Central Dogma: Teaching Students the Value of Using NCBI Tools like BLAST

Team Sublime Dynamite: Richard (Jake) Zimny, Brandon Boswell & Mary Knight

Background

In the typical 9th/10th grade Biology classroom, the following terms or concepts are not addressed:

 -Bioinformatics

 -Computorial Biology

 -Genomics

While in the AP classroom these concepts get touched upon, it is still a very full course from a content standpoint.

Objectives

– Students will understand some fundamental concepts of informatics, computational biology, and genomics

– This lesson is accessible for the students & teacher to use the NCBI tools (particularly BLAST).

– Aligns with IST-1.N.2: Genetic information flows from a sequence of nucleotides in DNA to a sequence of bases in an mRNA molecule to a sequence of amino acids in a protein

https://apstudents.collegeboard.org/ap/2019-05/ap-biology-course-and-exam-description.pdf

---

9th/10th Grade Biology
Student Handout

Name: _____   Date: _____   Period: _____

This is a screenshot that shows the accessing of the actual HBB gene using the NCBI databases. These databases are available to the public and contain all kinds of useful information.
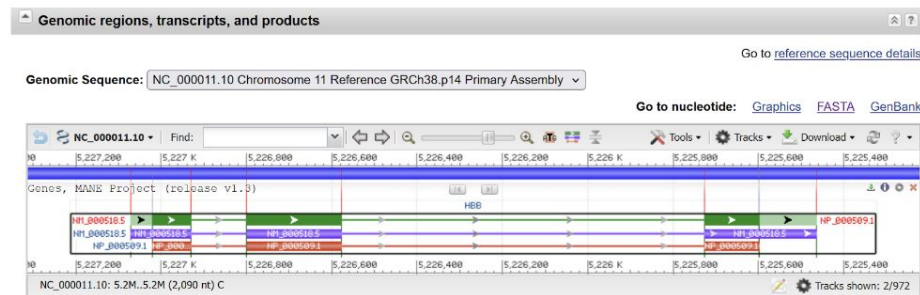
DIRECTIONS:

With the segments below, do the following:

1. Use the given DNA codons to transcribe the corresponding mRNA codons.
2. Use the mRNA codons and the codon chart to identify the amino acid sequence for that portion of the gene. (The codon chart is located on the last page)
3. Once you have finished the wild type transcription and translation, complete those processes for all four of the sample sequences.
4. Once the sample segments are completed, you then need to identify the types of mutations present in each sample, as well as identify the result of each mutation.

Wild Type HBB Gene Segment:

DNA: TAC   CAC   GTA   GAC   TGA   GGA   CTC   CTC   TTC   AGA   CGG   CAA

mRNA: ___   ___   ___   ___   ___   ___   ___   ___   ___   ___   ___   ___

AA Seq: ___   ___   ___   ___   ___   ___   ___   ___   ___   ___   ___   ___

Lesson Overview

Students start off the lesson by viewing a video by HHMI of a case study of 2 young ladies suffering from Sickle Cell Anemia

(Video linked on GitHub)

https://www.biointeractive.org/classroom-resources/genetic-treatment-sickle-cell-disease

Lesson Overview

DNA: TAC CAC GTA GAC TGA GGA CTC CTC TTC AGA CGG CAA

mRNA: ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____

AA Seq: ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____

After teaching the basic content for the Central Dogma, students should be able to complete this activity.

They begin by transcribing and translating the gene sequence as it has been done for the last 50+ years.

Educator Note: the sequence contains only coding regions for this protein; it is a sense strand that has a start codon at the beginning. Additionally, the nucleotide sequence is Wild Type HBB, which codes for the Hemoglobin β subunit.
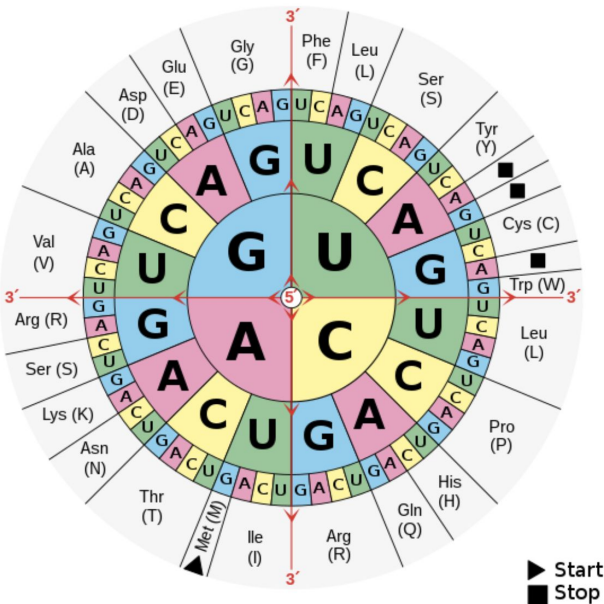
Lesson Overview

