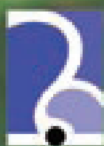
Genetic Variation and Nicotine Addiction

Investigating How Genes Influence Smoking Behavior



Education Outreach
Department of Genome Sciences
University of Washington

Teacher Edition

Genetic Variation and Nicotine Addiction:

Investigating How Genes Influence Smoking Behavior

A Human Genome Curriculum Supplement for
High School Biology

Teacher Edition

Education Outreach
Department of Genome Sciences
University of Washington

Copyright © 2004 by the University of Washington. All rights reserved. You have permission to reproduce items in this module for your classroom use. The copyright on this unit however, does not cover reproduction of these items for any other use. For permissions and other rights under this copyright, please contact:

Maureen Munn, Ph. D.,
University of Washington
Box 352145
Seattle, Washington 98195-2145
(206) 616-4538
mmunn@u.washington.edu

StarNet Administrative Staff

Debbie Nickerson, Ph.D., Principal Investigator, University of Washington, Seattle, Washington
Maureen Munn, Ph.D., Project Director, University of Washington, Seattle, Washington
Kristi Martinez, Project Manager, University of Washington, Seattle, Washington

Authors

Maureen Munn, Ph.D., University of Washington, Seattle, Washington
Kristi Martinez, University of Washington, Seattle, Washington

Contributing Author

Mark Bloom, Ph.D., BSCS, Colorado Springs, Colorado

Lead Teachers

The following teachers made major contributions to the development of the High School Human Genome Program and/or this curriculum unit:

Janet Bowersox, Nathan hale High School, Seattle, Washington
Bonnie dos Remedios, formerly of Roosevelt High School, Seattle, Washington
Patrick Ehrman, Institute for Systems Biology, Seattle, Washington
Edward Hopfner, formerly of Nathan Hale High School, Seattle, Washington
Mary Glodowski, Juanita High School, Kirkland, Washington
Sharon Masse, Garfield High School, Seattle, Washington
Peggy O'Neill Skinner, The Bush School, Seattle, Washington
Barbara Schulz, formerly of Lakeside School, Seattle, Washington
Mary Margaret Welch, Mercer Island High School, Mercer Island, Washington

Illustrations

Cindy Desmarais, University of Washington, Seattle, Washington
Kristi Martinez, University of Washington, Seattle, Washington
Justin Miller, Formerly University of Washington, Seattle, Washington

Web Site and Graphic Design

Megan Brown, Ph.D., University of Washington, Seattle, Washington
Cindy Desmarais, University of Washington, Seattle, Washington
Kristi Martinez, University of Washington, Seattle, Washington

BSCS Production Staff

Barbara Perrin, Production Manager, BSCS, Colorado Springs, Colorado
Lisa Rasmussen, Graphic Designer and Illustrator, BSCS, Colorado Springs, Colorado
Stacey Luce, Production Specialist, BSCS, Colorado Springs, Colorado
Pamela Warren, Project Assistant, BSCS, Colorado Springs, Colorado
Barbara Resch, Editor, Colorado Springs, Colorado

Thanks to

Maynard Olson, Director of the University of Washington Genome Center, for providing laboratory and office space for this project, and for his ongoing intellectual and technical support of project activities.

Participants of the 2003 and 2004 StarNet Summer Institutes for providing for valuable comments on this manuscript.

This project was supported by a Science Education Drug Abuse Partnership Award from the National Institute on Drug Abuse. Its content is solely the responsibility of the authors and does not necessarily represent the official views of NIDA or NIH.

Table of Contents

Teacher Edition

Introduction
- Overview 1
- Why Was This Unit Developed? 2
- What Background Do Students Need to Use This Unit? 2
- The BSCS 5E Instructional Model 3
  - Engage
  - Explore
  - Explain
  - Elaborate
  - Evaluate
- *Understanding by Design* 6
- Student Assessment 7
- Use of the Student Edition 8

Teacher Background
- A Genetic Basis for Nicotine Addiction 9
- Genetic Variation and Addiction 9
- What Genes Are Involved in Drug Addiction? 9
- What Are the Molecular Targets of Common Drugs of Abuse? 10
- How Is Genetic Variation Related to the Understanding of Drug Addiction? 11
- A Closer Look at Genes Involved in Nicotine Addiction 11
  - The Nicotinic Acetylcholine Receptor 11
  - The *CYP2A6* Gene 11
  - How do we make DNA sequencing templates across the *CYP2A6* gene? 13
- Single Nucleotide Polymorphisms 14

Glossary 17

References 21

Additional Web Resources for Teachers 23

Student Edition

Section I: Smoking and the *CYP2A6* Gene 25
- What Factors Influence Our Behaviors? 26
- No Accounting for Taste, or Is There? 27
- Exploring What We Know about Smoking and Addiction 28
- Reading: Nicotine, Neurotransmission, and the Reward Pathway 30

Reading: Metabolism of Nicotine 32
The Case of the Missing Enzyme 33
Variation in Proteins That Interact with Nicotine 34

Section II: Genetic Variation and the *CYP2A6* Gene 35
An Invitation from the StarNet Team 36
DNA Structure, Synthesis, and Sequencing
- A Review of DNA Structure 37
- A Review of DNA Synthesis 38
- Modeling DNA Synthesis Using Pop-It Beads 39
- Chain Termination DNA Sequencing 40
- Modeling DNA Sequencing Using Pop-It Beads 41

DNA Sequencing—Experimental Procedure
- Day 1: DNA Sequencing by Cycle Sequencing 44
- Day 2: Gel Electrophoresis 47
- Day 3: Visualizing the DNA 51

Identifying and Analyzing Single Nucleotide Polymorphisms
- Identifying Variants in the *CYP2A6* Gene 54
- Can a Single Nucleotide Change Affect the Enzyme (Protein) That Is Made? 55

DNA Sequencing—Experimental Procedure
- Day 4: Data Analysis 56

Modeling the Assembly of DNA Fragments 61
DNA Sequencing—Experimental Procedure
- Day 5: Building the Big Picture: Using the BLAST Search, Assembling Contigs, Amino Acid Translation, and Analysis of SNPs 62

Reading: The Role of the StarNet Project in Drug Addiction Research 67
Pulling It All Together 69
Appendix 1. Automated DNA Sequencing Protocol and Reagents 71

Introduction

Overview

The StarNet Project is a collaboration between the University of Washington and high schools across the country. Its goal is to discover genetic variation related to nicotine addiction. The project's central focus is on sequencing the gene that codes for the CYP2A6 enzyme, which is involved in clearing nicotine from the body. Scientists have hypothesized that people who make less or none of this enzyme will have altered smoking behavior compared with people with a normal amount of this enzyme.

This curriculum unit is organized into two sections. The first section deals with complex traits such as smoking and the genes associated with nicotine addiction. Students learn that behaviors are *multifactorial* traits because they are determined by both genetic and environmental factors. Students first investigate genetic variation as it applies to a single gene that influences our sense of taste. Next students turn their attention to the complex trait of smoking. They explore the influences of nicotine on different body systems. Activities focus on nicotine's ability to affect neurotransmission and on its metabolism in the liver. As the first section of the unit concludes, students narrow their focus to consider how variation in genes associated with nicotine addiction may influence people's smoking behaviors.

The second section of the unit begins with students reading a letter that invites them to participate in the StarNet Project. As members of the StarNet team, students work to identify previously unknown variations of the CYP2A6 protein by sequencing the *CYP2A6* gene. Collectively, students participating in the StarNet Project are working to sequence the *CYP2A6* gene from eight people who are part of the Human Polymorphism Discovery Panel. This panel was assembled by scientists at the National Institutes of Health to aid studies on genetic variation. The eight DNA donors represent different ethnic groups found within the United States. Students are using these DNA samples to look for single nucleotide polymorphisms (SNPs) in the *CYP2A6* gene.

In preparation for sequencing the *CYP2A6* gene, students review DNA structure and synthesis and carry out a modeling exercise that reinforces the major concepts. Before beginning the laboratory work, students are introduced to the chain termination method for sequencing DNA. They clarify their understanding of the technique using a modeling exercise similar to the one used to model DNA synthesis.

The DNA sequencing process requires about four days to complete. During downtimes, students complete print-based activities that prepare them for the data analysis portion of the project. Computers and bioinformatics software are used to analyze the DNA sequences obtained by the students. The sequences are examined for the presence of SNPs and are compared with other sequences contained in several large public databases.

The project web site (http://hsgp.genome.washington.edu) provides several important components, including DNA sequencing data and bioinformatics software. In addition, the site provides our DNA sequencing and bioethics units in PDF format, a list of schools participating in the DNA sequencing project, and a list of related web sites.

Why Was This Unit Developed?

In 1993, a partnership of University of Washington (UW) scientists and ethicists and Seattle area high school teachers initiated a project that enabled high school students to contribute to the Human Genome Project (HGP)[1] by sequencing small portions of human DNA. Called the High School Human Genome Program (HSHGP), this project developed classroom-friendly techniques for DNA sequencing and a bioethics unit for engaging students in thoughtful deliberations of issues related to genomic DNA sequencing. From 1993 to 2000, the HSHGP supported classroom sequencing and bioethical discussions and decision making in the Seattle area, mainly through funding from the Department of Energy, the National Institutes of Health, and the Howard Hughes Medical Institute.

Like the HGP, the scientific interests of the HSHGP evolved to investigate complex traits. The project is now focused on genes that are implicated in nicotine addiction. In 2000, the program was funded by the National Institute on Drug Abuse (NIDA) for the StarNet Project (Student-Scientist-Teacher Authentic Research Network). This grant enabled the dissemination of the classroom sequencing project to teachers and students throughout Washington State and the United States, as well as development of a unique program that enables teachers to engage in extended research experiences in UW laboratories investigating aspects of neurobiology, pharmacology, or drug abuse.

Among the many reasons for developing this unit, StarNet staff found the following to be most compelling:

- To help students gain a basic understanding of the biology of nicotine addiction
- To help students appreciate how DNA sequencing methods can contribute to our understanding of genetic variation and its influence on nicotine addiction
- To help students understand how and why computers are essential for analyzing sequence data
- To introduce teachers and students to some of the most common bioinformatics methods, including assembling contigs, searching for SNPs, and performing BLAST searches.

What Background Do Students Need to Use This Unit?

Typically, the classroom sequencing and bioethics units have been used in advanced high school biology classrooms, although they also have been used in sophomore high school biology and introductory college courses. Teachers report that successful integration of the unit depends on adequate student preparation (both conceptual knowledge and laboratory skills). It is also important to schedule sufficient time to complete the units (about one and a half weeks for DNA sequencing and one week for bioethics). The bioethics unit, which is based on presymptomatic genetic testing, can be downloaded from the High School Human Genome web site (http://hshgp.genome.washington.edu). To derive the greatest benefit from this unit, StarNet staff recommends that students be familiar with the following concepts:

- Mendel's laws of segregation and independent assortment
- The central dogma, which states that genetic information resides in DNA, passes through an RNA intermediate, and is ultimately expressed as protein
- The chemical nature of the gene, including the following:
    - o Genes are composed of DNA.
    - o DNA is a polymer composed of nucleotides.
    - o A nucleotide is composed of a deoxyribose sugar, a triphosphate group, and one of four nitrogen bases (A, adenine; C, cytosine, G, guanine, and T, thymine).
    - o DNA consists of two chains of nucleotides arranged in a double helix configuration.
    - o The two chains are held together by weak hydrogen bonds between the nitrogen bases.
    - o Hydrogen bonds form between adenine

  - and thymine and between cytosine and guanine.
  - o The two chains run in opposite directions.
  - o The direction of each chain is indicated by using numbered carbon atoms on the deoxyribose (5′ to 3′ and 3′ to 5′).
- DNA synthesis, including the following:
  - o The DNA unwinds in the area where synthesis will begin.
  - o A short piece of RNA called a primer binds to the DNA strand that will be copied (the template strand).
  - o The enzyme DNA polymerase adds nucleotides to the 3′ end of the growing DNA chain (new strand).

Students also will require the following laboratory skills:
- Measuring and using appropriate scientific units
- Micropipetting
- Conducting gel electrophoresis (using agarose or polyacrylamide gels)
- Following a scientific protocol

The 5E Instructional Model

Because learning does not occur by way of passive absorption, the lessons in this unit promote active learning. Students are involved in more than listening and reading. They are developing skills, analyzing and evaluating evidence, experiencing and discussing, and talking to their peers about their own understanding. Students work collaboratively with others to solve problems and plan investigations. Many students find that they learn better when they work with others in a collaborative environment than when they work alone in a competitive environment. When active, collaborative learning is directed toward scientific inquiry, students succeed in making their own discoveries. They ask questions, observe, analyze, explain, draw conclusions, and ask new questions. These inquiry-based experiences include both those that involve students in direct experimentation and those in which students develop explanations through critical and logical thinking.

The viewpoint that students are active thinkers who construct their own understanding from interactions with phenomena, the environment, and other individuals is based on the theory of *constructivism*. A constructivist view of learning recognizes that students need time to
- express their current thinking;
- interact with objects, organisms, substances, and equipment to develop a range of experiences on which to base their thinking;
- reflect on their thinking by writing and expressing themselves and comparing what they think with what others think; and
- make connections between their learning experiences and the real world.

This unit provides a built-in structure for creating a constructivist classroom: the 5E instructional model.[2] The 5E model sequences learning experiences so that students have the opportunity to construct their understanding of a concept over time. The model leads students through five phases of learning that are easily described using words that begin with the letter E: Engage, Explore, Explain, Elaborate, and Evaluate. The following paragraphs illustrate how the five Es are implemented across the lessons in this unit.

Engage

Students come to learning situations with prior knowledge. This knowledge may or may not be congruent with the concepts presented in this unit. The Engage phase provides the opportunity for teachers to find out what students already know, or think they know, about the topic and concepts to be covered. Engage lessons are designed to
- pique students' curiosity and generate interest,
- determine students' current understanding about the topic,
- invite students to raise their own questions,
- encourage students to compare their ideas with those of others, and
- enable teachers to assess what students do or do not understand about the stated outcomes of the lesson.

Explore

In the Explore phase, students begin their investigation of the topic. They ask questions and begin to clarify their understanding. This phase provides a common set of experiences within which students can begin to construct their understanding. Students

- interact with materials and ideas through classroom and small group discussions;
- consider different ways to solve a problem or frame a question;
- acquire a common set of experiences so that they can compare results and ideas with their classmates;
- observe, describe, record, compare, and share their ideas and experiences; and
- express their developing understanding of the topic.

Explain

The Explain phase provides opportunities for students to connect their previous experiences with current learning and to make conceptual sense of the main ideas of the unit. This stage also allows for the introduction of formal language, scientific terms, and content information that might make students' previous experiences easier to describe. Explain lessons encourage students to

- explain concepts and ideas (in their own words);
- listen to and compare the explanations of others with their own;
- become involved in student-to-student discourse in which they explain their thinking to others and debate their ideas;
- revise their ideas;
- record their ideas and current understanding,
- use labels, terminology, and formal language; and
- compare their current thinking with what they previously thought.

Elaborate

In the Elaborate phase, students apply or extend previously introduced concepts and experiences to new situations. In Elaborate lessons, students

- make conceptual connections between new and former experiences;
- connect ideas, solve problems, and apply their understanding to a new situation;
- use scientific terms and descriptions;
- draw reasonable conclusions from evidence and data;
- deepen their understanding of concepts and processes; and
- communicate their understanding to others.

Evaluate

The Evaluate phase is the final stage of the instructional model, but it only provides a "snapshot" of what the students understand and how far they have come from where they began. In reality, the evaluation of students' conceptual understanding and ability to use skills begins with the Engage phase and continues throughout each stage of the instructional model. When combined with the students' written work and performance of tasks throughout the unit, however, the Evaluate phase provides a summative assessment of what students know and can do.

Evaluate lessons provide an opportunity for students to

- demonstrate what they understand about the topic and how well they can apply their knowledge,
- share their current thinking with others,
- assess their own progress by comparing their current understanding with their prior knowledge, and
- ask questions that take them deeper into a concept.

When you use the 5E instructional model, you engage in practices that are different from those of a traditional teacher. In response, students learn in ways that are different from those they experience in a traditional classroom. Figures TE1 and TE2, What the Teacher Does and What the Students Do, outline these differences.

Figure TE1. What the Teacher Does

Stage of the instructional model	**That is *consistent* with the 5E instructional model**	**That is *inconsistent* with the 5E instructional model**
Engage	• Piques students' curiosity and generates interest • Determines students' current understanding (prior knowledge) of a concept or idea • Invites students to express what they think • Invites students to raise their own questions	• Introduces vocabulary • Explains concepts • Provides definitions and answers • Provides closure • Discourages students' ideas and questions
Explore	• Encourages student-to-student interaction • Observes and listens to the students as they interact • Asks probing questions to help students make sense of their experiences • Provides time for students to puzzle through problems	• Provides answers • Proceeds too rapidly for students to make sense of their experiences • Provides closure • Tells the students that they are wrong • Gives information and facts that solve the problem • Leads the students step-by-step to a solution
Explain	• Encourages students to use their common experiences and data from the Engage and Explore lessons to develop explanations • Asks questions that help students express understanding and explanations • Requests justification (evidence) for students' explanations • Provides time for students to compare their ideas with those of others and perhaps to revise their thinking • Introduces terminology and alternative explanations after students express their ideas	• Neglects to solicit students' explanations • Ignores data and information students gathered from previous lessons • Dismisses students' ideas • Accepts explanations that are not supported by evidence • Introduces unrelated concepts or skills
Elaborate	• Focuses students' attention on conceptual connections between new and former experiences • Encourages students to use what they have learned to explain a new event or idea • Reinforces students' use of scientific terms and descriptions previously introduced • Asks questions that help students draw reasonable conclusions from evidence and data	• Neglects to help students connect new and former experiences • Provides definitive answers • Tells the students that they are wrong • Leads students step-by-step to a solution
Evaluate	• Observes and records as students demonstrate their understanding of the concepts and performance of skills • Provides time for students to compare their ideas with those of others and perhaps to revise their thinking • Interviews students as a means of assessing their developing understanding • Encourages students to assess their own progress	• Tests vocabulary words, terms, and isolated facts • Introduces new ideas or concepts • Creates ambiguity • Promotes open-ended discussion unrelated to the concept or skill

Understanding by Design

The materials in this unit have been developed using core principles from *Understanding by Design* (*UbD*).[3] *Understanding by Design* provides an evidenced-based approach to materials development, which aligns well with the BSCS 5E instructional model. Briefly stated, this "backward design" process is as follows: in stage 1, developers articulate clearly and carefully what we want students to understand by the end of the chapter or unit. In stage 2, developers focus on what will serve as evidence of that learning. For this stage, writers fully develop the assessment component, the Evaluate lesson. After the Evaluate lesson is complete, in stage 3, developers go back to carefully craft the sequence of learning experiences that will help students develop the targeted understandings, that is, they develop the Engage, Explore, Explain, and Elaborate lessons.

