DNA decoded

ILLUMINA VIRTUAL GENOMIC SEQUENCING LAB

Genomic Sequencing: The Microbiome and Migration Mysteries

Project Details

GRADE RANGE

9–12

TIMING

4–5, 50 minutes class sessions

ACTIVITY SUMMARY

- In Illumina's Genomic Sequencing Virtual Laboratory, students will begin by identifying what genomic sequencing is and explore some of the many important and exciting fields that genomic sequencing is transforming, such as microbiology, agriculture, conservation, and human health.
- In the virtual lab, they will practice the techniques and the procedure that is required for genomic sequencing. In the virtual lab, they will practice the techniques and the procedure that is required for preparing DNA to be sequenced. They will see what is happening to segments of DNA at each phase of the process in a detailed animation. Then, they will practice using virtual lab equipment and proper lab techniques to amplify, mix, and prepare samples of DNA for whole-genome sequencing.
- Students will summarize what happens to the DNA at each stage of the protocol and research and discuss why each is an important step in genomic sequencing.
- As the results of the sequencing are revealed, students will be asked to identify ways that genomic sequencing is revealing important clues about our ancient human history and predict how the information gained from the genomic sequencing of our ancestors can help us to learn more about ourselves and create solutions to some of the human health issues we are facing today.

NEXT GENERATION SCIENCE STANDARDS (NGSS)

- **HS-LS1-1 From Molecules to Organisms: Structures and Processes**
    - Construct an explanation based on evidence for how the structure of DNA determines the structure of proteins which carry out the essential functions of life through systems of specialized cells.
- **HS-LS4-1 Biological Evolution: Unity and Diversity**
    - Communicate scientific information that common ancestry and biological evolution are supported by multiple lines of empirical evidence.
- **HS-ESS3-1 Earth and Human Activity**
    - Construct an explanation based on evidence for how the availability of natural resources, occurrence of natural hazards, and changes in climate have influenced human activity.
- **CTE Health Standards**
    - 4.3 Use information and communication technologies to synthesize, summarize, compare, and contrast information from multiple sources.
    - 4.5 Research past, present, and projected technological advances as they impact a particular pathway.
    - 9.7 Participate in interactive teamwork to solve real Health Science and Medical Technology sector issues and problems.
    - A1.1 Use data to explain how biotechnology fields such as pharmaceuticals, agriculture, diagnostics, industrial products, instrumentation, and research and development are impacting human life.

TEACHER NOTES:

- This virtual lab fits best into a biotechnology unit in a biology class or at the end of Unit 6: Gene Expression and Regulation of an AP Biology course.
- Prior to beginning the virtual lab, it is recommended that students have knowledge of cell types (prokaryotes), DNA structure, and polymerase chain reaction (PCR).
- Illumina's DNA Decoded "Microbiome and Mental Health" digital lesson is recommended for teachers to use prior to the virtual lab as a way to introduce the microbiome to students.
- Prior to beginning this activity, educators are invited to explore the other resources, lessons, and digital lesson bundles as well as an overview of genomics written for educators found on the DNA Decoded website.

- Suggested activities and lesson bundles for students who relate to topics surrounding human health and genetics include:
    - Genes or the Gym?
    - Race and Medicine
    - Medicine of the Future
    - The Microbiome and Mental Health
    - Targeting Cancer

INTRODUCTION AND PRE-LAB ACTIVITY—BACKGROUND ON GENOMIC SEQUENCING AND UNKNOWN DNA (30 MINUTES)

1. To begin the lesson, ask students to recall what a "genome" is and allow them to share their ideas with the class. Students should recall that the genome of an organism is the entirety of its genetic material. The instructor can clarify if necessary.
2. Explain to students that before they can begin the virtual lab, they will complete a pre-lab activity where they will learn about genome sequencing; the method of determining the order of letters or bases in the DNA.
3. Give each student a copy of the "Genome Sequencing Research Capture Sheet." Students should work individually using their student devices (laptop, iPad) and the links provided to answer the questions in each section of Part 1 of the research sheet. Students should have 15–20 minutes to complete their individual research.
4. Explain that in Part 2 of their research, they will each be assigned to a group of 3–4 students and given one of five fields of study that have been impacted by genomic sequencing. Each group will spend 20–25 minutes collaborating and using the website provided on their capture sheet to research and discover things that have been learned in those fields of study.
5. Assign each student group to one of the following fields of study:
    a. Agriculture
    b. Conservation
    c. Microbiology
    d. Cancer
    e. Human Health and Rare Diseases

