Biotechnology Explorer™

GAPDH PCR Module
Instruction Manual

Catalog #166-5010EDU

explorer.bio-rad.com

This kit is shipped at 4°C. Open immediately upon arrival and store reagents at –20°C within 2 weeks.

Duplication of any part of this document permitted for classroom use only.
Please visit explorer.bio-rad.com to access our selection of language translations for Biotechnology Explorer kit curricula.

BIO-RAD

Dear Educator:

Amplification as the path to visualization of DNA

Molecular biologists are faced with the classic needle in a haystack problem. Often we must find a few copies of a piece of DNA that code for a given gene in the haystack of DNA comprising the entire genome. Even if there are a thousand copies of the same piece of DNA it is often still difficult to locate them. However by selectively amplifying only that specific piece of DNA we can separate it from the rest of the DNA and visualize it. Once we can visualize the DNA, we can use the tools of molecular biology to open the whole vista of genetic engineering. We can work with the DNA to make discoveries in science, agriculture, and medicine. The method used to amplify the DNA is the polymerase chain reaction (PCR).

PCR is capable of repeatedly doubling the amount of specific DNA. After many PCR cycles a million-fold or billion-fold times as much DNA is generated. Because of the increasing use of PCR in science it is important to provide students with an understanding of the basic principles and applications of PCR. With this kit students can amplify the *GAPDH* gene from plant tissue. They will use an even more sensitive version of PCR known as nested PCR. Nested PCR requires two rounds of PCR, the first greatly amplifies the larger DNA region of interest, but not to the point that it can yet be detected. The second round more specifically amplifies a smaller (nested) piece of DNA to the point where it can be detected and used for further applications.
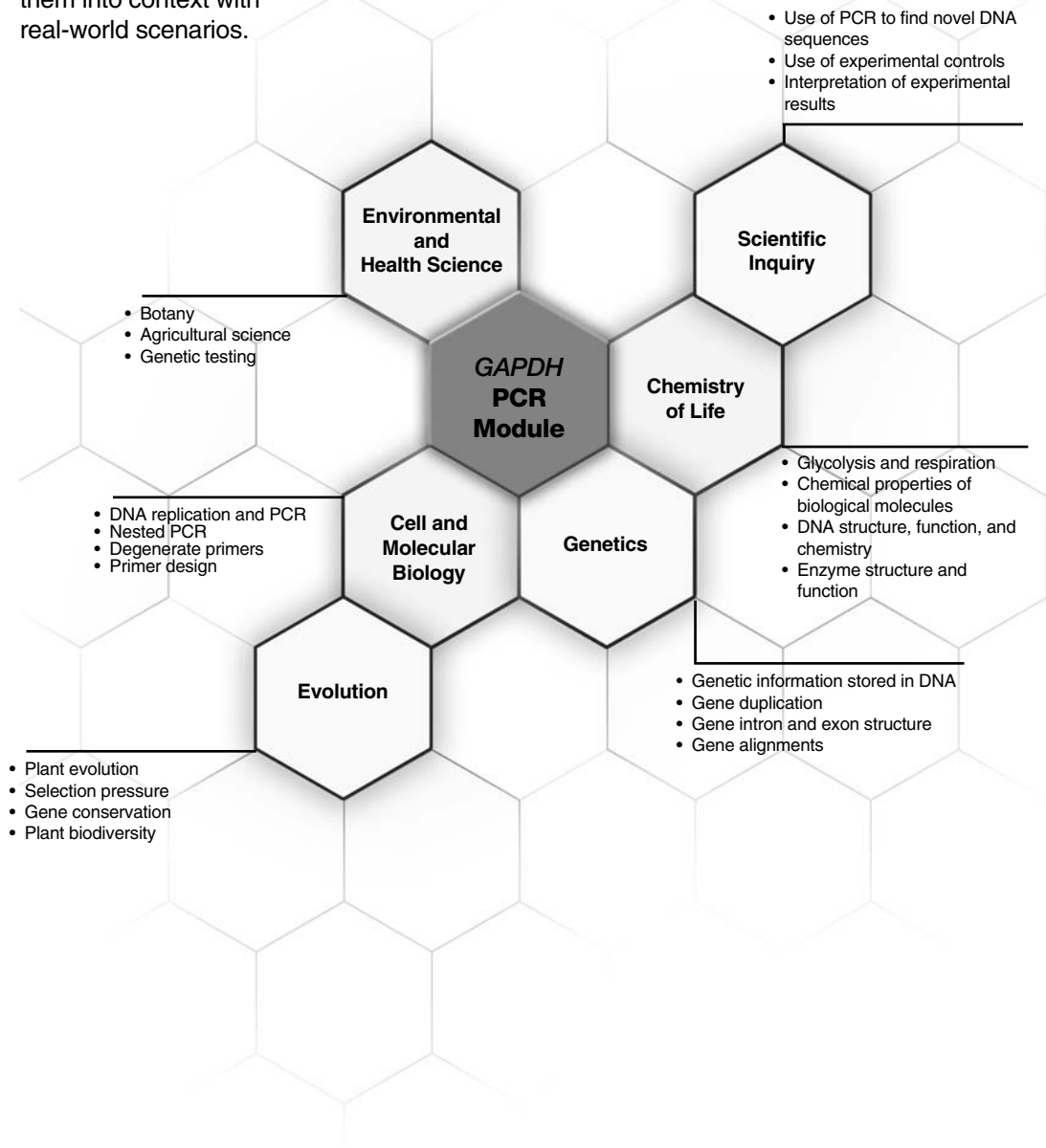
The *GAPDH* PCR module is part of the Bio-Rad Cloning and Sequencing Explorer Series (catalog #166-5000EDU) and can also be used as a stand alone kit to amplify *GAPDH* sequences from plant DNA. The Bio-Rad Nucleic Acid Extraction module (catalog #166-5005EDU) can be used to extract high quality DNA from plant tissue samples suitable for use in the *GAPDH* PCR module. The entire Cloning and Sequencing Explorer series was designed in collaboration with Dr. David L. Robinson and Dr. Joann M. Lau both of the Department of Biology, Bellarmine University, Louisville Kentucky. Bio-Rad thanks them both for their invaluable contributions.

For more biotechnology curricula PCR kits visit explorer.bio-rad.com. Look for the PV92 PCR Informatics kit (catalog #166-2100), the GMO Investigator kit (catalog #166-2500EDU), and the Crime Scene Investigator kit (catalog #166-2600EDU).

Biotechnology Explorer Program
Life Science Group
Bio-Rad Laboratories
6000 James Watson Dr.
Hercules, CA 94547

Create context. Reinforce learning. Stay current.
New scientific discoveries and technologies
create more content for you to teach,
but not more time. Biotechnology
Explorer kits help you teach more
effectively by integrating multiple
core content subjects into a
single lab. Connect concepts
with techniques and put
them into context with
real-world scenarios.

- Use of PCR to find novel DNA sequences
- Use of experimental controls
- Interpretation of experimental results

Environmental and Health Science

Scientific Inquiry

- Botany
- Agricultural science
- Genetic testing

GAPDH **PCR Module**

Chemistry of Life

- DNA replication and PCR
- Nested PCR
- Degenerate primers
- Primer design

Cell and Molecular Biology

Genetics

- Glycolysis and respiration
- Chemical properties of biological molecules
- DNA structure, function, and chemistry
- Enzyme structure and function

Evolution

- Genetic information stored in DNA
- Gene duplication
- Gene intron and exon structure
- Gene alignments

- Plant evolution
- Selection pressure
- Gene conservation
- Plant biodiversity

Table of Contents

Introduction

Polymerase chain reaction (PCR) has become an integral part of the biological sciences, widely used in research, forensic, and clinical laboratories. Included in the myriad of uses for PCR is its use to clone genes even when the DNA sequence of the gene is not known. By designing PCR primers based on gene sequences from related organisms, a method called nested PCR can find and amplify the gene of interest. Nested PCR is a two-step process. In the initial round of PCR, degenerate primers are used to amplify genes with sequences similar to the gene of interest. In the second round of PCR, more specific primers amplify the gene of interest from the products of the first PCR.

Lesson Time Line

Pre-lab activity	Setting the stage: DNA structure, DNA replication, PCR, primer design, degenerate primers, nested PCR	1–2 days, lecture and homework
	(Optional) Prepare genomic DNA using Nucleic Acid Extraction Module (catalog #166-5005EDU)	1 hour
Lesson 1	Set up PCR 1	45 minutes in lab
	Run first PCR	4 hours – overnight
	Analyze results with gel electrophoresis	45 minutes in lab
	Discussion of results	
Lesson 2	Exonuclease treatment of DNA	
	Set up PCR 2	90 minutes in lab*
	Run second PCR	4 hours – overnight
	Analyze results with gel electrophoresis	45 minutes in lab
Post-lab activity	Analysis and interpretation of results	1–2 days in lab and homework

* Can be split into two 45-minute lab sessions.

Storage Instructions

Open the kit as soon as it arrives and remove the bag of perishable components. Store these components in the freezer at –20°C.

Intended Audience

This kit is appropriate for advanced students with experience in molecular biology and PCR, and would fit in a molecular biology or biotechnology curriculum at the two- or four-year college level. It might also be appropriate for advanced high school students with prior experience with PCR.

Use of this kit to introduce students to PCR for the first time is not recommended. As an introductory PCR experiment, we recommend the Crime Scene Investigator PCR Basics Kit (catalog #166-2600EDU), PV92 PCR Informatics Kit (catalog #166-2100EDU), or GMO Investigator Kit (catalog #166-2500EDU).

Kit Inventory Checklist

This section lists the equipment and reagents needed for PCR amplification of the *GAPDH* gene in your classroom or teaching laboratory. The *GAPDH* PCR Kit (catalog #166-5010EDU) supports 12 student workstations, with 2–4 students per station. Open the kit as soon as it arrives and place the bag of perishable components in the freezer (–20°C).

It is recommended that PCR results be evaluated by agarose gel electrophoresis. The number of gel boxes and power supplies required depends on the number of student workstations you chose to set up. Information on pouring and running agarose gels are included in the manual (Appendix B), and you may wish to order a Bio-Rad Electrophoresis Module (catalog #166-0451EDU).

Kit Components	Number/Kit	(✔)
Initial *GAPDH* PCR primers, 50 µl	1	❏
Nested *GAPDH* PCR primers, 50 µl	1	❏
PCR master mix, 1.2 ml	3	❏
Plasmid control DNA, 1 ml	1	❏
Control *Arabidopsis* genomic DNA	1	❏
Exonuclease I, 50 µl	1	❏
Molecular weight ruler	1	❏
Sterile water, 2.5 ml	1	❏
PCR tubes, 0.2 ml	150	❏
Capless PCR tube adaptors, 1.5 ml	150	❏
Colored micro test tubes, 2.0 ml	120	❏
Foam micro test tube holders	12	❏
Curriculum, including Teacher's Guide, Student Manual, and graphic Quick Guide	1	❏

Required Accessories	Number/Kit	(✔)
Adjustable volume micropipets		
2–20 µl (catalog #166-0551EDU)	1–12	❏
20–200 µl (catalog #1660507EDU)	1	❏
Pipet tips, aerosol barrier		
2–20 µl, Xcluda™ B (catalog #211-2006EDU)	1	❏
20–200 µl, Xcluda D (catalog #211-2016EDU)	1	❏
Thermal cycler (catalog #170-9701EDU)	1	❏
Electrophoresis reagents and equipment	1–12	❏

Optional Accessories

Novel plant genomic DNA

Microcentrifuges (catalog #166-0602EDU or #166-0603EDU)

Vortexer (catalog #166-0610EDU)

Small Ethidium Bromide DNA Electrophoresis Reagent Pack (catalog #166-0451EDU)
   Includes 25 g agarose, 200 ml 50x TAE, 10 ml Ethidium Bromide

Small Fast Blast™ DNA Electrophoresis Reagent Pack (catalog #166-0450EDU)
   Includes 25 g agarose, 100 ml 50 x TAE, 100 ml 500x Fast Blast DNA stain

Refills Available Separately

GAPDH PCR Refill Pack (catalog #166-5011EDU)
   Includes initial *GAPDH* primers, nested *GAPDH* primers, PCR master mix, plasmid control DNA, control *Arabidopsis* genomic DNA, exonuclease I, molecular weight ruler, Orange G loading dye, and sterile water

2X master mix for PCR (catalog #166-5009EDU)

Curriculum Fit

Students develop abilities to conduct inquiry-based experiments

Students learn laboratory skills and techniques commonly used in research

Students formulate scientific explanations using data, logic, and evidence

Students develop an understanding of the molecular basis of heredity

Students develop an understanding of biological evolution

Students use advanced technology to solve a novel problem

Students develop an understanding of the chemistry of DNA

Background for Instructor

Polymerase chain reaction (PCR) has been a common technique for over 20 years. It has become standard procedure in research laboratories, medical diagnosis, and forensic science, among others:

- In "DNA fingerprinting," PCR is used to look at specific loci (regions of repetitive DNA). Variability at those loci can be used to distinguish one individual from another. This is the heart of current forensic science, and large DNA databases are maintained by law enforcement agencies at local, state, and national levels
- DNA fingerprinting can also used to identify individuals after a disaster, such as after hurricanes or terrorist attacks. DNA samples from victims can be compared to DNA amplified from hair follicles retrieved from a hairbrush, or the victim's DNA can be compared to that of family members. Kinship analysis can also be used to determine paternity
- In genetic testing, PCR can be used to detect gene mutations that might result in disease
- Infectious diseases can be detected using PCR. For example, viral DNA can be amplified to levels that can be detected by diagnostic tests, whereas the virus might not be detectable without amplification

All of these uses of PCR have one thing in common — they are all based on known DNA sequences.
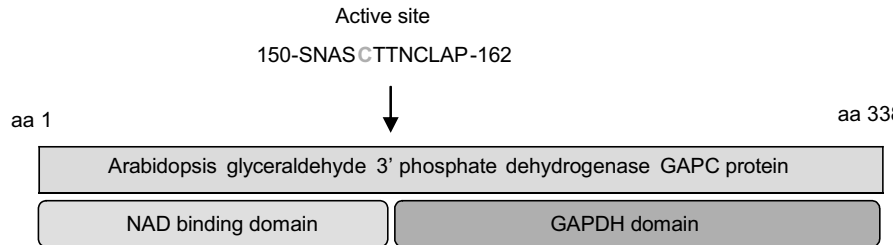
PCR has another important use in the laboratory. It can be used to clone DNA even when the DNA sequence is not known. There are two primary approaches for cloning without DNA sequence data. First, if DNA from the gene of interest has been sequenced in other species, particularly species closely related to the species of interest, then the DNA sequences from similar species can be used to design primers that are likely to amplify the target DNA. Second, the amino acid sequence of the protein coded by the gene of interest may be known. Using the amino acid sequence, researchers can work back to the DNA sequence that coded it, although because of the degeneracy of the amino acid code, there will be more than one possible DNA sequence for the gene.

The gene that will be studied in this experiment codes for the enzyme glyceraldehyde-3-phosphate dehydrogenase (GAPDH). There are several reasons why this gene was selected. First, GAPDH is a crucial enzyme in glycolysis. The gene is known as a housekeeping gene — a gene that is expressed constitutively and is necessary for cells to survive. Since GAPDH is abundant in cells and can be purified for study, much is known about the protein structure and function. GAPDH consists of four subunits (hence a tetramer) held together through non-covalent attachments. All four subunits may be identical (designated as A4, a homodimer) or they may consist of pairs of slightly different subunits (designated A2B2, a heterodimer). In both cases, each subunit has an active site and can bind one molecule of NAD$^+$ cofactor.
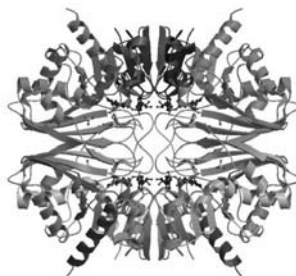
Structure of GAPDH bound to NAD$^+$ as determined by x-ray crystallography. Structure can be downloaded from the Protein Databank (www.rcsb.org) using the pdb identifier 1szj.

5

GAPDH protein has two major domains, the amino terminal has an NAD$^+$ binding domain and the carboxy terminal has glyceraldehyde 3' phosphate dehydrogenase activity.

Active site

150-SNAS**C**TTNCLAP-162

aa 1   ↓   aa 33

Arabidopsis glyceraldehyde 3' phosphate dehydrogenase GAPC protein	
NAD binding domain	GAPDH domain

GAPDH protein domain structure. The active cysteine is shaded grey.

In addition, recent research has found that GAPDH plays many other roles outside of glycolysis. For example, the human *GAPDH* gene is overexpressed (i.e., expressed at levels much higher than normal) in 21 different classes of cancer (Altenburg and Greulich, 2004). GAPDH has been shown to play roles in membrane fusion, endocytosis, microtubule bundling, and DNA repair. GAPDH is also involved in viral pathogenesis, regulation of apoptosis (programmed cell death), and human neuronal diseases including Alzheimer's and Huntington's disease (reviewed in Sirover 1999).

GAPDH catalyzes the sixth reaction of glycolysis, the pathway by which glucose is converted into pyruvate in a series of ten enzymatic reactions (see Glycolysis pathway figure). In mammals, most dietary polysaccharides are broken down to glucose in the bloodstream. In plants, glucose is synthesized from carbon dioxide in the Calvin cycle of photosynthesis. Glycolysis has a number of useable products:

- The production of ATP and NADH during glycolysis, providing energy for the cells
- Pyruvate, the end product of glycolysis, feeds into the citric acid cycle, producing more energy for the cells
- Many of the intermediate compounds of glycolysis (see Glycolysis pathway figure) are precursors for the formation of other biological molecules. For example, glucose-6-phosphate is a precursor for the synthesis of ADP, NAD$^+$, and coenzyme Q, and phosphoenolpyruvate is a precursor for the synthesis of the amino acids, tyrosine, phenylalanine, and tryptophan

The reaction catalyzed by GAPDH is:

Glyceraldehyde-3-phosphate + NAD$^+$ + Pi→ 1,3-bisphosphoglycerate + NADH + H$^+$

GAPDH oxidizes glyceraldehyde-3-phosphate (GAP) by removing a hydrogen ion (H$^+$) and transferring it to the acceptor molecule, NAD$^+$ (NAD$^+$ + H$^+$→NADH). In addition, GAPDH adds a second phosphate group to GAP. This reaction is catalyzed by a cysteine in the active site of the GAPDH protein.

