR reaction using a cDNA template and primers specific to your gene of interest.
- You ran thermal cycling conditions and monitored fluorescence using SYBR Green.
- You analyzed Ct values and created a bar graph to visualize expression across samples.
- You checked the specificity of your results using melting curve analysis.

Key Takeaways

- qPCR allows for **quantitative measurement** of gene expression.
- SYBR Green fluorescence is directly tied to **double-stranded DNA production**—more fluorescence = more gene product.
- A **sharp, single melting peak** indicates a specific PCR product; extra peaks could mean contamination or primer-dimers.
- Lower Ct values indicate **higher expression** of your target gene in the sample.

This page titled [Activity 4-3 - Real-Time Quantitative Polymerase Chain Reaction](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Victor Pham](#).

Activity 4-4 - DNA Sequencing with SeqStudio Genetic Analyzer

Learning Objectives

- Define the purpose and basic principles of Sanger sequencing.
- Explain how fluorescently labeled dideoxynucleotides are used to terminate DNA strands.
- Understand the role of capillary electrophoresis in sequence separation and detection.
- Describe the major components and workflow of the SeqStudio Genetic Analyzer.
- Outline the preparation steps for submitting samples for Sanger sequencing using SeqStudio.

Definition: Term

- **Sanger Sequencing:** A DNA sequencing method that uses chain-terminating dideoxynucleotides (ddNTPs) to generate DNA fragments of varying lengths, enabling the determination of the nucleotide sequence.
- **Dideoxynucleotide (ddNTP):** A nucleotide lacking a 3'-OH group, which prevents the addition of further nucleotides, thereby terminating DNA strand elongation.
- **Fluorescent Labeling:** A technique in which ddNTPs are tagged with fluorescent dyes, each color representing a different base (A, T, G, C) to allow detection.
- **Capillary Electrophoresis:** A method for separating DNA fragments by size using an electric field in a thin capillary tube filled with polymer.
- **Electropherogram:** A graphical output showing fluorescence intensity vs. time, which is interpreted to read the DNA sequence.
- **SeqStudio Genetic Analyzer:** A compact, benchtop DNA sequencer that automates Sanger sequencing using a four-color fluorescence system and capillary electrophoresis.
- **POP-1 Polymer:** A polymer used in the capillary to act as a sieving medium for fragment separation.
- **Hi-Di™ Formamide:** A denaturant that helps linearize DNA fragments for better resolution during capillary electrophoresis.

Note

Prepare your knowledge by searching the following to self-study:

- Animation: [How Sanger Sequencing Works (YouTube)]
- Thermo Fisher PDF Guide: *Getting Started with the SeqStudio Genetic Analyzer*
- Lab protocol: *Sample Preparation for Sanger Sequencing using SeqStudio*

Introduction to DNA Sequencing with SeqStudio

DNA sequencing is the process of determining the precise order of nucleotides—adenine (A), thymine (T), cytosine (C), and guanine (G)—in a DNA molecule. Understanding this sequence is essential in modern biology, biotechnology, and medicine because it allows researchers to identify genes, mutations, and genetic variations associated with traits or diseases. In this lab, we use a technique called **Sanger sequencing**, which is a widely used and highly accurate method suitable for sequencing single genes or short DNA fragments (usually up to 1000 base pairs). Sanger sequencing works by copying a single-stranded DNA template while randomly incorporating fluorescently labeled chain-terminating nucleotides (called **dideoxynucleotides** or ddNTPs). These special nucleotides stop DNA synthesis at random points, resulting in fragments of different lengths. Each ddNTP is tagged with a different fluorescent dye (corresponding to A, T, G, or C), and once the sequencing reaction is complete, the fragments are separated by size using **capillary electrophoresis**. The fluorescent labels are detected by a laser, allowing the instrument to "read" the DNA sequence based on the color of each terminal base in order from smallest to largest fragment.