6. Ask students to find their group. Explain that they should use the links provided on the capture sheet as they work together to research discoveries that have been made in their area of study through work in genomic sequencing. Students should record their thoughts on their capture sheet and then work with their group to record the group's findings on a piece of butcher block paper or poster board provided to the group to create a poster. Each group should be given 20 minutes to complete this.
7. Once the time is up, each group should choose a reporter who will share with the whole class what they have discovered in their research time using their poster as a visual aid. After each group has had a chance to report how genomic sequencing has impacted their field of study, students should answer the final question on the capture sheet: "Record your thoughts after hearing the other responses. What do you think are the most important or impactful things we've learned through genomic sequencing?"
8. If time allows, students can share their answers to the final question with the class, including what might have surprised them about all we can learn from genomic sequencing.

PRE-LAB DISCUSSION AND QUESTIONS

9. Next, inform students that in the Ilumina Virtual Lab, they will be playing the role of scientists who have been given the task of sequencing an unknown bacterial DNA sample to see what secrets it might reveal about human history.
10. To introduce the problem that students will be exploring, show the video clip "This 5,300-Year-Old Corpse Was Found by Accident" to the class, stopping at 2:02.

    ***Video Disclaimer (please read to students before showing the video):** "The video clip you will be watching contains graphic images and discussion of a human corpse. This may cause feelings of discomfort in viewers, and you should notify your teacher if you prefer not to view it. If you prefer to read a written summary of the relevant scientific and narrative facts instead, please ask for a copy of 'The Unsolved Case of Otzi the Iceman' article to read instead."*
11. Explain to students that one of the most amazing things about Otzi is that scientists have been able to extract and sequence ancient genetic material from the body, which has provided clues about his physical features, what he ate, and most recently, his microbiome. Provide students with a link to or a copy of the article "Paleomicrobiology and Microbial Ancient DNA Get to the Root of Disease Mysteries" and ask students to read the article.
12. Tell students that bacterium in Otzi's stomach has been recovered, and it is their job as a lab technician in a genomic sequencing lab to sequence the bacterial sample and use a genetic database to try to identify what type of bacteria it is by comparing it to known bacterial genomes.

13. Before they begin, give each student a copy of the "Background Reading: Analyzing the Results of Next-Generation Sequencing" capture sheet. Students should read through the information on the sheet individually and use the provided links to watch the video and explore the Krona charts.
14. Project the following pre-lab questions for students on the overhead screen and engage in a discussion before students begin the virtual lab.

Discussion Questions:

A. Based on what you learned in the "Background Reading: Analyzing the Results of Next-Generation Sequencing" capture sheet, create a sentence that explains the relationship between the following terms:

- GENOMIC MAPPING
- READS
- QUERY SEQUENCE
- REFERENCE GENOME

Discuss your thoughts with a person sitting beside you and then share your sentence with the instructor and the class.

B. BRAINSTORM: Knowing that the microbiome is an important part of human health, why might scientists be interested to see what type of bacteria could be found in the human gut 5,000 years ago?

C. Do you think that the bacteria found in the gut of modern-day humans is the same throughout the world? Why or why not?

D. Look at the eight steps in the procedure on the capture sheet. Predict what you will be doing to the DNA sample before sequencing. Can you recall what PCR is?

15. Students should follow the instructions in the virtual lab to complete each of the three levels while pausing after each level to record information on their summary sheet and discuss their answers to the end-of-level check-in questions (see at the end of each level educator instructions).
16. Students can work individually or in pairs to complete the virtual lab. Give each student or pair a copy of the Genomic Lab Report Capture Sheet.