When the source of carbohydrate for glycolysis is a sugar, glycolysis will occur in the cytosol, as it does in animal cells. When the carbohydrate source is starch however, glycolysis can occur in plastids (a group of organelles that includes chloroplasts).
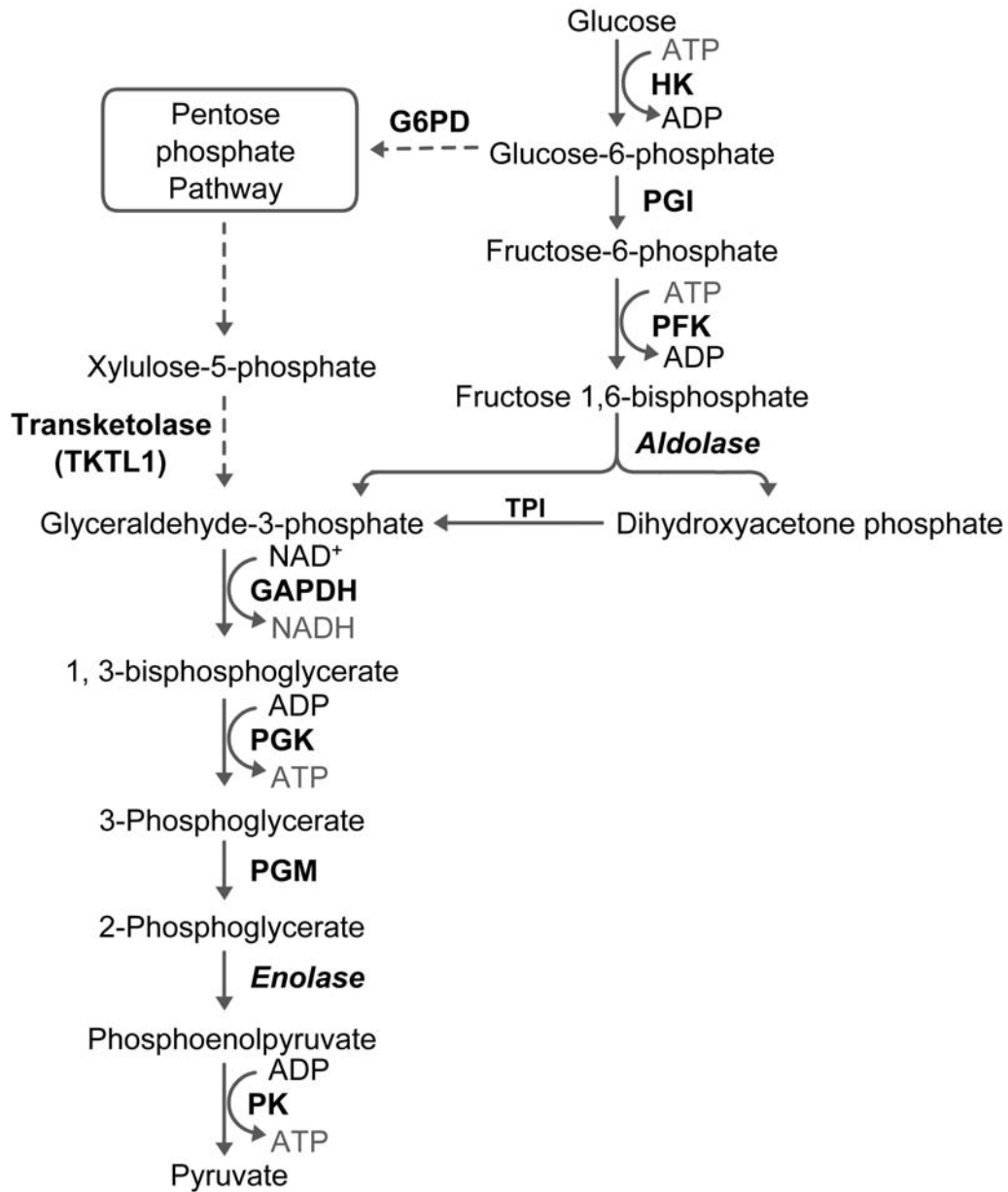
Glycolysis pathway. Glycolysis converts one molecule of glucose into two molecules of pyruvate and yields two molecules of both ATP and NADH. Abbreviations: HK (hexokinase); PGI (phosphoglucose isomerase); PFK (phosphofructokinase); TPI (triose phosphate isomerase); GAPDH (glyceraldehyde-3-phosphate dehydrogenase); PGK (phosphoglycerate kinase); PGM (phosphoglyceromutase) and PK (pyruvate kinase).

Origin of *GAPDH* Genes

In plants there are two metabolic pathways for carbohydrates: the Calvin Cycle in chloroplasts and glycolysis in the cytosol. The pathways share some enzymatic reactions (including the reaction catalyzed by GAPDH), but the enzymes in the two pathways are not identical even though they catalyze the same reactions in both pathways. The enzymes in the two pathways are isozymes or isoenzymes, homologous enzymes that catalyze the same reaction but differ in amino acid sequence. A separate gene encodes each isoenzyme, and all of the genes are nuclear (reviewed in Plaxton 1996). For example, the enzyme hexokinase phosphorylates glucose both in the chloroplast and in the cytosol, but two separate genes in the plant cell nucleus encode cytosolic hexokinase and chloroplastic hexokinase.

Isozymes are very common in plants and animals, and typically result from a gene duplication event that occurred millions of years ago. Sometimes the gene duplication event occurred within the nucleus itself. There are also genes located on chromosomal DNA that appear to have been transferred there from mitochondrial or plastid DNA. One of the observations about mitochondria and plastids that led to the endosymbiotic theory of evolution (that these organelles exist due to an ancient symbiotic event between prokaryotes and eukaryotes) is the fact that they contain DNA that is similar to bacterial DNA.

It was more than one billion years ago that photosynthetic cyanobacteria were engulfed by eukaryotic cells, becoming the antecedents of modern plastids. The resulting sub cellular organelles, plastids, have taken over many reactions for their host cells, including photosynthesis, carbohydrate metabolism, amino acid synthesis, lipid production, photorespiration, and nitrogen/sulfur reduction. At the same time, plastids still have their own DNA, as well as the machinery for replication, transcription, and translation. However, plastids retain only a fraction of the genome that their ancestors had. The plastid genome encodes between 120–135 genes (López-Juez 2007), whereas the closest living relative to the plastid ancestor, cyanobacteria of the genus *Nostoc* (Martin et al., 2002), have between 3,000–7,000 genes.

Most genes originally found in the symbiotic cyanobacteria are now found in the plant cell nucleus. Martin et al. (2002) report that about 18% of the protein-coding genes in *Arabidopsis thaliana* derive from cyanobacteria. However, the gene transfer was not one way. Genes that pre-existed in the nuclear genome have also been transferred to the plastid genome, but gene expression in the plastid is under nuclear control and most plastid proteins are encoded by nuclear DNA.

All GAPDH isozymes found in eukaryotes are nuclear-encoded and are believed to have originated in cyanobacteria (Martin et al., 2002). The duplication of *GAPDH* genes that gave rise to the chloroplastid form is believed to have occurred during the period when land plants first emerged (Teich et al., 2007), and subsequent gene duplications resulted in the multiple forms now present in modern plants.

GAPDH* Genes in *Arabidopsis thaliana

As the lab rat of plant research, *A. thaliana* has been studied extensively. *Arabidopsis* is a small flowering plant of the mustard family. There are several reasons why it is used as a model system, including its rapid life cycle (six weeks), small genome (125 Mb), and prolific seed production. Sequencing of the *Arabidopsis* genome was completed in the year 2000. More information about *Arabidopsis* genomics can be found on the TAIR (The *Arabidopsis* Information Resource) webpage (www.*arabidopsis*.org).

Since so much basic research is done using *Arabidopsis* as the model system, much is known about the seven *Arabidopsis GAPDH* known genes, including the chromosomes on which the genes are found (see the following page). The nomenclature for the *GAPDH* genes is determined by the enzyme localization and function:

- *GAPC* cytosolic
- *GAPCP* plastid
- *GAPA* chloroplast
- *GAPN* cytosolic, non-phosphorylating

Each gene represents a different GAPDH enzyme function, and several have multiple forms (isozymes). (Remember that GAPDH is a tetramer, and for each GAPDH, the four subunits may be identical or of two different types.).

Enzyme function	*GAPDH* protein subunits	EC designation	*Arabidopsis* gene	*Arabidopsis* mRNA accession number *	*Arabidopsis* chromosome	Position on chromosome
NAD$^+$-dependent *GAPDH* in cytosol	GAPC	EC 1.2.1.12	*GAPC*	NM_111283	3	1,080,957–1,083,537
	GAPC-2		*GAPC-2*	NM_101214	1	4,608,193–4,610,644
NAD$^+$-dependent *GAPDH* in plastids	GAPCP	EC 1.2.1.12	*GAPCP*	NM_106601	1	29,920,795–29,924,127
	GAPCP-2		*GAPCP-2*	NM_101496	1	5,574,304–5,574,616
NADP$^+$-dependent *GAPDH* in chloroplast	GAPA	EC 1.2.1.13	*GAPA*	NM_113576	3	9,796,308–9,798,282
	GAPA-2		*GAPA-2*	NM_101161	1	4,392,448–4,394,398
	GAPB		*GAPB*	NM_103456	1	16,132,283–16,132,283
Non-phosphorylating NADP$^+$ dependent *GAPDH* in cytosol	NP-GAP or ALD11A3	EC 1.2.1.9	*GAPN*	NM_201797	1	—

* The accession number is the identifier assigned to the sequence record when it is submitted to the National Institutes of Health National Center for Biotechnology Information nucleotide sequence database, GenBank. Accession number for *Arabidopsis* chromosome 1 is NC_003070 and chromosome 3 is NC_003074.

What is the EC designation? It is a numerical classification system from the Enzyme Commission (EC), assigning numbers to all enzymes based on the reactions that they catalyze. Each number of the EC designation has a particular meaning. For example, in the case of the phosphorylating GAPDH found in the cytosol:

EC 1._._._	designates enzymes that are oxidoreductases
EC 1.2._._	that act on the aldehyde or oxo group of donors
EC 1.2.1._	with NAD$^+$ or NADP$^+$ as acceptor
EC 1.2.1.12	GAPDH (phosphorylating) enzymes

Yet, where did all the *GAPDH* genes come from? They arose from a series of gene duplications over time. For example, *GAPB* originated by duplication of the *GAPA* gene, probably in early green algae (Brinkmann et al. 1989). Teich et al. (2007) developed a genetic phylogeny in plants using *GAPDH* from *Arabidopsis* and several other plants:
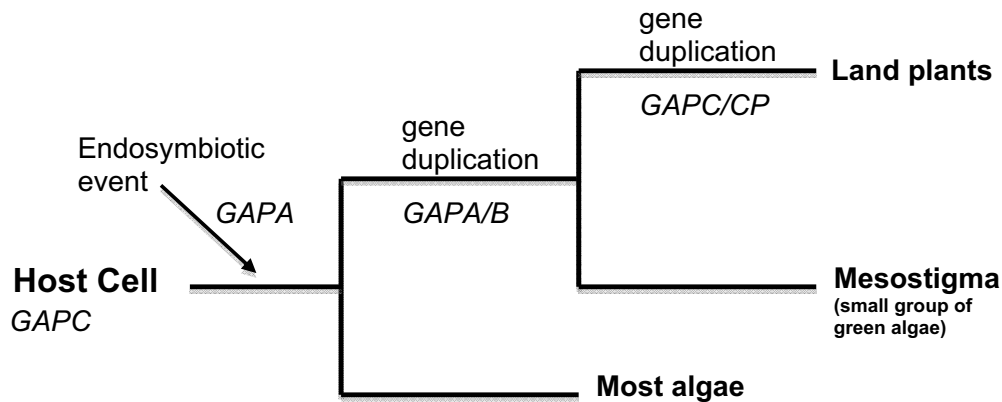
Phylogeny of *GAPDH* genes in plants. The original cells had a *GAPC* gene. When photosynthetic cyanobacteria were engulfed in the cells in an endosymbiotic event, the cells gained the *GAPA* gene. The next split was a gene duplication of *GAPA*, so all land plants (and *Mesostigma*, a small group of green algae) have the *GAPB* gene. The final gene duplication of *GAPC* separated the *Mesostigma* from land plants (which include dicots, monocots, clubmosses, mosses, and liverworts). Figure adapted from Teich et al. 2007.

PCR

Polymerase chain reaction (PCR) is a technique for rapidly creating multiple copies of a segment of DNA utilizing repeated cycles of DNA synthesis. PCR has revolutionized molecular biology and forensics, allowing amplification of small quantities of DNA into amounts that can be used for experimentation or forensic testing. Kary Mullis, who later won a Nobel Prize for his work, developed PCR in 1983. The subsequent discovery of a DNA polymerase that is stable at high temperatures and the introduction of thermal cyclers, instruments that automate the PCR process, brought the procedure into widespread use in the late 1980s.

From trace amounts of the DNA starting material (template), PCR produces exponentially larger amounts of a specific piece of DNA as targeted by primers. The template can be any form of DNA, and, theoretically, only a single molecule of DNA is needed to generate millions of copies. PCR makes use of two normal cellular activities: 1) binding of complementary strands of DNA, and 2) replication of DNA molecules by DNA polymerases.

DNA Structure

DNA strands are polymers of nucleotides, molecules comprised of a sugar, a phosphate group, and one of four bases: adenine, thymine, guanine, and cytosine (A, T, G, and C). The sugars and phosphates form the backbone of the DNA polymers. Each sugar has five carbons, making it a pentose. Each sugar is actually a deoxyribose because it has a hydrogen instead of a hydroxyl group (in RNA, the sugar is ribose, as it has the hydroxyl group). Each carbon in the sugar is numbered (see below), and the numbering is the source of the 3' and 5' nomenclature used for DNA. For example, the 5'-phosphate is the phosphate to which the next nucleotide will attach. It is called 5' because the phosphate group is attached to carbon number 5 of the sugar.

Each base forms hydrogen bonds with its complementary base, adenine with thymine (2 hydrogen bonds) and guanine with cytosine (3 hydrogen bonds). These pairings of A-T and C-G are called base pairs. Double-stranded DNA consists of two complementary strands of DNA held together by hydrogen bonding between the base pairs. The two strands are antiparallel, meaning that the strands are oriented in opposite directions. One strand, the sense or coding strand, has bases running 5' to 3', and the second strand, the antisense strand, has the complementary bases running 3' to 5'. When DNA is transcribed, the antisense strand serves as the template for synthesis of messenger RNA (mRNA). The mRNA will have the same sequence as the sense or coding strand of DNA (with uracil instead of thymine).
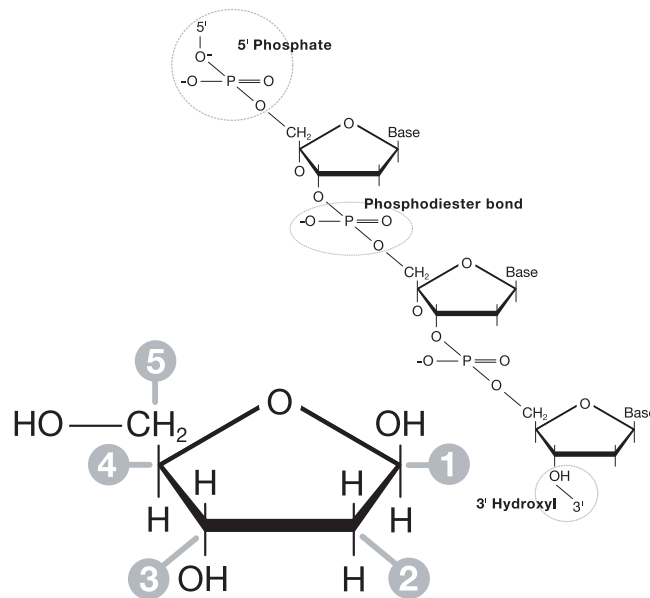

Chemical structure of deoxyribose sugar and deoxyribose nucleic acid (DNA).
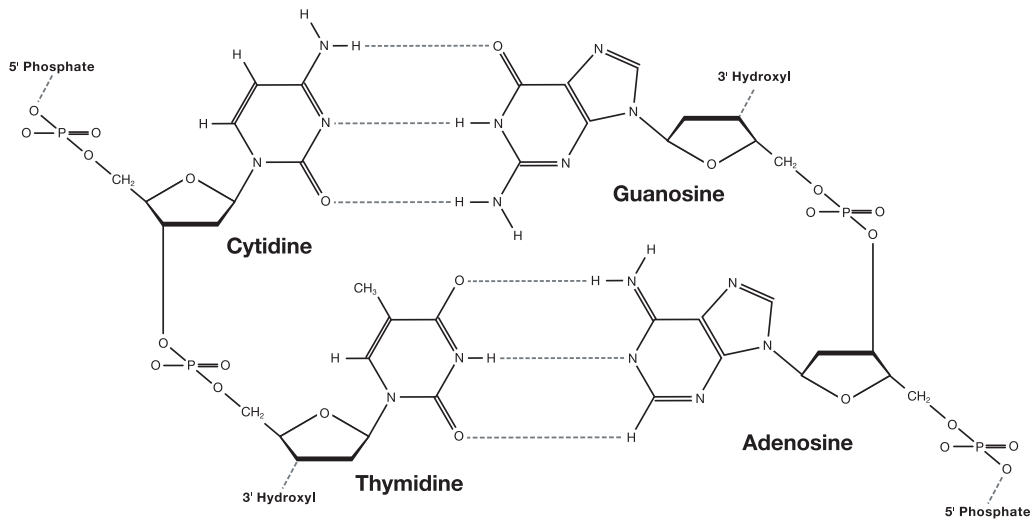
Base pairs of DNA. Cytidine base pairs with guanosine by three hydrogen bonds. Likewise thymidine base pairs with adenosine by two hydrogen bonds.