The instrument used in this lab is the **Applied Biosystems SeqStudio Genetic Analyzer**, a compact, automated machine that performs capillary electrophoresis and sequence detection. It uses a small **cartridge with four capillaries**, which means it can

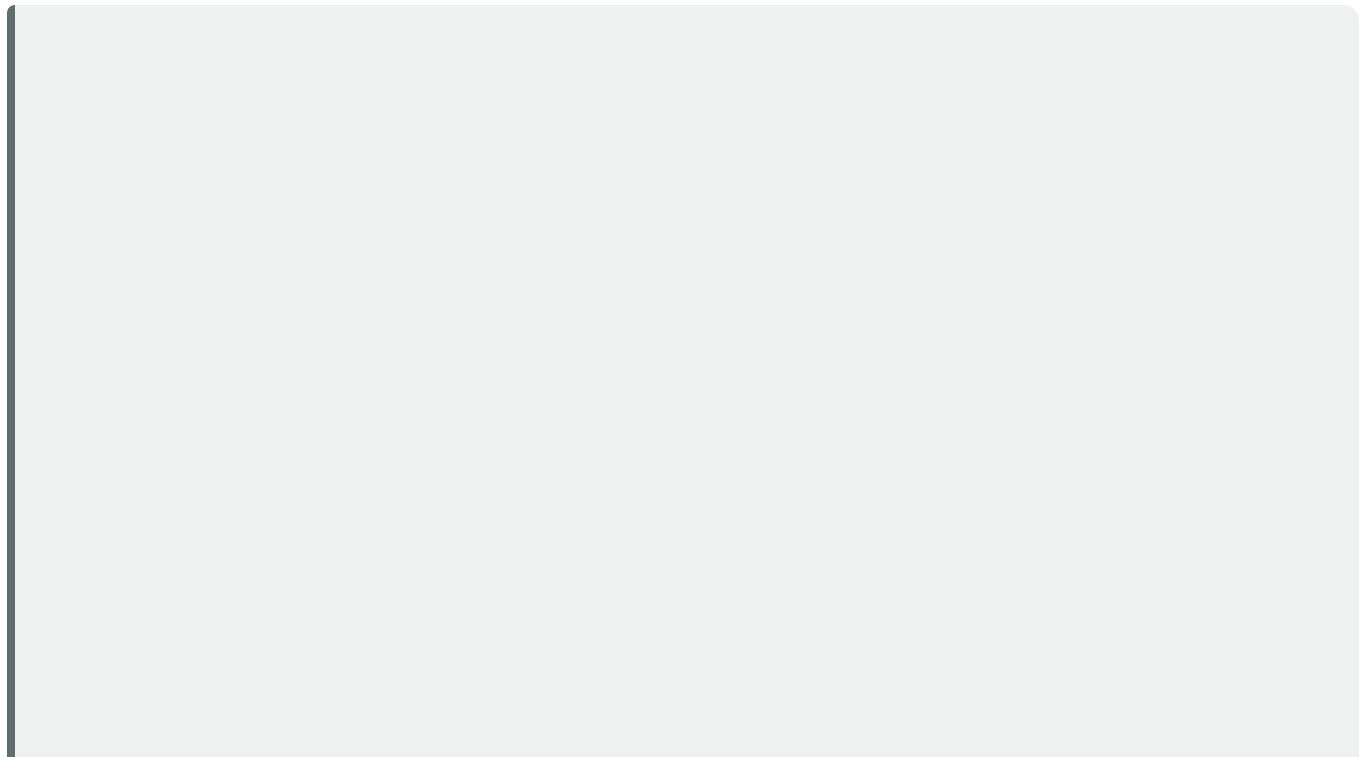
process up to four samples at a time. To run properly, the instrument requires a buffer-filled **Cathode Buffer Cartridge (CDC)** and carefully prepared samples that have been cleaned of contaminants such as excess primers, salts, and leftover PCR reagents.