As presented in *UbD*, essential questions go to the heart of the discipline, recur naturally, and raise other important questions. For developers,

Figure TE2. What the Students Do

Stage of the instructional model	That is *consistent* with the 5E instructional model	That is *inconsistent* with the 5E instructional model
Engage	• Become interested in and curious about the concept or topic • Express current understanding of a concept or idea • Raise questions such as, What do I already know about this? What do I want to know about this? How could I find out?	• Ask for the "right" answer • Offer the "right" answer • Insist on answers or explanations • Seek closure
Explore	• Explore materials and ideas • Conduct investigations in which they observe, describe, and record data • Try different ways to solve a problem or answer a question • Acquire a common set of experiences so they can compare results and ideas • Compare their ideas with those of others	• Let others do the thinking and exploring (passive involvement) • Work quietly with little or no interaction with others (only appropriate when exploring ideas or feelings) • Stop with one solution • Demand or seek closure
Explain	• Explain concepts and ideas in their own words • Base their explanations on evidence acquired during previous investigations • Record their ideas and current understanding • Reflect on and perhaps revise their ideas • Express their ideas using appropriate scientific language • Compare their ideas with what scientists know and understand	• Propose explanations from "thin air" with no relationship to previous experiences • Bring up irrelevant experiences and examples • Accept explanations without justification • Ignore or dismiss other plausible explanations • Propose explanations without evidence to support their ideas

Elaborate	• Focuses their attention on conceptual connections between new and former experiences • Use what they have learned to explain a new event or idea • Use scientific terms and descriptions previously introduced • Ask questions that help them draw reasonable conclusions from evidence and data	• Ignore previous information or evidence • Draw conclusions from "thin air" • Use terminology inappropriately and without understanding
Evaluate	• Make conceptual connections between new and former experiences • Use what they have learned to explain a new object, event, organism, or idea • Use scientific terms and descriptions • Draw reasonable conclusions from evidence and data • Communicate their understanding to others • Demonstrate what they understand about the concept(s) and how well they can implement a skill • Compare their current thinking with that of others and perhaps revise their ideas • Assess their own progress by comparing their current understanding with their prior knowledge • Ask new questions that take them deeper into a concept or topic area	• Disregard evidence or previously accepted explanations in drawing conclusions • Offer only yes or no answers or memorized definitions or explanations as answers • Fail to express satisfactory explanations in their own words • Introduce new, irrelevant topics

then, the implications are that essential questions be carefully written and logically sequenced. It also implies that there are fewer, more salient questions. Student understanding as a goal of *UbD* is multifaceted. Wiggins and McTighe suggest six facets of understanding, which, when taken together, represent a robust characterization of *understanding*. These facets include explanation, interpretation, perspective, application, empathy, and self-knowledge.

Student Assessment

Because teachers will use this unit in a variety of ways and at a variety of points in their curriculum, the most appropriate mechanism for assessing student learning is one that occurs informally throughout the unit, rather than something that happens more formally just once at the end of the unit. Accordingly, integrated within the student edition are specific assessment components. These "embedded" assessment opportunities include one or more of the following strategies:

- Performance-based activities (for example, using models or analyzing sequence data),
- Oral presentations (for example, presenting experimental results or participating in a discussion), and

- Written assignments (for example, answering questions or writing a report).

These strategies allow the teacher to assess a variety of aspects of the learning process, such as students' prior knowledge and current understanding, problem-solving and critical-thinking skills, and ability to synthesize ideas and apply understanding to a new situation.

Students should keep a detailed lab notebook that serves as a portfolio chronicling their experience with the unit. The lab notebook is needed for many of the student activities and is essential to complete the final evaluate activity. Student lab notebooks should contain the following items:

- All data and results of experiments,
- Answers to questions contained in the experimental procedures,
- Diagrams that they labeled,
- Photocopy of their sequencing ladder,
- Summary of their bioinformatics analysis, and
- Report from the final evaluate activity.

Use of the Student Edition

The student edition contains all the figures and tables necessary to carry out the activities. In some activities, students are directed to label the figures or write in the tables. If you intend to reuse copies of the student edition, instruct the students *not* to write on these figures and tables. All figures and tables that students will write on are presented on separate pages. This allows them to be easily photocopied for student use.

Teacher Background

A Genetic Basis for Nicotine Addiction

The abuse of psychoactive drugs (that is, drugs that affect the brain), particularly tobacco and alcohol, is a major health problem in the United States. The *2002 Monitoring the Future Study*,[4] which surveyed drug use among high school students across the United States, reported the following extent of lifetime drug use among 12th graders: alcohol: 78 percent, cigarettes: 57 percent, marijuana: 48 percent, inhalants: 12 percent, LSD: 8 percent, and cocaine: 8 percent. According to the surgeon general, tobacco smoking is the leading preventable cause of disease and death in the United States and accounts for nearly 20 percent of all deaths in developed countries.[5] The teen years are a critical time for making choices about smoking, since 80 percent of all smokers begin when they are teens.

Knowing the potentially deadly effects of psychoactive drugs, why do some people persist in using them? The answer, in part, is that their bodies have become *addicted* to a particular drug, so that quitting causes psychological or physical discomfort or both. Because of the important role of addiction in reinforcing drug use, understanding the processes that lead to drug addiction is a major focus of scientific research.

Genetic Variation and Addiction

What is genetic variation and how is it related to drug addiction? Each of us is strongly aware of how we are different from everyone else—our own uniqueness. People come in all different shapes, sizes, and colors, with a wide range of abilities, talents, and personalities. We even vary in the way we respond to drugs. What determines our traits? They are determined by a variety of factors, including genetics, our environment, and our culture.

An exciting revelation of the Human Genome Project is how similar all human beings are on the genetic level—we are all 99.9 percent the same![1, 6] This means that a comparison of the DNA sequence of two individuals would reveal approximately one different nucleotide for every thousand nucleotides of sequence. Some of these nucleotide differences do not have any effect at all, while others change a particular genetic trait. In rare cases, differences in the nucleotide sequence can have extreme consequences for the individual, but usually they result in the subtle differences that make each of us unique.

An important area in genomic research is characterizing variation in genes that are related to potential health risks. For example, scientists are looking for genes that predispose people to a higher risk of heart disease, which is the leading cause of death in the United States. Research shows that there are both genetic and environmental factors that determine susceptibility to heart disease, and many genes are involved. By comparing gene sequences in people from families with high and low incidence of heart disease, scientists can build a picture of the genetic factors that make people more susceptible and address better ways to prevent or treat this condition. A similar approach is being applied to drug addiction.

What Genes Are Involved in Drug Addiction?

To address this topic, we need to answer the following questions:

- What is drug abuse?

- What is drug addiction?
- How do addictive drugs work?

A drug is any chemical substance other than food or substances needed for normal life that causes a change in the structure or function of the body through its chemical actions. Under this very general definition, drugs include medications such as aspirin, antibiotics like penicillin, as well as psychoactive drugs.[7, 8]

In this discussion, we will focus on drugs that are commonly abused and have the potential to be addictive. Drug abuse is defined as taking a drug for any reason other than a medical one.[9] Common examples of drug abuse include taking drugs to get a feeling of euphoria (a high) and taking steroids to enhance athletic ability. Our study is concerned with the abuse of psychoactive drugs.

Commonly abused psychoactive drugs include legal substances like alcohol and nicotine and illegal substances like marijuana, cocaine, LSD, and heroin. As well as causing a sensation of pleasure in the user, these drugs have the potential to be addictive. According to the National Institute on Drug Abuse, a drug is said to be addictive if it causes “uncontrollable compulsive drug seeking and use, even in the face of negative health and social consequences.”[9] Addiction results from both psychological and physiological dependence on the drug. Psychological dependence is characterized by constant craving for the drug, in spite of its detrimental effects. Physical dependence results in unpleasant symptoms when the drug is withheld, also known as withdrawal.

How do drugs of addiction interact with our bodies at the molecular level? Addictive drugs exert their effect by interfering with a natural neural pathway in the brain called the reward pathway, which normally occurs in response to activities that promote survival, like eating and drinking.[10, 14] Our body reinforces these helpful behaviors through the stimulation of a specialized set of neurons in the brain that create the sensation of pleasure in response to these activities. One part of the reward pathway consists of specialized neurons in the ventral tegmental area (VTA) of the brain (just above the brain stem) that use the neurotransmitter dopamine to stimulate neurons in other parts of the brain. Stimulation of these neurons in the VTA results in an electrical impulse down the nerve axons. At the end of the nerve axons, vesicles containing dopamine fuse with the cell membrane, releasing dopamine into the synaptic cleft. The target cells at the cleft include nerve cells of the nucleus accumbens, a part of the emotional center of the brain (also called the limbic system), and neurons of the frontal region of the cerebral cortex (Figure TE3). After its release, dopamine is quickly reabsorbed by the cells that release it by a specialized pump called the dopamine transporter.[7, 9]

What Are the Molecular Targets of Common Drugs of Abuse?

Nicotine, cocaine, alcohol, and amphetamines all exert their addictive effects through the reward pathway.[9] Although each of these drugs interacts with the brain in a different way, their overall effect is similar. They all increase the amount of dopamine in the synaptic clefts within the reward pathway. Nicotine mimics the neurotransmitter acetylcholine and binds to specific acetylcholine receptors on neurons in the ventral tegmental area of the brain (discussed in more detail later). When nicotine binds to these receptors, an electrical impulse is sent down the nerve axon, resulting in the release of dopamine at the synapse. Through a different mechanism, alcohol also stimulates electrical impulses down nerve axons, resulting in release of dopamine at nerve endings. In contrast, cocaine and amphetamines bind to the dopamine transporter on the nerve endings and block the re-uptake of dopamine, which results in the accumulation of dopamine in the synapse. An additional effect of amphetamines is that they bind to dopamine vesicles in the nerve endings and stimulate the release of dopamine into the synapse.[9]

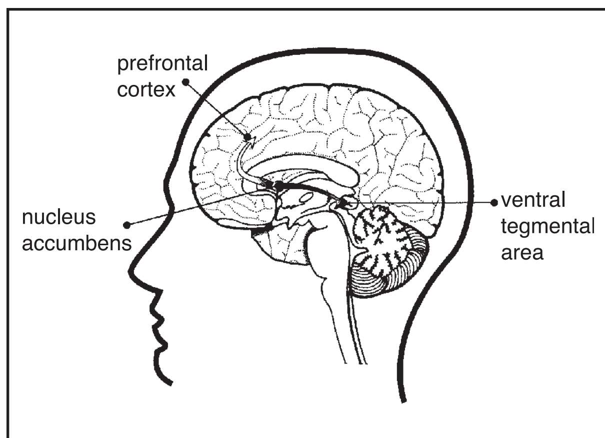


Figure TE3. The reward pathway in the brain[11]

How Is Genetic Variation Related to the Understanding of Drug Addiction?

Both research and everyday observation demonstrate that people react in different ways to drugs. For example, most of us know some people who struggle to quit smoking. They stop smoking and start again frequently. Other people are able to quit "cold turkey." The act of giving up nicotine has a strikingly different effect on these two groups of people, causing strong withdrawal symptoms in the former, and having a much less severe effect on the latter.[7] Is this difference genetic? In order to answer this question, researchers are sequencing genes implicated in addiction in many individuals and correlating their genotypes to their drug-taking profiles. Scientists decide which genes to study based on the known molecular mechanism of the drug. Several candidate genes that may have an effect on nicotine addiction include the genes for the dopamine receptor, nicotinic acetylcholine receptor, and enzymes involved in dopamine synthesis and metabolism and nicotine metabolism.[12, 13]

A Closer Look at Genes Involved in Nicotine Addiction

In 2000, the StarNet Project began to collaborate with Dr. Carl Ton at the University of Washington Genome Center. In his search for genetic variation that correlates with smoking behavior, Dr. Ton studied several different genes, including the genes that code for the subunits that make up the nicotinic acetylcholine receptor and the *CYP2A6* gene, which codes for an enzyme required for clearing nicotine from our bodies. Our current StarNet project is focused on analyzing genetic variation in the *CYP2A6* gene.

The Nicotinic Acetylcholine Receptor

When cigarette smoke is inhaled, the nicotine in the smoke is absorbed into the systemic circulation, reaching the brain within 10 seconds of the first puff.[14] Once in the brain, nicotine binds to receptors located on the cell bodies of neurons in the ventral tegmental area, as well as the terminals of these neurons, which are situated in the nucleus accumbens. Normally these receptors, called nicotinic acetylcholine receptors, bind the neurotransmitter acetylcholine (shown in Figure TE4). Nicotine is able to bind in place of acetylcholine on its receptor.

The nicotinic acetylcholine receptor (referred to as the nicotine receptor in the student edition) is a transmembrane protein made up of five subunits (shown in Figure TE5). The binding of either nicotine or acetylcholine to the receptor results in the transient opening of a cation-specific pore in the receptor, allowing cations to move into the neuron. This results in an electrical impulse down the nerve axon, which leads to the release of dopamine in the nucleus accumbens and the prefrontal cortex, creating a sensation of pleasure.[9, 15]

The CYP2A6 *Gene*

Our livers produce a family of enzymes called the cytochrome P450s. These enzymes are involved in the detoxification of fat-soluble molecules like certain by-products of metabolism, steroids, and drugs that would otherwise accumulate in the body. The cytochrome P450s carry out a series of chemical reactions that make these molecules water soluble so that the body can excrete them in the urine. One member of this family, called CYP2A6, converts nicotine to a

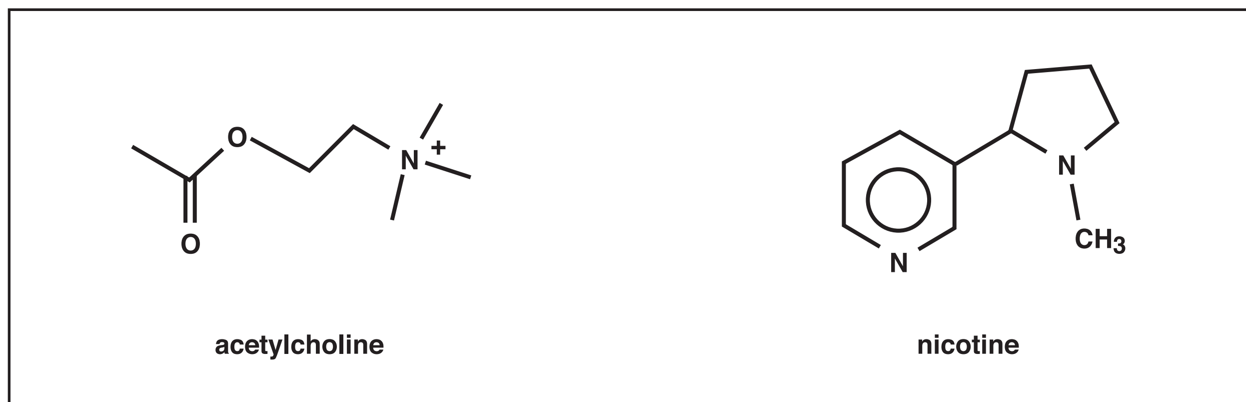


Figure TE4. Structures of acetylcholine and nicotine

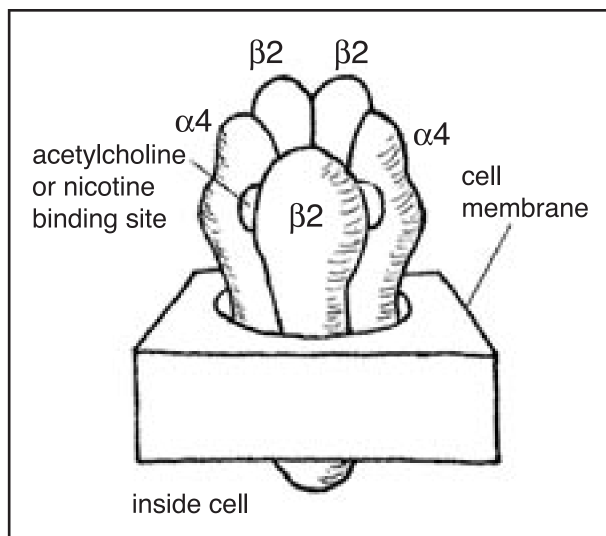


Figure TE5. Side view of the nicotinic acetylcholine receptor

chemical called cotinine, which is further modified by other enzymes and then excreted. This enzyme is the product of the *CYP2A6* gene.

Several different variations of the *CYP2A6* gene have already been identified.[13] These include single nucleotide mutations, deletions, and amplifications. The normal form of this gene codes for an active form of the CYP2A6 enzyme. One of the first variations to be identified consists of a single nucleotide change. The enzyme encoded by this variant has the amino acid histidine instead of leucine at position 160, and the enzyme is unstable and inactive. Several alleles in which the *CYP2A6* gene has been deleted also have been identified, and these result in reduced or no enzyme being produced, depending on whether the individual has one or no copies of the *CYP2A6* gene. One example of gene duplication has also been identified, and this results in increased enzyme production.

Is there a correlation between which two alleles of this gene people have (and thus how much of the CYP2A6 enzyme their bodies produce) and their smoking behaviors? Several studies have shown that there is. Pianezza, Sellers, and Tyndale demonstrated that people who make less CYP2A6 enzyme are less likely to be tobacco-dependent smokers.[16] Smokers with reduced CYP2A6 enzyme smoke significantly fewer cigarettes per day than people with normal levels of this enzyme. In a later study, they confirmed this result and showed that people who have a duplication of the gene in one of their alleles (and thus make more CYP2A6 enzyme) are heavier smokers than those who have two normal alleles.[17]

However, not all research studies support this hypothesis. What are some reasons for this conflict? In some of the earlier experiments, researchers overestimated the numbers of certain *CYP2A6* alleles because of technical difficulties. These experiments are being repeated using new procedures. A second reason is that different labs

use different methods to measure how much people smoke. Some studies rely on self-reporting of the number of cigarettes smoked, but this does not take into account how intensely the person smokes or how much of the cigarette is consumed. Other studies use additional measurements that are independent of self-reporting, including the level of carbon monoxide in the breath and levels of nicotine and cotinine in the urine and plasma. These techniques give a more accurate measurement of how much the person has smoked. Another consideration is that other genes may also be involved in determining how much a person smokes (e.g., dopamine receptor, dopamine-synthesizing and metabolizing enzymes, nicotinic acetylcholine receptor). This means that the contribution from CYP2A6 may be only part of the effect. Finally, certain alleles are more common in some ethnic groups than others, so results from different labs may be skewed depending on which populations were studied.