VIRTUAL LABORATORY—EDUCATOR BACKGROUND TIPS

Students should go to the following link to begin the virtual lab: https://www.illumina.com/content/dam/illumina-marketing/apps/dnaday/index.html

1. Remind students to read the instructions and information on the screen as they move through each phase of sequencing. Clarify for students that the process they are going through is called Library Prep. It prepares the DNA to be loaded onto the machine where it is sequenced.
2. It may be helpful to display the virtual lab on the overhead screen. Explain to students that they will first view an animation of what will be happening in the sequencing machine for that level of the lab before going through the lab procedure. This is where they will record information on the "Genomics Laboratory Report" capture sheet.
3. Students can pause the animation to write down information and can mute the sound if they choose.
4. Once students begin the lab procedure, they should click on "HINTS" on the bottom menu of the screen, which will help them navigate through the procedure.
5. It is recommended that the teacher complete the sequencing virtual lab on their own prior to assigning it to students to familiarize themselves with the lab and to help troubleshoot and answer technical questions for students. The teacher can also skip through various parts of the lab using the menu at the bottom of the screen.

LEVEL 1 (45 MINUTES)

DNA Library Prep

OVERVIEW:

- **A.** Students will travel inside the thermal cycler to view an animation that will allow them to view how DNA is cut, tagged, washed, and how polymerase chain reaction (PCR) is used to make copies of the DNA.
- **B.** After watching the animation, students will follow provided instructions to sit at their virtual worktable and mark two PCR tubes to be used in this part of the procedure.
- **C.** Next, they will select and set a pipette at 10.0µL. They will pipette DNA samples into tubes along with water, bead-linked transposomes (BLT), and BLT Master Mix (BLT MM).
- **D.** They will mix the tubes and then put them into a thermal cycler.
- **E.** Finally, they will connect the thermal cycler to a tablet, set up their protocol, and run the thermal cycler.
- **F.** In the thermal cycle, the DNA will be cut into fragments by enzymes. The DNA fragments will attach to bead-linked transposomes and are tagged; once this is achieved, tagmentation will end as a stop buffer is added, and the enzymes used for tagmentation are removed.

G. Next, the sample is returned to the thermal cycler, and primers are added to start the polymerase chain reaction (PCR) where many copies of the DNA fragments will be made using primers (short sequences of nucleic acids) and enzymes.

Ask students to **PAUSE** when they have completed Level 1—DNA Library Prep.

Note: Inform students that they will be asked to enter the teacher's email as they finish each level to confirm that they have completed the activity and that they should not skip this step!

Present the following end-of-level check-in questions:

1. **What is happening when the DNA is going through "tagmentation" at this level?**
2. **Why is PCR used in genomic sequencing?**
3. **Would it be especially important with a DNA sample like the one we are sequencing? Why?**

Ask students to share their ideas and answers to the questions with the whole group.

LEVEL 2 (45–60 MINUTES)

DNA Library Cleanup

OVERVIEW:

A. Students will again travel inside the thermal cycler to view an animation that will allow them to view how large DNA fragments are separated from the supernatant containing small and medium fragments using a magnet. Then, the medium fragments, which work best for sequencing, are attached to BLTs, and the remaining small DNA fragments are separated from them. Finally, the PCR reagents are removed with the addition of ethanol and a resuspension buffer (RSB), and magnets are used to remove the beads, leaving the DNA ready for sequencing.

B. After watching the animation, students will continue to follow provided instructions at their virtual worktable and tap the two PCR tubes on the table to get rid of any bubbles before placing them on a magnetic PCR rack for 5 minutes.

C. Next, they will select and set a pipette at 045µL. They will transfer the supernatant from the DNA samples into newly marked PCR tubes.

D. Then they will set the pipette to 200µL and pipette up and down 5 times to mix the diluted sample purification beads (DSPB).

E. After setting the pipette according to the instructions, they will transfer DSPB to each of the tubes and then pipette up and down to mix each of the samples and let them sit for 5 minutes.

F. The PCR tubes will be placed on the magnet for an additional 5 minutes, and the supernatant will be transferred to new tubes that will be placed into the thermal cycler.

G. Using a pipette set at 015µL, they will mix the sample purification beads (SPB) and then transfer 15µL into the PCR tube.

H. A pipette will be used to mix the samples, and they will sit for 5 minutes before being placed on the magnet PCR rack for another 5 minutes.

I. The supernatant from each PCR tube will be discarded, and the tubes will be placed in the thermal cycler.

J. Once out of the thermal cycler, ethanol will be added to each tube, and they will sit for 30 seconds.

K. Ethanol will be removed and discarded from each tube, and tubes will be moved to the thermal cycler where PCR reagents will be removed from the DNA and beads.

L. The tubes will leave the thermal cycler. A resuspension buffer (RSB) will be added to each of the tubes with a pipette and will be mixed to resuspend the pellet and sit for 2 minutes.

