

Sequencing DNA from Proteobacteria released from Melting Permafrost

SUMMARY

In this activity, students will be exposed to the concept of coding prior to using the [ncbi BLAST](#) program in the College Board AP Biology recommended activities. Part 1 of the lesson focuses on coding and databases. To appreciate the data science behind BLAST, students will first compare simple codes (i.e. 4-25 base pairs long) and run python programs in class that print the percent differences/similarities between sequences. This will provide students with the opportunity to appreciate what they are about to analyze, without simply being the end-user of already established programs (see [BLAST tutorial](#)).

Part 2 of the lesson focuses on choosing scientific questions to explore via data science. Ultimately, students will be using BLAST and UniProt in order to test hypotheses of phylogenetic tree drawing. This lesson demonstrates how many scientific concepts can overlap by relating the ecological importance of climate change and permafrost melting with sequence data of ancient organisms. Discussions of evolutionary biology, ancestry, and genetic engineering are tied to this topic to promote student engagement in choosing a sequence hypothesis.



Students will learn:

- What is permafrost?
- Why is it melting?
- What are proteobacteria?
- How overlapping areas of science can be used in investigations. Can DNA sequencing of bacteria be used as an indirect measure of climate change drivers? Can DNA sequencing be used to trace ancestry of modern bacteria in the area?

Part 1: Sequence and Databases

1. Do Now/Bell Ringer: How similar are these sequence? How would you determine % similarity? (Students will brainstorm/try methods to compare sequences. This will be used to show them that the program is doing the same thing. Seems easier to do without program BUT *what if the code was much longer?*)
2. Class notes and discussion: What is permafrost? Why is it melting? Why should we care?
3. Introduce Python code:

```

dna1 ='AGCAGAACTGACAGCTAAC'
dna2 ='ACGATCAAACCTATCTAGTC'
dna3 ='ACCCCCCGTCATCACATCC'
dna4 ='AGGGGCTACGAGCGCACGC'

sequences= [dna1, dna2, dna3, dna4]

for dna in sequences:
    print (dna)

for i, item in enumerate (sequences):
    for j in range (i +1, len(sequences)):
        print (i+1, j+1)
        print(item, sequences[j])
        count= 0
        for site in range(len(item)):
            if item[site] != sequences[j][site]:
                count += 1
        print (count)
        print ()
        similarity =(count) /len(dna1) * 100

print(f"The percentage similarity: {similarity:.2f}%")

```

- Students will try a simple simulation to get an idea about how sequences are compared by programs like Python, BLAST or UniProt.

Part 2: Hypothesis Making & Investigation

Lab [Activity](#) using BLAST and UniProt (adapted from College Board Lab and HHMI activity)

Extension: [Woolly Mammoth Revival](#)

- Watch this [video](#) to learn about this recent application of biotechnology
- Optional:* here is [an article](#) clarifying the genetic plan from the video. Contrast this to the [BBC video](#) which explains what the frozen zoo is (which is distinctly different from what's actually happening now).
- Summarize** your understanding of the video & be sure to use proper unit vocab.
- What is the **purpose** behind this reconstructed ecosystem? (minute 1:26-4:00). Is this **really** the woolly mammoth? (5:08-6:30). What is your **opinion**?

ASSESSMENT

Part 1: Students can run code and explain the meaning behind the percentages obtained

Part 2: Students pose a meaningful question of how sequence databases can be used in modern science and develop a hypothetical phylogenetic tree based on research to compare to a data-driven approach using UniProt.

Background Resources:

[From: Eight Metagenome-Assembled Genomes Provide Evidence for Microbial Adaptation in 20,000- to 1,000,000-Year-Old Siberian](#) *Appl Environ Microbiol.* 2021 Oct; 87(19): e00972-21. Published online 2021 Sep 10. Prepublished online 2021 Jul 21. doi: [10.1128/AEM.00972-21](#)

IMPORTANCE Permafrost around the world is thawing rapidly. Many scientists from a variety of disciplines have shown the importance of understanding what will happen to our ecosystem, commerce, and climate when permafrost thaws. The fate of permafrost microorganisms is connected to these predicted rapid environmental changes. Studying ancient permafrost with culture-independent techniques can give a glimpse into how these microorganisms function under these extreme low-temperature and low-energy conditions. This will facilitate understanding how they will change with the environment. This study presents genomic data from this unique environment ~20,000 to 1,000,000 years of age.