Before sequencing, DNA needs to be **amplified using Polymerase Chain Reaction (PCR)**. PCR makes millions of copies of a specific DNA region using primers—short DNA sequences that target the beginning and end of the region you want to amplify. For sequencing, two separate reactions are usually set up: one with a **forward primer** and one with a **reverse primer**, so that both strands of the DNA can be read independently. This gives you better sequence coverage and helps confirm accuracy. After PCR, the amplified DNA must be **purified**. This is crucial because any residual primers or nucleotides can interfere with the sequencing chemistry, resulting in poor-quality reads or background noise. The purification process includes both **enzymatic cleanup** (using enzymes that degrade leftover primers and dNTPs) and **column-based purification** (using silica membranes to bind and wash DNA). Once cleaned, the DNA is quantified, often using a **Nanodrop spectrophotometer**, to ensure there is enough material for sequencing and that it's free of contaminants.

The actual sequencing reaction involves mixing the purified DNA with the **BigDye Terminator v3.1 Cycle Sequencing Kit**, which contains DNA polymerase, normal dNTPs, fluorescently labeled ddNTPs, and necessary buffers. This reaction is cycled in a thermal cycler, similar to PCR, to produce the fluorescently labeled DNA fragments of varying lengths. Once samples are prepared, they are loaded into a 96-well plate designed for the SeqStudio instrument. The plate is covered with a **septa strip** to prevent evaporation during loading and electrophoresis. The SeqStudio performs the electrophoresis, separates the DNA fragments, and detects their fluorescence to generate a sequencing file. These files contain chromatograms—visual peaks representing each nucleotide—along with the base calls. The sequencing data is then transferred to a computer and analyzed using specialized software (such as **Sequencing Analysis 7, or SeqA 7**). The software allows users to visualize and interpret the sequence reads, identify mutations, and compare results with known sequences.

It is important to understand that Sanger sequencing is different from high-throughput or “next-generation” sequencing (NGS). While NGS can sequence millions of DNA fragments at once, it is more complex and expensive. Sanger sequencing remains the gold standard for smaller-scale applications, including verifying cloned genes, identifying point mutations, sequencing PCR products, or genotyping.

By the end of this lab, you will have a working understanding of how Sanger sequencing is performed from start to finish, including PCR setup, DNA purification, sequencing reaction, instrument operation, and data analysis. This hands-on experience provides a foundation for advanced studies in molecular biology, biotechnology, genetic engineering, and bioinformatics.