DNA Replication

DNA replication is an essential part of life. As cells divide, DNA must be duplicated, and the new DNA molecules must be exact copies of the original DNA. DNA polymerases are enzymes that synthesize the new DNA strands, and they are found in all cells. DNA polymerases link together free nucleotides in the order determined by the order of the template DNA that the polymerase follows. The new strand will be complementary to the template strand. In other words, each base of the new strand is the complement of the base in the template strand. For each A in the template, the new strand will have a T. For each G in the template, the new strand will have a C, etc. Since DNA polymerase can use only single-stranded DNA as a template (and since it can synthesize DNA only in one direction, 5' to 3'), double-stranded DNA must be uncoiled and the strands separated before the DNA can be replicated. DNA polymerase also needs a signal to tell the enzyme where to start synthesis and something to which the first base can be joined. This primer is a short strand of nucleotides that bind to the template DNA at the starting point and become the 5' end of the new DNA strand. In cellular DNA replication, the primers are small RNA molecules, but in PCR in the laboratory, the primers are DNA molecules.
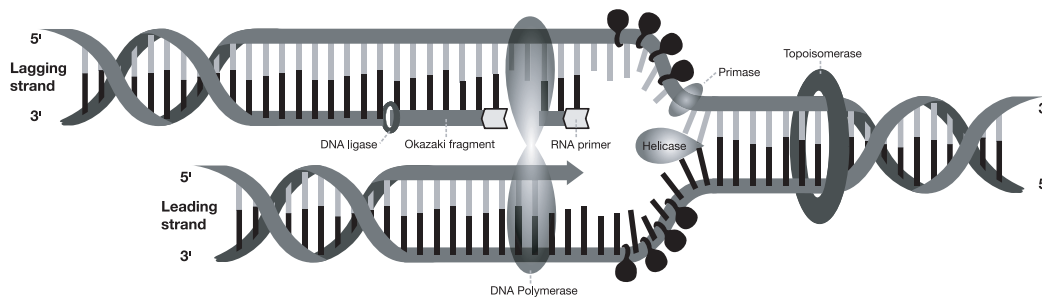

So, in its essentials, DNA replication sounds simple: unwind the double-stranded template, bind a primer to each strand to give the DNA polymerase a starting point, and the enzyme will produce replicated DNA strands. In reality, the process is much more complicated. There are as many as 40 proteins involved in DNA replication in eukaryotes. Without detailing all of the proteins involved, the steps of cellular DNA replication are:

1. Template DNA strands begin to separate at the **origin of replication**. An enzyme called DNA helicase breaks the hydrogen bonds between the base pairs to separate the strands. The point where the two strands separate is called the replication fork.

2. As the strands unwind and separate, the DNA ahead of the replication fork starts to form supercoils. An enzyme named **topoisomerase** moves ahead of the replication fork, nicking single strands of the double-stranded DNA and relaxing the supercoiled structure.

3. To keep the two strands from reannealing (binding to each other again), single-strand DNA binding proteins bind to each of the separated strands.

4. Since DNA polymerase can add nucleotides only to the 3' end of an existing nucleotide, an enzyme named primase binds to each of the template DNA strands and assembles a short primer of RNA. (The RNA primer will later be removed and replaced by DNA in the new strands.)

5. DNA polymerase begins to synthesize DNA by adding new nucleotides to the RNA primers.

6. Since DNA polymerase can synthesize DNA in only one direction, from 5' to 3' on the template strands, synthesis actually proceeds differently on the two template strands. On the 5' to 3' template strand, called the **leading strand**, DNA synthesis proceeds continuously, moving toward the replication fork.

   On the second strand, called the **lagging strand**, synthesis moves away from the replication fork and is discontinuous. DNA on the lagging strand is synthesized in short pieces (100 to 2000 bases) called Okazaki fragments. After the Okazaki fragments are synthesized, they are joined together by DNA ligase.

7. Although DNA polymerase is a **high-fidelity** enzyme, meaning that it makes few mistakes in replicating the bases, it does make some mistakes. In eukaryotic DNA replication, the error rate is one mistake in every 10,000 to 100,000 basepairs. Many DNA polymerases also have proofreading activity, which means that they can find mistakes and correct them as the enzyme moves along the template.

8. DNA replication in eukaryotes does not begin at a single origin of replication, but at numerous locations along a DNA molecule. Origins of replication are found about every 100 kilobases in eukaryotic cells. (Mammalian cells are estimated to have ~30,000 origins of replication.) In addition, at each origin of replication, there are actually two replication forks formed that head in opposite directions, and, although the description above referred only to replication at one fork, replication is occurring simultaneously (and in the opposite direction) at the other fork. So, replication moves along the template DNA molecule until the replication fork meets a fork coming in the opposite direction.
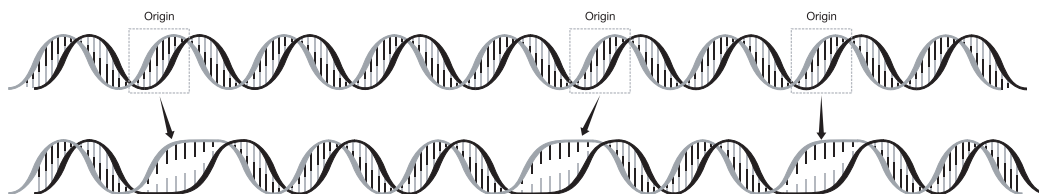
Each replication of DNA produces two strands of DNA, each identical to the original strand. Eukaryotic DNA replication is called **semiconservative** because each double-stranded product consists of one of the original strand and one newly synthesized strand.

PCR Step by Step

The strength of PCR lies in its ability to make many copies of (**amplify**) a single region (**target**) of a longer DNA molecule. For example, a scientist wanting to study a single human gene needs to amplify only that portion from the enormous human genome of approximately 3.3×10^9 base pairs! The first step is to identify and sequence areas of DNA upstream and downstream from the DNA of interest. Once this is done, short strands of DNA that are **complementary** to the upstream and downstream DNA are synthesized. As in cellular DNA replication these oligonucleotide **primers** are used as the starting point for copying the DNA of interest, but unlike cellular DNA replication, the primers used in PCR are DNA oligonucleotides not RNA.

Taq **DNA polymerase.** Originally, the DNA synthesis step of PCR was performed at 37°C using DNA polymerase from *E. coli*, but the enzyme was inactivated during the high-temperature denaturation step in each cycle. The enzyme thus had to be added anew during each cycle. The 1988 discovery of a thermally stable DNA polymerase brought PCR into the mainstream. *Taq* DNA polymerase was isolated from *Thermus aquaticus*, thermophilic bacteria that live in hot springs in Yellowstone National Park. Since the hot springs frequently approach boiling temperatures, *T. aquaticus* and other bacterial species that live in these waters must have enzymes that are functional at high temperatures. DNA polymerase from *T. aquaticus* is not inactivated by the denaturation step in PCR.

Since the discovery of *Taq*, several other thermostable DNA polymerases have been isolated. *Taq* has a drawback for DNA synthesis in PCR as it lacks a proofreading mechanism to catch and correct errors in the new DNA strand, so *Taq* is said to have low replication fidelity. In 1991, scientists discovered and characterized *Pfu* DNA polymerase from *Pyrococcus furiosus*, a thermophilic type of Archaebacteria. *Pfu* DNA polymerase has the proofreading capacity that Taq lacks, so use of *Pfu* means that there are fewer errors in the new DNA strands.

Subsequently, a number of companies have developed modified versions of DNA polymerases. For example, Bio-Rad's iProof™ polymerase, which is a DNA polymerase with proofreading capability similar to *Pfu*, fused to a protein that binds double-stranded DNA.

PCR involves a repetitive series of cycles, each of which consists of template denaturation, primer annealing (binding to the template DNA strand), and extension of the annealed primer by a heat-stable DNA polymerase.

All of the components needed for PCR are mixed in a micro test tube. They are:

- Template DNA

- *Taq* DNA polymerase (or another thermally stable DNA polymerase)

- Primers — synthesized to complement a specific region on the template DNA. The primers are added in excess (that is, there are many more primer molecules than template molecules in the reaction tube)

- Nucleotides — the four individual bases in the form of deoxynucleoside triphosphates (dNTPs), which allows them to be added to a DNA polymer. The dNTP mixture includes the same amounts of dATP, dTTP, dCTP, and dGTP

- Reaction buffer — prepared with the correct ionic strength of monovalent and divalent cations needed for the reaction, and buffered to maintain the pH needed for enzyme activity

The micro test tubes are specialized tubes used only for PCR. PCR tubes are plastic with very thin walls, allowing rapid transfer of heat through the plastic. The tubes usually hold only 0.2 or 0.5 ml. The PCR reaction tubes are placed in a thermal cycler, an instrument developed in 1987 that automates the heating and cooling cycles. Thermal cyclers contain a metal block with holes for the PCR tubes. The metal block can be heated or cooled very rapidly. Thermal cyclers are programmable, so the PCR reaction parameters (temperatures, time at each temperature, and number of cycles) can be stored by the instrument. This means that the user can just load the samples and push a button to run the reactions. Contrast this to the early researchers who had to sit by a series of water baths with a timer, switching the tubes from one temperature to another temperature manually for hours!

The first step of the PCR reaction is the denaturation step. Since DNA polymerase can use only single-stranded DNA as a template, the first step of PCR is uncoiling and separating the two strands of the template DNA. In cells, enzymes such as helicase and topoisomerase do this work, but, in PCR heat is used to separate the strands. When double-stranded DNA is heated to 94°C, the strands separate, or denature. Since complete denaturation of the template DNA is essential for successful PCR, the first step is frequently an extended denaturation period of 2–5 min. The initial denaturation is longer than subsequent denaturation steps because the template DNA molecules are longer than the PCR product molecules that must be denatured in subsequent cycles. Denaturation steps in subsequent PCR cycles are normally 30–60 sec.

The thermal cycler then rapidly cools the reactions to 40–60°C to allow the primers to anneal to the separated template strands. The temperature at which the primers will anneal to the template DNA depends on several factors, including primer length, G-C content of the primer, and the specificity of the primer for the template DNA. If the primer sequences complement the template sequences exactly, the primers will anneal to the template DNA at a higher temperature. As the annealing temperature is lowered, primers will bind to the template DNA at sites where the two strands are not exactly complementary. In many cases, these mismatches will cause the strands to dissociate as the temperature rises after the annealing step, but they can result in amplification of DNA other than the target.
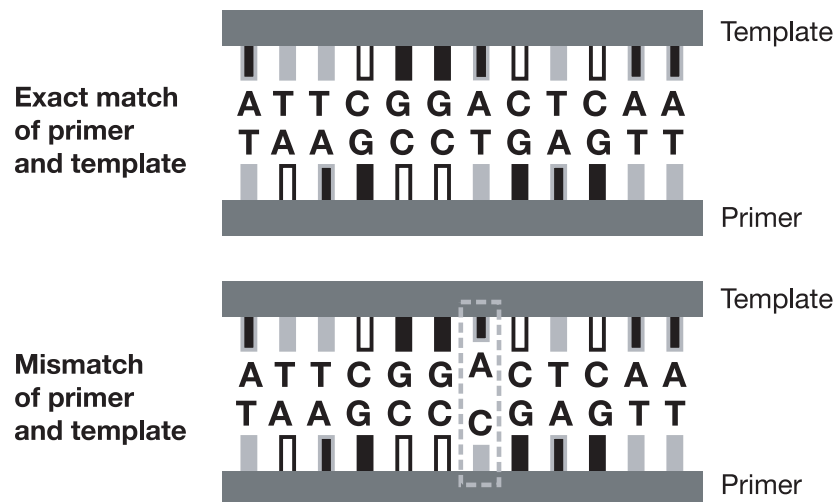
In the annealing step, the two original strands may re-anneal to each other, but the primers are in such excess that they out compete the original DNA strands for the binding sites.

The final step is extension, in which the reaction is heated to 72°C, the optimal temperature for *Taq* DNA polymerase to extend the primers and make complete copies of each template DNA strand.

At the end of the first PCR cycle (one round of denaturation, annealing, and extension steps define one cycle), there are two new strands for each original double-stranded template, which means there is twice as much template DNA for the second cycle of PCR. As the cycle is repeated, the number of strands doubles with each reaction. For example, after 35 cycles, there will be over 30 billion times more copies of the target sequence than at the beginning. The number of cycles needed for amplification depends on the amount of template DNA and the efficiency of the reaction, but reactions are frequently run for 30–40 cycles.
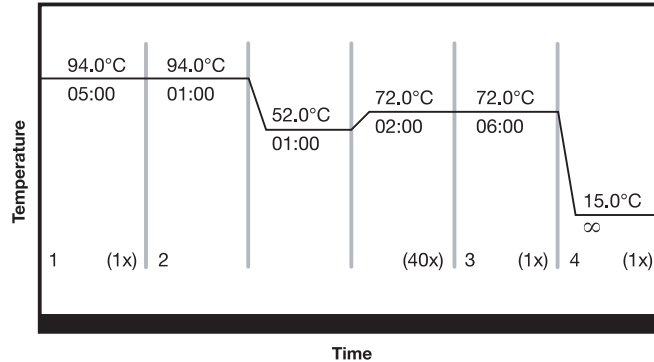
Example thermal cycling profile. In this profile an intial denaturation step of 95°C for 5 min is followed by 40 cycles of one minute denaturation, one minute annealing and two minutes extension. A final 6 minutes extension time is added to ensure completion of DNA synthesis. The final hold ensures samples are kept stable until the samples are retrieved.

PCR generates DNA of a precise length and sequence. During the first cycle primers anneal to the original template DNA strands at opposite ends and on opposite strands. After the first cycle, two new strands are generated that are shorter than the original template strands but still longer than the target DNA since the original template sequence continues past the location where the other primer binds. It isn't until the third PCR cycle that fragments of the precise target length are generated.
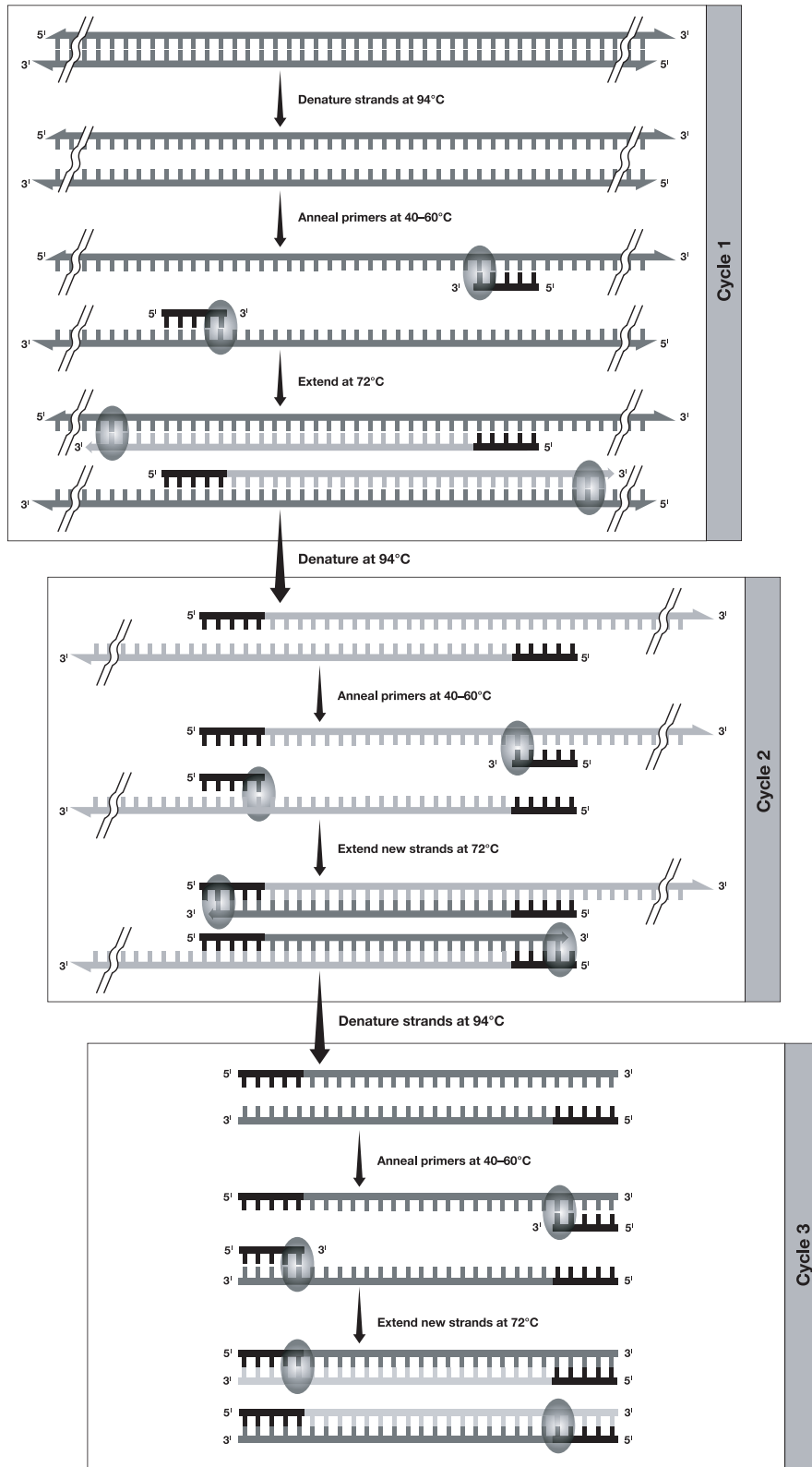
First three cycles of PCR. PCR takes three cycles before a product of the correct length is generated.