M. The mixed tubes will be placed on the magnet PCR rack for 2 minutes, and two new tubes will be marked before transferring the supernatant to the new PCR tubes.

Ask students to **PAUSE** and enter the teacher's email address when they have completed Level 2—DNA Library Cleanup. Present the following end-of-level check-in questions:

1. **Now that you have completed the experience, why do you think this level is called DNA Library Cleanup? What were you trying to isolate from the DNA needed for sequencing?**
2. **Why do we remove the large and small fragments?**
3. **What function, or role, does the magnetic tube rack serve in this process?**

LEVEL 3 (45–60 MINUTES)

DNA Sequencing

OVERVIEW:

A. Students will begin this level by traveling across the lab from their workspace to the sequencing machine.

B. They will mix the reagent cartridge 10 times before using the pipette to poke a hole in the foil of the cartridge.

C. Next, they will place the DNA sample in the library well of the cartridge and insert the flow cell into the cartridge.

D. They will place the cartridge in the sequencing machine and add the information provided to run the sequence before clicking "Start Run."

E. Once the procedure is complete, they will eject the cartridge to complete the lab!

Ask students to **PAUSE** and enter the teacher's email address when they have completed Level 3—DNA Sequencing. Present the following end-of-level check-in questions:

1. **What is a genomic library? (Click on the link: GENE LIBRARY to find the answer if you don't know!)**
2. **What might be some advantages of using technology such as a sequencer to help determine the origin of unknown DNA?**

17. Once students have completed all phases of the virtual lab, ask them to get into small groups and give them a copy of the "DNA Sequencing Analysis" sheet. Ask students to study the results and diagrams. (Teachers may also provide students with the published paper "The 5,300-year-old *Helicobacter pylori* genome of the Iceman" as a pdf as additional information if they choose.)
18. Groups should discuss the findings and read the linked article "Iceman's H. pylori Genome Hints at Ancient Migrations to Europe." They should complete the final two sections on the "DNA Sequencing Analysis" sheet and discuss their answer to the significance of the findings.
19. To conclude the activity, students will be given the question "Can we learn things about modern human health from sequencing ancient human genomes?" The instructor should give each student a copy of the Claim, Support, Question Capture Sheet.
20. In class, or as an out-of-class assignment, students will use the links provided on the capture sheet to answer the topic question (their claim), provide support from the articles and video for their claim, and think about what further questions they have about what we have and are yet to learn about ourselves and our health from genomic sequencing of ancient DNA.

PART 1—INDIVIDUAL RESEARCH

SECTION 1: Types of genomic sequencing

LINK: https://www.yourgenome.org/facts/types-of-genome-sequencing

Define the following:

- Whole Genome Sequencing: _____
    - DeNovo Sequencing: _____
    - Resequencing: _____
- Targeted Genome Sequencing: _____
    - Exome Pulldown: _____

SECTION 2: Timeline: Organisms that have had their genomes sequenced

LINK: https://www.yourgenome.org/facts/timeline-organisms-that-have-had-their-genomes-sequenced

Complete the table below:

Organism	Year sequenced	How many bases?	How many chromosomes?
Bacteriophage MS2			
Caenorhabditis elegans (Nematode worm)			
Drosophila melanogaster (fruit fly)			
Homo sapien (Human)			

SECTION 3: Next-generation sequencing

LINK: https://www.yourgenome.org/stories/next-generation-sequencing

- What is next-generation sequencing? _____

  _____

  _____
- What are some ways that next-generation sequencing improved on traditional methods of DNA sequencing? _____

- The human genome project took 10 years and cost nearly $3 billion using capillary sequencing. How has next-generation sequencing changed this?

PART 2—SMALL-GROUP RESEARCH

- What have we learned from genomic sequencing?

  LINK: https://www.illumina.com/areas-of-interest.html for your research!

- My group's assigned topic:

- Things we have learned/breakthroughs in my topic:

- Record your thoughts after hearing the other responses. What do you think are the most important or impactful things we've learned through genomic sequencing?

Next-generation sequencing provides DNA sequences from small DNA fragments called **reads**.

Most next-generation sequencing technologies fragment the genome prior to sequencing, and each sequenced fragment produces a read. The length of the read and how many are produced will depend on fragment size and the type of technology being used. As the fragments of DNA usually overlap, the reads can be pieced back together to reconstruct the genome. There are other next-generation sequencing methods that do not fragment the genome; these methods are called long-read sequencing as they produce very long reads.

