Bioinformatics: Through the Lens of COVID-19

ND EPSCoR Lesson Plan

Lesson Title: Bioinformatics: Through the Lens of COVID-19

Lesson Overview: In this lesson, the concept, technique and algorithm of bioinformatics will be introduced through a case study of COVID-19, which includes setting up a Linux operating system, downloading data and installing software in Linux commands, assembling SARS-Cov-2 genome.

Topic(s): Biology and Computer Science

Grade or Grade Band: 6-8

Lesson Objectives: After these lessons, the participants are supposed to

1) Define the concept and significance of bioinformatics
2) Develop a basic understanding of *Linux* operating system
3) Develop a basic understanding of *Linux* commands
4) Explore a de novo genome assembly

National Next Gen Standards:

- MS-LS3-1 Heredity: Inheritance and Variation of Traits: Develop and use a model to describe why structural changes to genes (mutations) located on chromosomes may affect proteins and may result in harmful, beneficial, or neutral effects to the structure and function of the organism.

North Dakota Standards:

- MS-LS3-1 Heredity: Inheritance and Variation of Traits: Develop and use a model to describe why structural changes to genes (mutations) located on chromosomes may affect proteins and may result in harmful, beneficial, or neutral effects to the structure and function of the organism.

Time Needed (estimate): Five 50 minute class periods

Lesson Author: Brittany Hagen
Dr. Brittany D. Hagen is an Associate Professor of Education and CAEP Accreditation Coordinator at Mayville State University in Mayville, ND. Dr. Hagen teaches courses related to foundations of education, educational technology, educational assessment, and elementary methods. Additionally, she has developed both online and classroom curriculums for a variety of age groups, including teach-the-teacher programs, assessment data modules, and high school aviation facilitator guides and interactive student activities. Dr. Hagen is also a proud Mayville State alumni, dedicated to developing highly effective teachers who share a passion for educating young learners.

Scientist/K12 Collaborator & University: Lu Liu, PhD, Department of Computer Science, North Dakota State University (email: lu.liu.2@ndsu.edu)

Scientist Bio/Research: Dr. Lu Liu holds a Ph.D. in Computer Science from The University of Texas at San Antonio (2016), M.S. and B.S. in Computer Science from Beijing University of Posts & Telecommunications (2011 and 2008 respectively). His research interests lie in the broad area of bioinformatics, data mining, machine learning and computational biology. More specifically, his current research focuses on two areas: (1) machine learning methods for modeling and analyzing biological networks; (2) data mining approaches to biomarker discovery and disease prediction via integrated analysis of heterogonous "omics" data.

Hello, my name is Lu Liu and I am an assistant professor at North Dakota State University where I teach computer science. Most of my students want to become programmers, software engineers, hardware engineers, system administrators, network administrators, database administrators, product managers, system architects, UX/UI designers, data scientists and bioinformaticians. Some want to become professors. I grew up in a rural area and we did not have access to a computer because computers were expensive. When I went to university, my elder sister picked computer science as my major because I was confused about so many majors and not sure which one to pick and she saw the opportunities in this subject. The Freshman Year of university was really a struggle for me because I found out my classmates from urban areas had learned programming in high school while I didn't even know how to start a computer. I felt programming was hard because I met a lot of unexpected problems when I started to learn programming. Most of the time, my program crashed. But I just hung in there. I asked for help from my classmates and teachers and searched for answers on the Internet. Until my Senior Year in university I realized that picking computer science was the right choice because during the study I not only mastered the skills of self-learning with the Internet resources but also entered into a field with abundant job opportunities. After I obtained my bachelor's degree, I decided to get a master's degree because I felt I was really getting into the subject. After I obtained my master's degree, I did some internships for a job and decided to pursue a doctorate degree in computer science in the USA. After I obtained my doctorate, I found a faculty job because I preferred to stay in academia. I started my own lab and taught machine learning, data mining and bioinformatics classes. My lab moved to North Dakota in 2018 and my current research lies in applying computer science technologies in analyzing big biological data such as genomic data.

Preparation/Materials

Background knowledge students must have to be successful: Basic understanding in DNA and how each species has their own genome and that has been a mission of scientists to understand the genome. For this lesson, students will need a background in computer operation.