Wild Type HBB Gene Segment:

DNA:	TAC	CAC	GTA	GAC	TGA	GGA	CTC	CTC	TTC	AGA	CGG	CAA
mRNA:	___	___	___	___	___	___	___	___	___	___	___	___
AA Seq:	___	___	___	___	___	___	___	___	___	___	___	___
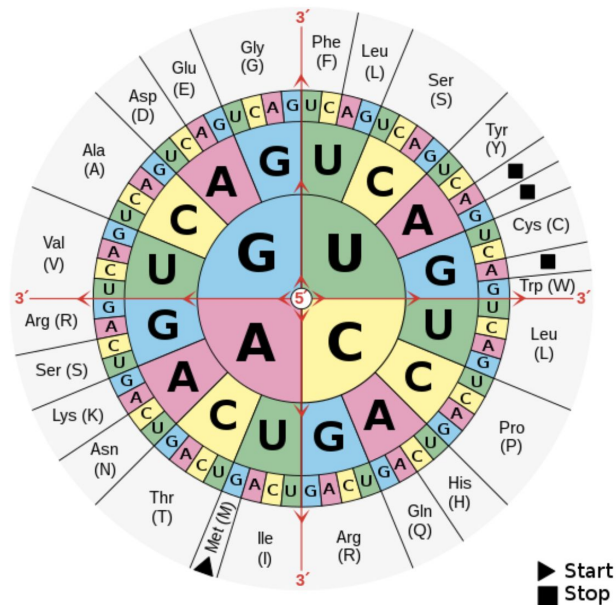
One resource teachers have access to is a Google Sheet (link on GitHub) which contains the actual HBB gene, but separates each triplet into a separate cell.

This additional resource would be helpful if students need more practice with transcription and translation.

Additionally this could be used to compare the class's speed to the speed of the BLAST program.
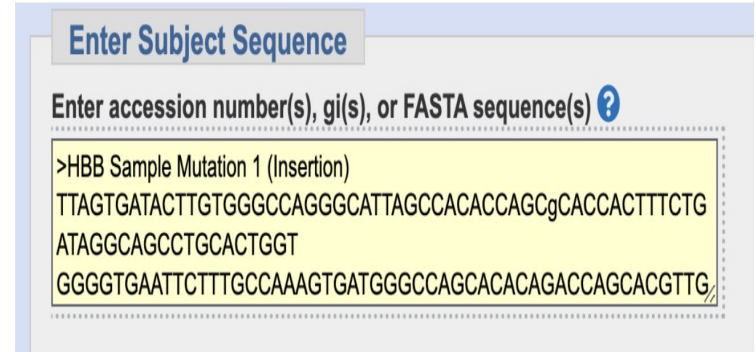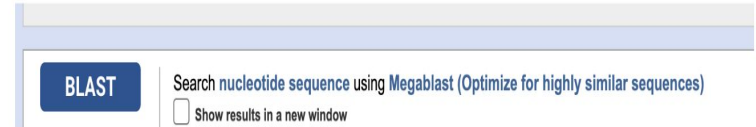
Lesson Overview

Following the completion of the traditional component of the activity, students can begin the BLAST portion of the lesson.

This portion of the activity can be modeled by the teacher with the whole class, OR, there is a student instruction sheet with step-by-step instructions and screenshots to walk the students through the process.