The reads from the DNA sample are then compared to known genomes of organisms (**reference genomes**). In mapping, the goal is to align each read in the set to the reference genome, allowing mismatches, insertions and deletion of bases, and clipping of short fragments on the ends of the reads.

Study the figure below to see a simplified explanation of genome mapping.

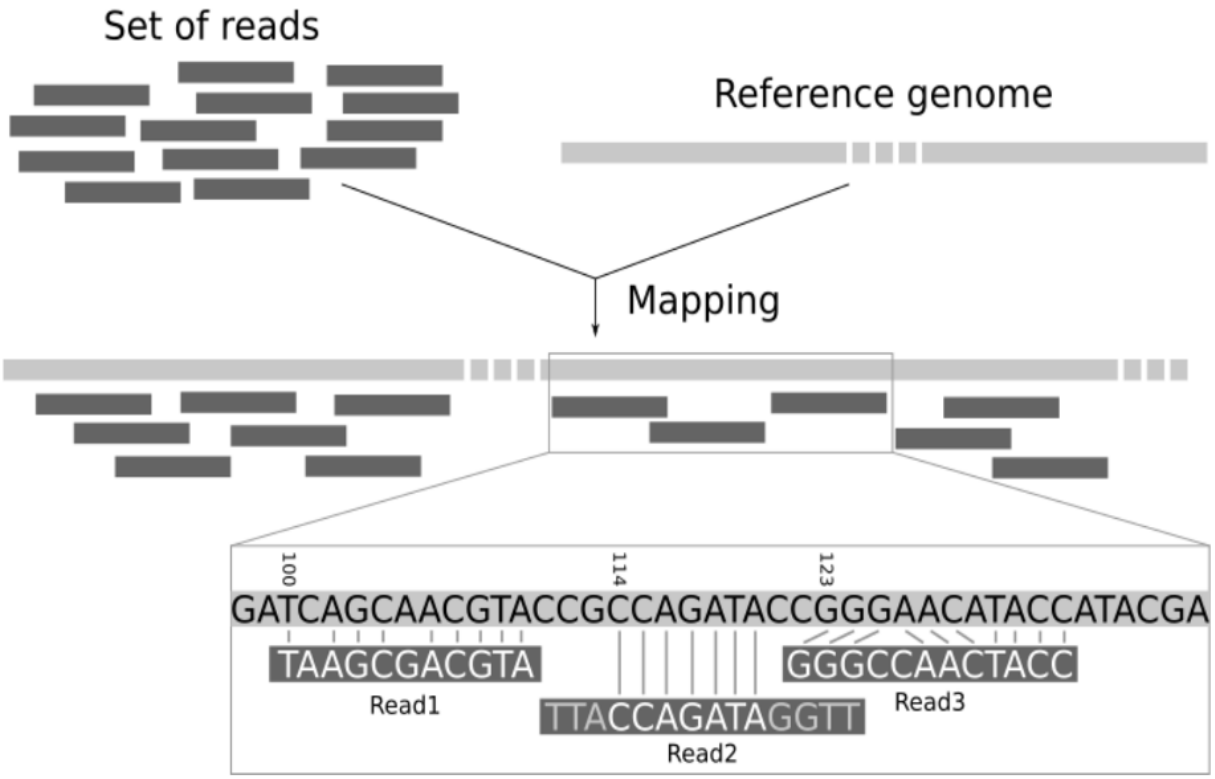


Figure 1: Illustration of the mapping process. The input consists of a set of reads and a reference genome. In the middle, it gives the results of mapping: the locations of the reads on the reference genome. The first read is aligned at position 100 and the alignment has two mismatches. The second read is aligned at position 114. It is a local alignment with clippings on the left and right. The third read is aligned at position 123. It consists of a 2-base insertion and a 1-base deletion.

BACKGROUND READING: ANALYZING THE RESULTS OF NEXT-GENERATION SEQUENCING

Scientists can use genetic databases, such as BLAST, to quickly compare and map genomes using DNA sequences from the unknown sample called **query sequences**.

Watch the following video on your student device to see how BLAST can help to identify and compare an unknown DNA to reference genomes—"Blast Tutorial."