DNA sequencing

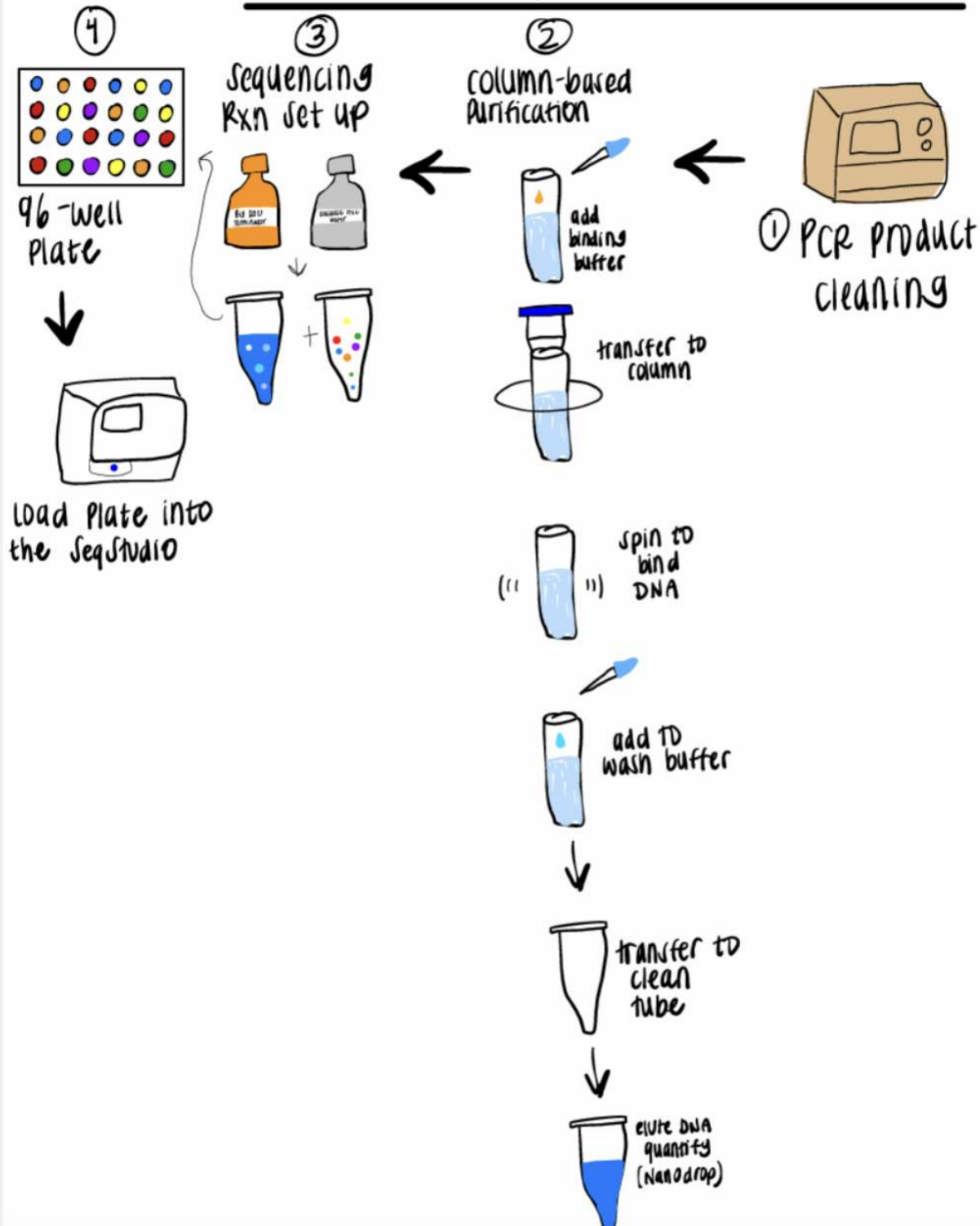


Image of a flow chart summarizing DNA fragment Sequencing. Image created by Dr. Victor Pham's student, Diana Valdovinos.

Reagents Functions and Purposes

SeqStudio Genetic Analyzer

- **Function:**
 - The SeqStudio is a modern **capillary electrophoresis instrument** designed for **automated Sanger sequencing**. It separates DNA fragments based on size and detects the fluorescent signal from labeled nucleotides at the end of each fragment.
- **Purpose:**
 - This instrument reads the DNA sequence by analyzing the light signals from the fluorescently labeled ddNTPs (dideoxynucleotides) in each fragment. It allows researchers to visualize the DNA sequence as a chromatogram (colored peaks representing A, T, G, and C). It's compact, user-friendly, and ideal for small to medium labs.

CDC (Cathode Buffer Cartridge)

- **Function:**
 - The **Cathode Buffer Cartridge (CDC)** delivers running buffer to the **capillaries** inside the SeqStudio. The buffer helps conduct electricity through the capillaries, which is essential for the separation of DNA fragments.
- **Purpose:**
 - The buffer allows DNA fragments of different sizes to migrate through the capillary tubes at different speeds, enabling size-based separation. A fresh CDC is essential for **consistent flow, proper electrical conductivity, and high-quality data**. Using an old or dry CDC can lead to poor resolution, failed runs, or instrument errors.

Hi-Di Formamide

- **Function:**
 - Hi-Di formamide is a **denaturing agent** that disrupts hydrogen bonding between DNA strands, keeping them in **single-stranded form**.
- **Purpose:**
 - Single-stranded DNA is necessary for accurate reading during capillary electrophoresis. If DNA strands re-anneal (stick back together), the sequencing machine cannot correctly interpret the sequence. Hi-Di formamide ensures the DNA remains separated and migrates uniformly through the capillaries for high-resolution detection.