A correlation also has been shown between the amount of the CYP2A6 enzyme produced and lung cancer. The CYP2A6 enzyme converts precarcinogenic compounds in tobacco smoke called nitrosamines to carcinogens. A study that analyzed the *CYP2A6* genes of individuals with lung cancer showed that these patients were less likely to have two copies of the deletion allele than were healthy subjects.[18] Thus, making less CYP2A6 enzyme results in a decreased risk for lung cancer. This may be due to the fact that these people are less likely to smoke or smoke fewer cigarettes than people with two normal alleles, or it may be due to the lower level of carcinogens in their bodies.

Understanding which genes predispose individuals to become smokers or continue to smoke may help in developing treatments and cessation programs. The drug methoxsalen inhibits CYP2A6 in vitro. When this drug was given to smokers, either with or without oral nicotine, they smoked less.[19] These findings are exciting because they suggest that it might be possible to help smokers quit smoking by inhibiting their CYP2A6 enzyme. More extensive studies are needed to determine whether this approach could be widely effective.

Our student sequencing project focuses on characterizing sequence variation in the *CYP2A6* gene among eight different individuals. Our eight subjects are part of the Human Polymorphism Discovery Panel, which is a panel of human subjects assembled by scientists at the National Institutes of Health for studying genetic polymorphisms. These individuals have provided their informed consent for use in genetic variation studies, and their identities are kept anonymous. The eight individuals in our study represent different ethnicities found in the population of the United States. Therefore, it is quite likely that we will identify some genetic differences. Within our panel of human subjects, we may find polymorphisms that have already been identified, and we may discover new ones.

CYP2A6 is located on chromosome 19 at position 19q13.2 (see Figure TE6). It occurs in a cluster of *CYP2A* genes that includes *CYP2A7* and *CYP2A13*, as well as two *CYP2A* pseudogenes. The *CYP2A* genes all have nine exons and almost the identical DNA sequences.[20] This can make it very difficult to distinguish them experimentally, and it has led to some contradictory results in the literature.[13]

How Do We Make DNA Sequencing Templates Across the CYP2A6 *Gene?*

The DNA templates used by students are made by the two-step polymerase chain reaction (PCR) process shown in Figure TE7. First, PCR is used to amplify a region (2,000–3,000 nucleotides) of the *CYP2A6* gene, using PCR primers that are specific for *CYP2A6* and do not match any other region in the genome, including two very similar genes, *CYP2A7* and *CYP2A13*. Then this primary PCR product is diluted and used as the DNA template in a secondary PCR reaction. In the second reaction, the PCR primers are about 200 to 300 nucleotides apart in order to make a DNA fragment that is about the right size for DNA sequencing. The secondary PCR primers each have an extra nonhuman DNA sequence on their 5′ ends that will bind to one of two differ-

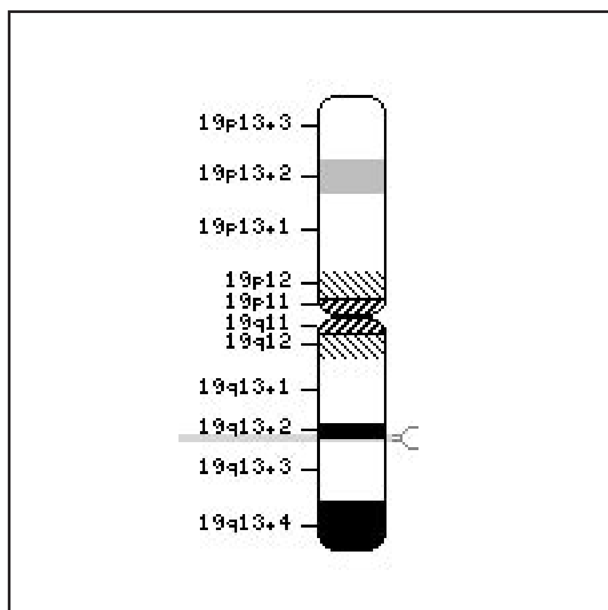


Figure TE6. Chromosomal location of the *CYP2A* gene family

ent DNA sequencing primers (referred to as the sequencing primer binding sites). This allows the PCR products to be sequenced from both ends, providing sequencing data that covers the entire fragment.

Single Nucleotide Polymorphisms

With the completion of the human genome sequence, many genomics labs are now focused on characterizing genetic variation among different individuals, especially in regions that affect human health. A variety of mutations lead to genetic variation, including insertions and deletions of a few or many nucleotides and single nucleotide changes. The latter are by far the most common. These are called single nucleotide polymorphisms, or SNPs ("poly" means several and "morphism" means form).[21]

In general, SNPs fall into two types, those that affect the production of proteins, and those that do not. Because less than 1.5 percent of the human genome is translated into proteins, most SNPs do not have a biological effect. However, those rarer SNPs that affect protein production are more interesting. SNPs can affect protein production in a number of ways. As students will learn from the activity *Identifying and Analyzing Single Nucleotide Polymorphisms,* the occurrence of a SNP in a coding region can change the amino acid sequence of the protein encoded by that gene (or if a stop codon is inserted, it can result in a shortened protein). This can dramatically alter how well that protein works, which in turn can have a dramatic effect on a particular genetic trait. SNPs also can occur in regions of the genome that affect gene expression, resulting in a change in the amount of a protein made. Finally, SNPs in RNA splice sites can change the processing of mRNA, leading to a number of different effects on protein production.

A SNP can result in a change in the amount of a protein made or in its function. Depending on the gene affected, this can have a subtle effect on the well-being of the individual, such as changing whether a person can detect a bitter taste, or it can have a dramatic effect, like causing a genetic disease. Scientists are particularly interested in SNPs that affect human health. The most common human diseases—heart disease, high blood pressure, cancer—have both a genetic and an environmental component. For each of these complex diseases, there may be one or more genes that contribute to a person's susceptibility, depending on what allele (form) of each gene the person has. Since SNPs are the most common type of mutation found in genes, scientists focus a lot of attention on discovering SNPs that correspond with certain human diseases.

Another area where SNP discovery is important is a field called pharmacogenetics.[21] People vary in the way that they react to therapeutic drugs, and yet patients with the same medical condition generally receive a uniform drug therapy. This can lead to adverse drug reactions. Pharmacogenetics is a growing field that investigates how drugs are metabolized in our bodies. Different forms of metabolizing enzymes affect the effectiveness and toxicities of drugs. Many enzymes in the cytochrome P450 family (which includes CYP2A6) are important in drug metabolism. They are a major focus of pharmacogenetic studies. In the future, a physician may

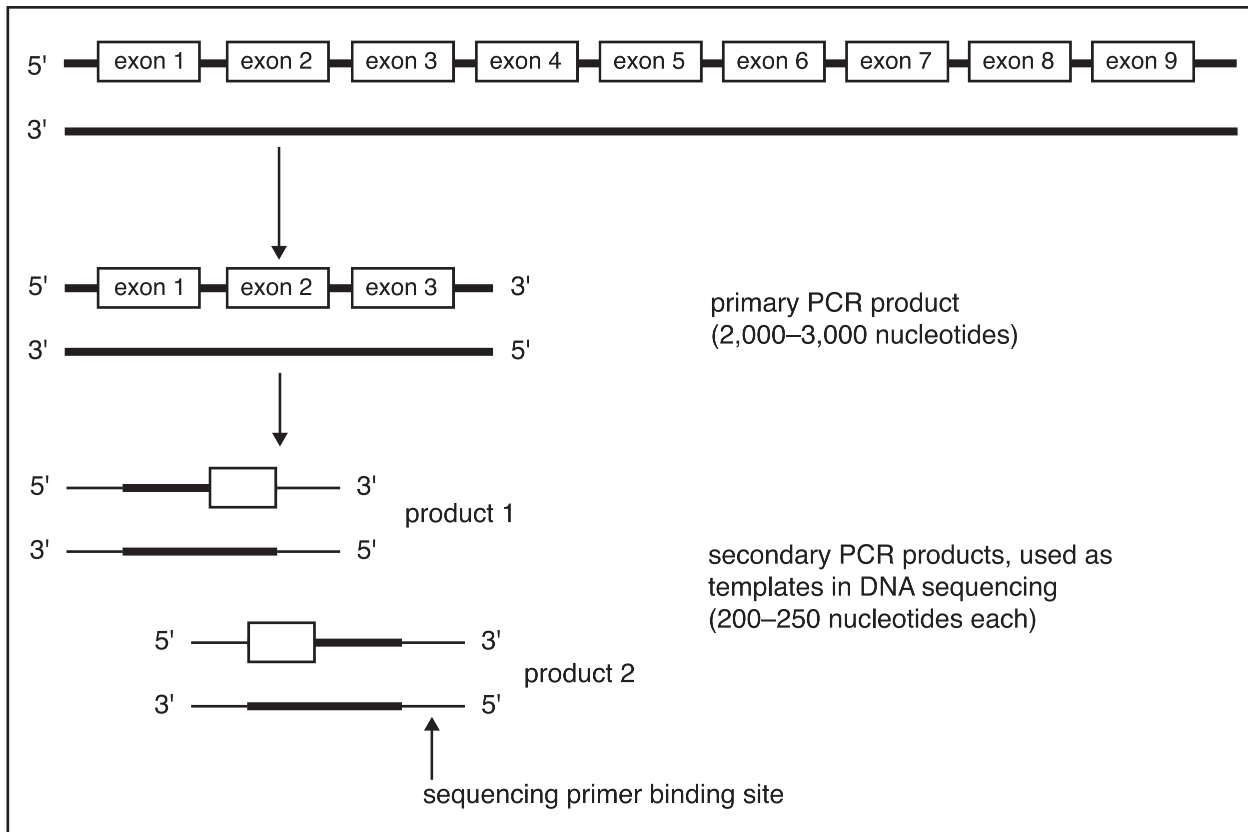


Figure TE7. PCR amplification of *CYP2A6* gene

first determine which genotype a patient has and then design a drug therapy that will work best for that person's drug metabolism.

Nearly 40 different genetic variants of the *CYP2A6* gene have been identified. (See http://www.imm.ki.se/CYPalleles for a complete list.) These include deletions, duplications, and SNPs. People who inherit two copies of a gene deletion have no active CYP2A6 enzyme, while people with one or more copies of the gene duplication have increased enzyme activity. Most SNPs that result in an amino acid substitution cause decreased or no enzyme activity. Several SNPs outside the coding region have been characterized, and some of these result in decreased enzyme activity as well. It is possible that students collaborating in the StarNet Project will identify an unknown SNP.

Students participating in the StarNet project will have an opportunity to sequence portions of *CYP2A6* alleles. After students have completed their data analysis, you will submit their data to StarNet Project staff. They will check the sequences and post them on the project Web site. When a SNP is identified, StarNet scientists verify the SNP by checking the sequencing data on both strands and having additional classrooms sequence the same samples. Once confirmed, the SNP is submitted to a designated database at the National Center for Biotechnology Information called dbSNP.

Glossary

acetylcholine: One of several chemicals used as a neurotransmitter in our bodies.

agarose: A polysaccharide derived from seaweed that is used to prepare gels for electrophoresis.

alkaline phosphatase (AP): An enzyme that removes a phosphate group from its substrate. It can be used as part of a DNA visualization process. Removal of the phosphate group produces a colored precipitate that indicates the presence of a DNA fragment on a nylon membrane.

assembly: Putting sequenced fragments of DNA into their correct order along the chromosome.

axon: Extensions of nerve cells that carry the nerve impulses away from the nerve cell body.

biotin: A vitamin molecule. In the context of DNA visualization, biotin specifically binds to the protein streptavidin, helping attach the enzyme alkaline phosphatase to the DNA fragments being visualized.

BLAST (Basic Local Alignment Search Tool): A computer program that searches for sequence similarities. It can be used to identify homologous genes and proteins in different organisms.

cell body: The portion of a neuron that contains the nucleus and most of the cytoplasm; distinct from the axon and dendrites.

chain termination: Refers to a type of DNA sequencing developed by Fred Sanger. It involves the incorporation of modified nucleotides that stop growing DNA chains at each nucleotide position.

codon: Three nucleotide sequence that codes for an amino acid during protein synthesis. There are 61 codons that code for the 20 different amino acids and three codons, called stop codons, which result in termination of a protein chain.

consensus sequence: A DNA sequence that is characteristic of a particular gene element. Consensus sequences for a particular gene element share most but not all nucleotides. In the context of DNA sequencing, the consensus sequence describes the most frequent sequence at each nucleotide position within a segment of an assembly of overlapping DNA sequences.

contig: 1. A contiguous sequence of DNA created by assembling shorter, overlapping sequenced fragments of a chromosome. 2. A list or diagram showing an ordered arrangement of cloned, overlapping fragments that collectively contain the sequence of an originally continuous piece of DNA.

cotinine: A breakdown product of nicotine, which is modified and eliminated from the body.

cycle sequencing: A DNA sequencing technique that combines the chain termination method with aspects of the polymerase chain reaction.

CYP2A6: An enzyme in the cytochrome P450 family. This enzyme converts nicotine to cotinine, which is then modified and eliminated from the body.

cytochrome P450: A family of enzymes produced in the liver that are involved with the detoxification of fat-soluble molecules including those associated with drug metabolism.

dendrite: A cluster of small fibers on the cell body of a neuron that receives chemical messages from neighboring neurons and transmits them to the cell body.

deoxynucleotide: A building block of DNA. A deoxynucleotide consists of a deoxyribose sugar, a triphosphate group, and one of four nitrogen bases (A, C, G, or T).

dideoxynucleotide: Synthetic nucleotides lacking both 2′ and 3′ hydroxyl groups. They act as chain terminators during DNA sequencing reactions.

DNA polymerase: An enzyme that adds nucleotides to a replicating DNA strand.

dopamine: A chemical messenger (neurotransmitter) that regulates brain processes such as those that control movements, emotions, pleasure, and pain.

dopamine transporter: A specialized pump located in the cell membrane of a neuron. It reabsorbs dopamine molecules after they have been released during neurotransmission.

drug: Any chemical substance other than food or substances needed for normal life that causes a change in the structure or function of the body through its chemical actions. In a medical context, this includes any substance used in the diagnosis, prevention, treatment, or cure of a disease. In an abuse context, this includes substances that alters consciousness and may be habit forming.

drug abuse: The continued use of a drug for any reason other than a medical one. This includes taking a psychoactive drug to get high or taking steroids to increase athletic ability.

drug addiction: Uncontrollable, compulsive drug seeking and use, even in the face of negative health and social consequences. Addiction results in a physical or psychological dependency on a drug.

eukaryotic cell: A cell that has a membrane-bound nucleus, as well as other membrane-bound structures. Organisms that have this cell type include animals, plants, protists, and fungi, and they are referred to as eukaryotes.

exon: A segment of mRNA in a eukaryotic cell that is translated into protein. During mRNA formation, segments of the mRNA are cut out, and the remaining pieces, called exons, are joined end to end (see **intron**).

FASTA format: A method of entering DNA or amino acid sequence data that is compatible with popular sequence comparison programs such as BLAST and FASTA.

genomic DNA: DNA isolated from the nucleus of a cell.

genotype: The genetic makeup of an individual. The expression of genotype as visible traits is called the phenotype.

intron: A segment of RNA that is removed from an mRNA before the mRNA is translated into protein (see **exon**).

multifactorial trait: A trait whose phenotype is influenced by the combined action of many genes and the environment.

neuron: A principal class of cells in the nervous system, composed of three parts: the cell body, dendrites, and axons. Neurons conduct electrical impulses and communicate with other cells through chemical messengers called neurotransmitters.

neurotransmission: The process by which neurons transmit messages to other neurons, muscle cells, or gland cells.

neurotransmitter: A chemical substance that transmits a nerve impulse across a synapse.

nicotine: An addictive chemical found in tobacco leaves.

nucleus accumbens: The part of the brain related to the limbic system that controls emotions.

phenotype: The externally or internally detectable characteristics of an organism that represent the influence of environment and genetic information (genotype).

polyacrylamide: A synthetic polymer, similar to that found in a soft contact lens, which is used to prepare gels for electrophoresis.

polymerase chain reaction (PCR): A laboratory technique that uses a heat-stable DNA polymerase to amplify a short DNA sequence.

primer: A short sequence of RNA or DNA that binds to a single-stranded region of DNA and serves as a binding site for DNA polymerase.

reward pathway: A specialized network of neurons in the brain that produce and regulate pleasure associated with eating, drinking, and sex. These neurons use dopamine as a neurotransmitter.

ribosome: A structure made of protein and RNA that is located in the cytoplasm of a cell and is the site of protein synthesis.

running buffer: A salt-containing solution that is used during electrophoresis to conduct electricity.

Sequencher: A computer program that is used to analyze DNA sequences.

single nucleotide polymorphism (SNP): A common single-base-pair variation in a DNA sequence.

streptavidin: A small bacterial protein. In the context of DNA visualization, streptavidin specifically binds to biotin, helping attach the enzyme alkaline phosphatase to the DNA fragments being visualized.

synapse: The tiny space between two nerve cells or between a nerve cell and a muscle or gland cell.

transcription: The process of making an RNA copy of a gene.

translation: The process of making a protein by joining amino acids together in the order specified by an mRNA.

ventral tegmental area (VTA): A region of the brain involved in the reward pathway; located near the top of the brain stem.

References

1. Cooper, N. G. (Ed.). 1994. *The Human Genome Project—Deciphering the blueprint of heredity.* Mill Valley, CA: University Science Books.

2. Bybee, R. W. 1997. *Achieving scientific literacy;* Portsmouth, NH: Heinemann.

3. Wiggins, G., and McTighe, J. 1998. *Understanding by design.* Alexandria, VA: Association for Supervision and Curriculum Development.

4. National Institute on Drug Abuse. 2004. *2002 monitoring the future study*. Retrieved January 2004 from http://www.nida.nih.gov/Infofax/HSYouthtrends.html

5. Department of Health and Human Services. 1988. *Public Health Service: The health consequences of smoking: Nicotine addiction. A report of the surgeon general.* (DHHS CDC Publication No. 88-8406). Washington, DC: U.S. Government Printing Office.

6. BSCS, Videodiscovery. 1999. *Human genetic variation* [NIH Curriculum Supplement Series Grades 9–12]. Rockville, MD: National Human Genome Research Institute and National Institutes of Health.

