

**Activity 1.1.3 Using DNA to Identify Pathogens**

Introduction

Samples of Sue’s blood, urine, and lymph are collected at the first infirmary visit and are sent off for diagnostic laboratory tests. As part of a pilot study, the college infirmary is working with the molecular biology department at the college to identify pathogens by their DNA sequences. The lab isolates *primers*, small segments of DNA, which attach to key genes in bacteria and viruses and allow amplification and sequencing of the DNA. Sue’s samples have been sent out for molecular testing. Little did the scientists know what they would find.

Computer advances now let researchers search through extensive sequence databases to find gene variations that could lead to disease, to track strains of various viruses or bacteria, and to design 3-D models of important human proteins. *Bioinformatics* is the field of science that combines biology, computers and information technology to store and analyze genetic data. In this course, you will observe many of the medical interventions that are made possible by the teaming of computers and biology. Bioinformatics can be used to help identify infectious agents, such as viruses and bacteria, involved in disease outbreaks.

In 1990, scientists working on the Human Genome Project embarked on a mission to map the human *genome*, the complete genetic layout of a human being. Through this project, completed in 2003, scientists have determined the complete nucleotide sequence of the DNA of each human chromosome. In addition to mapping human DNA, the Human Genome Project has helped map the genomes of other species, including mice, the fruit fly, and various bacteria. Knowing the nucleotide sequence of other living organisms, especially viruses and bacteria, allows scientists to identify and to research potential pathogens.

Modern techniques also allow scientists to sequence the DNA of living organisms and to store this information in large computer databases. DNA sequencing used to be an extremely tedious task. Computer technology, however, has automated the process. The sequence of A’s, C’s, G’s and T’s in a genome, all of the genetic information possessed by an organism, can now be visualized as fluorescent displays on a computer screen. This data is compiled and submitted to national molecular databases. This genetic information can be viewed and explored by scientists and by the general public.

Additional students are showing up at the infirmary. Time is running out and it is up to you to determine what is causing the outbreak on campus. Luckily, the laboratory on campus has finally returned DNA sequence data for patient Sue Smith. Explore the process of DNA sequencing, learn the techniques necessary to analyze DNA sequence data to identify agents of disease, and use your additional research and lab data to help diagnose Sue. The sooner you know what is making Sue ill, the sooner you can treat her, and take steps to stop the spread of disease on campus.

Equipment

* Computer with Internet access
* Activity 1.1.3 Student Resource Sheet
* Laboratory journal

Procedure

1. Obtain a Student Resource Sheet from your teacher.
2. Analyze the new patient information, as well as information about Sue Smith’s additional symptoms.

* Highlight or underline important information in the case history of each patient which appears to relate to a possible diagnosis. Be on the lookout for patient symptoms and any possible risk factors or lifestyle factors that affect a person’s chance of being infected with a particular illness.
* Use the Internet to generate a list of possible infectious agents or illnesses which may produce symptoms similar to what you are seeing in the patients.
* Begin to match information from your list with information presented in the case histories. Think about how this new information affects the conclusions you made at the end of Activity 1.1.2. Can you eliminate any possible pathogens? Have you added any new pathogens to your list?

1. Record new information in your laboratory journal and update your flow chart or graphic organizer showing student interactions.
2. Learn about using DNA sequences to identify pathogens by exploring the Howard Hughes Medical Institution Biointeractive virtual lab at <http://www.hhmi.org/biointeractive/vlabs/index.html>.
3. Click on the icon for the Bacterial Identification Lab.
4. Read the Introduction presented on the right side of the screen.
5. Copy the basic steps for isolating and analyzing pathogen DNA sequences into your laboratory journal.
6. Answer Conclusion question 1.
7. Click on the picture on the left side of the screen to enter the virtual lab.
8. Follow the directions on the screen to complete the virtual experiment.
9. Pay attention to the factual information presented on the right side of the screen. Use this information to complete the following tasks in your laboratory journal:

* Describe the limitations of traditional methods of identifying bacteria.
* Summarize the goal of each of the six parts of the lab.
* Sample Prep
* PCR Amplification
* PCR Purification
* Sequencing Prep
* DNA Sequencing
* Sequence Analysis
* Explain how fluorescent markers help determine a nucleotide sequence.

1. Answer Conclusion question 2.
2. In *Part 6 – DNA Sequence Analysis*, use the information on the screen to match the DNA sequence you have amplified to a sequence in the molecular database. You will be taken to the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) site.
3. Follow the directions provided to enter the relevant sequence data. Once you have completed the BLAST search, scroll down the results page to view the matches that are found. The descriptions of organisms or genes found at the top of the list have the most in common with the sequence that was entered.
4. Identify the most likely bacterial agent and then click on the picture of the computer screen shown on the left. Pick the appropriate multiple choice answer and see if your diagnosis is correct.
5. Compare your findings with another lab group.
6. Choose two other samples listed under the Samples tab and complete analysis as you did in step 13.
7. Click on the *Notebook* tab once you have selected a new sample.
8. Complete your analysis and list the name of the bacteria found in each of the samples you analyzed in your laboratory journal.
9. Answer conclusion question 3.
10. Visit the National Center for Biotechnology Information (NCBI) BLAST site at <http://blast.ncbi.nlm.nih.gov/Blast.cgi> to analyze the DNA sequence isolated from Sue’s lymph fluid.
11. From the menu under *Basic BLAST*, choose *nucleotide blast*.
12. Wait for the query page to open. Cut and paste the sequence data from the Student Resource Sheet into the *Enter Query Sequence* box.
13. Locate the heading *Choose Search Set*. Click on the down arrow to reveal more choices.
14. Under *Other Databases*, choose *Nucleotide Co*l*lection (nr/nt)*. Note that you can search information by amino acid sequences as well as nucleotide sequences. As you are looking strictly at DNA in this example, you will focus on nucleotides.
15. Locate the heading Program Selection. Choose somewhat similar sequences –blastn.
16. Once all the parameters are set, click the large blue BLAST button at the bottom of the page to search the genetic database for the entered sequence. This search may take a few moments.
17. Scroll down the results page to view the matches that are found. The descriptions of organisms or genes found at the top of the list have the most in common with the sequence that was entered.
18. Write down the name of the pathogen found in Sue’s system.
19. Use the Internet to search for more information on the disease agent. Use information you find to identify the disease that is affecting Sue.
20. Record all of your findings in your laboratory journal. Make sure to cite properly any sources you have used.
21. Discuss your findings with your colleagues (the other groups in the class).
22. Now that you have a diagnosis for Sue, review the patient information for all of the others in the infirmary and suggest who might be at risk for this disease. Record your observations in your laboratory journal.
23. Brainstorm strategies for how to proceed with the case in your laboratory journal.
24. Answer the remaining Conclusion questions.

Conclusion

1. How can scientists identify specific bacteria when they are amplifying and studying the same region of DNA in each species?
2. Why is PCR used in the process of DNA sequencing?
3. How can the DNA sequencing technique shown in the virtual lab be used to identify other classes of pathogens, such as viruses?
4. Explain how sequence data and information about patient symptoms led you to diagnose Sue’s illness.
5. How can DNA sequencing be used to identify genetic risk for certain diseases and disorders?