BigDye Terminator v3.1

- **Function:**
 - This is a **fluorescent sequencing reaction mix** used in the Sanger method. It contains:
 - DNA polymerase
 - Normal dNTPs (for elongation)
 - Fluorescently labeled **ddNTPs** (dideoxynucleotides), which terminate the growing DNA strand when incorporated
- **Purpose:**
 - When a ddNTP is added during DNA synthesis, the strand cannot be extended any further. This produces a collection of fragments of various lengths, each ending with a labeled ddNTP. The fluorescent tags are color-coded for A, T, G, and C, enabling the SeqStudio to "read" the sequence based on color and fragment length. This kit is essential for converting the DNA template into readable data.

Zymo ZR DNA Sequencing Cleanup Kit

- **Function:**
 - This is a **column-based DNA purification kit** that removes unwanted components from PCR reactions, including:
 - Salts
 - Residual primers
 - Unincorporated nucleotides
 - Enzymes
- **Purpose:**

- Impurities from PCR can interfere with the sequencing chemistry, causing noisy data or failed reads. The Zymo kit ensures that the DNA is **clean and concentrated**, making it suitable for the sensitive BigDye reaction and the SeqStudio instrument. It helps improve the accuracy and clarity of sequencing results.

96-well PCR Plates and 1.5 mL Microcentrifuge Tubes

- **Function:**
 - **96-well PCR plates** allow multiple sequencing or PCR reactions to be set up and run in parallel.
 - **1.5 mL microcentrifuge tubes** are used for sample transfers, purification steps, and storage.
- **Purpose:**
 - These plasticware items are essential for organizing and managing **multiple samples efficiently**, especially in high-throughput settings. The 96-well format is compatible with the SeqStudio tray, allowing students to load multiple samples for sequencing in one batch. Using standardized labware also reduces pipetting errors and ensures compatibility with centrifuges, thermal cyclers, and sequencing instruments.

Notes: DNA Sequencing (SeqStudio)

- **CDC = Cathode Buffer Cartridge**
 - Always use a **new** one for every sequencing run.
 - **Remove the cartridge before turning off** the instrument to prevent damage.
- **You'll prepare two samples per DNA:**
 - One with **forward primer**. Do not combine Forward and Reverse.
 - One with **reverse primer**. Do not combine Forward and Reverse.
- **SeqStudio uses 96-well plates**
 - Each plate allows for **~25 injections** (samples).
 - Fill all 4 lanes (machine runs 4 at a time), using **Hi-Di Formamide** in empty wells if needed.

Materials

- **PCR reagents and thermal cycler**
 - (Your target DNA, primers, dNTPs, polymerase, buffer)
- **ExoSAP-IT**
- **Zymo ZR DNA Sequencing Cleanup Kit**
- **DNA-free water**
- **Nanodrop Spectrophotometer**
- **BigDye Terminator v3.1 Cycle Sequencing Kit**
- **Forward and Reverse primers**
 - (Prepare separate tubes for each direction)
- **PCR tubes (0.2 mL)**
 - (For each sequencing reaction – 2 per sample)
- **96-well MicroAmp PCR plate**
- **Septa 8-strip caps**
- **Hi-Di Formamide**
- **Sequencing Standard (Optional Control)**
 - (Mixed with BigDye Terminator v3.1)
- **CDC Cartridge (Cathode Buffer Cartridge)**
 - (Always use a new one for each run)
- **SeqStudio Genetic Analyzer**
- **USB flash drive**
 - (To export .ab1 files)

- Lab computer with Sequencing Analysis Software v7 (SeqA 7)

Procedure:

Day 1: PCR Product Cleanup with ExoSAP-IT

1. Run a PCR reaction for your target DNA using specific forward and reverse primers (**This is the cDNA after RT-PCR**)

1. Obtain **one** PCR tubes with cap.

- Label the tube **clearly**:
 - Primer gene names on the **side**
 - Initial or name on the **top**