Primer design

Probably the most important variable in PCR is design of the primers. The primers will determine whether or not the correct piece of DNA is amplified. Primers are short, single-stranded oligonucleotides, synthesized in the laboratory and designed to bind to the DNA template strands at the ends of the sequence of interest. (Actually, very few laboratories prepare their own primers, as there are many companies that make primers to order — quickly, cheaply, and accurately.) Primer design is usually the responsibility of the researcher, and there are computer programs and websites that assist in primer design. Normally two different primers are needed, one for each of the complementary strands of template DNA.

Factors to be considered in designing primers include:

- Length: Primers of 18–30 nucleotides are likely to be specific for their target sequence. In other words, primers of this length are less likely than shorter primers to bind to sites on the template DNA other than the sites for which they were designed

- Melting temperature of primers (Tm): The Tm is the temperature at which half the primers dissociate from the target DNA. It is important that the two primers used in each PCR reaction have similar Tm (within 5°C). If the melting temperatures are very different, the primers will not bind equally during the annealing stage. Tm is determined by primer length and GC content. More energy is required to dissociate the 3 hydrogen bonds between G and C compared to the 2 hydrogen bonds joining A and T. T_m for primers around 18–24 bases in length can be estimated by their nucleotide content using the formula:

$$T_m = 2°C\ (A+T) + 4°C\ (G+C)$$

Another formula for Tm determination is:

$$T_m = 81.5 + 16.6\ (\log_{10} [I]\ ) + 0.41\ (\%G+C) - 600/n$$

where I is the molar concentration of monovalent cations and n is the number if bases in the primer. This formula will give accurate T_m in °C for primers from 20–100 bases long.

There are many website tools that will calculate exact Tm for primers. Free calculators include OligoCalc (hosted at Northwestern University, University of Pittsburgh, JustBio.com, and others) and PrimerFox (http://www.primerfox.com/). There are also calculators available on many biotech company websites, and some fee-based calculators.

- Annealing temperature: The temperature for the annealing step of PCR should be about 5°C below the T_m of the primers. A good target for annealing temperature is ~50–60°C

- GC content: The primer should have 40–60% guanine and cytosine. Higher GC content will give the primers too high a T_m for PCR. (If the T_m is too high, the annealing temperature can exceed the optimal temperature for *Taq* polymerase extension of the DNA strand.) Long stretches of any single base should be avoided. An ideal primer should have a random mix of bases with ~50% GC

- Intra- or inter-primer complementarity: Primers should not have any regions of complementarity longer than three bases. Otherwise, primers can form hairpins by internal annealing, or can form double-stranded structures that will interfere with PCR. Also it is very important that there not be complementarity at the 3'-ends of the two primers, as that can result in a PCR artifact called primer dimers. If primers hybridize at their 3'-ends, the hybrid molecule can act as a template for DNA polymerase, resulting in an unwanted PCR product. Primer dimers are more likely to be produced when the primers do not bind efficiently to the template DNA

- GC-clamp: The sequence of the primers at the 3'-end is important to ensure correct and strong binding of the primer to the template. If the primers contain GC clamps, one to three G or C bases at the 3'-end of the primer, the primer will form a more stable complex with the template DNA

Degenerate Primers

Normally PCR primers are unique sequences of nucleotides, designed based on the known sequence of the target DNA. When the sequence of the template DNA is not known, there are several alternate approaches for primer design. One approach is to take advantage of genetic homology among closely related organisms. For example, the target DNA may not have been sequenced in the species of interest, but the gene may have been sequenced in several other species. Genes that code for the same protein in different organisms are likely to have sequences that are conserved, very similar or even identical in the different species. These conserved sequences usually code for parts of the protein that are essential for function; in other words, mutations in these areas are likely to be detrimental to the organism, so evolution discourages changes.

If genomic DNA (gDNA) or mRNA sequences from similar species are aligned, a consensus sequence can be derived. The consensus sequence may be exactly the same in all species, or it may have one or more bases that vary among the species. For example, a consensus sequence could be represented by A-C-T-G-G-N-T-T-A-C-C-G, where A, C, G, and T represent the bases that are the same in all of the species compared, and N represents a base that varies. In other words, the base at the N position might be G in some species and C in others. See the table below for an example of determining a consensus sequence.

Since the goal of PCR is to amplify the DNA region of interest, primers are designed to bind to sequences on either side of that region. Once the two primers have been designed based on the consensus sequences derived from other organisms, it is possible that the primers will have enough complementarity with the target DNA to bind during the annealing step, and, in addition, binding of mismatched primers to template DNA can be encouraged by reducing the annealing temperature. However, to increase the probability that the primers will bind to the target DNA, one or more bases within the primers are substituted with the other three bases, introducing degeneracy to the primer sequences. (This is also described as introducing "wobbles" in the primer; the higher the degeneracy, the more "wobbles" in the primer.) In a simplified example, if the consensus sequence is NATC, the set of degenerate primers would be AATC, TATC, GATC, and CATC, assuming that all four bases are used.

However, in many cases, all of the bases are not used to substitute for the variable base. To increase the probability that the primer will anneal to the target DNA, the variable base is substituted with a similar base. For example, if the variable base is T, it might be replaced only with A (the other pyrimidine). There is a code to specify which bases to substitute at each variable position:

IUB Code	Bases	Derivation of IUB Code
N	A/G/C/T	any
K	G/T	keto
S	G/C	strong
Y	T/C	pyrimidine
M	A/C	amino
W	A/T	weak
R	G/A	purine
B	G/T/C	not A
D	G/A/T	not C
H	A/C/T	not G
V	G/C/A	not T

Oligonucleotide International Union of Biochemistry (IUB) codes for mixed (wobble) bases.

Plant	*Gene*	Accession number	Sequence
Arabidopsis	*GAPC1*	AT3G04120	GACTACGTTGTTGAGTCTACTGG
Arabidopsis	*GAPC2*	AT1G13440	GACTTTGTTGTTGAGTCTACTGG
Arabidopsis	*GAPCP1*	AT1G79530	GATTATGTTGTTGAGTCTTCCGG
Arabidopsis	*GAPCP2*	AT1G16300	GAGTATGTTGTTGAGTCTTCAGG
Pepper	*GAPCP*	CAN272042	GATTATGTTGTTGAATCTTCTGG
Liverwort	*GAPC*	AJ246023	GAGTACGTCGTCGAGTCTACCGG
Corn	*GAPC1*	ZMGPC1	GAGTACGTCGTGGAGTCCACCGG
Corn	*GAPC2*	U45855	GAGTATGTCGTGGAGTCCACCGG
Corn	*GAPC3*	U45856	GAATATGTTGTTGAGTCTACTGG
Corn	*GAPC4*	X73152	GAATATGTTGTTGAGTCTACTGG
Pea	*GAPC1*	L07500	GATATCATTGTTGAGTCTACTGG
Wheat	*GAPC*	EF592180	GAGTACGTTGTTGAGTCCACCGG
Rye Grass	*GAPC3*	EF463063	GACTACGTTGTTGAGTCCACTGG
Tobacco	*GAPC*	AJ133422	GATTACATTGTGGAGTCGACTGG
Tobacco	*GAPDH**	DQ682459	GATTTCGTTGTGGAATCCACTGG
Carrot	*GAPDH**	AY491512	GAGTACATTGTGGAGTCCACTGG
Blue Gem	*GAPDH**	X78307	GAGTACGTCGTTGAGTCGACTGG
Tomato	*GAPDH**	AB110609	GACTTCGTTGTTGAATCAACCGG
Snapdragon	*GAPDH**	X59517	GAGTATATTGTGGAGTCCACTGG
INITIAL FORWARD PRIMER			GABTATGTTGTTGARTCTTCWGG
Position of base			123456789

*Specific *GAPDH* gene not specified in GenBank

Example: Designing PCR primers by deriving a consensus sequence and introducing degeneracy.

The table above shows how alignment of *GAPC* genes from different plant species can be used to derive a consensus sequence that can then be used for primer design. The plant species are listed on the left and the *GAPC* genes are aligned on the right. The vertical blue highlighting in the sequences are bases conserved across all the species. Deriving the consensus sequence for the gene begins with the conserved bases. For example, all of the sequences begin with GA, so the consensus sequence will also begin with GA.

Although the other bases are not conserved, the differences between the species are not random. For example, the bases in positions 4 and 5 are always either A or T. (A and T are considered "weak" bases, as the base pairs they form are not as strong as those formed by G and C.) The base that is more commonly found in that position will be the one used in the consensus sequence, e.g., position 5 will be A, as A is found in 15 of the 19 sequences.

After most of the consensus sequence has been determined, degeneracy can be introduced at one or more positions, normally at the positions that show the most variability among species. For example, position 3 can be A (2 sequences), G (8 sequences), C (4 sequences), or T (5 sequences). Since A is less frequent, only G, C, or T will be substituted at position 3. Position 3 will be designated **B**, as in the code above, B means G/C/T. The second variable base will be designated **R**, as all the bases at that position are purines (G/A). The final variable base will be **W** (A/T).

Designing Primers Without DNA Sequence

Frequently a researcher may have purified a protein of interest and obtained some amino acid sequence data from the protein, but not have any of the DNA sequence for the protein. When that happens, there is another approach for designing primers. Since organisms use more than one codon for some amino acids (see table), primer mixtures can be synthesized that include all possible codons for each amino acid. Although it seems as though there would be huge numbers of oligonucleotides needed, the task can be simplified in several ways, such as choosing an area of protein sequence that is heavy in amino acids that are encoded by only one or two codons.

Ala	Arg	Asp	Asn	Cys	Glu	Gln	Gly	His	Ile
GCA	CGA	GAC	AAC	TGC	GAA	CAA	GGA	CAC	ATA
GCC	CGC	GAT	AAT	TGT	GAG	CAG	GGC	CAT	ATC
GCG	CGG						GGG		ATT
GCT	CGT						GGT		
	AGA								
	AGG								

Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
CTA	AAA	ATG	TTC	CCA	TCA	ACA	TGG	TAC	GTA
CTC	AAG		TTT	CCC	TCC	ACC		TAT	GTC
CTG				CCG	TCG	AGC			GTG
CTT				CCT	TCT	ACT			GTT
TTA					AGC				
TTG					AGT				

Amino acids and the DNA codons for each.

Note that when there are multiple codons for an amino acid, the codons are very similar. For all amino acids with up to four codons, only the third base is different between codons. For example, the four codons for valine all begin with GT and the third position changes. This third, changeable base is the "wobble" base. Also three codons code for a stop signal, these are TAG, TGA, and TAA.

By choosing amino acids with fewer codons, the number of degenerate primers can be minimized. For example, to make degenerate primers to the DNA that codes for the amino acid sequence Gly-Leu-Ser-Val, the mixture would include 576 different oligonucleotides:

Amino acid	Number of codons	Number of degenerate primers
Glycine	4	
Leucine	6	$4 * 6 * 6 * 4 = 576$
Serine	6	
Valine	4	

In comparison, degenerate primers to the amino acid sequence Asp-Trp-Cys-Glu would include only 8 different oligonucleotides:

Amino acid	Number of codons	Number of degenerate primers
Aspartic acid	2	
Tryptophan	1	$2 * 1 * 2 * 2 = 8$
Cysteine	2	
Glutamic acid	2	

Note that the sequence examples are only four amino acids in length, making the primers only twelve oligonucleotides long. Degenerate primers are usually longer, meaning that there will be more oligonucleotide combinations needed. To keep the number needed to a minimum, choose a target amino acid sequence containing amino acids coded by only one or two codons and try to avoid amino acids that have six codons.

The task is made somewhat simpler by the codon bias displayed by different organisms. Specific organisms have a tendency to favor particular codons. For example, humans use all six codons for arginine with relatively equal frequency, but the bacterium *E. coli* uses only two of the six codons most of the time. If the codon bias of the organism is known, then omitting those codons that are less likely to be used by the organism can reduce the number of degenerate primers needed.

Controls

PCR is notoriously sensitive to contamination by unwanted template DNA. This is a particularly serious problem when the amount of template DNA is limited. (Examples in which template DNA is limited are during amplification of ancient DNA and single-molecule PCR.) Practical measures to reduce contamination include use of pipet tips with filters, which act as barriers to prevent aerosol contamination of samples. All tubes and other materials should be handled with clean gloves. Most importantly, though, is that control reactions should be included in each PCR experiment. The negative control reaction should include exactly the same components as the experimental tubes but without any template DNA.

Analyzing Results

PCR products can be visualized by agarose gel electrophoresis, with the agarose concentration determined by the expected size of the products. In addition to the experimental samples, the negative control reaction and a size marker should be included on the gel. The size markers help determine if the size of the PCR product is as expected. The negative control reaction should not yield any amplified DNA. If it does, then the reactions may have been contaminated and the experimental results are suspect. However, small PCR products may be primer-dimers that have been amplified.
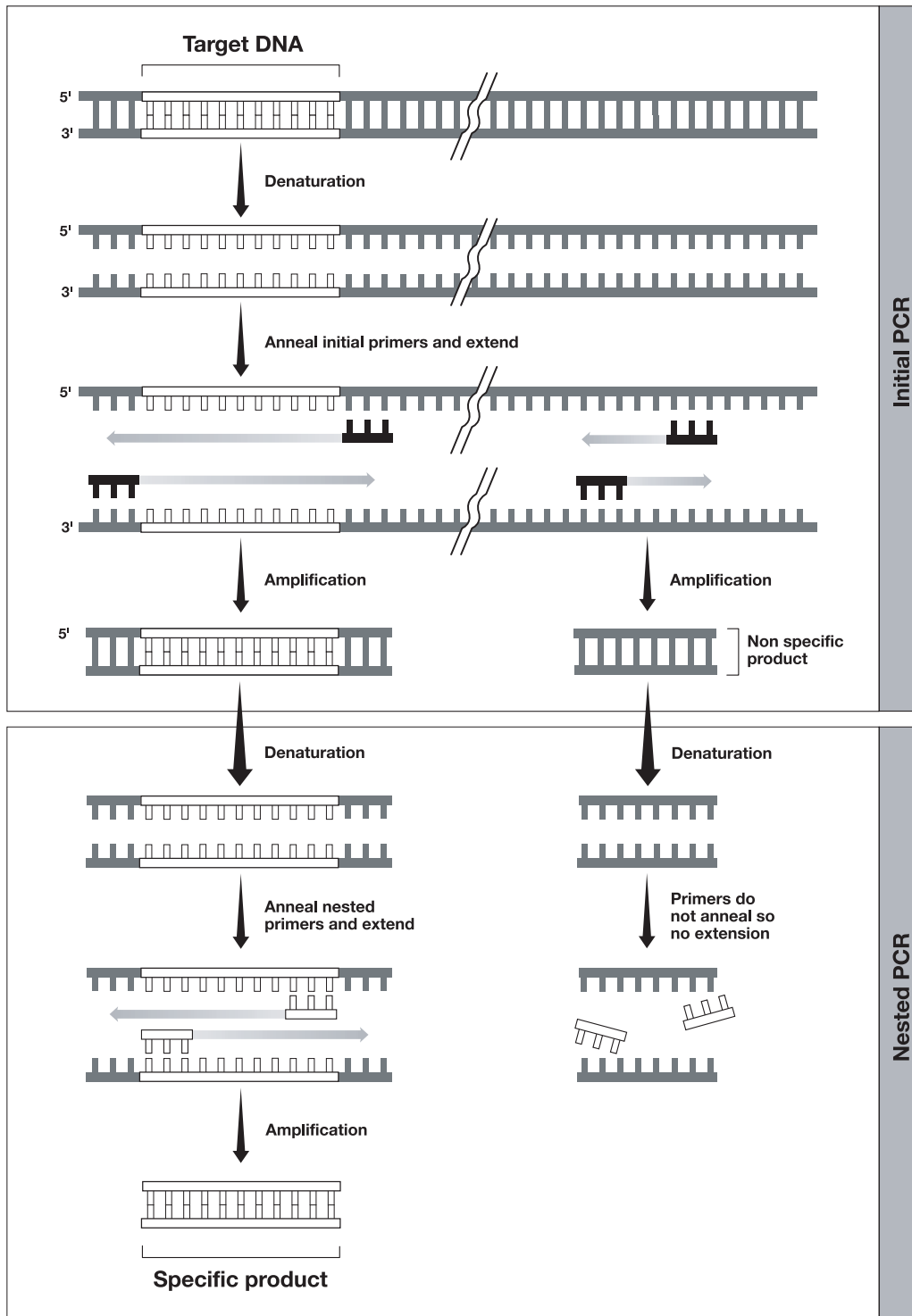
Nested PCR

A number of variations of PCR have been developed in the last 20 years to address specific research questions, including inverse PCR, *in situ* PCR, long PCR, real time PCR, and nested PCR. When there is the potential for primers to bind to sequences of the template DNA other than at the target area (for example, when using degenerate primers), nested PCR can increase the yield and specificity of amplification of the target DNA. Nested PCR uses two sequential sets of primers. The first primer set binds outside the target DNA, as expected in standard PCR, but it may also bind to other areas of the template. The second primer set binds to sequences in the target DNA that are within the region amplified by the first set (that is, the primers are nested). Thus, the second set of primers will bind and amplify target DNA within the products of the first reaction. The primary advantage of nested PCR is that if the first primers bind to and amplify an unwanted DNA sequence, it is very unlikely that the second set of primers will also bind within the unwanted region.