7. Benowitz, N.L. 1992. Cigarette smoking and nicotine addiction. *Medical Clinics of North America, 76,* 415–437.

8. *NIDA for teens: The science behind drug abuse.* 2004. Retrieved March 2004 from http://teens.drugabuse.gov/glossary.asp

9. BSCS, Videodiscovery. 2001. *The brain: Understanding neurobiology through the study of addiction.* [NIH Curriculum Supplement Series Grades 9–12]. Rockville, MD: National Institute on Drug Abuse and National Institutes of Health.

10. Balfour, D. J. K. 1994. Neural mechanisms underlying nicotine dependence. *Addiction, 89,* 1419–1423.

11. Illustration of the brain reproduced from National Institute on Drug Abuse Web Site. NIDA goes to school: Science-based drug abuse education. http://www.nida.nih.gov/

12. Arinami, T., Hishiguro, H., and Onaivi, E. S. 2000. Polymorphisms in genes involved in neurotransmission in relation to smoking. *European Journal of Pharmacology, 410,* 215–226.

13. Oscarson, M. 2001. Genetic polymorphisms in the cytochrome P450 2A6 (*CYP2A6*) gene: Implications for interindividual differences in nicotine metabolism. *Drug Metabolism and Disposition, 29,* 91–95.

14. National Institute on Drug Abuse. 1998. NIDA Research report series: *Nicotine addiction* (NIH Publication Number 01-4342). Rockville, MD.

15. Dani, J. A., and Heinemann, S. 1996. Molecular and cellular aspects of nicotine abuse. *Neuron, 16,* 905–908.

16. Pianezza, M. L., Sellers, E. M., and Tyndale, R. F. 1998. Nicotine metabolism defect reduces smoking. *Nature, 393,* 750.

17. Rao, Y., Hoffman, E., Zia, M., Bodin, L., Zeman, M., Sellers, and E., Tyndale, R. 2000. Duplications and defects in the CYP2A6 gene: Identification, genotyping, and in vivo effects on smoking. *Molecular Pharmacology, 58,* 747–755.

18. Miyamoto. M., Umetsu, Y., Dosaka-Akita, H., Sawamura, Y., Yokota, J., Kunitoh, H., Nemoto, N., Sato, K., Ariyoshi, N., and Kamataki, T. 1999. CYP2A6 gene deletion reduces susceptibility to lung cancer. *Biochemical and Biophysical Research Communications, 261,* 658–660.

19. Tyndale, R. F., and Sellers, E. M. 2001. Variable Cyp2A6-mediated nicotine metabolism alters smoking behavior and risk. *Drug Metabolism and Disposition, 29,* 548–552.

20. Miles, J. S., Bickmore, W., Brook, J. D., McLaren, A. W., Meehan, R., and Wolf, C. R. 1989. Close linkage of the human cytochrome P450IIA and P450IIB gene subfamilies: Implications for the assignment of substrate specificity. *Nucleic Acids Research, 17,* 2907–2917.

21. National Center for Biotechnology Information. 2002. *A science primer.* Retrieved March 2004 from http://www.ncbi.nlm.nih.gov/About/primer/

Additional Web Resources for Teachers

1. The GENETICS Project
   http://chroma.gs.washington.edu/outreach/genetics/

   A University of Washington K–12 outreach program that provides professional development for elementary, middle, and high school teachers and works to identify and develop outstanding genetics instructional materials.

2. Genetics Education Partnership
   http://chroma.gs.washington.edu/outreach/genetics/

   This site features an extensive collection of reviews of genetics curricular materials for grades K-12, genetics classroom activities and lesson plans, and links to many other genetics and biotechnology outreach programs throughout Washington state.

3. Cold Spring Harbor Dolan DNA Learning Center
   http://www.dnalc.org/

   One of the most well-designed and informative web sites geared towards science education, the DNA Learning Center is an excellent resource for students, teachers and the general public regarding classroom and computer-based learning resources.

4. Access Excellence
   http://www.accessexcellence.org/

   Access Excellence at the National Health Museum is the place to go to learn about biotechnology. This well-designed and well-organized site features teaching communities, plenty of activities, and opportunities for collaboration between teachers, scientists, and students.

5. Genetic Science Learning Center
   http://gslc.genetics.utah.edu/

   This University of Utah site features scientist profiles, separate sections for teachers, students, and families, and excellent overviews of introductory genetics as well as more specialized topics.

6. A Primer on Molecular Genetics
   http://www.ornl.gov/sci/techresources/Human_Genome/publicat/primer/toc.html

   This Department of Energy primer, published in 1992, contains a useful overview of many key genetic concepts, as well as valuable information about the Human Genome Project. Print-ready PDF versions of the document are also available.

7. Genomics and Its Impact on Medicine and Society: A 2001 Primer
   http://genome.gsc.riken.go.jp/hgmis/publicat/primer2001/

   This primer updates the 1992 Primer on Molecular Genetics. This edition covers basic science; the Human Genome Project; what we know so far; related ethical, legal, and societal issues; medicine; benefits; and more. Also available in PDF.

8. Genomics and Its Impact on Medicine and Society: The Human Genome Project and Beyond
   http://www.ornl.gov/sci/techresources/Human_Genome/publicat/primer2001/index.shtml

The 2003 update of the 1992 Primer on Molecular Genetics provides an overview of the Human Genome Project and discusses the next steps for genomic research. Information about microbial genomics also is available.

9. Genetics Education Center
   http://www.kumc.edu/gec/

   Based at the University of Kansas, the Genetics Education Center brings together information and links from a wide variety of different sources, making it an excellent starting point for investigation of genetics education resources.

10. National Human Genome Research Institute
    http://www.nhgri.nih.gov/

    The mission of the NHGRI is to direct the Human Genome Project for the National Institutes of Health (NIH). NHGRI is one of 27 institutes, centers, or divisions that make up the NIH, the federal government's primary agency for the support of biomedical research.

11. National Institute on Drug Abuse
    http://www.nida.nih.gov/

    A part the National Institutes of Health, NIDA is the primary funder of the High School Human Genome Program.

12. Neuroscience for Kids
    http://faculty.washington.edu/chudler/neurok.html

    Neuroscience site for K-12 students and teachers. Activities, trivia, and the latest in neuroscience, all at an accessible level for this audience.

13. The Wellcome Trust – Sanger Institute
    http://www.sanger.ac.uk/

    This site includes a historical account of the Human Genome Project and the challenges the project has faced, as well as current information on the progress of genomic research. Other unique pages are "20 Facts about the Human Genome" and a "Tour of the Chromosomes".

Student Edition

Section I: Smoking and the *CYP2A6* Gene

The unit begins by introducing students to the concept of smoking as a behavior that is influenced by many different factors. Student groups examine three photos of people engaged in different types of behaviors. Students discuss what factors may have influenced the people in the photos to participate in their behaviors. The next activity sets the stage for thinking about the many factors that influence smoking behavior. Students are introduced to the term multifactorial trait, which applies to characteristics that are influenced by several genes, as well as environmental factors. They examine whether or not they can taste the chemical phenylthiocarbamide (PTC), a trait that is determined by the action of a single gene. Then they explore whether there is a correlation between ability to taste PTC and the multifactorial trait, food preference.

Smoking, too, is a multifactorial trait. It is also an example of a potentially addictive behavior. Students begin their exploration of smoking by considering what they know, or think they know, about smoking and addiction. They examine the pathway that nicotine follows as it enters the body. They identify places along this pathway where genetic variation in response to nicotine might affect an individual's smoking behavior.

Students then focus their study on how nicotine affects brain function. They watch a video that explores how nerve cells communicate with each other and relate this process to the reward pathway that the nervous system uses to reinforce important behaviors. Nicotine is considered as a drug that interferes with neurotransmission and uses the reward pathway to reinforce smoking behavior.

The metabolism of nicotine and the role of the CYP2A6 enzyme are examined. Students model how the presence or absence of the CYP2A6 enzyme affects the amount of nicotine in the body.

Students identify where in the body proteins are found that interact with nicotine. They consider how variations in these proteins might affect smoking behavior. The concept of small genetic mutations called single nucleotide polymorphisms (SNPs) is introduced. In preparation for "Section II: Genetic Variation and the *CYP2A6* Gene", students speculate on methods that can be used to identify SNPs in the genes that influence nicotine addiction.

Engage
What Factors Influence Our Behaviors?

Major Concepts
- There are many different types of behaviors.
- Individual behaviors can be influenced by many different factors.
- Acceptance by peers can reinforce behaviors.

Materials
- 1 copy of Figure 1 per student

Preparation
- Make photocopies of Figure 1 (1 per student).

Procedure Notes
- Step 2. After student groups have had a chance to discuss factors that influence the behaviors depicted in the photos, ask each group to share its thoughts with the class. Make a list on the board of the different factors identified by the class.

Answers to Questions
- Step 3a. Are there similarities in the factors that contribute to these different behaviors?

  Students will suggest some factors that may contribute to these different behaviors such as enjoyment and acceptance from peers. They may also point out that some people are naturally good at something, but they also need to practice.

- Step 3b. Consider people who smoke. Do people vary in their smoking habits? In what ways do they vary? What factors contribute to this variation?

  Students will recognize that people vary in their smoking habits. They may comment that some people smoke a little and others smoke a lot. At their age, students may attribute acceptance by their friends as a major reason for smoking. Students who live in households where others smoke may or may not smoke themselves in response. Ask students why some people smoke when they are alone. If not mentioned, bring up the idea of addiction.

Explore

No Accounting for Taste, or Is There?

Major Concepts

- A person's genetic instructions are called genotype.
- A person's physical traits are called phenotype.
- Some traits are determined by the action of a single gene.
- Behaviors are called multifactorial traits because they are influenced by many different genes and environmental factors.

Materials

Part A

- 1 piece of phenylthiocarbamide paper per student
- 1 piece of control paper per student
- Pink Post-it notes (1 per student) (or another color, such as blue, that can be distinguished from yellow by someone with red-green color blindness)
- Yellow Post-it notes (1 per student)
- 1 copy of Figures 2 and 3 per student

Part B

- 1 piece of black licorice per student
- Pink Post-it notes (1 per student)
- Yellow Post-it notes (1 per student)
- 1 copy of Figures 4 and 5 per student

Preparation

- Assemble materials.

Procedure Notes

- Part A: Step 2. After students taste the PTC paper and record their reactions, collect the PTC paper strips for disposal. This helps avoid getting the PTC and control papers confused with each other.

Answers to Questions

- Part B: Step 6. Which of these factors are genetic? Which are environmental?

  Genetic factors mentioned may include taste and smell. Environmental factors mentioned may include preference or sensitivity to a particular flavor because of lifetime exposure or associations with specific experiences.

- Part B: Step 7. In what ways does it [sense of taste] contribute to your health? Does it ever work against good health?

  Taste helps determine a person's diet. Unhealthy diets can lead to obesity and various types of malnutrition. The ability to taste bitter foods may have an evolutionary advantage by helping people to avoid substances that are bitter and toxic. Conversely, some bitter-tasting foods are nutritious, so it would be a disadvantage to avoid those foods.

Explore/Explain

Exploring What We Know about Smoking and Addiction

Major Concepts

- Smoking behavior is a multifactorial trait.
- Smoking affects many different body systems.
- Nicotine is a drug of addiction.
- Smoking poses health hazards.

Materials

- 1 copy of Figure 6 per student

Preparation

- Make photocopies of Figure 6 (1 per student).

Procedure Notes

- Step 2. After students have traced their pathways, ask several students to share their ideas with the class. Before proceeding, make sure that students realize that nicotine enters the lungs; is absorbed into the blood; reaches all parts of the body, including the brain; is carried by the blood to the liver, where it is metabolized; and is excreted via the kidneys.

Answers to Questions

- Step 1a. What is a drug?

  A drug is any substance that causes a change in the body through its chemical action. In a medical context, a drug is a chemical that is taken to prevent or treat a disease. In an abuse context, a drug is chemical that is taken to achieve a pleasurable state or (in the case of addiction) to prevent the appearance of withdrawal symptoms.

- Step 1b. What is drug abuse?

  Drug abuse refers to the use of a drug for non-medical reasons that results in negative medical, social, or legal consequences.

- Step 1c. What is drug addiction?

  Drug addiction is "uncontrollable compulsive drug seeking and use in the face of negative health and social consequences" (National Institute on Drug Abuse).

- Step 1d. What is nicotine?

  Nicotine is an addictive drug contained in tobacco.

- Step 1e. What evidence do you have that smoking tobacco is addictive?

  Students may comment that some people who smoke are unable to stop despite their best efforts. Others

may note that some smokers experience strong cravings for tobacco.

- Step 1f. Discuss the health hazards of using tobacco (smoking cigarettes, cigars, or pipes or chewing tobacco).

  Hazards mentioned may include
  - *a. cancers of the mouth, throat, and lungs;*
  - *b. emphysema;*
  - *c. coughing;*
  - *d. risks to pregnant women who smoke;*
  - *e. secondhand smoke; and*
  - *f. heart disease.*

- Step 1g. What factors contribute to smoking being addictive?

  Students may mention genetic factors such as genes associated with taste, nicotine metabolism, and brain function. Environmental factors may include peer pressure, family environment, and influence of media and films.

- Step 4. When nicotine reaches the brain, does it have any influence on the brain or does the nicotine simply pass through the brain on its way out of the body?

  Students should mention that nicotine can influence brain function. They may not understand the details but they should know that cravings for nicotine originate in the brain.

Elaborate

Reading: Nicotine, Neurotransmission, and the Reward Pathway

Major Concepts

- The process by which nerve cells communicate with each other is called neurotransmission.
- Neurotransmission is mediated through the action of chemical messengers called neurotransmitters.
- The reward pathway reinforces behaviors that are important for survival.
- Nicotine interferes with neurotransmission and reinforces smoking behavior through the reward pathway.
- Nicotine mimics the action of the neurotransmitter acetylcholine.

Materials

- Video of *Animated Neuroscience and the Action of Nicotine, Cocaine, and Marijuana in the Brain*
  It can be ordered from:
  Films for the Humanities and Sciences
  http://www.films.com
  Item number BVL8352
  Available on DVD and VHS
- 1 copy of Figure 12 per student

Preparation

- Make sure that a video player and television are available.
- Make photocopies of Figure 12 (1 per student).

Answers to Questions

1. What kind of signal is used to transmit a message from one end of a neuron to its other end?

   An electrical impulse transmits the message from one end of the neuron to the other.

2. What kind of signal is used to transmit a message between two neurons?

   A chemical signal is used to transmit a message between two neurons.

3. What is the molecular target of each of the following drugs of abuse:

   a. Nicotine:
   the acetylcholine receptor

   b. Cocaine:
   the dopamine transporter (called the dopamine uptake pump in the video)

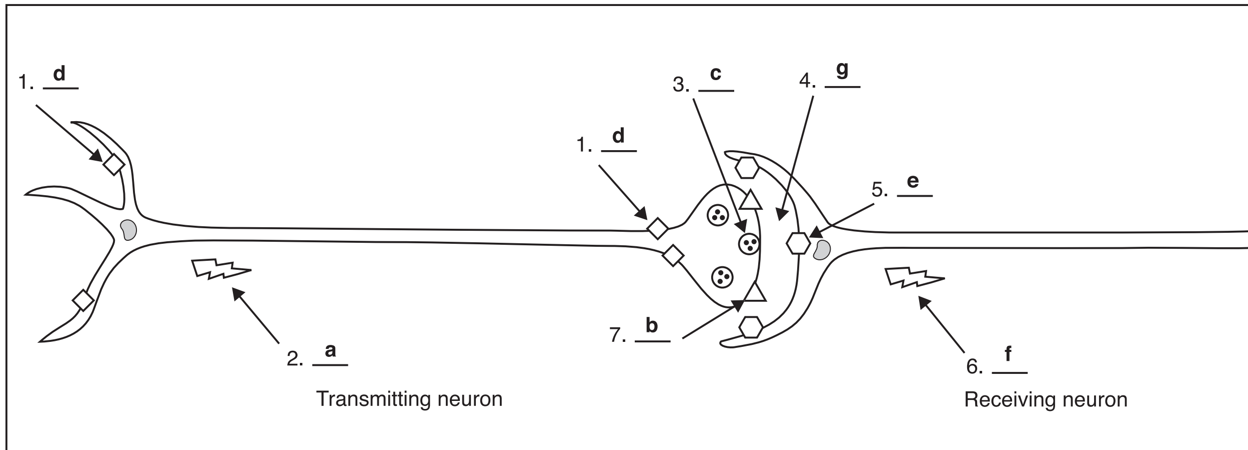
   c. Marijuana:
   the THC receptor on neurons in the hippocampus

Opportunity for Assessment

Answer is given in Figure TE8.

Figure TE8. Effect of nicotine on neurotransmission (answer key)

Match steps 1 through 7 with their correct descriptions (a through g).



◇ = nicotine receptor

= electric impulse

= vesicle of dopamine

⬡ = dopamine receptor

△ = dopamine transporter

a. Electric impulse travels along axon away from cell body.
b. Dopamine transporter takes up dopamine from synapse.
c. Vesicles of dopamine fuse with cell membrane.
d. Nicotine binds to receptors on transmitting neuron.
e. Dopamine binds to receptors on receiving neuron.
f. Electric impulse traveling away from receiving neuron increases or decreases.
g. Dopamine released into synapse.

Elaborate

Reading: Metabolism of Nicotine

Major Concepts

- Nicotine is processed in the liver and excreted in the urine.
- The liver enzyme CYP2A6 coverts nicotine to cotinine.
- The CYP2A6 enzyme processes about 70 percent of the nicotine in the body.

Materials

- 1 copy of Figure 13 per student

Preparation

- Make photocopies of Figure 13 (1 per student).

Procedure Notes

- After students complete the reading and look at Figure 13, give them an opportunity to ask questions and clarify their understanding before proceeding to the next activity.

Elaborate

The Case of the Missing Enzyme

Major Concepts

- The CYP2A6 enzyme processes most (about 70 percent) of the nicotine in the body.
- A person lacking the CYP2A6 enzyme maintains a higher level of nicotine in his or her body after smoking as compared with a person with the enzyme.

Materials

- 1 cup beans or beads (navy beans or kidney beans or beads of a similar size) per group
- 2 similar boxes, one with 2 circular holes in bottom (3 cm and 2 cm in diameter) per group
- 1 timer or watch with a second hand per group
- 1-250 mL graduated cylinder or a calibrated beaker per group
- 1 roll of cellophane tape per group

Preparation

- Cut 2- and 3-cm holes in half of the boxes.
- Place beans or beads in a cup for each student group.

Procedure Notes

- Part B: Step 1. To save time, have one roll of tape available per student team. Instruct students to tape over the larger (3 cm) hole in the box as shown below.