2. Add ____ **μL of DNA (you need 200ng)** into each PCR tube.

- Volume calculation for DNA Example: If you have DNA is 50 ng/μL → $200\text{ng} \div 50\text{ng}/\mu\text{L} = 4 \mu\text{L}$ needed

3. Add **enough water** to bring the total to **12 μL** (i.e., If 4 μL DNA, add 8 μL water)

4. Into each PCR tube, use DIFFERENT pipet tips to add:

- **10 μL of 2X Master Mix**
- **8 μL of your 2 μM primer Master (contains Reverse and Forward primers)**
 - You should have 6 tubes in total, each having different primers

5. Please place them in the thermal cycler (PCR machine), and set the machine to the following cycle:

Step	Temperature	Time
Initial denaturation	95°C	2 min
Denaturation	95°C	1 min
Annealing	59°C	1 min
Extension	72°C	2 min
Repeat steps 2–4 for 40 cycles		
Final extension	72°C	10 min
Hold	4°C	∞

Day 2: Part 1: Column-Based DNA Purification via Zymo ZR DNA sequencing Cleanup kit

1. After PCR, add 2 μL of ExoSAP-IT directly to each PCR tube.

- Exonuclease removes leftover primers, and Shrimp Alkaline Phosphatase (SAP) inactivates excess nucleotides.

2. Incubate at 37°C (Waterbath) for 5 minutes to allow the enzymes to work.

3. Freeze at -80°C for at least 1 minute to stop the reaction and preserve the sample.

4. Add 240 μL of sequencing binding buffer to the PCR product.

5. Transfer this mix to the Zymo-Spin IB column over a collection tube.

6. Spin for 30 seconds to allow DNA to bind to the silica membrane.

7. Add 300 μL of sequencing wash buffer, spin again, and discard the flow-through in the sink.

8. Place the spin column in a clean 1.5 mL microcentrifuge tube.

9. Elute the DNA by adding 20 μL of DNA-free water and spin for 1 minute. Discard your spin column in the trash.

- Now you have purified DNA in your centrifuge tube, ready for quantification.
- Use a Nanodrop spectrophotometer to determine DNA concentration and record the value in nM/μL.
- For sequencing, we aim for up to 100 ng of DNA per reaction.

Instrument Setup: Instructor-Only

Only trained personnel should prepare and initialize the SeqStudio machine.

1. Turn on the SeqStudio.
 - Login: **Admin**, Password: **1234**.
2. **Hit the eject button** (top right) and lift open the door.
3. Retrieve the **CDC cartridge** from 4°C storage. Remove the plastic and protective capillaries and insert it into the instrument.
4. Close the door and wait for it to initialize.
5. Retrieve **Hi-Di Formamide** and **Sequencing Standard BigDye Terminator v3.1 (BDT)** from -20°C storage.
6. Mix **300 µL of Hi-Di Formamide into the sequencing standard BDT vial**, if not already done.
 - Use a Sharpie and write "Hi-Di" on the BDT vial for future reference
7. During run setup: Choose "**Sequencing → Plate**", then select "**ShortSeq (non-BDX)**" with **Dye Set: Z.Bigdye Terminator v3.1**.
8. Hit eject (top right) to eject the plate
9. In one well, Instructor should add **20 µL of BDT + Hi-Di Formamide** as a control.
10. All unused wells in the 96-well plate **must be filled with 20 µL of Hi-Di Formamide OR the 20 µL of student's PCR sample** to avoid run errors. The instrument injects 4 wells at a time, and it cannot process an empty one.
11. See the next step to prepare the student's PCR sample.

Day 2: Part 2: Sequencing Reaction Setup (Student's PCR Samples)

Each student will prepare **two reactions per sample**: one tube using the forward primer only and one tube with the reverse primer only.

1. Transfer up to **10 µL (or 100 ng) of your DNA** to two new PCR tubes.
 - The maximum volume can be **10 µL**, even if your DNA concentration is too low.
 - The minimum volume should be **2 µL**, even if your DNA concentration is too high.
2. Add **4 µL of BigDye Terminator v3.1** to each tube.
 - The BigDye reagents incorporate labeled terminators that create a set of terminated fragments of various lengths, which the instrument can detect via fluorescence.
3. Add **4 µL of forward primer** to one tube only, and **4 µL of reverse primer** to the other tube only.
4. Adjust with **nuclease-free water** to bring the total volume to **20 µL**.