Nested PCR of plant *GAPC* gene

In this experiment, nested PCR will be used to amplify a portion of *GAPC*, the plant gene for NAD$^+$-dependent, cytosolic *GAPDH*. The section of the gene to be amplified encodes around two thirds of the protein including the active site of the enzyme. The primer annealing sites and the position of the primer annealing sites relative to the intron-exon structure of the *Arabidopsis* genes are shown below.

```
GACTACGTTGTTGAGTCTACTGGTGTCTTCACTGACAAAGACAAGGCTGCAGCTCACTTGAAGGTTTGTCT
TATTTGAATTGGTTATTTTTGTCTTGTAATGATATAAATAGTTTATGTGCTAGAATTTGCTTAGTATCATT
CAACTAAATTTGTGACTTGTTGTATTTTCAGGGTGGTGCCAAGAAGGTTGTTATCTCTGCCCCCAGCAAAG
ACGCTCCAATGTTTGTTGTTGGTGTCAACGAGCACGAATACAAGTCCGACCTTGACATTGTCTCCAACGCT
AGCTGCACCACTAACTGCCTTGCTCCCCTTGCCAAGGTAAAATATCTGATATTCTATATGATCAAATTTGA
CTTTGTATTTCAAGTTGAAGTGACTAATTTCATTTAACGTTCTTTGATTTCATTGTGTAGGTTATCAATGA
CAGATTTGGAATTGTTGAGGGTCTTATGACTACAGTCCACTCAATCACTGGTAAATTTATCAATCAGTTAG
AAGTTTATTACAAACTTGCTTGCCTATAGGTGGAAAATTTGTGATTTAATGGGGTTTGCTTTATGATTTCA
GCTACTCAGAAGACTGTTGATGGGCCTTCAATGAAGGACTGGAGAGGTGGAAGAGCTGCTTCATTCAACAT
TATTCCCAGCAGCACTGGAGCTGCCAAGGCTGTCGGAAAGGTGCTTCCAGCTCTTAACGGAAAGTTGACTG
GAATGTCTTTCCGTGTCCCAACCGTTGATGTCTCAGTTGTTGACCTTACTGTCAGACTCGAGAAAGCTGCT
ACCTACGATGAAATCAAAAAGGCTATCAAGTAAGCTTTTGAGCAATGACAGATTAAGTTTACTTATATTCC
AGTAGTGATCAAATTACTCACCAAGTGTTTTTACCACCAATACATAGGGAGGAATCCGAAGGCAAACTCAA
GGGAATCCTTGGATACACCGAGGATGATGTTGTCTCAACTGACTTCGTTGGCGACAACAGGTCGAGCATTT
TTGACGCCAAGGCTGGAATTGCATTGAGCGACAAGTTTGTGAAATTGGTGTCATGGTACGACAACGAATGG
```

Fig. *GAPDH* **PCR Primers.** Positions of initial *GAPDH* PCR primers (bold) and nested PCR primers (italics underlined and highlighted) on *Arabidopsis GAPC* genomic DNA are indicated. Note: reverse primers are complimentary to the reverse compliment DNA strand.

Target DNA

5'
3'

Denaturation

5'

3'

Anneal initial primers and extend

5'

3'

Amplification

Amplification

5'

Non specific product

Initial PCR

Denaturation

Denaturation

Anneal nested primers and extend

Primers do not anneal so no extension

Amplification

Specific product

Nested PCR

Nested PCR. Nested PCR involves two rounds of PCR with the product of the first round acting as a template for the second round.

24
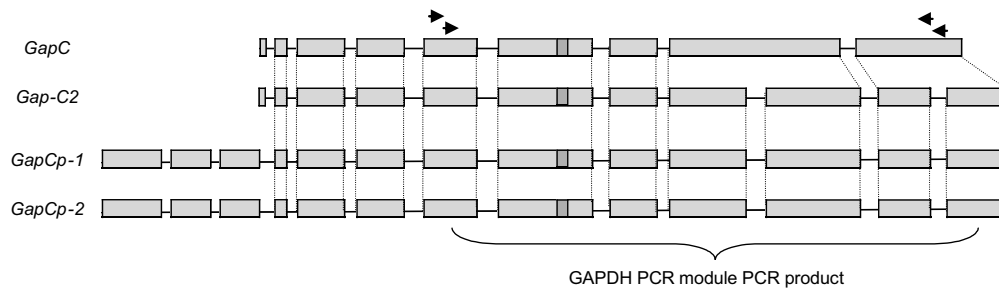
Intron-exon schematic of *Arabidopsis GAPC* family of genes. *GAPC* differs from the rest of the family by the absence of 2 introns which shortens the gene. The *GAPCP* sub-family of genes have a signal peptide at their N termini. Arrows indicate annealing positions of initial (outer arrows) and nested (inner arrows) *GAPDH* PCR primers on exon-intron structure of *GAPC* family genes. Darker bar indicates location encoding enzyme active site. Note: figure is not to scale.

While the *GAPDH* protein sequence is highly homologous between family members and between species, the gene structure and intronic sequences are more variable. This can be observed in the differences in gene structure between the *Arabidopsis GAPC* gene family, where *GAPC* is missing 2 introns present in the other family members, resulting in a shorter PCR product- see below. This results in PCR products of different lengths which can be identified by agarose gel electrophoresis. Moreover in our studies investigating other plant species during the development of this lab, we have found numerous other instances of missing introns. Indeed the houseplant pathos had no introns in the amplified region resulting in a very short PCR product of only 602 bp.

Since the initial primers are degenerate and were designed based on a consensus sequence derived from a number of *GAPDH* genes (including isozymes of *GAPC* and *GAPCP*, among others; see table above), the initial primers may bind to the target DNA at several locations. These locations may be sequences of *GAPDH* genes other than *GAPC*, or they may be unrelated sequences that have a high degree of complementarity to the one or more of the degenerate primers. So, it is likely that multiple bands of amplified DNA may be seen on an agarose gel after the initial round of PCR. The sizes of PCR products expected from *Arabidopsis GAPC* family genes using the primers in this lab are:

Arabidopsis GAPC Gene	Length of PCR product (bp)	
	Initial primers	**Nested primers**
GAPC	1065	993
GAPC-2	1216	1145
GAPCP-1	1303	1231
GAPCP-2	1205	1133

Note: The pGAP plasmid contains the sequence for the *GAPC* gene Initial PCR product.

The nested primers were designed to be more specific to *GAPC* (rather than the *GAPCP* sub-family) and are not degenerate, so in the second round of PCR, only the *GAPC* genes from the initial PCR will be amplified. For example, if your gDNA is from *Arabidopsis*, then the nested PCR should amplify only *GAPC* and *GAPC-2*. The nested primers should not bind to DNA coding for *GAPDH* isozymes or to unrelated DNA sequences, so that DNA should not be amplified.

What does it mean if no DNA is visible on an agarose gel after the nested PCR? If the experimental controls worked (meaning that the problem was not with the reagents or the thermal cycler), then it is likely that no *GAPC* was amplified from your genomic DNA sample. The most probable reason is that the initial primers did not bind to any target DNA because there was too little complementarity between the primers and the target. Even if one or more of the primers did bind to the target DNA, if the primer is not bound tightly enough, it can detach from the template as the temperature rises for strand extension.

Instructor's Advanced Preparation

Background

In this *GAPDH* PCR Module, students will perform two rounds of polymerase chain reaction (PCR) to amplify a portion of the *GAPC* gene from genomic DNA. The genomic DNA: 1) may be purified from plant material using the Bio-Rad Nucleic Extraction Module (catalog #166-5005EDU); 2) may be purchased from a science supply house; or 3) may be the *Arabidopsis* genomic DNA included in this kit. The initial PCR uses degenerate primers to amplify a pool of DNA fragments. Whereas primers are usually synthesized with one nucleotide at each position, degenerate primers have one or more positions that may have different nucleotides synthesized at that position. In this experiment, three positions in the initial primer nucleotide sequence have more than one nucleotide synthesized at that position. The initial forward primer is:

GA<u>B</u>TATGTTGTTGA<u>R</u>TCTTC<u>W</u>GG

The degeneracy is introduced at positions 3, 15, and 21.

Position	IUB code	Bases
3	B	G/T/C
15	R (purine)	G/A
21	W (weak)	A/T

The initial reverse primer is prepared in the same way.

The second round of nested PCR uses more specific primers to amplify a specific *GAPC* gene from the pool. The nested primers are not degenerate. They are located internal on the sequence from the initial primers. Here is the alignment of the two forward primers:

Initial forward primer	GABTATGTTGTTGARTCTTCWGG
Nested forward primer	CTACTGGTGTCTTCACTGACAA

The primers will not bind equally well to all plant species, so different plants will result in different reaction efficiencies in PCR. Students may well obtain a clonable fragment after the first round of PCR. In this case, clonable means that a PCR product of sufficient quantity, (generally visible on an agarose gel is sufficient), quality (means few or single bands), and of the expected size. However, the PCR could also be so inefficient that students cannot detect PCR product after the first round. In order to obtain target DNA of sufficient quantity and purity to clone, we highly recommend performing both rounds of PCR. However, even then it is possible that no band may be visible following nested PCR, either because the DNA extraction was unsuccessful or because the primers did not bind well to the target DNA. (Some plants known to amplify include leaves from cabbage, sage, parsley, petunia, spinach, grass, dandelion, and thyme.)

If students are using genomic DNA that they have prepared or that has been purchased, the genomic DNA from *Arabidopsis thaliana* included with the kit should be used as a control. It should be treated as an additional genomic DNA sample and amplified in both rounds of PCR. In addition, plasmid DNA encoding the targeted portion of the *GAPDH* gene from *Arabidopsis* is included with this kit and should be used as a template in each round of PCR to ensure the PCR is working. By performing these controls, students will not only understand the use of controls in experiments, but control reactions can also serve as a source of clonable PCR product should their own samples not amplify. A negative (no template) control is also recommended.

Tasks to perform prior to lab

1. Dilute required amount of 5x genomic *Arabidopsis* DNA solution (25 ng/µl) in sterile water to 5 ng/µl. For 12 student teams, combine 20 µl of DNA with 80 µl sterile water and mix well. Each student team requires 6 µl.

2. Optional: Prepare PCR master mix with primers. Depending on the level of your students, you may wish to prepare this mix ahead of time. The student protocol includes an option, with molarity calculations, for the students to prepare their own master mix with primers. The preparation method is the same for each different primer set. However, add primers to master mix a maximum of 30 min prior to use. Each student team requires 120 µl of 2x master mix with primers for 5 PCR reactions. For 12 student teams, 30 min prior to use, add 30 µl of specific primers (blue initial primers for the first round of PCR and yellow nested primers for the second round) to 1,500 µl of Bio-Rad 2x master mix and mix thoroughly. Store on ice.

3. For the second round of PCR, place the exonuclease I on ice.

4. Prepare 500 bp molecular weight ruler. Add 100 µl of 5x Orange G Loading Dye to the 500 bp molecular weight ruler provided.

5. If desired, instructors may prepare 1% agarose gels and running buffer for students according to standard protocols, alternatively students can prepare their own electrophoresis reagents (see Appendix B). Each student team will require 2 gels

6. Program thermal cycler.

    For the **first** round of PCR, program the thermal cycler with the initial *GAPDH* PCR program:

Initial denaturation	95°C for 5 minutes

    Then 40 cycles of:

Denaturation	95°C for 1 min
Annealing	52°C for 1 min
Extension	72°C for 2 min
Final extension	72°C for 6 min

Hold	15°C forever

    For the **second** round of PCR, program the thermal cycler with the nested *GAPDH* PCR program:

Initial denaturation	95°C for 5 min

    Then 40 cycles of:

Denaturation	95°C for 1 min
Annealing	46°C for 1 min
Extension	72°C for 2 min

 | Final extension | 72°C for 6 min |
    | --- | --- |
    | Hold | 15°C forever |

GAPDH PCR – Quick Guide

Stage 2A

1. Label five PCR tubes with your initials and the following labels:
   - 1 – Negative control (sterile water)
   - 2 – *Arabidopsis* gDNA
   - 3 – Positive control pGAP plasmid
   - 4 – Plant 1 gDNA (optional)
   - 5 – Plant 2 gDNA (optional)

2. Place each PCR tube into a tube adaptor, cap each tube. Place the adaptor tube with PCR tube on ice.
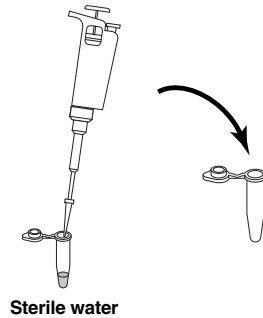
   PCR tube   **Capless tube**   **Ice**
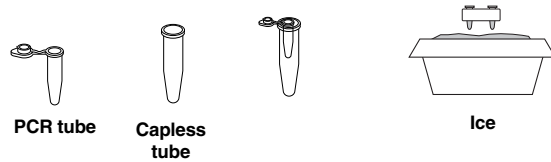
3. Flick to thoroughly mix reagent tubes and centrifuge 10 sec to force contents to bottom of tubes.
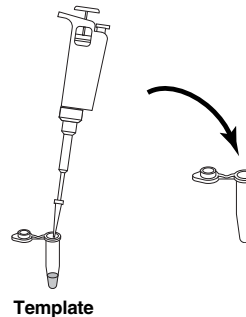
   Flick to mix
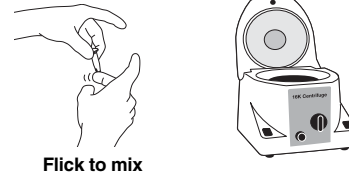
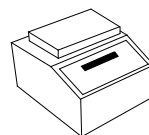4. Pipet 20 µl of 2x blue mastermix with Initial primers (2x MMIP) into each PCR tube.

   MasterMix Blue
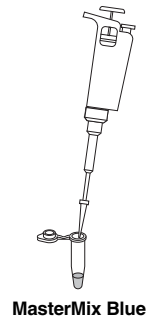
5. Using a fresh tip, pipet 15 µl of sterile water to each tube.

   Sterile water

6. Using a fresh tip every time, add 5 µl of the appropriate DNA template to each tube. Gently pipet up and down to mix reagents. Recap tubes.
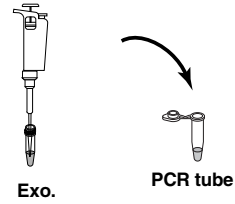
   Template
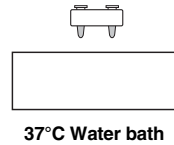
7. When instructed, place PCR tubes in thermal cycler.

Stage 2B

1. Using a fresh tip each time pipet 1 µl of exonuclease I into each blue initial PCR reaction that amplified genomic DNA (i.e. tube # 2, 4 and 5). Mix well by pipetting up and down.

Exo.   **PCR tube**

2. Incubate exonuclease reactions at 37°C for 15 min.
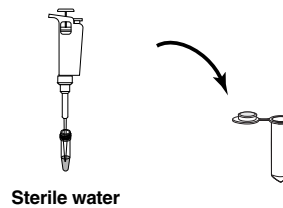
37°C Water bath

3. Heat inactivate exonuclease by incubating at 80°C for 15 min.

80°C Water bath
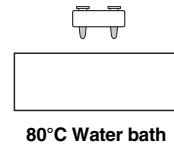
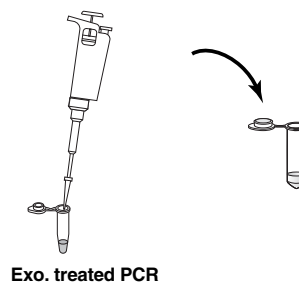4. Label three microcentrifuge tubes for the exonuclease treated PCR reactions.
    -Exo *Arabidopsis* gDNA
    -Exo plant 1 gDNA (optional)
    -Exo plant 2 gDNA (optional)
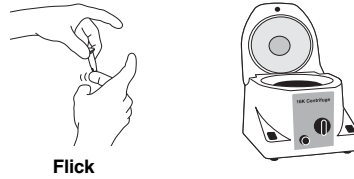
Sterile water

5. Add 98 µl of sterile water to each microcentrifuge tube.

6. Using a fresh tip each time, add 2 µl of each exonuclease treated PCR reaction into the appropriate microcentrifuge tube.

Exo. treated PCR

7. Vortex or flick each tube to mix. Spin briefly in a microcentrifuge to force the contents to the bottom of the tube.

Flick

8. Label five PCR tubes with your initials and the following labels.
   - 6 – Negative control (sterile water)
   - 7 – Exo *Arabidopsis* gDNA
   - 8 – Positive control pGAP plasmid
   - 9 – Exo plant 1 (optional)
   - 10 – Exo plant 2 (optional)

9. Pipet 20 µl of 2x yellow MasterMix with nested primers (2x MMNP) into each PCR tube.
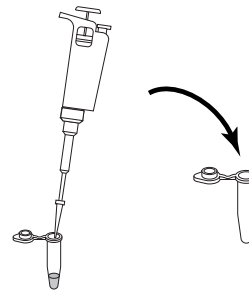
master mix yellow

10. Using a fresh tip each time, pipet 20 µl of template (according to your tube labels) into the appropriate PCR tube. Gently pipet up and down to mix reagents. Recap tubes.
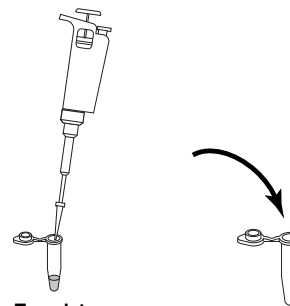
Template

11. When instructed, place PCR tubes in thermal cycler.

Student Protocol

Introduction

In this experiment, a portion of the glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) gene will be cloned. Because it is a vital metabolic enzyme involved in glycolysis, one of the most basic of biological processes, the *GAPDH* protein is highly conserved between organisms, especially in vital domains of the enzyme, such as the active site. However, this does not mean that the gene and DNA sequence are identical between organisms. A large 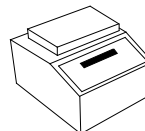percentage of a gene's DNA does not code for protein — this "intronic" DNA is not subject to the same selective pressures as DNA that does code for protein. In addition, in the gene sections that do encode protein (the exons), there is degeneracy of the genetic code, which means that multiple DNA triplet codons encode the same amino acid. (For example, GCA, GCC, GCG, and GCT all code for alanine.) Also, some regions of the protein less vital to enzyme function do not have the same degree of selective pressure as the regions that are essential to enzyme function, such as the enzyme active sites or cofactor binding sites.

