Connect the dots…DNA to DISEASE

Developed by: M. Oltmann & A. James.

Teachers Material

California State Standards

Cell Biology

- **1.d.** Students know the central dogma of molecular biology outlines the flow of information from transcription of ribonucleic acid (RNA) in the nucleus to translation of proteins on ribosomes in the cytoplasm.

Genetics

- **4.b.** Students know how mutations in the DNA sequence of a gene may or may not affect the expression of the gene or the sequence of amino acids in an encoded protein.
- **4.c.** Students know specialization of cells in multicellular organisms is usually due to different patterns of gene expression rather than to differences of the genes themselves.
- **4.e.** Students know proteins can differ from one another in the number and sequence of amino acids.

Investigation & Experimentation

- **1.a.** Select and use appropriate tools and technology (such as computer-linked probes, spreadsheets, and graphing calculators) to perform tests, collect data, analyze relationships, and display data.
- **1.m**. Investigate a science-based societal issue by researching the literature, analyzing data, and communicating the findings. Examples of issues include irradiation of food, cloning of animals by somatic cell nuclear transfer, choice of energy sources, and land and water use decisions in California.

Synopsis

Students transcribe and translate a given sequence of DNA and perform a BLAST search against a database of known proteins to determine which protein their sequence encodes. The goal is to show students that genes encode proteins, which in turn can cause disease if mutated or function improperly.

Background

Unfortunately, most students fail to make the connection between DNA sequence, proteins, and protein function to sustain life. The goal of this activity is to allow students to discover that DNA sequences, while the A, T, G, and C's seem like non-sense, do in fact encode very important proteins that help us sustain life. They will also discover that some of these proteins are involved in causing diseases.

The program used to search the available database of proteins is called BLAST. It is provided by the National Institutes of Health and is widely used by research scientists to search for DNA and protein sequences. Be sure to emphasize to students that this is a very advanced program that is usually only used by experienced scientists.

Objectives

The goal of this activity is to allow students to discover on their own that the seemingly "nonsense" DNA sequences they often hear about actually encode for meaningful proteins which have vital functions in the body. They will learn that defective proteins are usually the cause of most diseases.

Suggested Timeline

This lab can be completed in 1.5 hours.

Materials

1 DNA sequence per group
1 computer with an internet connection per group

Advanced Teacher Preparation

Cut the DNA sequences into strips, or copy each to a separate piece of paper. You should also research the various proteins (google search) so students' questions can be answered.

Experimentation

Objectives

Students will be able to:
1. Transcribe and translate a DNA sequence
2. Use the internet and a specialized database search engine to identify the protein encoded by a certain DNA sequence.
3. Research a certain protein and the disease that can result from an abnormality in the protein

Engagement

The teacher can emphasize that the techniques and technology the students are using is usually only used by experienced scientists – this will encourage student participation and enthusiasm. The teacher should let the students make the discoveries guided by this lab on their own, with little coaching.

Exploration

The students will learn to use an advanced protein database search program on the computer, which will enhance their technology skills. They will also research specific diseases and their causes, which all relate to protein expression.

Exploration Phase

The exploration phase should be based on the students' discoveries after searching the protein database. They will find that the DNA sequence that they were originally given actually encodes a real protein that when abnormally expressed causes a common disease. Questions such the following should be asked to ensure understanding of the lab:

1. How is DNA inherited?
2. What are the types of mutations that can occur during transcription and translation?
3. What are the consequences of these mutations?

Concept Application Phase

The students will apply what they have learned by answering the questions in the student hand-out.

Feedback Questions

1. Did this lab teach you something that you previously didn't know?
2. Are you more interested in learning about proteins and how they contribute to diseases?

TEACHER INFORMATION (KEY) -- give students only the DNA sequence…only one per group.

Template DNA SEQUENCE	PROTEIN	gene	DISEASE
tacgagtgtaagtaccggagactgtcgctccttcttcacacacta	Presenilin 2	PS2	Alzheimer's
tacctacataagtactttcctgaaagtttccggttcctccctcaa	Synuclein	SNCA	Parkinson's
tacgcgaaggcgaaaccccaccaccacggtgggcggcaccggccg	Laforin	EPM2A	Epilepsy
tacgtaacccctttgggacacgcctaagaacaccgaaaccgggata	Leptin	OB	Obesity
tacggataacctaggtttctctccggttgtaaaaaactttaaaaa	BRCA 2	BRCA2	Breast cancer
tacttttatagtaccgacctaacgttgtttggttgtcacttttc	Dystrophin	DMD	Duchenne Muscular Dystrophy
tacttccaagacacccgacgcaacgaccagtgtaaggaccgtcc t	Apolipoprotein E	APOE	Atherosclerosis

Do not give to students

Cut out and give to students (1 per group)

(1) Presenilin 2	tacgagtgtaagtaccggagactgtcgctccttcttcacacacta	Group 1
(2) Synuclein	tacctacataagtactttcctgaaagtttccggttcctccctcaa	Group 2
(3) Laforin	tacgcgaaggcgaaaccccaccaccacggtgggcggcaccggccg	Group 3
(4) Leptin	tacgtaacccccttgggacacgcctaagaacaccgaaaccgggata	Group 4

(5) BRCA 2	tacggataacctaggtttctctccggttgtaaaaaactttaaaaa	Group 5

(6) Dystrophin	tacttttatagtaccgacctaacgttgtttggttgtcacttttc	Group 6

(7) Apolipoprotein E	tacttccaagacacccgacgcaacgaccagtgtaaggaccgtcct	Group 7

(STUDENT HANDOUT BEGINS)

Names: ____
Group: ______________________

Connect the dots…DNA to DISEASE

Introduction
We've learned that DNA is the genetic material that organisms inherit from their parents, but have you ever thought about what exactly this DNA encodes for? How do our cells use DNA as a set of instructions for life? How is the information in our DNA/genes used by our bodies? And what happens when the DNA is mutated or not used properly?