Differentiation and accommodation to support learning for all students:
When designing any lesson, it is important to address the needs of all learners. Please refer to the following resource for ideas on how to adjust your lesson to accommodate your students' particular learning needs: https://www.understood.org/en/learning-thinking-differences/treatments-approaches/educational-strategies/common-classroom-accommodations-and-modifications

Essential Terminology:

- Bioinformatics- interdisciplinary field (such as biology, computer science, information engineering, mathematics and statistics) that develops methods and software tools for understanding biological data.
- DNA- a self-replicating material which is present in nearly all living organisms as the main constituent of chromosomes. It is the carrier of genetic information.
- Nucleotide- a compound consisting of a nucleoside linked to a phosphate group. Nucleotides form the basic structural unit of nucleic acids such as DNA.
- Operating System- is system software that manages computer hardware, software resources, and provides common services for computer programs.
- Genome - the complete set of genes or genetic material present in a cell or organism.
- Genome assembly- aligning and merging fragments from a longer DNA sequence in order to reconstruct the original sequence

Resources:

- Ubuntu walk through
- What is Bioinformatics video
- EPSCoR Website
- Gizmo

Websites:

- https://ubuntu.com/tutorials/create-a-usb-stick-on-ubuntu#1-overview
- https://www.youtube.com/watch?v=K9MqyU298uE
- https://www.ncbi.nlm.nih.gov/genbank/sars-cov-2-seqs/
- https://www.ndepscor.ndus.edu/ndep/nature/sunday-academy/stem-module-topics/
- https://www.explorelearning.com/

Materials needed:

- USB sticks- OPTIONAL: all downloaded and set up (Lesson 2/Activity 1)
- Jigsaw puzzles for each group
- Computer
- Projector
- Bubbles or Nerf balls
- Spoons for each student
- Ping pong balls

PowerPoint – found as separate attachment

Procedure/Activities

Lesson 1: Disease Spread Simulation (50 minutes)

Engage:

1. To engage students in learning about how germs are passed during an outbreak, complete Pre-Lesson Activities 1, 2, and 3. Each of these activities, including materials and step-by-step instructions, can be found in the accompanying documentation.

Explore:

2. Using the Disease Spread Lab simulation worksheet, have students complete the simulation using the Gizmo software. More information can be found on the worksheet. This activity should take about 20 minutes and can be accessed here with a free account: https://www.explorelearning.com/index.cfm?method=cResource.dspDetail&ResourceID=379

Explain:

3. Review the Disease Spread Lab Simulation results as a whole group. Review each step to help explain how the diseases spread (predict, identify, experiment, interpret, experiment, interpret, analyze, apply).

Extensions for learning more about this topic:

4. There are several extensions of the Gizmo Disease Spread Lab simulation so you may consider exploring those to further extend this lesson for students.

Evaluation of learning (formative or summative task)

5. Ask students to come up to the board and write one thing they learned about the spread of diseases. Work together to highlight common ideas and discuss key points you want students to remember. Also use this opportunity to clear up misconceptions.

Lesson 2: Setting up the Game (50 minutes)

Engage:

1. There are numerous options for this lesson for the teacher to find what works best for their class. Have groups made prior as teacher knows students best. Try and pair up students who are strong with computers with those who aren't.
2. To introduce students to the activity, share information about the game and the case study (PPT Slide 2). Information about each game and materials needed found in the "Resources" portion of this document.

Explore:

3. Review PPT Slide 3 to discuss the rules of the game.
4. Hand out the Pre-Lesson Assessment questions to partners or groups you have assigned. Have students work through the questions together. Provide guidance as necessary. The answer key is located after the student copy in the attached document. Once students or groups have completed the Pre-Lesson Assessment, discuss the answers. The activity is not graded, it's designed to help introduce the topic**.**

Explain:

5. To provide context and support student learning, review important vocabulary, including: nucleotide, DNA, genome, human genome, and bioinformatics (PPT Slides 4-9). Be sure to show students the video on slide 8 to develop students' understanding of bioinformatics.

Extensions for learning more about this topic:

6. To further extend the lesson, have students discuss a list of bioinformatics applications. They can discuss and list them in small groups and then discuss as a whole group. Share the applications of bioinformatics found on PPT Slide 10.
7. Review the notes within the PPT Slide 10 for more applications of how bioinformatics are used in today's world. Notes include:
    a. disease diagnosis, personalized treatment
    b. food fermentation (pickle, beer, cheese)
    c. forensic (Golden State Killer)
    d. select plants with desired properties (high yield, nutrition, insect resistance, environment friendly, use less water)

Evaluation of learning (formative or summative task)

8. Ask students to list aloud one thing they learned about bioinformatics in this lesson. Share responses with the whole group, highlighting important information and clearing up misconceptions.

Lesson 3: Popular Software (50 minutes)

For middle schoolers, it is recommended that the instructor sets up the USB, Activity 1, before hand so it is ready to go if you have a group who is not strong with computers or the time for it. It will allow students to jump right into the lesson.

Engage:

1. Display the "Popular software" graphic on PPT Slide 11. Ask students to identify any of the softwares they recognize.

Explore:

2. Review the list of popular softwares and ask students how those are related to bioinformatics that were discussed in the first lesson.

Explain:

3. Using PPT Slides 12-16, review vocabulary related to bioinformatics, including: computing resources, computer hardware, and operating systems. Make students aware that there are different forms of computers including supercomputers, desktops, laptops, smart phones, etc.