5.) Put a ">" in front of the text that is not part of the nucleic acid sequence, this makes it a title

Enter Subject Sequence

Enter accession number(s), gi(s), or FASTA sequence(s) ❓

>HBB Sample Mutation 1 (Insertion)
TTAGTGATACTTGTGGGCCAGGGCATTAGCCACACCAGCgCACCACTTTCTG
ATAGGCAGCCTGCACTGGT
GGGGTGAATTCTTTGCCAAAGTGATGGGCCAGCACACAGACCAGCACGTTG

6.) Push the BLAST button

BLAST Search nucleotide sequence using Megablast (Optimize for highly similar sequences)
☐ Show results in a new window

7.) Wait for the computation to run, it should only take a few seconds since it is a small sequence
8.) Once the computations are complete click on the alignment tab down at the bottom of the retrieval

ary   Alignments

Lesson Overview

Students will use BLAST to compare up to 4 different sequences to the Reference Sequence for the HBB gene. There are 3 types of mutations represented in the sample sequences:

-insertion

-deletion

-point mutation

The fourth sample has multiple mutations within the one sequence - a challenge for students and a good option for group work.

HBB Sample Mutation 1 (Insertion)

TTAGTGATACTTGTGGGCCAGGGCATTAGCCACACCAGCgCACCACTTTCTGATAGGCAGCCTGCACTGGT
GGGGTGAATTCTTTGCCAAAGTGATGGGCCAGCACACAGACCAGCACGTTGCCCAGGAGCTGTGGGAGGA
AGATAAGAGGTATGAACATGATTAGCAAAAGGGCCTAGCTTGGACTCAGAATAATCCAGCCTTATCCCAA
CCATAAAATAAAAGCAGAATGGTAGCTGGATTGTAGCTGCTATTAGCAATATGAAACCTCTTACATCAGT
TACAATTTATATGCAGAAATATTTATATGCAGAGATATTGCTATTGCCTTAACCCAGAAATTATCACTGT
TATTCTTTAGAATGGTGCAAAGAGGCATGATACATTGTATCATTATTGCCCTGAAAGAAAGAGATTAGGG
AAAGTATTAGAAATAAGATAAACAAAAAGTATATTAAAAGAAGAAAGCATTTTTTAAAATTACAAATGC
AAAATTACCCTGATTTGGTCAATATGTGTACACATATTAAAACATTACACTTTAACCCATAAATATGTAT
AATGATTATGTATCAATTAAAAATAAAAGAAAATAAAGTAGGGAGATTATGAATATGCAAATAAGCACAC
ATATATTCCAAATAGTAATGTACTAGGCAGACTGTGTAAAGTTTTTTTTTAAGTTACTTAATGTATCTCA
GAGATATTTCCTTTTGTTATACACAATGTTAAGGCATTAAGTATAATAGTAAAAATTGCGGAGAAGAAAA
AAAAAGAAAGCAAGAATTAAACAAAAGAAAACAATTGTTATGAACAGCAAATAAAAGAAACTAAAACGAT
CCTGAGACTTCCACACTGATGCAATCATTCGTCTGTTTCCCATTCTAAACTGTACCCTGTTACTTATCCC
CTTCCTATGACATGAACTTAACCATAGAAAAGAAGGGGAAAGAAAACATCAAGCGTCCCATAGACTCACC
CTGAAGTTCTCAGGATCCACGTGCAGCTTGTCACAGTGCAGCTCACTCAGTGTGGCAAAGGTGCCCTTGA
GGTTGTCCAGGTGAGCCAGGCCATCACTAAAGGCACCGAGCACTTTCTTGCCATGAGCCTTCACCTTAGG
GTTGCCCATAACAGCATCAGGAGTGGACAGATCCCCAAAGGACTCAAAGAACCTCTGGGTCCAAGGGTAG
ACCACCAGCAGCCTAAGGGTGGGAAAATAGACCAATAGGCAGAGAGAGTCAGTGCCTATCAGAAACCCAA
GAGTCTTCTCTGTCTCCACATGCCCAGTTTCTATTGGTCTCCTTAAACCTGTCTTGTAACCTTGATACCA
ACCTGCCCAGGGCCTCACCACCAACTTCATCCACGTTCACCTTGCCCCACAGGGCAGTAACGGCAGACTT
CTCCTCAGGAGTCAGATGCACCAT

Lesson Overview

The teacher can decide whether to hide the location and/or identity of the mutations.

Teachers find these mutated sequences on the teacher instruction document or the student instruction document, and then they copy and paste them into the BLAST browser.

We created these by downloading the gene file & manually making changes. The gene was also converted from antisense to sense using Colab.