To clone a known gene from a less well-characterized organism, primers must be designed that are complimentary to conserved regions of the *GAPDH* gene. However, even conserved regions are not identical between organisms, so a best guess of the gene sequence is made using a comparison alignment of the sequences of the gene from different, related organisms, with the understanding that the primers will not be an exact match to the sequence and may amplify non-specific sections of DNA in addition to the target sequence. A second set of primers is then designed, interior to the first set of primers, and used to amplify the PCR products from the first round of PCR. This is called nested PCR. Nested PCR works because there is an extremely slim chance that any DNA that was amplified non-specifically in the first round of PCR will also contain these interior sequences, while the specifically amplified DNA (the target sequence) should contain these sequences, even if the match is not perfect.

The Experiment

The strategy for this experiment uses nested PCR to amplify portions of the *GAPC* gene from genomic DNA (gDNA) of the plant of interest. In the initial round of PCR, a set of primers using degenerate (less specific) sequences will amplify the *GAPC* gene from genomic DNA. Then, in the second round of PCR (the nested PCR), a more specific set of primers will amplify *GAPC* from the initial PCR products. The primers are color-coded; the initial PCR primers are blue and the nested PCR primers are yellow. It is very important not to reverse the order in which the primers are used or to mix the two primer sets together in the PCR reactions.

Before performing the nested PCR, the primers that were not incorporated into PCR products in the first round must be removed so that they do not amplify target DNA is the second round of PCR. To remove the primers, exonuclease I, an enzyme that specifically digests single-stranded DNA, will be added to the PCR products from the first round. After exonuclease I digests the primers, the enzyme must be inactivated to prevent the exonuclease from digesting the nested primers that will be added for the next round of PCR.

Following exonuclease treatment and inactivation of the enzyme, the PCR products from the gDNA template(s) generated in the first round of PCR need to be diluted. The products from the first round of PCR contain both the original template gDNA and the PCR products, *GAPC*-like sequences. The *GAPC*-like sequences predominate, so they are most likely to be amplified during nested PCR, but diluting the gDNA makes it even less likely that is will be a template as well.

As each PCR reaction takes approximately 3–4 hours to run, it is most practical to run the PCR reactions on separate days. Since the reagents used in these experiments function optimally when prepared fresh, it is highly recommended that the reagents be prepared just prior to setting up the reactions and running the PCR.

Note: PCR is extremely sensitive to contamination by DNA from many sources. All reagents to be used for PCR should be handled with care to minimize contamination. It is recommended that either PCR should be set up in an area of the laboratory or classroom that is separate from the DNA extraction area or the lab benches used should be thoroughly swabbed down with a commercial cleaner or 10% bleach (ethanol does not destroy DNA). In addition, aerosol barrier filter pipet tips should always be used to set up PCR. The micropipets should be carefully cleaned with a 10% solution of bleach before performing PCR. In research labs, PCR flow hoods are frequently used to prevent contamination.

Pre-lab Focus Questions

1. Why was the discovery of *Taq* DNA polymerase important for the development of PCR?

2. The protein glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) is necessary for cellular function and is highly conserved between organisms. Why is it probable that proteins needed for cell to survival will be very similar (= highly conserved) in many different organisms?

3. What are degenerate primers?

4. What is a consensus sequence?

What would be the consensus sequence for the following aligned sequences?

```
A T T G C T T C
A A T G C T A C
A T T C C T A C
A T T G C T A C
```

Student Workstations

Each student team will require the following items to set up 5 of the initial PCR reactions:

Material Needed for Each Workstation	Quantity
Bio-Rad 2x master mix	120 µl
Blue initial primers	4 µl
(Optional) gDNA from 2 plants — previously extracted	2
1x *Arabidopsis* gDNA (diluted to 5 ng/µl)	6 µl
Control pGAP plasmid DNA	6 µl
Sterile water	100 µl
20 µl adjustable-volume micropipet and filter tips	1
PCR tubes	5
PCR tubes adaptors	5
Microcentrifuge tube	1
Marking pen	1
Ice bath	1
Tube rack	1

Each student team will require the following items to set 5 of the nested PCR reactions:

Material Needed for Each Workstation	Quantity
Bio-Rad 2x master mix	120 µl
Yellow nested primers	4 µl
PCR reactions from initial PCR	1–3
Control pGAP plasmid DNA	25 µl
Sterile water	350 µl
20 µl adjustable-volume micropipet and filter tips	1
200 µl adjustable-volume micropipet and filter tips	1
PCR tubes	5
PCR tube adaptors	5
Microcentrifuge tubes	4
Marking pen	1
Ice bath	1
Tube rack	1

Each student team will require the following items to electrophorese their PCR reactions:

Material Needed for Each Workstation	Quantity
500 bp molecular weight ruler	12 ul
5x Orange G Loading Dye	50 ul
1% agarose gels	1
Agarose gel electrophoresis equipment	1
Power supply	1
20 µl adjustable-volume micropipets and filter tips	1
Microcentrifuge tubes for preparing samples	10
Gel visualization and documentation system	
(optional) Agarose gel casting system and reagents	1

Common Workstation

Material Required	Quantity
Thermal cycler	1
Water bath, incubator, or heating block (optional — thermal cycler may be used)	1
Microcentrifuge (optional)	1

Student Protocol

Preparation for Initial PCR (First Round of PCR)

1. Plan your initial PCR experiment. You will perform one initial PCR reaction on *Arabidopsis* gDNA, one on pGAP plasmid DNA (positive control), and one or more genomic DNA samples (optional), plus one negative control with sterile water instead of DNA. Generate a table to record the label for each PCR tube, the DNA template, and the primers used to amplify the DNA (see below).

Tube Label	Template	Primers

2. Prepare a master mix. (**Note**: Master mix should be prepared no more than 30 min prior to performing PCR. Your instructor may have prepared it just prior to the lab.) A master mix is a mixture of all the reagents required for PCR except the template DNA. Making a single mixture reduces potential pipetting errors, so your PCR will be more consistent.

   Bio-Rad's 2x master mix is provided as a 2x (double strength) colorless reagent; when mixed with an equal amount of DNA template, all components are at optimal concentrations in the final reaction. Bio-Rad 2x master mix contains *Taq* polymerase, dNTPs, buffer, and salt, but it does not contain primers. It is necessary to add primers to Bio-Rad 2x master mix to form a complete 2x mastermix with initial primers- "2xMMIP".

   Why are primers not added to commercial master mixes?

   _____

   _____

   Each PCR volume is 40 µl. Thus, each reaction requires 20 µl of 2xMMIP (Bio-Rad 2x master mix with blue initial primers) plus 20 µl of DNA template. The amount of 2xMMIP to prepare is calculated by multiplying the number of reactions plus one (to allow for pipetting errors) by the volume of 2xMMIP required for each reaction (20 µl).

   Referring to your table, calculate how much 2xMMIP is required.

   Volume of 2xMMIP required = (# PCR reactions +1) x 20 µl = _____ µl

   For the initial round of PCR, blue initial primers will be used. These primers are targeted for a section of DNA outside of the target sequence. The initial primers are supplied at a 100 µM concentration. For this PCR, the concentration of primers in the 2xMMIP should be 2 µM.

Calculate the volume of initial primers required in the 2xMMIP (typically this equates to 2 µl of primers for every 100 µl of the 2xMMIP). Remember the formula $M_1V_1 = M_2V_2$. Rearranged, the formula is:

$$M_1 \times V_1 / M_2 = V_2$$

Substituting in the equation, you get:

Required volume of initial primers in µl =

$$\frac{\text{(required concentration of primers in µM) x (required volume of 2xMMIP in µl)}}{\text{(given concentration of primers in µM)}}$$

Required volume of initial primers: _____ µl

Label a microcentrifuge tube 2xMMIP. Not more than 30 min before use, add the calculated volume of initial primers to the required volume of Bio-Rad 2x master mix in a labeled screwcap tube. Mix well by pipetting up and down several times or vortexing. If a microcentrifuge is available, spin tube briefly to collect the contents at the bottom of the tube. Keep on ice.

Experimental Procedure for Initial PCR

1. Referring to your label, label your PCR tubes with your initials and the tube label.

2. Place each PCR tube into a tube adaptor and cap each tube. Place the adaptor tubes containing PCR tubes on ice.

3. Ensure all the reagents are thoroughly mixed, especially the gDNA. Mix tubes containing reagents thoroughly by vortexing or flicking to ensure the DNA is homogeneously distributed. Before opening the tubes, spin in a microcentrifuge for 5–10 sec to force contents to the bottom of the tube (to prevent contamination).

   Each PCR needs to be set up with the following reagents

Blue 2xMMIP	20 µl
Sterile water	15 µl
DNA template or negative control	5 µl
Total	40 µl

4. Pipet 20 µl of 2xMMIP into each PCR tube.

5. Add 15 µl of sterile water to each tube.

6. Referring to your table, use a fresh pipet tip to add 5 µl of the appropriate DNA template to each tube and gently pipet up and down to mix reagents. Use a fresh filter tip each time. Recap tubes.

7. When your instructor tells you to do so, place your PCR tubes into the thermal cycler.

   The PCR reaction will run for the next several hours using the following initial *GAPDH* PCR program:

Initial Denature:	95°C for 5 min
Then 40 cycles of:	
Denaturation:	95°C for 1 min
Annealing:	52°C for 1 min
Extension:	72°C for 2 min
Final Extension:	72°C for 6 min
Hold:	Forever at 15°C

   Note: Store any remaining genomic DNA at –20°C.

8. If you have not already done so, prepare agarose gels to analyze the results of your experiment. You will need 1% agarose gels with sufficient wells for each of your samples plus an additional well for your molecular size marker. See Appendix B for detailed instructions on preparing agarose gels.

9. Retrieve your PCR samples from the thermal cycler.

10. Label microcentrifuge tubes for each of your PCR reactions and pipet 5 µl of 5x Orange G loading dye into each tube. Do not add loading dye directly to your PCR tubes — this can inhibit the second round of PCR.

11. Pipet 20 µl of your completed PCR reactions into their respective tube and pipet up and down to mix.

12. Load 10 µl of the 500 bp molecular weight ruler and 20 µl of each initial PCR reaction into your gel, record your gel loading order and electrophorese your samples.

    Note: the bands in the 500 bp molecular weight ruler are 500, 1,000, 1,500, 2,000, 2,500, 3,000 3,500, 4,000, 4,500 and 5,000 bp.

13. Visualize your PCR products on your agarose gel and record results.

    Note: The pGAP plasmid control should yield a visible band of ~ 1 kb. It is relatively common for plant genomic DNA to yield multiple bands, or even very faint or no bands in the first round of PCR, and yet still be amplified in the second round of PCR.

Initial PCR Focus Questions

1. Did the negative control generate a PCR product? If yes, what size is the DNA band and what does this mean for the experiment? If no, what does this mean for the experiment?

2. Did the pGAP control plasmid generate a PCR product? If yes, what size is the DNA band and what does this mean for the experiment? If no, what does this mean for the experiment?

3. Did the *Arabidopsis* genomic DNA generate a PCR product? If yes, what size is/are the DNA bands and what does this mean for the experiment? If no, what does this mean for the experiment?

4. If used, did your novel genomic DNAs generate PCR products? If yes, what size is/are the DNA bands and what does this mean for the experiment? If no, what does this mean for the experiment?

Preparation for Nested PCR (Second Round of PCR)

1. Plan your nested PCR experiment. One nested PCR will be performed for each gDNA sample amplified in the initial round of PCR. In addition, you will perform one positive control using the control pGAP plasmid DNA and one negative control with sterile water instead of DNA. Generate a table to record the label for each nested PCR tube, the DNA template (be sure to state whether this is a PCR product from the initial PCR and its dilution factor), and the primers used to amplify the DNA (see below).

Tube Label	Template	Dilution Factor	Primers

2. Prepare a master mix. (**Note**: Your instructor may have done this just prior to the lab.) Each nested PCR volume is 40 µl. Thus, each nested PCR requires 20 µl of 2xMMNP composed of Bio-Rad 2x mastermix and yellow nested primers). Referring to your table, and calculations from the initial round of PCR if necessary, calculate how much 2xMMNP is required.

   Volume of 2xMMNP required = (# nested PCR reactions +1) x 20 µl = _____ µl

   For the nested PCR, nested primers will be used that specifically target DNA interior to the locations where the initial PCR primers bound. The yellow nested primers are supplied at a 25 µM concentration. For this PCR, the concentration of primers in the 2xMMNP should be 0.5 µM. Calculate the volume of nested primers required in the 2xMMNP (typically this equates to 2 µl of primers for every 100 µl of the 2xMMNP). Remember the formula $M_1V_1 = M_2V_2$. Rearranged, the formula is:

$$M_1 \times V_1 / M_2 = V_2$$

   Substituting in the equation, you get:

   Required volume of yellow nested primers in µl =

$$\frac{\text{(required concentration of primers in µM) x (required volume of 2xMMNP in µl)}}{\text{(given concentration of primers in µM)}}$$

   Required volume of nested primers: _____ µl

Label a microcentrifuge tube 2xMMIP. Add the calculated volume of yellow nested primers to the required volume of Bio-Rad 2x master mix in a labeled screwcap tube. Mix well by pipetting up and down several times or vortexing. If a microcentrifuge is available, spin tube briefly to collect the contents at the bottom of the tube. Keep on ice.

Experimental Procedure for Nested PCR

1. Prepare template DNA. Remove unincorporated primers from initial PCR tubes using exonuclease I. Using a fresh tip each time, pipet 1 µl of exonuclease I into each initial PCR tube (gDNA samples only, not the pGAP or negative control samples). Mix well by pipetting up and down,

2. Incubate at 37°C for 15 min.

3. Heat-inactivate the enzyme by incubating at 80°C for 15 min.

   Why is it necessary to inactivate the exonuclease?

   _____

4. Label a microcentrifuge tube for each exonuclease-treated initial PCR tube.

5. The initial PCR will be diluted 100-fold for the nested PCR, Since the final reaction volume is 40 µl, rather than pipetting 0.4 µl of the initial PCR (40 µl / 100 = 0.4 µl), the DNA will be diluted to allow pipetting a larger volume. To dilute each initial PCR to 1/50 the original concentration, first pipet 98 µl of sterile water into each of the labeled microcentrifuge tubes.

6. Using a fresh tip each time, pipet 2 µl of the appropriate initial PCR into the appropriate microcentrifuge tube. Close the cap.

7. Vortex or flick the tube with your finger to mix. Spin briefly in a microcentrifuge to collect the liquid at the bottom of the tube.

8. Label PCR tubes according to your plan and place in PCR tube adaptors on ice.

   Each PCR needs to be set up with the following reagents

Yellow 2xMMIP	20 µl
DNA template or negative control	20 µl
Total	40 µl

9. Pipet 20 µl of 2xMMIP into each PCR tube.

10. Referring to your table, use a fresh pipet tip to add 20 µl of the appropriate DNA template (diluted initial PCR, pGAP plasmid DNA, or sterile water for the negative control) to each PCR tube. Gently pipet up and down to mix reagents. Recap tubes.

11. When your instructor tells you to do so, place your PCR tubes into the thermal cycler.

    The PCR reaction will run for the next several hours using the following Nested *GAPDH* PCR program:

Initial Denature:	95°C for 5 min
Then 40 cycles of:	
Denaturation:	95°C for 1 min
Annealing:	46°C for 1 min
Extension:	72°C for 2 min
Final Extension:	72°C for 6 min
Hold:	Forever at 15°C

13. Retrieve your PCR samples from the thermal cycler.

14. Add 10 µl of 5x loading dye to each of your PCR tubes.

    Note: if using the PCR product for a further downstream application- such as ligation or restriction digestion analysis- do not add loading dye directly to your PCR reaction, as the loading dye may inhibit downstream applications- combine PCR reaction for analysis and loading dye in a separate tube.

15. Load 10 µl of the 500 bp molecular weight ruler and 20 µl of each nested PCR reaction into your gel, record your gel loading order and electrophorese your samples.

    Note: if loading both initial and nested PCR reactions on the same gel- load only 5 µl of the nested PCR reactions since the nested PCR is usually much more efficient than the initial PCR and the intense bands can obscure faint initial PCR reactions.

16. Visualize your PCR products on your agarose gel and record results.

The fragment of *GAPC* that has been targeted varies in size between plant species. The expected size of the fragment from the initial round of PCR is 0.5 to 2.5 kb. The expected size of the fragment from the second round of nested PCR is slightly smaller than the PCR product from the initial round. It is possible doublets (2 DNA fragments of similar sizes) may be amplified from the genomic DNA (gDNA) of some plants, most likely due to amplification of two *GAPC* genes that are highly homologous.

Nested PCR Focus Questions

1. What does the enzyme exonuclease I do? Why did you treat your initial PCR product with exonuclease I before using the DNA as a template for Nested PCR? What would have been the result if you had not done an exonuclease I treatment?

2. What would have happened if the exonuclease I had not been heat-inactivated after use?

3. In the initial PCR, you used degenerate primers. In nested PCR, you used non-degenerate primers. In which round of PCR would you expect to see more PCR products of different sizes?

4. Did the pGAP control plasmid generate a PCR product? If yes, what size is the DNA band and what does this mean for the experiment? If no, what does this mean for the experiment?

5. Did the *Arabidopsis* genomic DNA generate a PCR product? If yes, what size is/are the DNA bands and what does this mean for the experiment? If no, what does this mean for the experiment?

6. If used, did your novel genomic DNAs generate PCR products? If yes, what size is/are the DNA bands and what does this mean for the experiment? If no, what does this mean for the experiment?