Extensions for learning more about this topic:

4. Using PPT Slide 17 and Activity 1 Sheet, have students create a bootable USB stick. The following website has step by step instructions for the students to follow: https://ubuntu.com/tutorials/tutorial-create-a-usb-stick-on-windows#1-overview. The goal of this activity is to create a bootable USB stick which will take about 30-40 minutes. Creating the bootable stick will prepare students for the following lessons.
5. The teacher can choose to share the activity 1 sheet with students and display PPT Slide 17 to help them along. Once they have the USB stick ready to go, review PPT Slides 18-19 with students. This will help them understand the purpose of the terminal.
6. Then, have them practice Linux commands (found on PPT Slides 20-23 and Activity Sheet 1).

Evaluation of learning (formative or summative task)

7. Allow students to practice creating the Linux commands for as long as time allows. When they are done practicing, gather them together as a whole group and discuss the top three things they learned through creating the bootable USB stick and practicing the Linux commands. Preview tomorrow's lesson by explaiind to students that they will be learning to download data and install bioinformatics tools.

Lesson 4: Downloading data and installing bioinformatics (50 minutes)

Engage:

1. Pass out a notecard or sticky note to each student. Ask them to write down one thing they remember from the previous lesson. Have them crumple their notes and throw them in a pile at the front of the room. Once all ideas are collected in a pile, select a few of them and read aloud to students. Use this discussion to highlight key points and clear up misconceptions.

Explore:

2. In this lesson, the teacher and students will use Linux commands to download SARS-Cov-2 sequence data and install bioinformatics tools, such as SRA toolkit.
3. To review Linux commands from the previous lesson, review Activity Sheet 1 and PPT Slides 20-23.

Explain:

4. Using PPT Slides 24-26 and Activity 2 Sheet, have students download data and install bioinformatics tools.

Extensions for learning more about this topic:

5. The Activity 2 Sheet contains an extension activity for those students who finish early or who are ready for an additional challeng.

Evaluation of learning (formative or summative task)

6. Allow students to complete the Activity 2 Sheet. Review results as a class to check for understanding.

Lesson 5: Playing a Jigsaw Puzzle (50 minutes)

Engage:

1. Cut up a blank piece of paper into puzzle piece shapes. Cut the paper into enough pieces so that every student gets one and there are a few extras. Withhold the few extra pieces so the puzzle will never be complete. Hand out one puzzle piece per student and ask them to assemble the puzzle as a group.
2. After they've had time to put the puzzle together, ask students, "what happens when you don't have a picture to follow and not all the pieces are here?" Students may respond with, "you can't complete the puzzle".
3. Discuss the importance of having all the pieces to the "genome puzzle" and explain that students will be discovering genomes during the activity today.

Explore:

4. Ask students to think back to the previous four lessons and recall one vocabulary term they learned. If not stated, highlight key vocabulary for this lesson: DNA and genome.
5. Review PPT Slides 27 – 30 and Activity Sheet 3 to complete the "playing a jigsaw puzzle" activity.

Explain:

6. Allow students to explore the SARS-CoV-2 Genome website on PPT Slide 30 on their own. Discuss key things they discover on the website and how that relates to genomes and bioinformatics.

Extensions for learning more about this topic:

7. To extend the lesson, have students research about different diseases or pandemics in the past. Upon completion of their research, they can create a presentation to share with the class. The overarching questions asked of all students should be:
    a. What about the viruses or bacteria make them so deadly?
    b. What do we know about their genome?
    c. What is done to treat and prevent these diseases?
8. To conclude the lesson, review the main points found on PPT Slide 31.

Evaluation of learning (formative or summative task)

9. As an evaluation of learning, review the three discussion questions at the bottom of Activity 3 Sheet as a class. Highlight key information and clear up misconceptions as necessary.

Additional Lesson Resources / Materials

References:

Liu, L. "Bioinformatics through the lens of COVID19". 29 July, 2020.

Saskatchewan Science Centre. "Classroom Activity- Disease Transmission". 29 July, 2020. https://static1.squarespace.com/static/563a8427e4b02d05f44d829d/t/564ce8d2e4b0e4c59118f59a/1447880914071/classroom+activity%3B+science+1+-4+Heath+disease+transmission.pdf

Websites for purchasing materials

- 16 GB USB Stick
    - https://www.amazon.com/SanDisk-Flash-Cruzer-Glide-SDCZ60-016G-B35/dp/B007YX9O9O/ref=sr_1_3?dchild=1&keywords=usb+stick&qid=1596135559&sr=8-3

If need general supplies:

- Nasco (https://www.enasco.com/c/Education-Supplies/Science)
- Flinn (https://www.flinnsci.com/)
- Carolina (https://www.carolina.com/lab-supplies-and-equipment/science-lab-supplies/science-lab-classroom-supplies/10300.ct)
- School Specialty (https://www.schoolspecialty.com/science-supplies-and-products)
- Amazon (www.amazon.com)