HBB Sample Mutation 1 (Insertion)

TTAGTGATACTTGTGGGCCAGGGCATTAGCCACACCAGCgCACCACTTTCTGATAGGCAGCCTGCACTGGT
GGGGTGAATTCTTTGCCAAAGTGATGGGCCAGCACACAGACCAGCACGTTGCCCAGGAGCTGTGGGAGGA
AGATAAGAGGTATGAACATGATTAGCAAAAGGGCCTAGCTTGGACTCAGAATAATCCAGCCTTATCCCAA
CCATAAAATAAAAGCAGAATGGTAGCTGGATTGTAGCTGCTATTAGCAATATGAAACCTCTTACATCAGT
TACAATTTATATGCAGAAATATTTATATGCAGAGATATTGCTATTGCCTTAACCCAGAAATTATCACTGT
TATTCTTTAGAATGGTGCAAAGAGGCATGATACATTGTATCATTATTGCCCTGAAAGAAAGAGATTAGGG
AAAGTATTAGAAATAAGATAAACAAAAAAGTATATTAAAAGAAGAAAGCATTTTTTAAAATTACAAATGC
AAAATTACCCTGATTTGGTCAATATGTGTACACATATTAAAACATTACACTTTAACCCATAAATATGTAT
AATGATTATGTATCAATTAAAAATAAAAGAAAATAAAGTAGGGAGATTATGAATATGCAAATAAGCACAC
ATATATTCCAAATAGTAATGTACTAGGCAGACTGTGTAAAGTTTTTTTTTTAAGTTACTTAATGTATCTCA
GAGATATTTCCTTTTGTTATACACAATGTTAAGGCATTAAGTATAATAGTAAAAATTGCGGAGAAGAAAA
AAAAGAAAGCAAGAATTAAACAAAAGAAAACAATTGTTATGAACAGCAAATAAAAGAAACTAAAACGAT
CCTGAGACTTCCACACTGATGCAATCATTCGTCTGTTTCCCATTCTAAACTGTACCCTGTTACTTATCCC
CTTCCTATGACATGAACTTAACCATAGAAAAGAAGGGGAAAGAAAACATCAAGCGTCCCATAGACTCACC
CTGAAGTTCTCAGGATCCACGTGCAGCTTGTCACAGTGCAGCTCACTCAGTGTGGCAAAGGTGCCCTTGA
GGTTGTCCAGGTGAGCCAGGCCATCACTAAAGGCACCGAGCACTTTCTTGCCATGAGCCTTCACCTTAGG
GTTGCCCATAACAGCATCAGGAGTGGACAGATCCCCAAAGGACTCAAAGAACCTCTGGGTCCAAGGGTAG
ACCACCAGCAGCCTAAGGGTGGGAAAATAGACCAATAGGCAGAGAGAGTCAGTGCCTATCAGAAACCCAA
GAGTCTTCTCTGTCTCCACATGCCCAGTTTCTATTGGTCTCCTTAAACCTGTCTTGTAACCTTGATACCA
ACCTGCCCAGGGCCTCACCACCAACTTCATCCACGTTCACCTTGCCCCACAGGGCAGTAACGGCAGACTT
CTCCTCAGGAGTCAGATGCACCAT

Lesson Overview

We modeled the mutations from actual known mutations to increase the authenticity of the activity.

We attempted to get actual variants from Clinvar, but most of them weren't available. We assumed it was for the purposes of masking patient info/HIPAA.

HBB Sample Mutation 1 (Insertion)