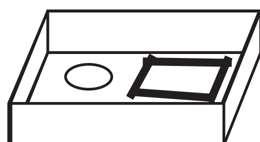


Figure TE9. Box with larger hole taped over

Answers to Questions

1. How did covering one hole affect the number of beans (or beads) that passed through the top box into the lower box?

   Covering the hole allowed fewer beans or beads to fall into the lower box.

2. Relate your results with this simple model system to nicotine metabolism in the liver of an individual who has the CYP2A6 enzyme and one who lacks the enzyme.

   A person who lacks the CYP2A6 enzyme metabolizes nicotine a slower rate than a person who has the enzyme. This means that a person lacking the enzyme will maintain a higher level of nicotine in his or her body as compared with a person with the enzyme.

3. How might a difference in CYP2A6 level affect a person's smoking behavior?

   A smoker who lacks CYP2A6, has less of it, or produces a less-active version may smoke less because it takes longer for nicotine to break down. People without this enzyme might be less likely to smoke because their nicotine levels reach a point where they feel sick.

Elaborate

Variation in Proteins That Interact with Nicotine

Major Concepts

- There are many different proteins that interact with nicotine in the body.
- Variations in these proteins may influence an individual's smoking behavior.
- The most common type of genetic variation is the single nucleotide polymorphism (SNP).

Materials

- 1 copy of Figure 12 per student
- 1 copy of Figure 13 per student

Preparation

- Make photocopies of Figures 12 and 13 (1 copy of each per student), or ask students to refer to the copy given to them previously.

Procedure Notes

- After students complete steps 1 and 2, have a few student groups report their results to the class. Guide the discussion to reach a consensus as to which proteins are most likely to influence smoking behavior.

  Students may suggest the following proteins: nicotine receptor, dopamine receptor, dopamine transporter, CYP2A6, and other nicotine-metabolizing enzymes.

Answers to Questions

Consider this. If you were a scientist, how would you propose to find SNPs in genes related to nicotine addiction?

Students may suggest that they would first identify genes that may influence smoking behavior (such as those that code for the proteins discussed in this activity). Then they would sequence those genes from two populations, one that included smokers and another that included nonsmokers. If certain SNPs are present significantly more frequently in smokers than nonsmokers, then those SNPs may come from genes that influence nicotine addiction.

Student Edition

Section II: Genetic Variation and the *CYP2A6* Gene

The second section of this curriculum unit begins with an invitation to the students to join the StarNet Project. The invitation letter explains that some studies suggest that variations in the gene for the CYP2A6 enzyme may be associated with differences in people's smoking patterns. The hypothesis is that certain alleles will characterize individuals who are more likely to become addicted to nicotine. This would allow DNA-based tests to identify such individuals and target prevention strategies to those who most need it. As part of this effort, students are asked to participate in a collaboration that will identify previously unknown SNPs in the *CYP2A6* gene. The DNA samples come from eight individuals who agreed to donate samples for this project.

Students begin their investigation of the *CYP2A6* gene by reviewing what they know about DNA structure and synthesis. DNA synthesis is modeled using pop-it beads. Students are then introduced to the chain termination method for sequencing DNA. This is the same technique used by scientists involved in the Human Genome Project. To reinforce their understanding, students model the sequencing process using pop-it beads.

At this point, students begin the four-day process of actually sequencing DNA from one of the eight individuals who are part of the StarNet Project. During the sequencing process, there are times when students must wait for one step to be completed before moving on. We take advantage of such downtimes by having students perform print-based activities that enhance their understanding of the data analysis.

Such print-based activities help students learn to identify SNPs in DNA sequences and to determine if individual SNPs cause changes to be made in the amino acid sequences of the corresponding proteins. Another activity models the process by which computers are used to assemble long DNA fragments called contigs from a series of shorter, overlapping DNA fragments.

To analyze the real DNA sequences obtained by students as part of the StarNet Project, computers and bioinformatics software are used. Students read the DNA ladders on their group's membranes and enter the sequences into the computer. A program called Sequencher is first used to identify potential SNPs in the DNA sequences. This program also helps students determine if the observed SNPs result in changes to the corresponding amino acid sequences. Next the program assembles the short, overlapping DNA sequences into a longer one (contig). Another program called BLAST (Basic Local Alignment Search Tool) is used to compare the sequence of the contig with all the sequences contained in the GenBank database hosted by the National Center for Biotechnology Information. The BLAST search results allow students to confirm that they have indeed sequenced a portion of the *CYP2A6* gene as intended.

Engage

An Invitation from the StarNet Team

Major Concepts

- Some studies suggest that different alleles of the *CYP2A6* gene are associated with different smoking patterns.
- Eight people from different ethnic groups have donated DNA samples to be sequenced by the StarNet team.

Materials

- StarNet invitation letter

Preparation

- Have students read the StarNet invitation letter in the student materials or provide each student with a photocopy of the letter.

Procedure Notes

- Students may have questions after reading the invitation letter. Be prepared to answer their questions and clarify their understanding.

Explore/Explain
DNA Structure, Synthesis, and Sequencing

A Review of DNA Structure

Major Concepts
- DNA consists of two nucleotide chains that interact to form a double helix.
- Each nucleotide consists of a deoxyribose sugar, a phosphate group, and a nitrogen base (adenine, cytosine, guanine, or thymine).
- Each DNA strand is held together by covalent bonds between the sugar of one nucleotide and the phosphate group of the adjoining nucleotide.
- The two DNA strands are held together by hydrogen bonds that form between the nitrogen bases. Adenine bonds with thymine, and cytosine bonds with guanine.
- The two DNA strands run in opposite directions (as indicated by the 5′ and 3′ notations).

Materials
- 1 copy of Figures 15 and 16 per student

Preparation
- Make photocopies of Figures 15 and 16 (1 per student).

Procedure Notes
- Listen to the various student groups as they discuss their understandings. Make sure that they can use terms such as nucleotide, complementary bases, and 5′ and 3′ correctly.

Explore/Explain
DNA Structure, Synthesis, and Sequencing

A Review of DNA Synthesis

Major Concepts
- DNA polymerase is an enzyme that uses a DNA strand as a template to assemble a new complementary strand.
- DNA polymerase begins to synthesize a new strand from a position where a short RNA primer binds to the template strand.
- DNA polymerase adds nucleotides to the 3′ end of the growing strand.
- Each nucleotide added to the new strand is complementary to the nucleotide across from it on the template strand.
- DNA synthesis can be performed in a test tube using just DNA polymerase, a mixture of the four nucleotides, a DNA primer, and a DNA template strand, all within an appropriately buffered solution.

Materials
- 1 copy of Figure 17 per student

Preparation
- Make photocopies of Figure 17 (1 per student).

Procedure Notes
- Listen to the various student groups as they discuss their understandings. Make sure that they know all the ingredients necessary to synthesize DNA in a test tube and can describe their functions.

Explore

DNA Structure, Synthesis, and Sequencing

Modeling DNA Synthesis Using Pop-It Beads

Major Concepts

- The DNA primer hydrogen bonds to its complementary sequence on the template strand.
- DNA polymerase adds new nucleotides to the 3′ end of the growing strand.

Materials (per student)

- 1 paper strip with DNA sequence = DNA template
- 4 joined beads = DNA primer
- Individual pop-it beads of each color = nucleotides (A is green, C is blue, G is yellow, T is red)
- 1 roll of cellophane tape = hydrogen bonds
- 1 roll of masking tape and pen to label primers
- 1 pair of scissors

Preparation

- Make photocopies of Figure 20 (1 per student).
- Provide students with scissors to cut out the DNA templates from each copy of Figure 20.
- Assemble the 4-bead primers in this sequence: 5′ CAGT 3′ (1 per student).
- Provide students with tape.

Procedure Notes

- After the students have finished the activity, hold a brief discussion that summarizes the major concepts of DNA synthesis. Be prepared to answer questions and to clarify students' understandings.

Explore/Explain
DNA Structure, Synthesis, and Sequencing

Chain Termination DNA Sequencing

Major Concepts
- Chain termination DNA sequencing is based on the process of DNA synthesis that takes place in the cell.
- Dideoxynucleotides differ from deoxynucleotides in that they are missing the oxygen atom on the 3′ carbon of the deoxyribose.
- When a dideoxynucleotide is incorporated into a growing chain, the chain growth stops because there is no oxygen atom for the incoming deoxynucleotide to bind to.
- Chain termination DNA sequencing relies on incorporating a dideoxynucleotide (chain terminator) at every position in the growing strand.

Materials
- None needed.

Preparation
- None needed.

Procedure Notes
- After students finish the reading and look over Figures 21 and 22, be prepared to answer questions and clarify students' understandings.

Explore/Explain

DNA Structure, Synthesis, and Sequencing

Modeling DNA Sequencing Using Pop-It Beads

Major Concepts

- Chain termination DNA sequencing is based on the process of DNA synthesis that takes place in the cell.
- Dideoxynucleotides differ from deoxynucleotides in that they are missing the oxygen atom on the 3′ carbon of the deoxyribose.
- When a dideoxynucleotide is incorporated into a growing chain, the chain growth stops because there is no oxygen atom for the incoming deoxynucleotide to bond to.
- Chain termination DNA sequencing relies on incorporating a dideoxynucleotide (chain terminator) at every position in the growing strand.
- Gel electrophoresis is used to separate the products of the DNA sequencing reactions. The gel acts as a molecular sieve, separating DNA fragments by size, shape, and charge.
- DNA is negatively charged and during electrophoresis moves toward the positive electrode.

Materials (per student group)

- 1 paper strip with DNA sequence = DNA template
- 4 joined beads = DNA primer
- Pop-it beads of each color = nucleotides (A is green, C is blue, G is yellow, and T is red)
- A few pop-it beads of the same color that have the pop-it end removed = dideoxynucleotide = chain terminator
- 1 roll of cellophane tape = hydrogen bonds
- 1 roll of masking tape and pen to label primers
- 1 pair of scissors

Preparation

- Make photocopies of Figure 25 (4 copies per student group).
- Provide students with scissors to cut out the DNA templates from each copy of Figure 25 (8 templates per student group)
- Assemble the 4-bead primers in this sequence: 5′ CAGT 3′ (8 primers per student group).
- Prepare 4 different reaction mixes or check that you have 2 sets of each of the 4 reaction mixtures:
  - o The “A” mix contains a mixture of 30 normal green beads, 10 green beads that have the pop-it end cut off, and 40 each of the blue, yellow, and red beads.
  - o The “C” mix contains a mixture of 30 normal blue beads, 10 blue beads that have the pop-it end cut off, and 40 each of the green, yellow, and red beads.
  - o The “G” mix contains a mixture of 30 normal yellow beads, 10 yellow beads that have the pop-it end cut off, and 40 each of the green, blue, and red beads.
  - o The “T” mix contains a mixture of 30 normal red beads, 10 red beads that have the pop-it ends cut off, and 40 each of the green, blue, and yellow beads.
- Use colored tape to mark the outline of a “gel” on the floor or a table as shown in Figure TE10.

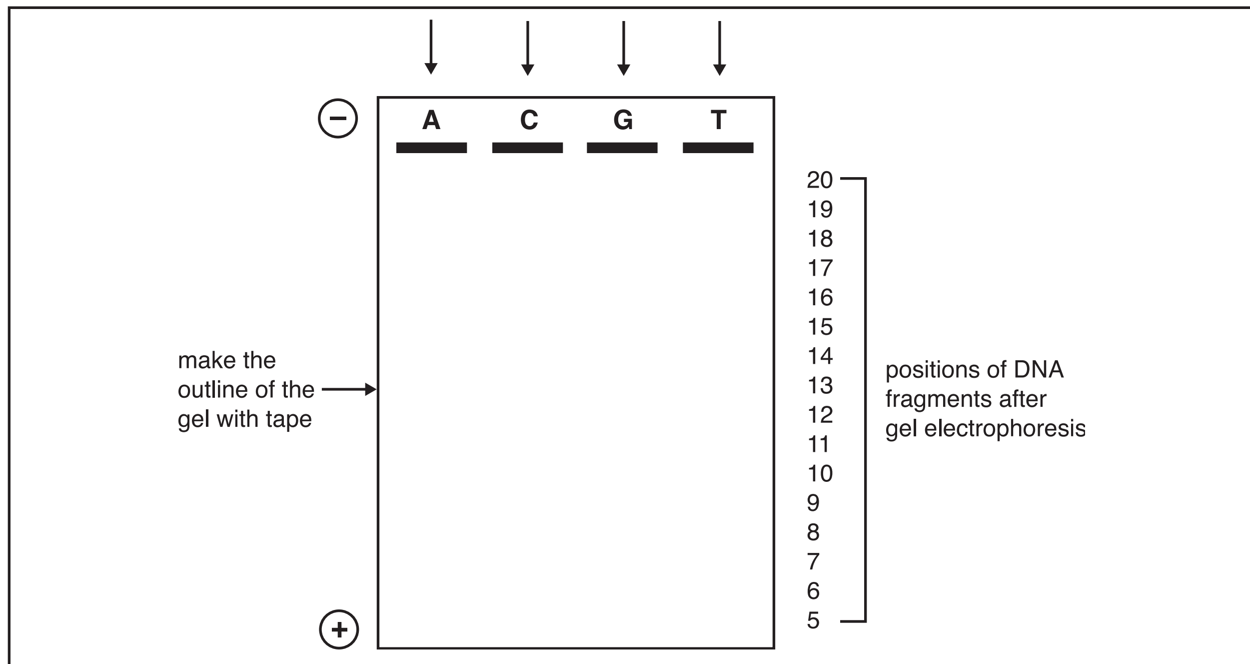


Figure TE10. DNA strands are placed in the correct lane and lined up at the position that corresponds to their length

Procedure Notes

- Part A: Step 4. Make sure that students extend enough primers so that there is at least one DNA fragment of each size between 5 and 20 bases.
- Part B. During the assembly of the gel pattern, watch for the misconception that each band on the gel corresponds to a single DNA fragment. In an actual gel, each band contains many DNA fragments of the same size.
- After the students have finished the activity, hold a brief discussion that reviews the major concepts of DNA synthesis and sequencing. Be prepared to answer questions and to clarify students' understandings.

Answers to Questions

- Part C: Step 1. Identify the smallest "band" (it should be at position 5). Which lane is it in: A, C, G, or T?

  As seen in Figure TE11, the smallest band is in the G lane.

- Part C: Step 2. Now move up the band pattern to the next lowest band (it can be in any of the four lanes). What lane is it in?

  As seen in Figure TE11, the next nucleotide is in the A lane.

- Part C: Step 3. Continue reading the sequence until the top of the gel is reached. Why is this kind of gel often called a "sequencing ladder"?

  The results from a sequencing gel are often referred to as a sequencing ladder because the individual bands that make up the pattern resemble the individual rungs of a ladder.

Opportunity for Assessment

To help you assess students' understanding, assign each student to write a brief article for the school newspaper that describes how DNA molecules are sequenced.

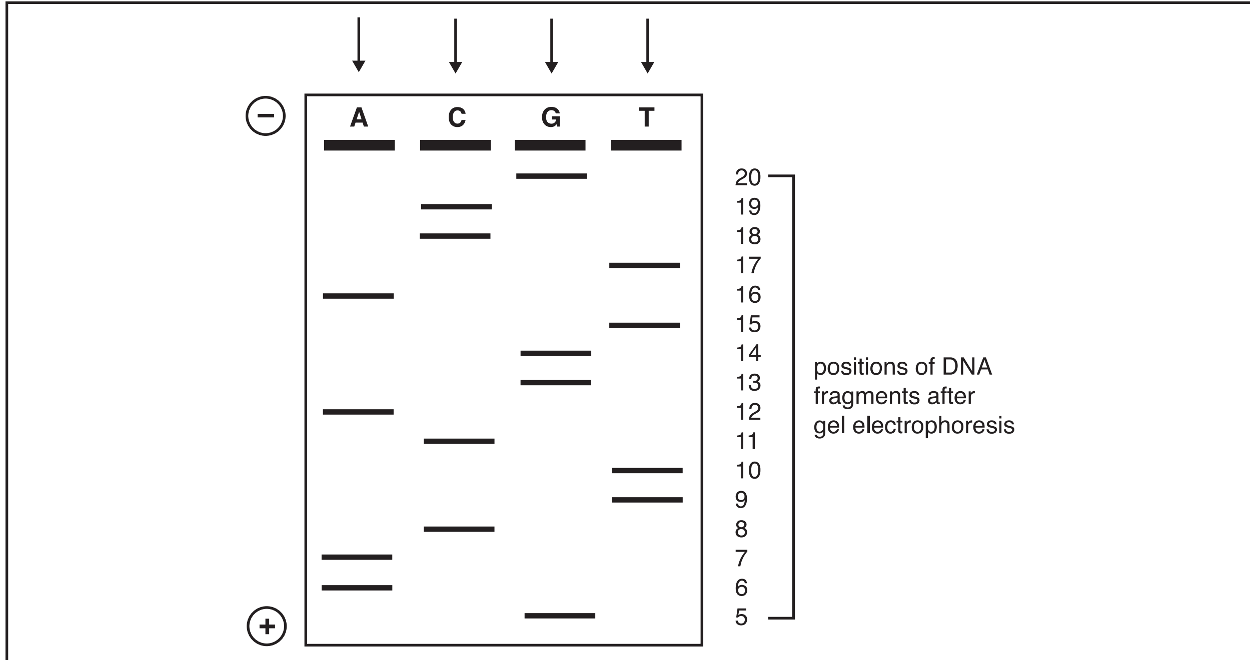


Figure TE11. Results of DNA sequencing

Explore/Explain

DNA Sequencing—Experimental Procedure

Day 1: DNA Sequencing by Cycle Sequencing

Major Concepts

- Cycle sequencing uses temperature to wind and unwind the double helix:
  - o 95°C is used to denature the DNA.
  - o 45–50°C is used to allow the DNA primer to bind to the template DNA.
  - o 70–72°C is used to allow the heat-stable DNA polymerase to extend a new strand from the primer.
- Cooling the reaction mixtures to 4°C inactivates the DNA polymerase.

Materials

- You are responsible for picking up the equipment and reagent kit from the University of Washington StarNet lab or your partner site lab. Be sure to communicate with the program manager at your site regarding the number of classrooms you'll be sequencing with and the date and time of pickup.
- When you return to the classroom with the kit, put all reagents away at the *correct temperatures*, as labeled and directed at the time of kit pickup.
- Check all supplies at least one day in advance of beginning to ensure that nothing is missing from the kit.
- Make arrangements for a scientist volunteer to attend your class during the experiment days. You are responsible for communicating your schedule with the volunteer. Contact the program manager at your site for names of potential volunteers.
- Provide each student with disposable gloves.