Appendix A
Genomic DNA from Plants

The genomic DNA (gDNA) needed for this lab can be obtained from several sources.

1. Bio-Rad Nucleic Acid Extraction Module (catalog #160-5005EDU) contains instructions and materials with which your students can extract DNA from their choice of plants. The kit contains a list of plants from which *GAPDH* is known to amplify using this kit.

2. This kit contains controls (gDNA from *Arabidopsis thaliana* and an *A. thaliana GAPDH* gene fragment cloned into the plasmid vector) to assess the validity of the results. Alternatively, the control *Arabidopsis* DNA can be used as the primary gDNA for the PCR.

> *Arabidopsis thaliana* is a member of the mustard family, which includes species such as cabbage and radish. It is a small flowering plant that is widely used as a model organism in plant biology.

3. Plant gDNA can be purchased from a variety of companies.

4. Researchers in plant bioscience at nearby colleges, universities, or research centers are another possible source of gDNA.

Purifying Genomic DNA From Plants

Isolation of DNA from plants has complications not seen in other organisms, such as animals and bacteria. Plants are multicellular eukaryotic organisms with rigid cell walls composed of multiple layers of cellulose. The cell wall is needed to maintain shape and give rigidity to plant cells, but it complicates DNA extraction. Plant polysaccharides, such as the cellulose from the cell wall, colocalize with DNA during the purification process. In addition, individual plant types have their own problems. For example, many conifers and fruit trees contain a high concentration of compounds called polyphenols.

RER
Nucleolus
Nucleus
Golgi apparatus
Vacuole
Peroxisome
Chloroplast
Mitochondria
Cell wall
Cell membrane

When plant cells containing polyphenols are lysed, the polyphenols bind to the DNA and make it useless for further experiments. In addition, extracting DNA from plant cells has the same issues as extracting DNA from any other cell types. There are enzymes in the cells that can digest the DNA during extraction and membrane-bound compartments within the cells have acidic contents that can damage DNA when the contents are spilled into the cytosol during lysis.
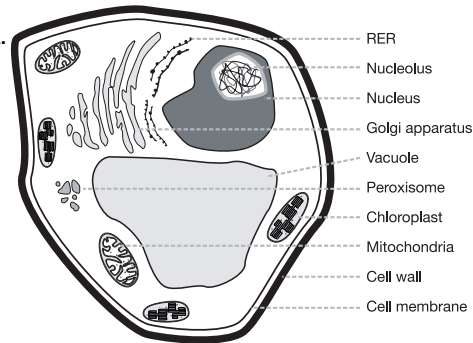
The steps in isolating genomic DNA from plants are:

1. **Harvesting plant tissue**. There are several options that will increase DNA yield. Young leaves are the best source for genomic DNA. Young leaves are still growing and have a greater ratio of nuclear volume to cytoplasmic volume, and younger leaves have fewer chemicals that interfere with DNA isolation. Also, fresh samples will have higher yields of DNA than will preserved or dried leaves.

2. **Grinding and lysing plant tissue**. Because of the rigid cell walls, plant tissue must be physically ground or crushed prior to DNA isolation. The most common way to grind plant tissue is with a mortar and pestle, though blenders and mechanical tissue grinders are also used with some plant species. Grinding of plant tissue is frequently done in the presence of lysis buffer. Lysis buffer has several important components:

- EDTA (ethylenediaminetetraacetate). EDTA serves a dual function, removing (chelating) magnesium ions (destabilizing the cell wall and the cell membrane) and inhibiting nucleases, enzymes that could digest the DNA

- Buffering capacity. When cells are lysed, acidic compounds are released from subcellular organelles and the lysis buffer must be such that the overall pH of the lysate does not change dramatically. Most lysis buffers are prepared at pH 8.0, made with Tris [tris (hydroxymethyl) aminomethane], a commonly used buffer in molecular biology

- Detergent. Detergent breaks up cell membranes by removing lipid molecules from the membranes. The choice of detergent depends on the application. Ionic detergents, such as SDS (sodium dodecyl sulfate) or Sarkosyl (N-laurylsarcosine), are commonly used, but non-ionic detergents such as Triton X may also be used (Non-ionic detergents are milder than ionic detergents and usually leave proteins intact and functional.)

The "classical" method of DNA extraction from plants uses a nonionic detergent, CTAB (cetyltrimethyl ammonium bromide), to lyse the plant cells and precipitate the DNA, leaving most polysaccharides in solution.

Because of special problems found in some plants, specialized techniques have been developed for DNA extraction. For example, for plants with high levels of polyphenols, including grape species, many fruit trees, and conifers, lysis buffer containing PVP (polyvinylpyrrolidone) and a high concentration of salt works best for DNA isolation. PVP binds the polyphenols, preventing them from complexing with the DNA, and the high salt reduces the coprecipitation of polysaccharides with the DNA.

3. **Purifying DNA**. After cell lysis and centrifugation to remove debris, there will be proteins and RNA in the supernatant with the DNA. There are several methods used to remove these contaminants. It is important to remember that during the purification of DNA the lysate should never be vortexed or mixed vigorously, as that would result in breaking the large molecules of genomic DNA into smaller pieces (known as shearing of DNA). There are two reasons that intact, full-length DNA molecules are important. First, the procedures are designed to purify large molecules and small DNA molecules may be lost in the process.

Second, having intact DNA may be important for subsequent steps in the experiment, for example, to use as a template for PCR.

The most common methods of separating genomic DNA from other cellular components include:

- DNA binds strongly to silica in the presence of high concentrations of salts that disrupt hydrophobic interactions (chaotropic salts). Although not clearly understood, the binding is thought to be due to the exposure of phosphate residues on the DNA as a result of dehydration by the salts. The exposed phosphates adhere (adsorb) to the silica. The DNA can be released from the silica by reducing the salt concentration. This chemistry is the basis for many of the commercially available kits for DNA purification.

- Purification of DNA with ion exchange chromatography uses a positively charged resin or other matrix, usually in a glass or plastic chromatography column. DNA and RNA are both negatively charged and will bind to the positively charged matrix through ionic interactions. Other components of the cell lysate, those that are not negatively charged, will not bind to the matrix and will be discarded. Molecules bound to the matrix can be removed (eluted) by increasing the salt concentration of the solution on the column. Molecules elute based on the strength of their binding to the column, with more weakly bound molecules eluting at lower salt concentration than more strongly bound molecules. Using this differential elution, DNA can be separated from proteins and RNA.

- DNA can be purified in a two-step process of enzyme treatment and organic extraction. Proteins in the cell lysate are degraded by treatment with proteases, enzymes that break the proteins into small pieces. Besides getting rid of protein contaminants, this step also degrades enzymes called nucleases that might destroy the target DNA. Proteinase K is one of the most commonly used proteases, degrading most proteins and inactivating enzymes under a broad range of conditions. After the proteins have been degraded, organic extraction is used to precipitate proteins in the lysate. Phenol or phenol mixed with chloroform will cause proteins to coagulate at the interface of the organic and the aqueous solutions. The proteins are removed by centrifugation, leaving RNA and DNA in the aqueous solution. The RNA is removed by treating the solution with enzymes called ribonucleases (RNases) that degrade RNA but do not damage DNA. Ribonuclease A, a nuclease that cleaves only single-stranded RNA molecules, is commonly used

4. **Concentrating DNA**. Many DNA preparations result in a solution that is too dilute for experimental purposes, so the DNA must be concentrated. The most common methods use alcohol. In the presence of high salt concentrations (for example, NaCl), ethanol or isopropanol will precipitate DNA. The cations neutralize the charge on the phosphate backbone of the DNA and allow the DNA molecules to get close together. (Normally, in aqueous solution, the strong negative charges of the DNA molecules repulse each other.) The ions do not bind strongly to DNA in aqueous solution, but when an organic solvent like alcohol is added, the binding becomes much stronger and the DNA-cation complexes precipitate out of solution. If there is enough DNA present, a white precipitate should be visible, but DNA may very well be present even if a precipitate is not observed. The precipitated DNA is pelleted by centrifugation, after which the pellet should be washed at least once with 70–80% ethanol to remove any remaining salt. After the washes, the DNA pellet should be dried to evaporate any remaining ethanol and resuspended in water or the desired buffer.

Appendix B
Prepare Running Buffer and Agarose Gels*

The recommended agarose concentration for gels in this classroom application is 1%. This concentration provides good resolution and minimizes run time required for electrophoretic separation of DNA fragments. The recommended thickness for the gel is 0.75–1.0 cm for easy sample loading and gel handling.

Visualization of DNA

The most widely used method to visualize DNA in agarose gels is detection of ethidium bromide staining of the DNA using a UV transilluminator and documentation system. The following protocol uses ethidium bromide (available in Bio-Rad's Ethidium Bromide DNA electrophoresis Reagent Pack catalog #166-0451EDU). An alternative means of visualizing DNA is Bio-Rad's Fast Blast™ DNA staining solution (catalog #166-0420), which is a biologically safe DNA stain that does not require any documentation system. Gels are stained with Fast Blast after electrophoresis either with a quick 15 minute protocol or overnight. Fast Blast is around 5x less sensitive than ethidium bromide, which may mean that some faint DNA bands that might be visible with ethidium bromide may not be visible with Fast Blast staining. Another alternative is SYBR® Green I, a DNA stain that also requires a visualization and documentation system.

Be sure to use electrophoresis buffer, not water, to prepare agarose gels.

1. **Electrophoresis buffer for preparing gels**. TAE (Tris-acetate-EDTA) electrophoresis buffer can be purchased as a 50x concentrated solution (catalog #161-0743EDU) that must be diluted to 1x TAE for preparing agarose gels. To prepare 8 agarose gels, 500 ml of 1x TAE will be adequate. To make 500 ml of 1x TAE from 50x TAE concentrate, add 10 ml of 50x concentrate to 490 ml of distilled water. (In addition, diluted TAE is needed to fill the electrophoresis chamber. See below.)

2. **Agarose gel preparation***. These procedures may be carried out 1 to 2 days ahead of time by the teacher or done during class by the individual student teams.

    a. To make a 1% agarose solution, use 1 g of agarose for each 100 ml of 1x TAE electrophoresis buffer. Be sure to use electrophoresis buffer, not water. If electrophoresis chambers are limiting, you can use a 7 x 10 cm tray and two 8-well combs to pour a gel that can be used to run two sets of student samples. 12- well combs are also available and allow for more samples to be run per gel.

    b. Use this table as a guide for gel volume requirements when casting single or multiple gels.

    Volume of 1% agarose for:

Number of gels	7 x 7 cm tray	7 x 10 cm tray
1	40 ml	50 ml
2	80 ml	100 ml
4	160 ml	200 ml
8	320 ml	400 ml

c.   Add the agarose powder to a suitable container (e.g., 500-ml Erlenmeyer flask for 200 ml or less). Add the appropriate amount of 1x TAE electrophoresis buffer and swirl to suspend the agarose powder in the buffer. If using an Erlenmeyer flask, invert a smaller Erlenmeyer flask into the open end of the 500-ml Erlenmeyer flask containing the agarose. The small flask acts as a reflux chamber, thus allowing long or vigorous boiling without much evaporation. The agarose solution can be prepared for gel casting by boiling until agarose has melted completely on a hot plate, in a hot water bath, or in a microwave oven.

Caution: Always wear protective gloves, goggles, and lab coat while preparing and casting agarose gels. Boiling agarose solution or the flasks containing hot agarose can cause severe burns if allowed to contact skin.

Microwave oven method. This technique is the fastest and safest way to dissolve agarose. Place the gel solution in an appropriate bottle or flask into the microwave. LOOSEN THE CAP IF YOU ARE USING A BOTTLE. Use a medium setting and set to 3 min. Stop the microwave oven every 30 sec and swirl the flask to suspend any undissolved agarose. Boil and swirl the solution until all of the small transparent agarose particles are dissolved.

d.   Cool agarose solution to 55–60°C (a water bath is useful for this step). Add 5 µl of ethidium bromide (10 mg/ml) to each 100 ml of molten agarose to final concentration of 0.5 µg/ml and swirl to mix. Solutions or gels containing ethidium bromide should never be heated or microwaved.

e.   Prepare the gel casting apparatus and pour the molten agarose into the gel casting tray containing the comb/s and allow to solidify at room temperature for 15–20 min.

f.   Carefully remove the combs from the solidified gel. Agarose gels can be stored wrapped in plastic wrap, sealed in plastic bags for up to 2 weeks at 4°C.

* Convenient precast agarose gels are available from Bio-Rad. These are 1% TAE gels with ethidium bromide and fit into Bio-Rad's Mini-Sub Cell GT cell or any horizontal gel electrophoresis system that fits 7 x 10 cm gels. The gel have 8 wells (catalog #161-3016EDU) or 12 wells (catalog #161-3022EDU).

3.   **Electrophoresis buffer for filling chambers**. Concentrated TAE buffer must also be diluted for use as running buffer. There are two way in which to do this:

a.   Conventional electrophoresis uses running buffer at 1x concentration and the gels are run at 100V for 30 min. Each Bio-Rad Mini-Sub Cell electrophoresis chamber requires ~275 ml of buffer. To prepare enough 1x TAE for 12 chambers, add 70 ml of 50x TAE to 3.43 L of distilled water and mix. If adding ethidium bromide to running buffer add 175 µl of ethidium bromide (10 mg/ml) and mix- see note below on adding ethidium bromide to running buffer.

b.   The Fast Gel protocol uses a reduced concentration of running buffer (0.25x TAE), so gels can be run at 200 V for under 20 min. To prepare enough 0.25x TAE for 12 chambers, add 17.5 ml of 50x TAE to 3.48 L of distilled water and mix. If adding ethidium bromide to running buffer add 175 µl of ethidium bromide (10 mg/ml) and mix- see note below on adding ethidium bromide to running buffer. Monitor gel loading dye migration during electrophoresis to prevent running the gel for too long. **Note**: Do not use 0.25x TAE to prepare agarose gels; doing so will result in loss of DNA resolution.

Note on addition of ethidium bromide to running buffer: Many researchers add ethidium bromide to both gel and running buffer at a concentration of 0.5 µg/ml. We have found that for most applications where gels are electrophoresed for less than 30 min addition of ethidium bromide to just the gel and not the running buffer produces good results, this also reduces the disposal requirements for the lab. However, because ethidium bromide migrates to the negative electrode, in the absence of ethidium bromide in the running buffer, ethidium stain will gradually be lost from the bottom of the gel which, if gels are run for longer than 30 min may result in DNA bands being missed. If this occurs, or if a post-staining method is preferred, gels can be post stained with a 0.5 µg/ml solution of ethidium bromide for 20 min and then rinsed and washed two times for 20 min with distilled water.

Safety Note: Ethidium bromide may cause heritable genetic damage and should be treated with special care as a toxic substance. Wear suitable protective clothing, gloves and eye/face protection and follow good laboratory practices during use. Please refer to MSDS for further information.

Appendix C
Instructor's Answer Guide for Focus Questions

Pre-lab Questions

1. **Why was the discovery of *Taq* DNA polymerase important for the development of PCR?**

   Taq DNA polymerase, initially isolated from the thermophilic bacteria *Thermus aquaticus* (found in a hot pool in Yellowstone National Park in Wyoming), is active at 72°C. This is a much higher temperature than other DNA polymerases, so it is not destroyed by the high temperatures in PCR. Prior to the discovery of *Taq*, fresh DNA polymerase had to be added after each denaturation cycle of PCR.

2. **The protein glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) is necessary for cellular function and is highly conserved between organisms. Why is it probable that proteins needed for cell to survival will be very similar (= highly conserved) in many different organisms?**

   When mutations occur in the DNA coding for a protein that is needed for cell viability, it is very unusual for the mutation to improve the functioning of the enzyme. In most cases, the mutation will have a negative effect, and the organism with the mutation will have lower survival and reproductive rates. This selective pressure means that mutations in exonic (coding) DNA are not likely to be passed on to subsequent generations. Important proteins are commonly very similar across a wide variety of organisms, particularly in regions of the proteins important for function, such as active sites and cofactor binding sites.

3. **What are degenerate primers?**

   Degenerate primers are a mixture of PCR primers that are similar, but not identical. They may be designed based on a consensus sequence derived from similar organisms, with substitutions of different bases at one or more locations in the primers. Degenerate primers are often used when the exact sequence of the target DNA is not known.

4. **What is a consensus sequence?**

   A consensus sequence is one derived from the alignment of DNA or RNA sequences from similar organisms. Normally, each position in the consensus sequence is determined by the base that predominates in that position in the majority of the aligned sequences. Consensus sequences can be used to design primers for PCR.

   What would be the consensus sequence for the following aligned sequences?

	A T T G C T T C
	A A T G C T A C
	A T T C C T A C
	A T T G C T A C
Consensus	**A T T G C T A C**

The consensus sequence is simple at the positions in which all sequences have the same base. For example, the first position would be A. But at positions in which there are choices, the base found most frequently at that position will determine the base in the consensus sequence. For example here, the second position is T because three of the sequences have T at that position.

Initial PCR Focus Questions

1. **Did the negative control generate a PCR product? If yes, what size is the DNA band and what does this mean for the experiment? If no, what does this mean for the experiment?**

   Yes- The water or master mix was contaminated with plant DNA and thus the remaining reactions may similarly be contaminated so the results cannot be trusted. In a research lab the experiment would be repeated. Alternatively, if small, a product may be a primer dimer band.

   No- The water and master mix were not contaminated with amplifiable DNA. This does not mean that the samples are also free of contamination, but makes the probability less likely.

2. **Did the pGAP control plasmid generate a PCR product? If yes, what size is the DNA band and what does this mean for the experiment? If no, what does this mean for the experiment?**

   Yes- The band should be approximately 1 kb. The reaction was set up properly and the master mix, primers, and thermal cycler are working as expected.

   No- Something was wrong with the experiment. The problem could be the reaction was not set up properly and the master mix, primers, or thermal cycler are not working as expected, or an inhibitor could have been introduced into the PCR reaction.