TTAGTGATACTTGTGGGCCAGGGCATTAGCCACACCAGCgCACCACTTTCTGATAGGCAGCCTGCACTGGT
GGGGTGAATTCTTTGCCAAAGTGATGGGCCAGCACACAGACCAGCACGTTGCCCAGGAGCTGTGGGAGGA
AGATAAGAGGTATGAACATGATTAGCAAAAGGGCCTAGCTTGGACTCAGAATAATCCAGCCTTATCCCAA
CCATAAAATAAAAGCAGAATGGTAGCTGGATTGTAGCTGCTATTAGCAATATGAAACCTCTTACATCAGT
TACAATTTATATGCAGAAATATTTATATGCAGAGATATTGCTATTGCCTTAACCCAGAAATTATCACTGT
TATTCTTTAGAATGGTGCAAAGAGGCATGATACATTGTATCATTATTGCCCTGAAAGAAAGAGATTAGGG
AAAGTATTAGAAATAAGATAAACAAAAAGTATATTAAAAGAAGAAAGCATTTTTAAAATTACAAATGC
AAAATTACCCTGATTTGGTCAATATGTGTACACATATTAAAACATTACACTTTAACCCATAAATATGTAT
AATGATTATGTATCAATTAAAAATAAAAGAAAATAAAGTAGGGAGATTATGAATATGCAAATAAGCACAC
ATATATTCCAAATAGTAATGTACTAGGCAGACTGTGTAAAGTTTTTTTTTAAGTTACTTAATGTATCTCA
GAGATATTTCCTTTTGTTATACACAATGTTAAGGCATTAAGTATAATAGTAAAAATTGCGGAGAAGAAAA
AAAAGAAAGCAAGAATTAAACAAAAGAAAACAATTGTTATGAACAGCAAATAAAAGAAACTAAAACGAT
CCTGAGACTTCCACACTGATGCAATCATTCGTCTGTTTCCCATTCTAAACTGTACCCTGTTACTTATCCC
CTTCCTATGACATGAACTTAACCATAGAAAAGAAGGGGAAAGAAAACATCAAGCGTCCCATAGACTCACC
CTGAAGTTCTCAGGATCCACGTGCAGCTTGTCACAGTGCAGCTCACTCAGTGTGGCAAAGGTGCCCTTGA
GGTTGTCCAGGTGAGCCAGGCCATCACTAAAGGCACCGAGCACTTTCTTGCCATGAGCCTTCACCTTAGG
GTTGCCCATAACAGCATCAGGAGTGGACAGATCCCCAAAGGACTCAAAGAACCTCTGGGTCCAAGGGTAG
ACCACCAGCAGCCTAAGGGTGGGAAAATAGACCAATAGGCAGAGAGAGTCAGTGCCTATCAGAAACCCAA
GAGTCTTCTCTGTCTCCACATGCCCAGTTTCTATTGGTCTCCTTAAACCTGTCTTGTAACCTTGATACCA
ACCTGCCCAGGGCCTCACCACCAACTTCATCCACGTTCACCTTGCCCCACAGGGCAGTAACGGCAGACTT
CTCCTCAGGAGTCAGATGCACCAT

Extensions

A discussion about what students thought about the BLAST program should follow the activity. It would be useful for students to consider other ways these tools can be used.

Other possible options include:

~Dividing the gene as it is on the Google Sheet (already in triplets), and then have students each transcribe and translate their portion. They can then make an actual, physical chain with paper shapes or beads, to give them an idea of what an entire gene looks like as an amino acid chain.

~Students make their own mutant variant to see how the number mutations and percent identity or E value are related. In other words, how many mutations are necessary so that the BLAST program does not recognize the sequence as HBB?

```
TTAGTG-ATACTTGTGGGCCAGGGCATTAGCCACACCAGCCACCACTTTCTGATAGGCAG
.......A....................................................
CCTGCACTGGTGGGGTGAATTCTTTGCCAAAGTGATGGGCCAGCACACAGACCAGCACGT
............................................................
TGCCCAGGAGCTGTGGGA-GGAAGATAAGAGGTATGAACATGATTAGCAAAAGGGCCTAG
.......-.........C..........................................
CTTGGACTCAGAATAATCCAGCCTTATCCCAACCATAAAATAAAAGCAGAATGGTAGCTG
............................................................
GATTGTAGCTGCTATTAGCAATATGAAACCTCT-TACATCAGTTACAATTTATATGCAGA
.................................C..........................
AATATTTATATGCAGAGATATTGCTATTGCCTTAAC-CCAGAAATTATCACTGTTATTCT
....................................T.......................
TTAGAATGGTGCAAAGAGGCATGATACATTGTATCATTATTGCCCTgaaagaaagagatt
.......-....................................................
agggaaagtattagaaataagataaacaaaaaagtatattaaaagaagaaagcatttttt
...............................G............................
aaaattacaaatgcaaaattaCCCTGATTTGGTCAATATGTGTACACATATTAAAACATT
............................................................
ACACTTTAACCCATAAATATGTATAATGAT-TATGTATCAATTAAAAATAAAAGAAAATA
.............................C..............................
AAGTAGGGAGATTATGAATATGCAAATAAGCACACATATATTCCAAATAGTAATGTACTA
............................................................
```

Future Work

Making corresponding protein sequences for the same mutated HBB genes, so that students can run a Protein BLAST and then compare amino acid sequences.

Adding resources so that students can visualize 3D structures from the HBB protein with the mutations they analyzed.