Day 2: Part 3: Loading the SeqStudio

1. Place the samples in a **MicroAmp 96-well tray retainer**.
2. Cover with a **septa 8-strip (single-use cap)**.
3. Load the plate into the SeqStudio.
4. **Check the cartridge** again to ensure it has enough buffer.
5. Retract the plate and **start the run** through the software.

Day 3: Data Export and Analysis

Once the run is complete:

1. **Export data to USB**: Use the machine interface.
2. Plug USB into the lab laptop (Login: **Username: Administrator, Password: Administrator**).
3. Transfer data files from USB to the laptop.

Using Sequencing Analysis Software (SeqA 7):

1. Open **SeqA 7**. Login Password is **SeqUser**
2. Click **File → Add Samples**, navigate to USB, and add selected samples to import files.
3. After loading, click "**Show Box**" → go to the "**Sequence**" tab.
4. Copy your sequence into Notepad or another text editor for review.

 Note

- Always **label forward and reverse reactions clearly**, as mix-ups are common.
- **Ensure no bubbles** are present in the PCR plate before placing it in the instrument.
- If sequencing fails, check DNA concentration, primer quality, and cleanup steps.
- The SeqStudio uses **capillary electrophoresis**, so small contaminants or incomplete purification can affect peak quality.

 Post-Lecture Objectives

After completing the lab and lecture, students should be able to:

- Describe how the sequencing reaction and capillary electrophoresis steps are performed on the SeqStudio.
- Troubleshoot common issues in Sanger sequencing (e.g., weak signal, mixed peaks).
- Interpret an electropherogram and use software (e.g., FinchTV or Chromas) to read a DNA sequence.
- Apply their sequence data to identify genes or organisms using BLAST.
- Explain the significance of clean PCR products, primer concentration, and sample volume in successful sequencing.

This page titled [Activity 4-4 - DNA Sequencing with SeqStudio Genetic Analyzer](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Victor Pham](#).

Index

D

dire

Glossary

Sample Word 1 | Sample Definition 1

Detailed Licensing

Overview

Title: Lab manual: Molecular biology and Genetic Engineering (Biot 275)

Webpages: 26

Applicable Restrictions: Noncommercial

All licenses found:

- [CC BY-NC 4.0](#): 96.2% (25 pages)
- [Undeclared](#): 3.8% (1 page)

By Page

- Lab manual: Molecular biology and Genetic Engineering (Biot 275) - [CC BY-NC 4.0](#)
 - Front Matter - [CC BY-NC 4.0](#)
 - TitlePage - [CC BY-NC 4.0](#)
 - InfoPage - [CC BY-NC 4.0](#)
 - Table of Contents - [Undeclared](#)
 - Licensing - [CC BY-NC 4.0](#)
 - Activity 1-1 – Data on Human Traits - [CC BY-NC 4.0](#)
 - Activity 1-2 - Human Genetic Inheritance and Diseases - [CC BY-NC 4.0](#)
 - Activity 1-3 - Genetic evolution and Identifying Homologs - [CC BY-NC 4.0](#)
 - Activity 1-4 - PowerPoint Presentation on the Bioinformatic Evolution and function of proteins - [CC BY-NC 4.0](#)
 - Activity 2-0 - Introduction to Plasmid and Enzyme p450 - [CC BY-NC 4.0](#)
 - Activity 2-1 - Plasmids Isolation - [CC BY-NC 4.0](#)
 - Activity 2-2 - DNA Analysis via Restriction Enzymes - [CC BY-NC 4.0](#)
 - Project 2-3 - Identifying Plasmid Restriction Mapping Using Agarose Gel Electrophoresis - [CC BY-NC 4.0](#)
 - Activity 2-4 - Bacterial Transformation with Engineered Plasmid - [CC BY-NC 4.0](#)
 - Activity 3-1 - Identifying Genetically Modified Organisms Through DNA Analysis - [CC BY-NC 4.0](#)
 - Activity 3-2 - Primer Design and Barcoding via Bioinformatics - [CC BY-NC 4.0](#)
 - Activity 4-1 - RNA Extraction via Guanidine Thiocyanate and Phenol - [CC BY-NC 4.0](#)
 - Activity 4-2 - Reverse Transcription Polymerase Chain Reaction - [CC BY-NC 4.0](#)
 - Activity 4-3 - Real-Time Quantitative Polymerase Chain Reaction - [CC BY-NC 4.0](#)
 - Activity 4-4 - DNA Sequencing with SeqStudio Genetic Analyzer - [CC BY-NC 4.0](#)
 - Back Matter - [CC BY-NC 4.0](#)
 - Index - [CC BY-NC 4.0](#)
 - Glossary - [CC BY-NC 4.0](#)
 - Detailed Licensing - [CC BY-NC 4.0](#)
 - Detailed Licensing - [CC BY-NC 4.0](#)