Preparation

- Provide each student group with a group supply kit.
- Prepare the reaction mixtures for each class no more than 2 hours before class begins:
  1. Thaw 1 aliquot each of the A, C, G, and T reaction mixes without enzyme for both the forward and reverse reactions (45 µL each). Keep cold.
  2. Spin the tubes containing the reaction mixes and those containing the DNA polymerase for a few seconds to bring the liquids to the bottoms of the tubes.
  3. Add 3 µL of DNA polymerase (3.8 units/µL ThermoSequenase DNA polymerase containing thermostable pyrophosphatase) to each reaction mix. There are a total of 8 reaction mixtures per class, 4 containing the forward primer and 4 containing the reverse primer.

  Note: Be careful pipetting the enzyme. It is very viscous and requires proper technique. Pipette slowly to avoid introducing bubbles. Pipette the solutions up and down slowly to mix. Change tips after each addition to prevent cross-contamination of reaction mixes.
- Prepare four 8-µL aliquots of each reaction mix in the color-coded 0.5-mL tubes. Label tubes "F" if they contain the forward primer or "R" if they contain the reverse primer. Use the following colors:

  A reaction mix = clear tube
  C reaction mix = blue tube
  G reaction mix = yellow tube
  T reaction mix = pink tube

- Store your reaction mixtures on ice or in the fridge until ready to use (they can be prepared up to 2 hours before use).
- Give each student group an A, C, G, and T reaction mix for either the forward or reverse reaction.
- If your students are temperature cycling by hand, fill the water bath to a level just below the top of the metal tube rack. Allow at least 1 hour for the water baths to reach the correct temperature.
- Determine the group number assignments for your class(es). This is important information that needs to be included in the data file name when students enter their sequence into the computer later on. If you will be sequencing with more than one class, a double-digit group number assignment is useful in creating a unique DNA template name (that is, period 1 group 2 would be assigned "12," while period 3 group 2 would be assigned "32"). When the data is returned to the University of Washington and analyzed, each template needs to have a unique name. If one is not given in class, a unique template name will be assigned before the data are posted.
- You may find it useful to divide your classroom in half, with one side performing the forward reactions and the other side performing the reverse reactions.
- Prepare student groups for manual thermal cycling by asking them to self-assign roles (cycler, timekeeper, water bath watcher) ahead of time.

Potential Volunteer Activities

- If the volunteer is present prior to class time, then he or she can help you set up water baths and set out student lab supplies.
- During the class, the volunteer can rotate among student groups and ensure that students are pipetting correctly and can answer questions.
- During cycle sequencing, the volunteer can oversee the water baths and make sure that cycling progresses in a timely fashion.

Procedure Notes

- Manual temperature cycle sequencing is a low-tech, low-cost alternative that has been adapted for use in the classroom. Scientists use a machine called a thermal cycler instead of water baths to cycle the reaction mixtures through the appropriate temperatures. Some thermal cyclers work by heating and cooling a metal block, while others work by using air convection. Regardless, this technology makes cycle sequencing much easier and more precise.
- If students are unfamiliar with the operation of the micropipettes, give them an opportunity to practice using colored water.
- Remind students that to prevent cross-contamination, they must change tips after every transfer of a reagent in or out of a reaction tube.
- Make sure that students wear gloves when carrying out DNA sequencing procedures. This is necessary to help avoid contaminating the DNA samples and because the loading mix contains formamide, which is potentially hazardous.
- Part A: Step 1. Make sure the tubes containing the reaction mixes are color coded.
- Part B: Step 1. Make sure that the water baths are at the correct temperatures and that the water levels are adequate to cover the sides of the reaction tubes (without covering their lids) when placed in the sample racks. Check the level of the 95ºC water bath partway through the cycling procedure. If more water is needed, add some immediately after moving a sample rack to the 45–50ºC water bath. If the 95ºC water bath is not up to temperature at the end of the cycle, leave the sample rack in the 70–72ºC water bath until the 95ºC temperature is reached.
- Be careful removing the lid of the sample rack after cycling because tubes can stick to it.
- After students have completed Part C, collect all of the student samples, make sure that they are properly labeled, and place them in a –20ºC freezer.

- Have students properly dispose of waste material, clean their work areas, and wash hands before leaving the laboratory.

Answers to Questions

1. Draw a diagram that shows what is happening in the reaction tubes at each of the three temperatures used during thermal cycling.

2. Describe the set of DNA fragments synthesized in each of the four reaction mixes.

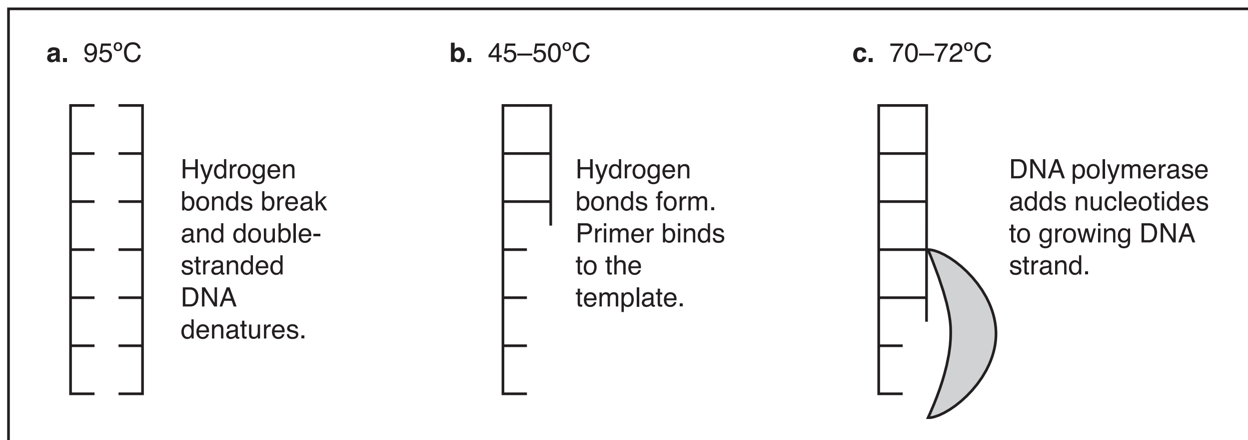


Figure TE12. Thermal cycling reactions

In the A reaction mix, a set of fragments is synthesized that begins with the primer and ends at each successive position in the sequence where an adenine base has been incorporated. Similar sets of fragments are obtained from the other three reaction mixes. Altogether, the four reaction mixes produce a series of DNA fragments with every possible length represented.

Explore/Explain
DNA Sequencing—Experimental Procedure

Day 2: Gel Electrophoresis

While scientists still use gel electrophoresis equipment such as that described here to sequence DNA, the demands of the Human Genome Project pushed the development of automated DNA sequencing techniques. Facilities that must produce large amounts of sequence data employ sophisticated technology that includes use of fluorescent DNA tags, capillary gels, laser detection, and computer analysis. In this type of setup, the reaction mixture is pumped into a thin, glass capillary that contains polyacrylamide gel. Hundreds of reactions can be electrophoresed at one time, one reaction per capillary tube. Together, the many capillary tubes required to run multiple reactions is called a capillary array. To read the DNA sequence, a laser detector positioned at the end of the capillaries reads the color of the fluorescent tags on the DNA fragments as they exit the tube. This information is entered directly into a computer for further analysis. More information on how the fluorescent tags differ from the detection methods used here is provided in "Appendix 1, Automated DNA Sequencing Protocol and Reagents".

Major Concepts

- Polyacrylamide gel electrophoresis is a method of separating DNA fragments that result from cycle sequencing in order to resolve fragments that differ in length by just a single nucleotide.
- DNA fragments are separated in the gel according to size with the smallest DNA fragments at the bottom of the gel and the largest fragments toward the top; each band in the gel is composed of multiple copies of DNA fragments that are the same size.
- DNA samples are electrophoresed under denaturing conditions to prevent the DNA molecules from forming areas of double helix that can interfere with their rate of migration.
- After the DNA fragments have been separated in the gel, they are adsorbed onto a nylon membrane, which is stained using a DNA detection system.

Materials

- The gel electrophoresis apparatus and running buffer are provided with the sequencing kit.
- You will pick up 1 prepoured sequencing gel per class 1–3 days prior to the day you will use the gel in class. Communicate with the program manager at your site regarding the days you will need a gel.
- Students will need a P-20 micropipette, timer, floating tube rack, and flat tips. They will also need an electric frying pan containing water heated to 95ºC.
- Provide each student with disposable gloves.

Preparation

- Provide each student group with a group supply kit.
- Begin setting the gel up at least 1.5 hours before loading, or if time allows, it can be set up as early as the previous evening. Always remember to wear gloves when setting up the gel.
    1. Wipe the outside of the gel with a damp paper towel to clean it. The surface of the gel may have a thin coating of urea on it (urea is one of the components of the gel mixture), and this can make the gel plates foggy.
    2. Place the gel sandwich upright in the lower buffer chamber so that the glass plate with

temperature indicator faces you. The electrode on the lower buffer chamber should be at the back of the plates.
3. Position the stabilizing bar at the front.
4. Add approximately 500–600 mL of buffer to the lower chamber. The buffer level should submerge the bottom of the plates, but rise no higher than 1 cm above the bottom of the black clamps.
5. Attach the lower electrodes to the base, making sure that they are firmly in place.
6. Fill the upper buffer chamber with approximately 1200 mL of buffer. The buffer should completely submerge the wells and rise within 1 cm of the top of the front plate.
7. Mark off the outer boundaries of the wells with a marker and remove the comb. Note, if you are setting up the gel the night before, leave the comb in place until you are ready to start running the gel. Fit the top electrodes over the gel, making sure that they are firmly connected.
8. Prerun the gel until it reaches a temperature near 45°C (about 1 hour).

Bio-Rad PowerPac 3000 Power Supply

1. Turn on the power supply and allow 10 minutes for it to warm up.

2. Turn off the power supply. Plug the electrical leads from the gel box into the power supply. Make sure everything is firmly connected. Turn the power supply back on.

3. Select Manual
   Select ConstantV
   Enter 2000
   Select Time/V-hour
   Press Time
   Enter 2:30

4. To start, press {person running icon}

5. To pause during loading, press {v in a circle icon}.

6. To stop, press {stop sign icon}.

Note: If you find that the gel is not running and all connections are firm, then please call the StarNet Project team at the University of Washington (1-877-897-6229, 206-616-2009, or 206-616-4538).

- Just prior to class, fill the frying pan with 2 to 3 cm of water and heat to 95ºC.
- Have ice on hand for cooling samples after heating.
- Make sure that you have an activity available for the students when they are not loading the gel (e.g., questions for understanding, reading the next day’s protocol).

Potential Volunteer Activities

- It is very helpful to have a scientist volunteer in the class on gel day so that one person can monitor the gel loading while the other helps students who are not loading the gel. Decide what role you want the volunteer to play.
- Whoever is monitoring the gel needs to pause and start the power supply at the appropriate times, guide students as they rinse the wells and load their samples, and call each group up to heat and ice their samples prior to loading.
- You may also want to have your volunteer stay beyond the class period to monitor the gel as it runs.

Procedure Notes

- Make sure that students wear gloves when carrying out DNA sequencing procedures. This is necessary to help avoid contaminating the DNA samples. Also, the loading mix contains formamide and the gel solution contains urea, both of which are potentially hazardous.

Part A: Preparing and Loading the Samples

- If you use a floating sample rack, you may need to secure the tubes with aluminum foil. The metal thermal cycling rack may also be used.
- Inform students that sample loading is the key to getting high-resolution gels. To achieve the best results, students must clean their four sample wells immediately before loading their samples (that is, rinse all four wells at once and then load the samples). This allows the samples to sink to the bottom of the well and sit directly on top of the acrylamide gel. Students should also load the gel as quickly as possible so that it does not cool down.
- Warn students to be careful when flushing the wells as the polyacrylamide that forms the wells is fragile and can easily be broken.
- Remind students that to prevent cross-contamination, they will need to use a clean, flat-tipped microcapillary tip every time they load a reaction mixture.
- Make sure that you or a scientist volunteer is available to supervise students when they load their samples and to start and stop the power supply. Have students leave an empty lane between their samples and those of the previous group. Do not use the outermost lanes.
- Using a waterproof felt pen, label the four lanes that each lab group uses with its group number and “A”, “C”, “G”, or “T”.

Part B: Running the Gel

- During electrophoresis, make sure that the temperature indicator on the outer gel plate reads between 45°C and 50°C.
- Remember to periodically check the level of the buffer in the upper chamber to make sure it is 1 cm from the top edge of the glass plate.
- Periodically during the run, use a marking pen to indicate the left and right edges of each set of samples. This will help students to position the membranes during the blotting procedure.

Part C: Disassembling the Gel

- You will conduct all of the steps for this portion of the procedure, as it is somewhat delicate, only requires one person to complete, and often occurs after the students have left for the day.
- Follow the detailed protocol provided in the Student Edition.

Part D: Blotting Procedure

- As this procedure usually occurs after the students have left for the day, you will probably com-

plete this portion of the experiment. Follow the protocol in the Student Edition.
- If textbooks are not available, two buffer containers of equal weight work well. Be sure to distribute the weight evenly across the Plexiglas.
- Timing is critical. Do not leave gels blotting for more than 40 minutes!
- After the blotting and crosslinking are complete, make sure that the membranes are individually placed inside folded pieces of paper, labeled with the student group names, and stored in the dark until you are ready to proceed.

Answers to Questions

1. Name two gel substances that are commonly used for electrophoresis. What are the advantages of each?

   Agarose and polyacrylamide are used to make gels for electrophoresis. Agarose gels are easy to prepare and can be used to separate DNA fragments of widely different size ranges. Polyacrylamide gels can be used for high resolution, separating DNA fragments that differ by just a single base.

2. During electrophoresis of the sequencing gel, the DNA is denatured.

   a. What does “denatured” mean?

   “Denatured” means that the DNA molecules have been unwound into single strands (the hydrogen bonds are broken).

   b. Why does the DNA need to be denatured when it is running on a sequencing gel?

   DNA must be maintained in a single-stranded form during electrophoresis. Otherwise, it may bind to template DNA or form areas of double helix that alter the fragments’ rate of migration through the gel. This can cause different fragments to run together, interfering with the reading of the DNA sequence.

   c. How is the DNA kept in the denatured form as the gel runs?

   The DNA is made single stranded before being loaded on the gel by heating to 95ºC in the presence of formamide, which is part of the loading mix. The gel solution contains 7 M urea. The urea maintains the DNA in denatured form during electrophoresis. The gel heats up during electrophoresis (to about 45ºC), and this also helps to keep the DNA denatured.

3. Why did we transfer DNA from the polyacrylamide gel to the nylon membrane?

   The DNA fragments must be taken out of the gel and attached to the surface of a nylon membrane so that the DNA is accessible to the reagents used for visualization.

Explore/Explain

DNA Sequencing—Experimental Procedure

Day 3: Visualizing the DNA

The protocol that your students are using is called a colorometric detection system because the DNA is visualized by the appearance of a colored precipitate. This is not the detection system usually used by scientists who sequence DNA. In the past, scientists favored the use of radioactively labeled nucleotides. Today, with the increasing use of automated DNA sequencing techniques, scientists use four nucleotides, each labeled with a different color fluorescent dye. During electrophoresis, the DNA fragments exit the gel and are scanned by a laser. The color of the light emitted by the fluorescent dye identifies each base. This approach allows one reaction mixture to be loaded on the gel instead of four.

Major Concepts

- DNA fragments on a membrane can be visualized using a colorometric detection system.
- The detection system used in this experiment exemplifies how molecular biologists have adapted molecular interactions as part of experimental techniques. For example, the high binding specificity of streptavidin for biotin is used in a variety of procedures. Similarly, alkaline phosphatase is used in a variety of assays because of its durability and wide substrate specificity.
- The DNA primer used in the cycle sequencing reactions was labeled with a chemical tag on its 5′ end called biotin (a vitamin molecule). Therefore, the resulting DNA fragments are all biotin labeled.
- Streptavidin (a bacterial protein) specifically binds to biotin; each streptavidin protein can bind four biotin molecules.
- Alkaline phosphatase is an enzyme that removes phosphates; some substrates for alkaline phosphatase produce a purple precipitate when a phosphate is removed. The alkaline phosphatase used in the detection system is covalently linked to a biotin molecule.
- Streptavidin serves as a bridge to connect the biotin labeled DNA fragments on the nylon membrane with biotin labeled alkaline phosphatase, which in turn, reacts with a substrate to produce a purple precipitate at the position of each biotin labeled DNA band on the membrane.

Materials

- The buffers and supplies required for day 3 are provided in the kit. Students will need group and class supply kits, washing trays, buffer solutions and detection reagents, timers, cylinders, disposable gloves, Saran Wrap, and aluminum foil.

Preparation

- Provide student groups with a group supply kit, detection tray, and a 250-mL graduated cylinder.
- At least 1 hour before beginning, check to make sure that there is no precipitate at the bottom of the blocking solution. If you find precipitate, place the entire container in a sink full of hot water and shake periodically. To avoid precipitate, store blocking solution in a warm location (approximately 25°C).
- You may want to divide the buffers into 2 aliquots and establish 2 stations in your room so students can access solutions more easily.
- Remind students that they will need to assign 2 people in their group to prepare the washes while the other 2 people shake the membrane while it is being washed.

Procedure Notes

- Make sure that students wear gloves when carrying out the washing procedures. Some of the chemicals in the washing solutions are potentially hazardous.
- Step 1. The blocking solution neutralizes the surface of the membrane that has not already bound DNA fragments. The blocking step is needed to prevent streptavidin from binding to the membrane in the next wash. If streptavidin binds nonspecifically to the membrane, then during the final wash step purple precipitate will form all over the membrane, obscuring the pattern of DNA fragments.
- If you run low on wash solution I, mix one part blocking solution with nine parts distilled water (that is, 1 mL blocking solution to 9 mL distilled water).
- If your classroom is very cool at night, you may want to move the trays to a warmer location. Alkaline phosphatase functions optimally at approximately 25°C.
- If you class will not meet within 16 hours of incubating the membranes with the color solution, you may need to add the stop solution and rinse with distilled water.
- After students have completed the washing procedures, make sure that the membranes are completely dry and place them inside plastic page protectors. Store membranes in the dark until you are ready to proceed. Be aware that the DNA bands will begin to fade if exposed to light for more than one hour.