Diagrams called **Krona charts**, zooming, and multilevel pie charts can be created from genomic data and used to analyze genomic data. Go to the following links to view sample interactive Krona charts that show the abundance of species of bacteria in samples of the human microbiome.

- **LINK 1:** Abundance and taxonomy of species in a human gut sample (http://marbl.github.io/Krona/examples/phymmbl.krona.html?collapse=false&color=true&key=false).
- **LINK 2:** Comparison of the abundance of bacteria species in skin microbiome samples from two people on the same day (http://marbl.github.io/Krona/examples/rdp.comp.krona.html?color=true).

  Note: You can view the Krona chart for each sample by clicking on "M3 Skin Day 3" and "F4 Skin Day 3" in the upper left corner menu.

SOURCES:

- https://www.genomicseducation.hee.nhs.uk/glossary/read/#:~:text=In%20next%2Dgeneration%20sequencing%2C%20a,a%20small%20section%20of%20DNA)
- https://training.galaxyproject.org/training-material/topics/sequence-analysis/tutorials/mapping/tutorial.html
- https://github.com/marbl/Krona/wiki
- https://blast.ncbi.nlm.nih.gov/Blast.cgi

CAPTURE SHEET

Genomics Lab Technician Name:

Date:

Description/History of the Sample:

Bacterial DNA sample taken from the stomach of the mummy, Özti

Approximate age of sample ~5,000 yrs

Location of sample—Ötztal Alps, mountains on the border of Austria and Italy

Reason for Sequencing:

Searching for microbes similar to ones found in the modern-day human microbiome

Test Type:

Microbial whole-genome sequencing

Equipment Used:

Illumina iSeq 100 Sequencer

Description of Procedure/Test Methodology:

(Write a summary of what happens in each of the 8 steps and sequencing.)

LEVEL 1—DNA LIBRARY PREP (FROM EXTRACTED DNA SAMPLES)

1. Cut and tagment DNA: _____

_____

2. Stop tagmentation: _____

_____

3. Wash tagmented DNA: _____

_____

4. Add indexes & PCR: _____

_____

LEVEL 2—DNA LIBRARY CLEANUP

5. Remove large fragments: _____

_____

6. Remove small fragments: _____

_____

7. Remove PCR reagents: _____

_____

8. Isolate clean DNA libraries: _____

_____

LEVEL 3—DNA SEQUENCING

Summarize the steps for Level 3

Comparative analysis using BLAST shows the bacteria most closely matches *Helicobacter pylori* (H. pylori); a common human pathogen found in up to 50% of the modern human population.

Helicobacter pylori strain Hpfe077 chromosome, complete genome

Sequence ID: CP094103.1 **Length:** 1624315 **Number of Matches:** 1

Range 1: 1292696 to 1292719 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
48.1 bits(24)	0.007	24/24(100%)	0/24(0%)	Plus/Minus

```
Query  1        AGATCGGAAGAGCACACGTCTGAA  24
               ||||||||||||||||||||||||
Sbjct  1292719  AGATCGGAAGAGCACACGTCTGAA  1292696
```