Detailed Licensing

Overview

Title: Lab manual: Molecular biology and Genetic Engineering (Biot 275)

Webpages: 26

Applicable Restrictions: Noncommercial

All licenses found:

- [CC BY-NC 4.0](#): 96.2% (25 pages)
- [Undeclared](#): 3.8% (1 page)

By Page

- Lab manual: Molecular biology and Genetic Engineering (Biot 275) - [CC BY-NC 4.0](#)
 - Front Matter - [CC BY-NC 4.0](#)
 - TitlePage - [CC BY-NC 4.0](#)
 - InfoPage - [CC BY-NC 4.0](#)
 - Table of Contents - [Undeclared](#)
 - Licensing - [CC BY-NC 4.0](#)
 - Activity 1-1 – Data on Human Traits - [CC BY-NC 4.0](#)
 - Activity 1-2 - Human Genetic Inheritance and Diseases - [CC BY-NC 4.0](#)
 - Activity 1-3 - Genetic evolution and Identifying Homologs - [CC BY-NC 4.0](#)
 - Activity 1-4 - PowerPoint Presentation on the Bioinformatic Evolution and function of proteins - [CC BY-NC 4.0](#)
 - Activity 2-0 - Introduction to Plasmid and Enzyme p450 - [CC BY-NC 4.0](#)
 - Activity 2-1 - Plasmids Isolation - [CC BY-NC 4.0](#)
 - Activity 2-2 - DNA Analysis via Restriction Enzymes - [CC BY-NC 4.0](#)
 - Project 2-3 - Identifying Plasmid Restriction Mapping Using Agarose Gel Electrophoresis - [CC BY-NC 4.0](#)
 - Activity 2-4 - Bacterial Transformation with Engineered Plasmid - [CC BY-NC 4.0](#)
 - Activity 3-1 - Identifying Genetically Modified Organisms Through DNA Analysis - [CC BY-NC 4.0](#)
 - Activity 3-2 - Primer Design and Barcoding via Bioinformatics - [CC BY-NC 4.0](#)
 - Activity 4-1 - RNA Extraction via Guanidine Thiocyanate and Phenol - [CC BY-NC 4.0](#)
 - Activity 4-2 - Reverse Transcription Polymerase Chain Reaction - [CC BY-NC 4.0](#)
 - Activity 4-3 - Real-Time Quantitative Polymerase Chain Reaction - [CC BY-NC 4.0](#)
 - Activity 4-4 - DNA Sequencing with SeqStudio Genetic Analyzer - [CC BY-NC 4.0](#)
 - Back Matter - [CC BY-NC 4.0](#)
 - Index - [CC BY-NC 4.0](#)
 - Glossary - [CC BY-NC 4.0](#)
 - Detailed Licensing - [CC BY-NC 4.0](#)
 - Detailed Licensing - [CC BY-NC 4.0](#)