Materials (per group)
DNA sequence
Computer with an internet connection

Procedure

1. Obtain your DNA sequence from your teacher.

2. Convert your DNA sequence into a complementary mRNA sequence.

   EXAMPLE: DNA: T A C G G C T A G
   ↓
   mRNA: A U G C C G A U C

   Your DNA sequence:

   mRNA sequence:

3. Determine the codons.

   EXAMPLE: mRNA: A U G C C G A U C

Codons: AUG ↓ CCG AUC

Codons:

_____ _____ _____ _____ _____ _____ _____ _____ _____ _____ _____ _____ _____ _____ _____

4. Translate the codon sequence into an amino sequence. Use the chart provided.

Codons: AUG CCG AUC

↓

Amino Acids: Methionine Proline Isoleucine

Amino Acid Sequence:

_____ _____ _____ _____ _____

_____ _____ _____ _____ _____

_____ _____ _____ _____ _____

5. Write out the one-letter abbreviations for the amino acids in the sequence. Use the chart provided.

____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____ ____

6. Go to http://www.ncbi.nlm.nih.gov/BLAST/ and choose Protein-Protein BLAST (top of the second column).
7. Enter the one-letter abbreviations for your amino acid sequence in the SEARCH box – be sure to enter them in the correct order!
8. Click on the "BLAST" button.
9. At the next page, click on the "FORMAT" button. It may take a few minutes to process your sequence.

10. At the next page, scroll down to the list of proteins that matched your sequence. Choose one that matches one on the list of possible proteins that was given to you.

11. The protein our DNA sequence encodes is (should be in the list provided): _____

12. Now go to https://ghr.nlm.nih.gov/gene with the name of your protein to find out the disease your protein is involved in. Use either the alphabet list or the search box.

12. This protein is involved in the following disease: _____

13. Write a brief paragraph explaining the disease caused by this protein or a mutation in this protein.

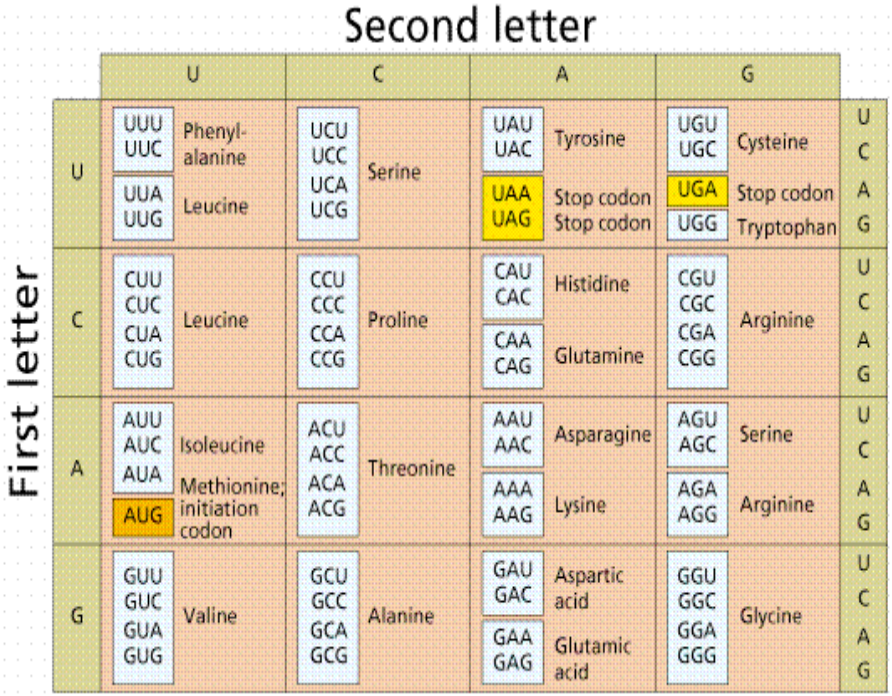
14. List 3 things you learned in this activity (either technical concepts, such as using the computer or scientific concepts).

(1)

(2)

(3)

AMINO ACID CHARTS AND PROTEIN NAMES



AMINO ACID	abbreviation
Alanine	A
Arginine	R
Asparagine	N
Aspartic acid	D
Cysteine	C
Glutamine	Q
Glutamic acid	E
Glycine	G
Histidine	H
Isoleucine	I
Leucine	L
Lysine	K
Methionine	M
Phenylalanine	F
Proline	P
Serine	S
Threonine	T
Tryptophan	W
Tyrosine	Y
Valine	V

Possible proteins
Presenilin 2
Synuclein
Laforin
Leptin
BRCA 2
Dystrophin
Apolipoprotein E