Helicobacter pylori strain Hpfe073 chromosome, complete genome

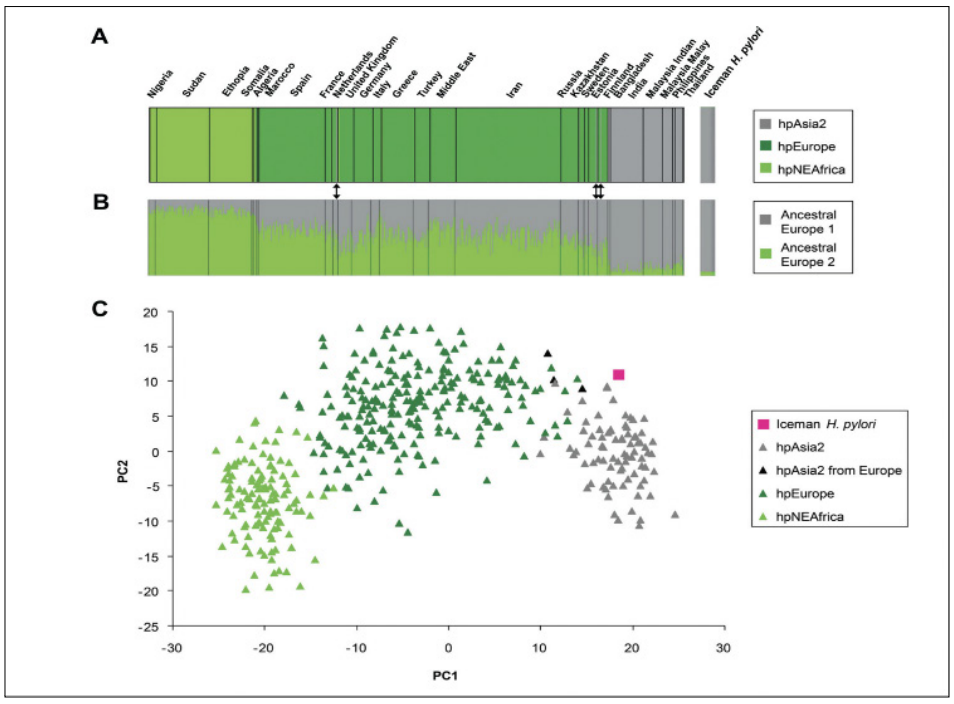
Sequence ID: CP094108.1 **Length:** 1607927 **Number of Matches:** 1

Range 1: 685157 to 685180 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
48.1 bits(24)	0.007	24/24(100%)	0/24(0%)	Plus/Minus

```
Query  1       AGATCGGAAGAGCGTCGTGTAGGG  24
              ||||||||||||||||||||||||
Sbjct  685180  AGATCGGAAGAGCGTCGTGTAGGG  685157
```

When later compared to various known strains of *H. pylori*, the sequence of the Iceman bacteria shares the highest levels of ancestry with the hpAsia2 strain and shares low ancestry with the hpNEAfrica strain as seen in the figures below.



Why are these results significant?

Go to the following link to find out: https://www.pbs.org/wgbh/nova/article/ancient-icemans-h-pylori-genome-hints-at-ancient-migrations-to-europe/

What new information have we learned about Ozti through sequencing? What further questions does this lead you to ask and want to explore?

TOPIC: What can ancient human genomes tell us about human history and health?

Use the following links for your investigation:

- The Revolution of Ancient DNA—What Does Genetics Tell Us About the Past?
- The Ancestral Human Microbiome—Adventures in Genomics
- What Can Ancient DNA Teach the Modern World About Mental Health?
- Got health problems? Blame it on Neanderthal DNA
- Archaeology of the Invisible—Adventures in Genomics
- Piles of ancient poop reveal 'Extinction Event' in human gut bacteria

Drawing on your research, prior knowledge, or reading, complete the table:

CLAIM	SUPPORT	QUESTION
Make a claim (explanation, interpretation) about the topic.	Identify support (things you see, feel, know) for your claim.	Ask a question related to your claim. What isn't explained?

End-of-Level Check-In Questions

LEVEL 1

1. **What is happening when the DNA is going through "tagmentation" at this level?**
    - Unfragmented DNA is cleaved and tagged for analysis.
2. **Why is PCR used in genomic sequencing?**
    - PCR is used to make many copies of the DNA fragments.
3. **Would it be especially important with a DNA sample like the one we are sequencing? Why?**
    - Sometimes a sample can be very small or very old, and having more copies makes sequencing easier and more accurate.

LEVEL 2

1. **Now that you have completed the experience, why do you think this level is called DNA Library Cleanup? What were you trying to isolate from the DNA needed for sequencing?**
    - This level separates the medium DNA fragments from large and small fragments.
2. **Why do we remove the large and small fragments?**
    - iSeq 100 is designed to read a specific fragment size. Much like "the Goldilocks Scenario" some are too big, some are too small, and some are just right.
3. **What function, or role, does the magnetic tube rack serve in this process?**
    - It pulls the beads and DNA to the bottom of the tube, separating them from the supernatant, ethanol, and PCR reagent (which will be discarded).

LEVEL 3

1. **What is a genomic library? (Do some research to find the answer if you don't know!)**
    - A genomic library is a collection of DNA fragments that represent an entire genome (or nearly the entire genome) of an organism.
2. **What are the advantages of using a sequencer over other methods of DNA analysis?**
    - Sequencers (such as next-generation sequencers) reduce the time and cost it takes to sequence genetic material and can process and store large amounts of information efficiently.