Answers to Questions

1. What is the role of each of the following molecules in the DNA detection technique used in this experiment?

   Refer to Figure TE13.

   a. Blocking solution
   Blocking solution is used to neutralize the membrane surface and prevent nonspecific binding of streptavidin.

   b. Biotin
   Biotin helps form the top and bottom of the detection system sandwich. It is used to label the DNA primer. It is also attached to the alkaline phophatase enzyme. Both of these bind to streptavidin.

   c. Streptavidin
   Streptavidin forms the middle part of the detection system sandwich. It first binds to the biotin labeled DNA primer, then to the biotin labeled alkaline phosphatase.

   d. Alkaline phosphatase
   Alkaline phosphatase is attached to biotin. It binds to the streptavidin-biotin primer complex. When exposed to its substrate, it catalyzes a reaction that produces a purple precipitate.

   e. Color solution
   The color solution provides substrates (NBT and X-phosphate) for the alkaline phosphatase enzyme.

2. What is the advantage of using this visualization technique instead of attaching a colored dye directly to the DNA primer?

This detection system uses an alkaline phosphatase enzyme that serves to amplify the color signal.

3. What result would you expect if you forgot to use the blocking solution in the first step?

   If you forget to use blocking solution, streptavidin will bind all over the membrane, resulting in color being deposited all over the membrane, obscuring the DNA fragment pattern.

4. What result would you expect if you switched the order that streptavidin and alkaline phosphatase are added?

   If alkaline phosphatase was added before the streptavidin, then the enzyme could not bind to the DNA fragments. The resulting membrane would appear blank.

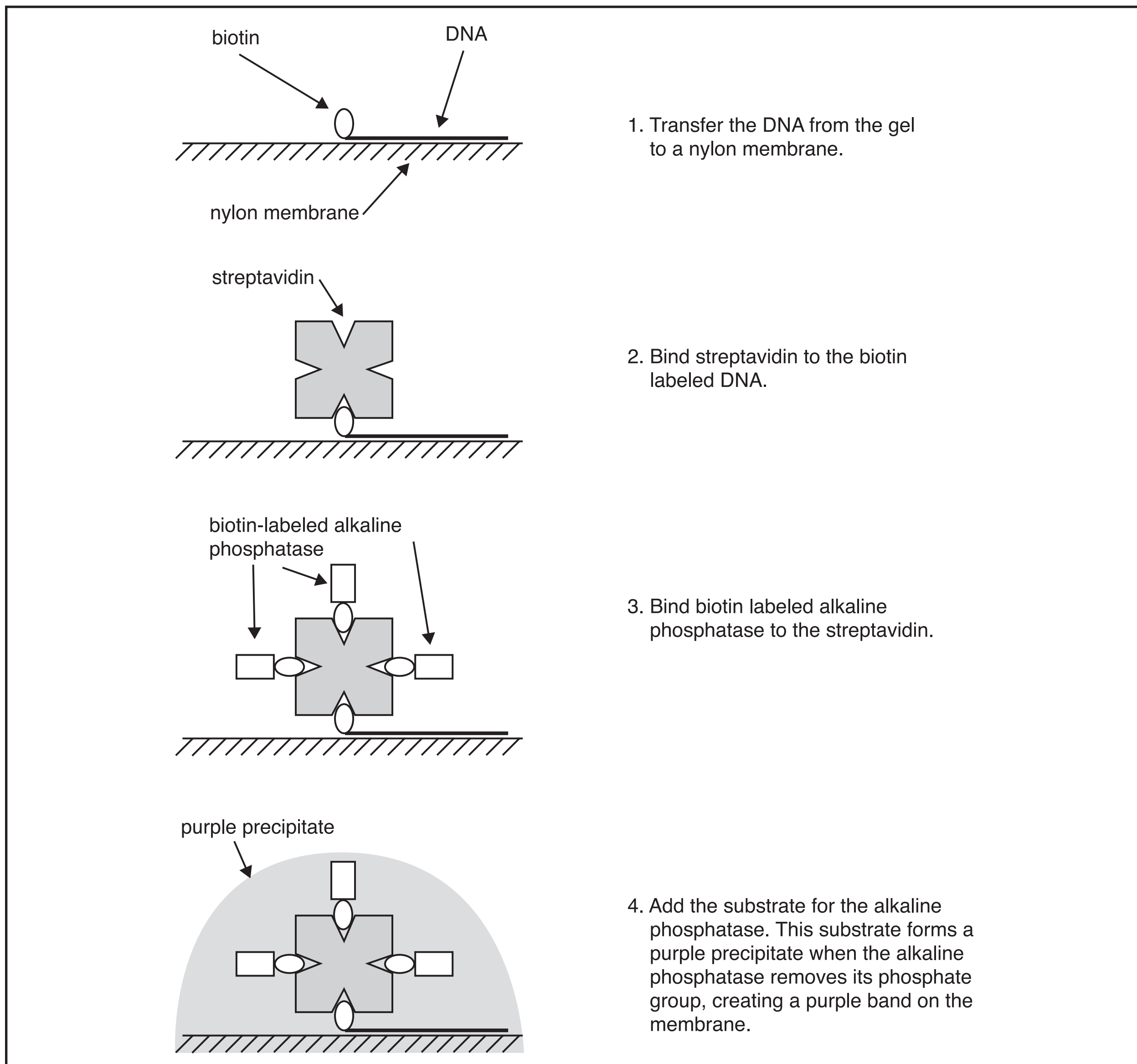


Figure TE13. Steps in visualizing the DNA

Explain

Identifying and Analyzing Single Nucleotide Polymorphisms

Identifying Variants in the *CYP2A6* Gene

Major Concepts

- DNA sequencing is useful in identifying single nucleotide changes within a population.
- A DNA template derived from genomic DNA will result in a mixture of two copies of the CYP2A6 gene, which may be the same or different alleles.
- If the DNA template used was derived from an individual who is homozygous for the CYP2A6 gene, then the DNA template will be the same for both alleles of the gene.
- If the DNA template used was derived from an individual who is heterozygous for the CYP2A6 gene, then the DNA sequence will reflect places where the two alleles differ, usually by a single nucleotide. On the sequencing ladder, this type of single nucleotide change appears as two bands with equal intensity at the same position.

Materials

- 1 copy each of Figures 35, 36, and 37 per student

Preparation

- Make photocopies of Figures 35, 36, and 37 (1 per student).

Procedure Notes

- The sequence of the primer does not show up on the membrane because sequencing occurs by adding nucleotides onto the primer. Thus, the first nucleotide that can be detected by sequencing is the first one added onto the primer.

Answers to Questions

1. Describe what the membrane will look like if there is a SNP.

   If a heterozygous SNP is present at a given base position, then two different bands of equal intensity will be seen at that base position on the membrane.

2. If you think you see a SNP on a DNA sequence using the forward primer, what do you predict you will see when you sequence the same DNA template using the reverse primer?

   If the reverse primer is used, the SNP will display a base complementary to that seen using the forward primer.

Explain
Identifying and Analyzing SNPs

Can a Single Nucleotide Change Affect the Enzyme (Protein) That Is Made?

Major Concepts

- During transcription, only one strand of the DNA double helix is copied to produce an mRNA molecule.
- A genetic code table can be used to translate an mRNA sequence into an amino acid sequence.
- A single nucleotide change in a gene can result in a different amino acid being added to the growing protein chain.

Materials

- 1 copy each of Figures 38, 39, and 40 per student

Preparation

- Make photocopies of Figures 38, 39, and 40 (1 per student).

Answers to Questions

1. Do the two *CYP2A6* alleles code for the same amino acid sequence? If not, how are they different?

   The third amino acid of the CYP2A6*1 *allele is arginine, while the third amino acid of the* CYP2A6*6 *allele is glutamine.*

Explore/Explain

DNA Sequencing—Experimental Procedure

Day 4: Data Analysis

Major Concepts

- DNA templates from genomic DNA contain sequences from both the maternal and paternal chromosomes.
- Heterozygous bases result in two different bands at the same position on the membrane.

Materials

- 1 copy each of Figures 42 and 43 per student

Preparation

- Make photocopies of Figures 42 and 43 (1 per student). Note that the photocopies of Figure 43 should be on pink paper.
- Make 1 photocopy of each student group's nylon membrane for each student in the group.

Procedure Notes

- After students have completed their analyses, please submit the class data to the StarNet Project. From each student group collect the following:
  a. the original nylon membrane, clearly labeled with the group's data file name;
  b. a clean photocopy of the membrane, also labeled with the group's data file name; and
  c. the pink data record sheet with the group's consensus sequence and labeled with the date, your school (see Figure TE14 for the two-letter school designation), class section, lab group, DNA template number, and the names of all the students in the lab group, including their signatures.
- Place the original membrane and its photocopy into a single plastic page protector.
- Make a record that lists for each class the various student groups and the names of the students within each group.
- Mail or hand deliver these items to the Genome Center as soon as possible for further analysis.

StarNet Project
University of Washington Genome Center
Box 352145
Fluke Hall, Room 225
Mason Road
Seattle, WA 98195-2145

Figure TE14. Two Letter School Designations

School	Code		School	Code
Adv. Technologies Academy	at		Montesano HS	mo
Air Academy HS	aa		Mount Hood CC	mh
Appomattox Regional GS	ap		Mountain View HS	mv
Arlington HS	ar		Mountlake Terrace HS	mt
Ballard HS	bl		Naches HS	na
Bates Technical College	ba		Nathan Hale HS	nh
Bellevue CC	bv		North Central HS	nc
Bemidji HS	be		North Kitsap HS	nk
Bethel HS	bt		North Seattle CC	ns
Broomfield HS	br		Olympic College	oc
Bush School	bu		Open HS	op
Camas HS	ca		Pasadena City College	pc
Cascade Senior HS	cc		Pierce College (CA)	pe
Central Virginia CC	cv		Pierce College (WA)	pi
Central Virginia Governor's School	cg		Pittsburg HS	pb
Centralia HS	ce		Point Grey HS	pg
Chief Sealth HS	cs		Port Angeles HS	pa
Columbia Basin College	cb		Portland CC	pl
Columbia River HS	cr		Prairie HS	pr
Curtis Senior HS	cu		Redmond HS	rm
Davis HS	da		Renton HS	re
Denver School of the Arts	sr		Richard Bland HS	rb
Dinwiddie HS	dw		Ridgefield HS	rf
Eastlake HS	el		Ridley College	rc
Eastside Catholic HS	es		Roanoke Valley GS	ro
Englewood HS	ew		Rogers HS	rg
Enumclaw HS	en		Roosevelt HS	rv
Evergreen HS (Seattle)	ev		Sacred Heart Academy	sh
Evergreen HS (Vancouver)	eg		Scarborough HS	sc
Evergreen State College	ec		St. Stephen's Episcopal School	ss
Ferris HS	fe		Franklin HS	fl
Foster HS	fo		Garfield HS	gf
Framingham HS	fr		George Washington HS	gw
Meadowdale MS	md		Gig Harbor HS	gi
Mercer Island HS	mi		Gray's Harbor College	gh

Green River CC	gr		Sheridan HS	sr
Harrisonberg HS	ha		Shorecrest HS	sc
Honaker HS	hn		Shorewood HS	sw
Horace Mann	hm		Skagit Valley College	sk
Hudson's Bay HS	hb		Skyview HS	sv
Inglemoor HS	im		Southwest Virginia CC	so
John Oliver HS	jo		Spokane Falls CC	sf
Juanita HS	ju		Spokane Skills Center	sl
Kamehameha Schools	ka		Squalicum HS	sq
Kamiak HS	km		St. Michael's University School	st
King's College	kc		Summer Institute	si
Lake Stevens HS	la		Timberline HS	tl
Lake Washington HS	lw		Vancouver Schools	vs
Lakeside School	ls		Vashon Island HS	vi
LBJ HS	lb		Virginia Tech	vt
Lewis and Clark HS	le		UW Undergraduates	ug
Los Angeles Valley College	lo		UW Upward Bound	ub
Lower Cape May Regional HS	lc		Waubonsie Valley HS	wv
Lynnwood HS	ly		West Seattle HS	ws
Marysville-Pilchuch HS	mp		Western Baptist College	wb
Sammamish HS	sa		White River HS	wr
San Mateo HS	sm		Woodinville HS	wo
Science for Success	su		Yorkhouse School	yh
Shadle Park HS	sp			

Answers to Questions

1. What are the bands on a sequencing gel?

   Each band on a sequencing gel represents many copies of a DNA fragment that ends with a base whose identity is known (A, C, G, or T).

2. Focus on the T lane as an example. Why do you get a band for each T in the DNA sequence?

   A set of DNA fragments is produced that stop at every position where a T was incorporated into the new strand. The T reaction mix contains many copies of the DNA template. This ensures that there will be multiple copies of every possible fragment that ends in a T base.

3. Think about how we read a DNA sequence from the bottom to the top of a membrane.

   a. Are we reading the newly synthesized DNA or the template strand?

      It is the newly synthesized DNA strand that is read off the membrane.

b. What direction are we reading on the DNA strand, 5′ to 3′ or 3′ to 5′?

The DNA strand is read in the 5′ to 3′ direction going from the bottom to the top of the membrane.

c. Why is it important to always read and record the sequence in the same direction?

By convention scientists record DNA sequences in the 5′ to 3′ direction, and they assume that a DNA sequence is written 5′ to 3′ unless it is clearly labeled as 3′ to 5′. Most software programs, like Sequencher, are designed to read DNA in the 5′ to 3′ direction.

4. Provide explanations for the following experimental results:

a. No bands are visible anywhere on the membrane.

1. The DNA synthesis reactions failed.
2. The products of the DNA synthesis reactions failed to attach to the nylon membrane.
3. During detection, the biotin-labelled alkaline phosphatase was added before the streptavidin, or one of these components was omitted.
4. The alkaline phosphatase enzyme was inactive.

b. No bands are visible in one lane of the membrane.

1. DNA synthesis failed in that reaction tube.
2. The DNA fragments in that lane failed to attach to the nylon membrane.
3. Not enough of that reaction mixture was loaded on the gel.

c. Two bands of equal intensity are visible at the same position on the membrane.

The DNA template comes from genomic DNA and therefore contains two alleles that differ at that base position. This is evidence of a SNP being present at this position.

d. Bands are visible at the same position across all four lanes of the membrane.

The DNA polymerase "stalled" at this position because of secondary structure in the DNA template.

e. Severe blotching on the membrane makes it hard to read.

1. The membrane may not have been properly washed with blocking solution.
2. The membrane may not have been rinsed properly with wash solution I or II.
3. The color solution may contain substrate that is old and not working properly.

5. Discuss how a membrane would be affected by each of the following scenarios:

a. A lab group forgets to heat and ice its DNA samples prior to loading the gel.

The DNA fragments would be double stranded and therefore would not enter and migrate through the gel as expected.

b. Dideoxynucleotide mixtures are mistakenly used that contain dideoxynucleotides at a much higher ratio (1:1), for instance, instead of the correct 1:9 ratio. What about a lower ratio like 1:999?

A higher ratio of dideoxy to normal nucleotides would result in the loss of some of the larger DNA fragments. A lower ratio of dideoxy to normal nucleotides would result in the loss of the smaller fragments.

c. A lab group accidentally loads both its A and C tube mixtures into the same lane on the sequencing gel.

There would be A- and C-terminated fragments (bands) in the same lane, so you couldn't tell whether a given band was an A or a C.

6. A commonly asked question is, How do we know that every possible nucleotide will be represented in our sequencing ladder? Convince yourself that there will be DNA bands at every possible position on the gel by doing the following set of calculations:

a. Calculate the maximum number of DNA molecules that can be synthesized in each reaction.

Total molecules of DNA template per reaction:
$4\ \mu L \times 0.11\ pmole/\mu L = 0.44\ pmole$

Total molecules of DNA primer per reaction:
$8\ \mu L \times 0.13\ pmole/\mu L = 1.04\ pmole$

In 10 cycles, we could theoretically make 10 x 0.44 pmoles or 4.4 pmoles of DNA. However, there are only 1.04 pmoles of primer available. This means that we can only make 1.04 pmoles of DNA molecules.

The maximum number of DNA molecules synthesized is therefore
$(1.04 \times 10^{-12}\ moles) \times (6.022 \times 10^{23}\ molecules/mole) =$
$6.26 \times 10^{11}\ molecules$

b. How many of these newly synthesized DNA molecules were loaded in each lane of the gel?

After adding 10 µL stop mix to the reaction, the volume is
$12\ \mu L + 10\ \mu L = 22\ \mu L.$

The volume of sample loaded on the gel is 5 µL.

The number of DNA molecules loaded on the gel is therefore
$(5 \div 22) \times (6.26 \times 10^{11}\ molecules) = 1.42 \times 10^{11}\ molecules.$

c. How many DNA molecules are there in each band on the gel? (Assume that there is the same number in each band and in any one reaction. The DNA template is 300 nucleotides long. DNA synthesis could terminate at approximately one-fourth of these to make 75 fragments.)

There are 1.42×10^{11} *molecules loaded per gel lane. Assuming that there are 75 DNA fragments per gel lane and the same number of molecules per band, then there are*
$1.42 \times 10^{11}\ molecules \div 75 = 1.89 \times 10^{9}$ *DNA molecules per band.*

Explain

Modeling the Assembly of DNA Fragments

Major Concepts

- A contig is a long DNA fragment that is assembled from a series of shorter overlapping fragments.
- A computer program such as Sequencher can use DNA sequence data to assemble contigs.

Materials (per student)

- Figure 45, Fragments for Contig Assembly
- 1 pair of scissors
- 1 roll of cellophane tape

Preparation

- Make photocopies of Figure 45 (1 per student).

Procedure Notes

- To save time you can cut the DNA fragments out of Figure 45 before class and provide each student with an envelope containing the fragments ready to assemble.

Elaborate

DNA Sequencing—Experimental Procedure

Day 5: Building the Big Picture: Using the BLAST Search, Assembling Contigs, Amino Acid Translation, and Analysis of SNPs

Preparation for Building the Big Picture: Making a Folder of Your Class's DNA Data Files

Preparation

- Create a folder on your classroom computer that is labeled with the class name and the date. The DNA sequences from each of the student groups will be entered into separate data files within this folder.

Procedure Notes

- Remind students to save their data as text only (*.txt). They can do so via the pull-down menu at the bottom of the save window. Sequencher does not recognize Word files.

Elaborate

BLAST Search

Major Concepts

- DNA and protein sequence data is collected worldwide in large public databases.
- BLAST (Basic Local Alignment Search Tool) is a software tool that is used to compare a DNA or amino acid sequence with those contained in a database.

Materials

- Computers with access to the Internet
- Classroom folder of DNA data files

Preparation

- Place a bookmark for the National Center for Biotechnology Information Web site (http://www.ncbi.nlm.nih.gov/BLAST) on each computer.
- Load the classroom folder of DNA data files onto each computer.