3. **Did the *Arabidopsis* genomic DNA generate a PCR product? If yes, what size is/are the DNA bands and what does this mean for the experiment? If no, what does this mean for the experiment?**

   Yes- There may be up to 3 bands visible for the control gDNA in the 1–1.2 kb range. These bands represent different *GAPC* genes of *Arabidopsis*, of which there are four. Bands in this reaction indicate the PCR reaction progressed as expected on a control genomic template.

   No- If a band was not generated by the control gDNA, this does not mean that the experiment has failed. There may be amplified DNA that is below the threshold of detection by agarose gel electrophoresis but that will amplify in the next round of PCR.

4. **Did your novel genomic DNAs generate PCR products? If yes, what size is/are the DNA bands and what does this mean for the experiment? If no, what does this mean for the experiment?**

   Yes- Plant *GAPC* genes range from 500 bp to 2.5 kb. A band here indicates successful amplification from the genomic DNA.

   No- If a band was not generated by PCR of your genomic DNA, this does not mean that the experiment has failed. There may be amplified DNA that is below the threshold of detection by agarose gel electrophoresis but that will amplify in the next round of PCR, hence the requirement for the nested PCR approach. However, it is possible that the initial primers had too little complementarity with the target DNA and, hence, either did not bind to the target or the primer detached from the template during denaturation.

Lesson 2: Nested PCR Focus Questions

1. **What does the enzyme exonuclease I do? Why did you treat your initial PCR product with exonuclease I before using the DNA as a template for nested PCR? What would have been the result if you had not done an exonuclease I treatment?**

   Exonuclease I removes nucleotides from the 3'-hydroxyl end of single stranded DNA.

   Exonuclease I digested any remaining initial PCR primers that were not incorporated into double-stranded DNA products.

   Without treatment with exonuclease I, any leftover initial PCR primers would bind to the template DNA during the second round of PCR. During the nested PCR, you want only the nested PCR primers to bind to the template.

2. **What would have happened if the exonuclease I had not been heat-inactivated after use?**

   If not heat-inactivated, the exonuclease I would digest the nested PCR primers just as it did the remaining initial primers.

3. **In the initial PCR, you used degenerate primers. In nested PCR, you used non-degenerate primers. In which round of PCR would you expect to see more PCR products of different sizes?**

   Degenerate primers are more likely to bind to multiple locations on the template DNA, whereas the nested PCR primers are more specific and should bind only to the target DNA. More products are expected in the first round of PCR than in the second.

4. **Did the pGAP control plasmid generate a PCR product? If yes, what size is the DNA band and what does this mean for the experiment? If no, what does this mean for the experiment?**

   Yes- The band should be approximately 1 kb and means the master mix, primers, and thermal cycler are working as expected.

   No- Something was wrong with the experiment. The problem could be the reaction was not set up properly or the master mix, primers, or thermal cycler are not working as expected, or an inhibitor could have been introduced into the PCR reaction.

5. **Did the *Arabidopsis* genomic DNA generate a PCR product? If yes, what size is/are the DNA bands and what does this mean for the experiment? If no, what does this mean for the experiment?**

   Yes- There may be 1–2 bands visible for the control gDNA in the 1–1.2 kb range. These bands represent the *GAPC* and *GAPC-2* genes of *Arabidopsis*. Bands in this reaction indicate the PCR reaction progressed as expected on a control genomic template.

   No- If a band was not generated by the control gDNA, and the experimental controls worked, then it is likely that no *GAPC* was amplified in the initial PCR or that there has been some experimental error (such as forgetting to add template DNA to the reaction).

6. **Did your novel genomic DNAs generate PCR products? If yes, what size is/are the DNA bands and what does this mean for the experiment? If no, what does this mean for the experiment?**

    Yes- Plant *GAPC* genes range from 500 bp to 2.5 kb. A band here indicates successful amplification of the DNA amplified in the first round of PCR.

    No- If you saw a band or bands after the first round of PCR, but did not see any bands after nested PCR, it may be that none of the DNA amplified in the first PCR was the target DNA. Degenerate primers may bind at more than one location on the template, and those locations might not include the target area.

    If you did not see a band or bands in the sample after the first round of PCR, and you do not see bands here, it is probable that no DNA was amplified from your genomic DNA. This does not mean that you did anything wrong. The initial primers may not have bound to the template DNA because the primers did not have enough complementarity to the template or the genomic DNA may not have been of sufficient quality or quantity to yield a PCR product.

Appendix D
Glossary of Terms

Annealing – binding of single-stranded DNA to complementary DNA sequences. Oligonucleotide primers bind to single-stranded (denatured) template DNA

Antiparallel strands – DNA strands oriented in opposite directions, such that the 5'-phosphate end of one strand is aligned with the 3'-hydroxyl of the other strand

Base pairs – complementary nucleotides held together by hydrogen bonds. In DNA, adenine is bonded by two hydrogen bonds with thymine (A-T) and guanine with cytosine by three hydrogen bonds (G-C). Because of the three H-bonds between GC (compared to the two between AT), GC bonding is stronger than the AT bonding

Chloroplast – an organelle found in plant cells, responsible for photosynthesis. A type of plastid

Chromatography – a technique for separating molecules based on their physical characteristics, such as size (size exclusion chromatography), charge (ion exchange chromatography), or hydrophobicity (hydrophobic interaction chromatography). Another type, affinity chromatography, separates molecules based on their specific activity, cofactor, or chemistry (such as sugars on a glycoprotein)

Codon – set of 3 nucleotides (in DNA or mRNA) that code for a single amino acid. Many amino acids have multiple codons

Codon bias – although there are multiple codons for many amino acids, specific organisms tend to favor particular codons instead of using all indiscriminately

Complementarity – binding of two strands of nucleotides derived from the base pairing of the nucleotides (AT and GC). The two DNA strands of the double helix have complementarity, as do DNA or RNA primers bound to template DNA

Consensus sequence – a sequence derived from the alignment of DNA or RNA sequences from similar organisms. Normally, each position in the consensus sequence is determined by the base that predominates in that position in the majority of the aligned sequences. Consensus sequences can be used to design primers for PCR

Constitutive gene expression – refers to genes that are transcribed continuously, such as housekeeping genes

Degenerate primers – a mixture of PCR primers that are similar, but not identical. They may be designed based on a consensus sequence derived from similar organisms, with substitutions of different bases at one or more locations in the primers

Denaturation – separation of complementary strands of DNA into single stranded DNA. Denaturation of DNA is also sometimes referred to as melting. *In vivo*, DNA is denatured by enzymes, but in PCR, DNA is denatured by heat

DNA ligase – an enzyme that repairs single strand breaks in double-stranded DNA

DNA polymerase – an enzyme involved in DNA replication. DNA polymerase links a new nucleotide to a growing strand of DNA. The new nucleotide is connected via a phosphodiester bond to the 3'-hydroxyl group of the most recently incorporated nucleotide on the new strand

DNA replication – the process of copying a DNA molecule. A double-stranded DNA molecule is replicated to form two identical double-stranded DNA molecules

Electrophoresis – a technique for separating molecules based on their relative migrations in an electric field. DNA and RNA are usually separated using agarose gel electrophoresis and proteins are separated using a polyacrylamide matrix (PAGE or SDS-PAGE)

Endosymbiotic theory – the theory that plastids and mitochondria exist as subcellular organelles in modern eukaryotic cells as a result of an ancient symbiotic event between eukaryotes and prokaryotes. For example, plastids are believed to derive from photosynthetic cyanobacteria that were engulfed by eukaryotic cells more than a billion years ago

Exon – a eukaryotic gene segment that is transcribed to RNA, retained after RNA processing, and will be (with other exons) part of the mRNA that will be translated to protein. Exon can refer to either the DNA sequence or the RNA transcript. Exons are separated in DNA and in the primary RNA transcript by introns (see below). Exons are also known as the protein coding sequences of genes, and introns as non-coding regions

Exonuclease – an enzyme that removes nucleotides from the ends of DNA strands. Exonuclease I, used in this experiment, removes nucleotides in a stepwise manner from the 3'-hydroxyl end of single-stranded DNA

Extension – the step in PCR in which the new strand is extended (or elongated) from the primer by addition of dNTPs to the 3'-end of the growing strand

GAPDH (glyceraldehyde-3-phosphate dehydrogenase) – the enzyme that catalyzes the sixth reaction of glycolysis, oxidizing glyceraldehyde-3-phosphate into 1,3-bisphosphoglycerate

GC clamp – one to three G or C bases at the 3'-end of a primer. If there is a GC clamp at the 3'-end of the primer, the primer will form a more stable complex with the template DNA

Gene duplication – the duplication of a region of DNA containing a gene. Gene duplication events have been important during evolution, as once the genes are duplicated, the gene copy can mutate to create a different gene. Groups of genes that have resulted from gene duplication events, such as the genes that code for *GAPDH*, are called gene families

Genomic DNA (gDNA) – all of the DNA found in a cell or organism

Glycolysis – the pathway by which glucose is converted into pyruvate in a series of ten enzymatic reactions, producing energy for the cell as well as precursors for many biological molecules

Housekeeping gene – a gene that is expressed constitutively in cells and is necessary for the cell to survive

Intron – eukaryotic gene segment that is transcribed to RNA, but is spliced out from the primary transcript during RNA processing. Introns are interspersed between exons (see above) and are also known as non-coding regions of DNA

Isozyme – isoenzymes. Enzymes that catalyze the same reaction, but differ in amino acid sequence or other physical characteristics. Isozymes are coded at different loci in the organism's genome and result from gene duplication events during evolution

Lysis – the process of rupturing a cell to release its contents. Once lysed, the solution is called a lysate

Master mix – a premixed reagent solution for chemical or biological reactions. A PCR master mix containing all components needed for PCR except for the template DNA. Master mix contains dNTPs, primer(s), buffer, salts, DNA polymerase, and $Mg2^+$

Melting temperature (T_m) – the temperature at which the two strands of a DNA molecule dissociate

Nested PCR – a variation of PCR in which two sequential rounds of PCR are performed, each with a different set of primers. The first set of primers binds outside of the region of interest, and the second set binds within the region amplified by the first set of primers

Nucleotide – a fundamental unit of DNA and RNA. Molecules comprised of a sugar, a phosphate group, and one of four bases: adenine, guanine, cytosine, and thymine (DNA) or uracil (RNA)

PCR – polymerase chain reaction. A technique for rapidly creating multiple copies of a segment of DNA utilizing repeated cycles of DNA synthesis

Okazaki fragment – short pieces (100–2,000 bases) of DNA synthesized on the lagging strand during DNA synthesis. The fragments are later linked by DNA ligase

Organic extraction – a technique used to separate molecules based on their differential solubility in solutions that do not mix, such as an aqueous solution and an organic solution. For extraction of DNA, DNA will remain in the aqueous solution, and the contaminants such as proteins will be in the organic phase or at the aqueous/organic interface

Origin of replication – a particular sequence on a molecule of DNA at which DNA replication is started. Circular plasmid DNA normally has a single origin of replication (ORI) but eukaryotic genomic DNA has many origins on each molecule

Plastid – a double membrane-bound organelle found in the cytoplasm of plants. Different plastids perform different roles in plants, including photosynthesis and storage of metabolites and pigments. Plastids include chloroplasts, leucoplasts, and chromoplasts

Primer – short, single-stranded oligonucleotide (usually 18–24 bases in length) designed to bind to DNA template strands at the end of the sequence of interest and serve as the starting point for DNA synthesis. Primers can be either single-stranded DNA or RNA

Primer specificity – the degree to which the primer sequence complements the template sequence. The more specific the primer is for the template, the higher the temperature at which the primer and template will anneal to each other. Conversely, if there are mismatches between primer and template, the annealing temperature will be lower; in fact, the primer may dissociate from the template prior the amplification

Replication fidelity – the number of mistakes made by DNA polymerase as it replicates the bases during DNA replication. High-fidelity DNA polymerase makes few errors. For example, eukaryotic DNA polymerase makes one mistake every 10,000–100,000 base pairs replicated, and it is considered high-fidelity. *Taq* DNA polymerase has low replication fidelity

***Taq* DNA polymerase** – a DNA polymerase that is stable at high temperatures and is commonly used in PCR. The enzyme was originally isolated from the thermophilic bacterium, *Thermus aquaticus*

Template DNA – the "target DNA" that contains the sequence to be amplified by PCR

Thermal cycler – an instrument used in PCR that automates the repeated cycles of heating and cooling

Transcription – synthesis of mRNA from a DNA template

Translation – the synthesis of amino acids from mRNA, producing proteins

Appendix E
References

Allison, L.A. 2007. *Fundamental Molecular Biology*. Blackwell Publishing, Malden.

Altenberg, B. &Greulich, K.O. 2004. Genes of glycolysis are ubiquitously overexpressed in 24 cancer classes. Genomics **84**: 1014–1020.

Brinkmann, H, Cerff, R., Salomon, M., &Soll, J. 1989. *Cloning and sequence analysis of cDNAs encoding the cytosolic precursors of subunits GAPA and GAPB of chloroplast glyceraldehyde-3-phosphate dehydrogenase from pea and spinach*. Plant Mol. Biol. **13**: 81–94.

Brown, T.A. 2006. *Gene Cloning and DNA Analysis, An Introduction*. Blackwell Publishing: Oxford.

Brown, T.A. 2007. *Genomes 3*. Garland Science: New York.

Fore, J., I.R. Wiechers, & Cook-Deegan, R. 2006. *The effects of business practices, licensing, and intellectual property on development and dissemination of the polymerase chain reaction: case study*. J Biomed DiscovCollab. **1**: 7–24.

Glick, B.R., and Pasternak, J.J. 1998. *Molecular Biotechnology*. ASM Press, Washington.

Lodge, J., Lund, P., and Minchin, S. 2007. *Gene Cloning, Principles and Applications*. Taylor & Francis: New York.

López-Juez, E. 2007. *Plastid biogenesis, between light and shadows*. J. Exper. Bot. **58**: 11–26.

Martin, W., Rujan, T., Richly, E., Hansen, A., Cornelsen, S., Lins, T., Leister, D., Stoebe, B., Hasegawa, M, & Penny, D. 2002. *Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastic phylogeny and thousands of cyanobacterial genes in the nucleus*. PNAS **99**: 12246–12251.

McPherson, M., & Møller, S. 2006. *PCR*. Taylor & Francis, New York.

Metzenberg, S. 2007. *Working with DNA*. Taylor & Francis Group, New York.

Petersen, J., Brinkmann, H., &Cerff, R. 2003. *Origin, evolution, and metabolic role of a novel glycolytic GAPDH enzyme recruited by land plant plastids*. J. Mol. Evol. **57**: 16–26.

Plaxton, W.C. 1996. *The organization and regulation of plant glycolysis. Annu. Rev. Plant Physiol*. Plant Mol. Biol. **47**: 185–214.

Primrose, S.B., and Twyman, R.M. 2006. *Principles of Gene Manipulation and Genomics*. Blackwell Publishing, Malden.

Sambrook, J & Russell, D.W. 2001. *Molecular Cloning, A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Sirover, M.A. 1999. *New insights into an old protein: the functional diversity of mammalian glyceraldehyde-3-phosphate dehydrogenase*. Bioch. Biophys. Acta **1432**: 159–184.

Teich, R., Grauvogel, C., & Petersen, J. 2007. *Intron distribution in Plantae: 500 million years of stasis during land plant evolution*. Gene **394**: 96–104.

Weissensteiner, T., Griffin, H.G., and Griffin, A. (Eds.) 2004 *PCR Technology: Current Innovations*. CRC Press, New York.

Yuryev, A. (Ed.) 2007. *PCR Primer Design*. Humana Press, Totowa.

On-line Resources

PCR Station: http://www.pcrstation.com/learn-about-pcr/

PrimerFox http://www.primerfox.com/

Legal Notices

Notice regarding Bio-Rad thermal cyclers and real-time systems: Purchase of this instrument conveys a limited, non-transferable immunity from suit for the purchaser's own internal research and development and for use in applied fields other than Human In Vitro Diagnostics under one or more of U.S. Patents Nos. 5,656,493, 5,333,675, 5,475,610 (claims 1, 44, 158, 160–163, and 167 only), and 6,703,236 (claims 1–7 only), or corresponding claims in their non-U.S. counterparts, owned by Applera Corporation. No right is conveyed expressly, by implication, or by estoppel under any other patent claim, such as claims to apparatus, reagents, kits, or methods such as 5' nuclease methods. Further information on purchasing licenses may be obtained by contacting the Director of Licensing, Applied Biosystems, 850 Lincoln Centre Drive, Foster City, California 94404, USA. Bio-Rad's real-time thermal cyclers are licensed real-time thermal cyclers under Applera's United States Patent No. 6,814,934 B1 for use in research and for all other fields except the fields of human diagnostics and veterinary diagnostics. Purchase of iTaq™ DNA polymerase includes an immunity from suit under patents specified in the product insert to use only the amount purchased for the purchaser's own internal research. No other patent rights (such as 5' Nuclease Process patent rights) are conveyed expressly, by implication, or by estoppel. Further information on purchasing licenses may be obtained by contacting the Director of Licensing, Applied Biosystems, 850 Lincoln Centre Drive, Foster City, California 94404, USA

Bio-Rad Laboratories, Inc. is licensed by Invitrogen Corporation to sell reagents containing SYBR Green I for use in real-time PCR, for research purposes only. Purchase of this product includes an immunity from suit under patents specified in the product insert to use only the amount purchased for the purchaser's own internal research. No other patent rights are conveyed expressly, by implication, or by estoppel. Further information on purchasing licenses may be obtained by contacting the Director of Licensing, Applied Biosystems, 850 Lincoln Centre Drive, Foster City, California 94404, USA.

SYBR is a registered trademark of Invitrogen Corporation.

BIO-RAD

**Bio-Rad
Laboratories, Inc.**

Sig 0308

10011369 Rev A