Procedure Notes

- Remind students that T is replaced by U when DNA is translated into mRNA.

Answers to Questions

1. Does your BLAST search confirm that you sequenced part of the *CYP2A6* gene? How do you know?

   Students should see that their BLAST search results include matching sequences that are identified as coming from the CYP2A6 *gene.*

2. Compare your DNA sequence with the best match from the BLAST search. How similar are they? Do you observe any differences? If so, what do you think causes these differences?

   Student answers will vary. Some students will see search results that match their sequence exactly. Others will see an occasional base change. Causes for such differences might be sequencing errors or possible SNPs.

3. As you look down the list of sequences that matched your sequence, what patterns do you observe in the types of sequences that match yours? Discuss the patterns you observe.

   Students will notice that search results that come from human sequences are very similar to their sequence. Sequences that come from animals show similarities that reflect evolutionary relatedness to humans. They also will see matches with other members of the CYP2A *subfamily (mainly* CYP2A7 *and* CYP2A13*).*

Elaborate

Assembling Contigs, Amino Acid Translation, and Analysis of SNPs

Major Concepts

- Computers are necessary for analyzing large amounts of sequence data.
- Computer programs can assemble contigs from a series of shorter, overlapping fragments.
- A consensus sequence is that sequence which best agrees with each of the individual sequences that were used in the assembly.
- Computer programs can provide additional information about DNA sequences, such as the amino acid translation.

Materials

- 1 computer per student group
- 1 copy of the Sequencher program loaded onto each student computer
- Your classroom folder of sequence data
- The files of the exons for the DNA fragments sequenced by the student groups (for example, if a student group sequenced DNA from exon 1, then they need exon 1)
- 1 copy of the CYP2A6 amino acid sequence per student group
- 1 copy of Figure 40, the Genetic Code table, per student group
- 1 good quality-photocopy of each student group's sequencing membrane

Preparation

- Download onto each student computer the most current version of the Sequencher program from http://hshgp.genome.washington.edu/teacher_resources/download.htm. Familiarize yourself with the software.
- Download onto each student computer the subject data file for the subject they have sequenced (from http://hshgp.genome.washington.edu/teacher_resources/download.htm).
- Load the classroom folder of DNA data files onto each computer.
- Make a file of each exon sequenced in your classroom. These sequences are available at http://hshgp.genome.washington.edu/teacher_resources/download.htm. Save each exon as a separate simple text file and place in the classroom data folder.

Procedure Notes

- We highly recommended that you run through the steps that use the Sequencher program before class. This will help you to troubleshoot any problems that the students encounter. Additionally, updates to the Sequencher program may require that you make slight changes in the student directions from those printed in the unit.
- To introduce students to the concept of assembling DNA data files, you may want to begin with just a few files, such as their own classroom data. Then carry out the larger assembly as described in the student unit.

Answers to Questions

1. Figure 54 shows a part of an assembly file for some of our nicotine receptor data. The column at the left lists the names of the DNA sequence files. To the right, the corresponding nucleotide sequences are lined up. The consensus sequence appears at the bottom of the screen.

   a. What do the symbols • and + below the consensus sequence mean?

   The dot denotes that one or more of the sequences does not agree with the consensus sequence at that nucleotide position. The plus sign indicates that one or more of the sequences contains an unidentifiable nucleotide denoted by an N.

   b. Do you think that we have good-quality data in this region? Explain.

   The top sequence is very different from all the others. This is why the consensus sequence shows many disagreements with the others. If this top sequence is excluded, then the data appear to be good.

   c. Suggest two things that we could do to improve the quality of the data in this region.

   1. Remove the top sequence data.
   2. Obtain additional sequence data for this region.

2. In *Part C*, why did you remove the intron sequences from your consensus sequence before carrying out the amino acid translation?

   The intron sequences are not translated into amino acids inside an organism, so they need to be removed.

3. How would the following changes in the nucleotide sequence of an exon affect amino acid sequence it encodes?

   a. Insertion of one nucleotide

      Insertion of one nucleotide would cause all of the codons after the insertion to change, so the amino acid sequence would change. You may want to point out that this is called a frameshift mutation.

   b. Deletion of two nucleotides in a row

      Deletion of two nucleotides in a row would also cause all of the codons after the deletion to change, resulting in a different amino acid sequence. This is also an example of a frameshift mutation.

   c. Insertion of three nucleotides in a row

      Insertion of three nucleotides in a row would add one new codon to the sequence and may also change the codons on both sides of the insertion. This would result in one additional amino acid in the protein chain and may change the amino acids on both sides of this additional one.

4. Complete the amino acid translation for trial 1 and 2 below, and then answer the questions that follow.

In trial 1 they tested A in Position 6 and A in Position 14

GTG	GT**A**	TTC	AGC	A**A**C	GGG	GAG
Val	*Val*	*Phe*	*Ser*	*Asn*	*Gly*	*Glu*

In Trial 2 they tested G in Position 6 and T in Position 14

GTG	GT**G**	TTC	AGC	A**T**C	GGG	GAG
Val	*Val*	*Phe*	*Ser*	*Ile*	*Gly*	*Glu*

   a. Did the SNPs at positions 6 and 14 of exon 3 result in a change in the amino acid sequence of CYP2A6?

      The two possible nucleotides at position 6, A and G, both resulted in valine (Val) being inserted. At position 14, having an A resulted in a glycine (Gly) being inserted, and having a T resulted in an asparagines (Asn) being inserted.

   b. Discuss in general terms why some SNPs result in changes in the encoded protein and others do not. Refer to the Genetic Code table to develop and support your arguments. Are certain positions in the codon (positions 1, 2, or 3) more or less likely to result in a change?

      All except one of the amino acids (methionine or Met) are coded for by more than one codon. Changing one letter in the codon sometimes changes it to another codon for the same amino acid. A SNP of this type would not cause an amino acid change. Students might also mention that not all SNPs are in exons. A SNP in an intron usually would not affect the protein because introns aren't

translated. However, if the SNP was right at the place where the intron is cut out (called the splice junction), it might affect the cutting and rejoining of exons (called splicing). Other SNPs are outside of genes, and usually don't have an effect on protein production.

Students should notice that different codons for the same amino acid often vary at the third position. For the three amino acids that have six codons, there is also variability in the first position. Thus, when a SNP occurs in the third position of a codon it often doesn't cause a change in the amino acid, and some SNPs in the first position doesn't cause a change.

Elaborate

Reading: The Role of the StarNet Project in Drug Addiction Research

What will happen to your sequencing data? After you have analyzed your DNA sequencing membrane, your teacher will submit your data to the StarNet Project staff, who will check the sequences and post them on the project web site. If you think you have found a SNP, the StarNet scientists will first check your sequencing ladders to make sure they agree with your analysis. If the data appear conclusive, then they will send the same DNA template to additional classrooms for re-sequencing. Once confirmed, the SNP will be submitted to a SNP database at the National Center for Biotechnology Information called dbSNP.

How will the sequencing data from the StarNet Project contribute to scientific research, especially the understanding of drug addiction? The SNPs discovered by students will be added to the growing list of known mutations in the *CYP2A6* gene. So far nearly 40 different genetic variants of the *CYP2A6* gene have been identified. These include single nucleotide mutations, deletions, duplications, and amplications. The amount of CYP2A6 enzyme a person makes depends on which two forms of the *CYP2A6* gene the person has inherited.

Why is there so much interest in genetic variation in the *CYP2A6* gene? One reason is that this enzyme is involved in the metabolism of a number of therapeutic drugs, either converting an inactive form of a drug to the active form or making the active form inactive. Understanding how our genes contribute to the way our bodies metabolize drugs is a new field called pharmacogenetics. In the future, physicians may check what forms of certain drug-metabolizing genes a patient has before prescribing a drug that best fits his/her needs. *CYP2A6* is also involved in the metabolism of environmental chemicals. It activates several pre-carcinogens to the carcinogenic form, including aflotoxin B1, a compound produced by a fungus that commonly infects peanuts, and several compounds found in tobacco smoke. And finally, different variants in the *CYP2A6* gene either increase or decrease nicotine metabolism, which may have an effect on people's smoking behavior.

Research and our own observations show that smokers adjust their smoking to keep nicotine at a certain level. For example, when switched to low nicotine cigarettes, smokers tend to smoke more cigarettes. Also, when smokers use a nicotine patch or are given intravenous nicotine, they smoke less. These observations led researchers to hypothesize that variation in the amount of CYP2A6 enzyme made by an individual would affect smoking behavior. Several studies have shown that people who make lower amounts of CYP2A6 are either non-smokers or they smoke less than people who make normal amounts of the CYP2A6 enzyme.

However, not all research studies support this hypothesis. What are some reasons for this conflict? In some of the earlier experiments, researchers overestimated the number of certain *CYP2A6* alleles because of technical difficulties. These experiments are being repeated using new procedures. A second reason is that different labs use different methods to measure how much people smoke. Some studies rely on self-reporting of number of cigarettes smoked, but this does not take into account how intensely the person smokes or how much of the cigarette is consumed. Other studies use additional measurements that are independent of self-reporting, including the level of carbon monoxide in the breath, and levels of nicotine and cotinine in the urine and plasma. These techniques give a more accurate measurement of how much the person has smoked. Another consideration is that other genes may also be involved in determining how much a person smokes (e.g. dopamine receptor, dopamine synthesizing and metabolizing enzymes, nicotinic acetylcholine receptor). This means that the contri-

bution from CYP2A6 may be only part of the effect. Finally, certain alleles are more common in some ethnic groups than others, so results from different labs may be skewed depending on which populations were studied.

Understanding which genes predispose individuals to become smokers or continue to smoke may help in developing treatments and cessation programs. The drug, methoxsalen, inhibits CYP2A6 in vitro. When this drug was given to smokers, they smoked less.[17] These findings are exciting because they suggest that it might be possible to help smokers quit smoking by inhibiting their CYP2A6 enzyme. More extensive studies are needed to determine whether this approach could be widely effective.

Once we know that a gene is associated with a genetic condition, it is possible to develop a DNA-based test for it. However, it is unlikely that genetic testing for the *CYP2A6* gene would be offered widely, because of the possible harm such a test could cause. For example, some people might choose to smoke based on their genotype, even though they would still be at risk for the many other diseases associated with smoking. Other people might face genetic discrimination if their genotype suggests that they are more likely to smoke or develop cancer based on their genetic profile. Additional studies are still needed to determine whether there is a correlation between the amount of CYP2A6 enzyme that a person makes, smoking behavior, and susceptibility to lung cancer.

Evaluate

Pulling It All Together

Major Concepts

- Behavioral traits are multifactorial (affected by genes and the environment).
- Drugs of addiction interfere with the process of neurotransmission.
- Addictive behaviors are reinforced by the brain's reward pathway.
- Sequence variations within a gene may or may not affect the sequence and function of the encoded protein.
- The chain termination method for sequencing DNA involves synthesis reactions, electrophoresis, and detection.
- Useful information can be obtained from DNA and amino acid sequence data using the tools of bioinformatics.
- Understanding how genes influence complex traits such as addiction can help people make rational lifestyle choices.

Materials

- Copies of case study 1 or 2 for each student

Preparation

- Have students read a case study from the student edition or alternatively provide each student with a photocopy of a case study.

Procedure Notes

- You have the option of using one or both of the case studies. Each case study is designed to encourage students to reflect on what they have learned during the unit and to prepare a brief report that summarizes that information.
- To guide the class you may provide them with the list of major concepts listed at the top of the page. Students' presentations should use information from their lab notebooks to specifically address each major concept.

Appendix 1.

Automated DNA Sequencing Protocol and Reagents

Background

Please review the section Chain Termination DNA Sequencing.

In principle, manual and automated sequencing are very similar and follow the same basic steps:

1. Synthesis of DNA fragments that are partial copies of the DNA piece being sequenced
2. Separation of DNA fragments according to size by gel electrophoresis
3. Detection of DNA fragments

As you'll see below, automated sequencing uses a fluorescent label on the DNA so that it can be detected automatically, thus enabling high throughput sequencing of many long fragments at one time. Each step is outlined below.

Part A: DNA Sequencing by Thermal Cycling

Cycle sequencing is very similar to the polymerase chain reaction. Like PCR, cycle sequencing uses a heat-stable DNA polymerase that functions at a high temperature and is resistant to near boiling temperatures for many hours. This process has two main advantages: it requires much less DNA template than other techniques, and the high temperatures help to "melt out" secondary structure in the DNA template (like hairpins) that could otherwise inhibit the DNA polymerase. It is also much easier!

First, all of the components needed for DNA synthesis are mixed in a tube:

- DNA template (the DNA molecule being sequenced)
- A DNA primer
- The four deoxynucleotides (A, C, G, and T)
- The four dideoxynucleotides (dideoxyA, dideoxyC, dideoxyG, and dideoxyT)
- Reaction buffer
- Heat-stable DNA polymerase

For each DNA template being sequenced, we need to prepare just one reaction. Each of the dideoxynucleotides has a different colored fluorescent label (A is green, C is blue, G is yellow, and T is red). This means that the DNA fragments synthesized during the sequencing reactions are color coded.

DNA synthesis is carried out by incubating the reaction at three different temperatures. DNA samples are placed into an automated heating block called a thermal cycler. The thermal cycler is programmed to quickly reach and hold the samples at the desired temperatures for as many cycles as needed. First, the reactions are heated to 95°C to break apart hydrogen bonds (that is, the base pairing is disrupted). Then the DNA is cooled to 45–55°C to allow the primer to base pair to the DNA template. Finally, the reaction tube is raised to 70–72°C, the optimum temperature for the DNA polymerase to function. After 1 minute at 70–72°C, the reaction is cycled through the same three temperatures again, allowing another round of DNA synthesis to take place. Every time the cycle is repeated, a new DNA strand is made. The DNA templates are re-used during each cycle because the 95°C incubation releases the

template strands from the newly synthesized DNA strands, making them available to be copied again. Typically, about 25 cycles are needed to ensure that enough DNA is synthesized to be analyzed. Following the last cycle, the reaction is cooled to 4°C and then stored at –20°C. Before being sequenced, the reaction is purified using ethanol precipitation which removes the unused nucleotides and polymerase molecules.

To simplify the sample preparation, the four deoxynucleotides and dideoxynucleotides are premixed with DNA polymerase and reaction buffer by the manufacturer. All you need to do is put some of the reaction mix in a color-coded tube and add the DNA template, a primer, and sterile water. A low concentration of DNA template can be used because it is recycled each time. The primer is often 50 times more concentrated than the template.

Parts B and C: DNA Separation and Detection

As in manual sequencing, the DNA fragments are separated by gel electrophoresis in a denaturing polymer. The main difference in automated sequencing is that the gels are placed inside glass capillary tubes. The arrangement of glass tubes (typically 96 are used) is called a capillary array. One sample is loaded into each glass tube. The capillary array is placed inside a light-proof container with a laser detector positioned at one end of the capillary gels. The detector is on a slider so that it can move back and forth across the width of the gel. The entire system is attached to a computer, which coordinates the loading of the reactions, running of the fragments, movement of the detector, and storing and processing of the data collected by the detector.

The polymer that separates the DNA fragments is automatically loaded into the glass tubes. Once the computer is programmed with the appropriate sample names, a loading arm takes some of each reaction and loads it into the corresponding capillary gel. After all the samples are loaded, the electrophoresis run is started.

The detection of the DNA sequence also differs from the method used during manual sequencing. In automated sequencing, the colored labels on the DNA fragments are detected by the laser detector at the end of the capillary gels. As each DNA fragment migrates past the laser beam, the detector determines the color of the fluorescent label on the fragment and sends this information to the computer. The detector samples each lane in rapid succession sending information from all 96 capillaries simultaneously to the computer. This information is integrated and processed by a sequencing software program. The sequence data for each DNA sample is presented as a four-color chromatograph.

Procedure

Check off each step as you complete it.

Pre-experiment setup

______ Write down your full DNA template name in your lab notebook.

2A6E1P1S1_F1SC_0504
E1 = exon 1
P1 = primer pair set 1
S1 = subject 1 (the individual whose DNA is being sequenced)
F = forward reaction (use R if you did the reverse reaction)
1 = lab group number (assigned by your teacher)
SC = school code
0504 = date (month and year)

Figure A1. The data file name. For two-letter school designations, see DNA Sequencing–Experimental Procedure, Day 4: Data Analysis on page 56. The DNA templates you receive are labeled with the information up to the first underscore. You will need to provide the additional information.

Part A

______ 1. Label the top of a purple or orange 0.5 mL tube with your DNA subject number and group number. Label orange tubes "F"—you will be using the forward sequencing primer. Label purple tubes with "R" and use the reverse sequencing primer.

______ 2. Put the following in the appropriate colored tube in the order listed:

Sterile water	4 µL
Sequencing Primer (Forward or Reverse)	1 µL
BDT Reaction Mix (buffer, deoxynucleotides, dideoxynucleotides, DNA polymerase)	4 µL
DNA Template	1 µL
TOTAL	10 µL

Figure A2. Setting up sequencing reactions

______ 3. Close the lids and spin in the centrifuge for a few seconds.

______ 4. Place your reaction in the thermal cycler. Note the position of your tube in case the label wipes off. When everyone's sample is in place, close the lid.

Cycle sequence, use the following program:

Hold at 95°C	3 minutes	
95°C	10 seconds	
45°C	10 seconds	25 cycles
70°C	1 minutes	
Hold at 4°C	indefinitely	

Setting Up and Using the Thermal Cycler
(MJ Research Model PTC-100/60)

A. Plug the cord from the Hot Bonnet to the back of the thermal cycler (cylindrical plug). Plug the thermal cycler into an electrical outlet and turn on power (switch on back).

B. On the display you will see SELF TEST for a few seconds and then

RUN PROGRAM   ENTER PROGRAM

C. Select RUN PROGRAM and press PROCEED.

D. The screen will say:
RUN HS1?   Press PROCEED   to indicate that this is the program you want.

E. The screen will ask whether you want to use the Hot Bonnet by asking

ENABLE   DISABLE   HEATED LID?

Select ENABLE and press PROCEED.

F. The program will now run. At the end of the run, or if you need to stop the run at any time, press STOP or CANCEL and then PROCEED.

Consult the manual if you have any other questions.

______ 5. Collect the sample with a quick spin. Check to make sure that each tube is still clearly labeled.

______ 6. Place samples from the class in a rack and put in a –20°C freezer until your visit to the University of Washington Genome Center or the sequencing center at your partner site.

______ 7. Prior to loading your samples, the StarNet staff will precipitate them to remove the remaining nucleotides and polymerase molecules.

______ 8. Place plate in sequencer and program computer.