(STUDENT HANDOUT BEGINS)

Names: ____
Group: ______________________

Connect the dots…DNA to DISEASE

Introduction

Assume the role of a molecular biologist who specializes in protein function. Studying a group of diseased individuals you have discovered that all of the people afflicted with the disease share a gene that others (those without the disease) do not have in their genes. The next step in your research is to transcribe and translate the base sequence to figure out the amino acid sequence of the resulting protein. Afterwards you will enter the sequence of amino acids into a database of proteins to discover the name of the faulty protein that seems to be causing the mysterious malady. Lastly, conduct independent research about the protein to describe the function of the protein and explain why the malformed molecule causes medical problems.

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:

____________________ ____________________ ____________________ ____________________ ____________________

____________________ ____________________ ____________________ ____________________ ____________________

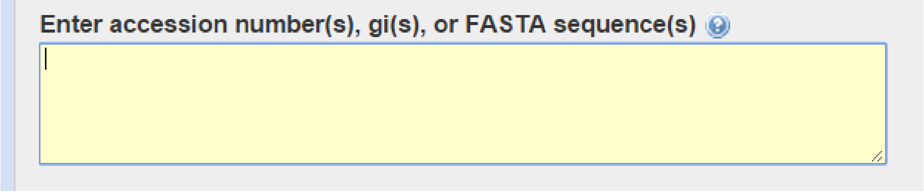
____________________ ____________________ ____________________ ____________________ ____________________

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

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

5. Go to http://www.ncbi.nlm.nih.gov/BLAST/ and choose Protein Blast Under the Basic Blast heading.

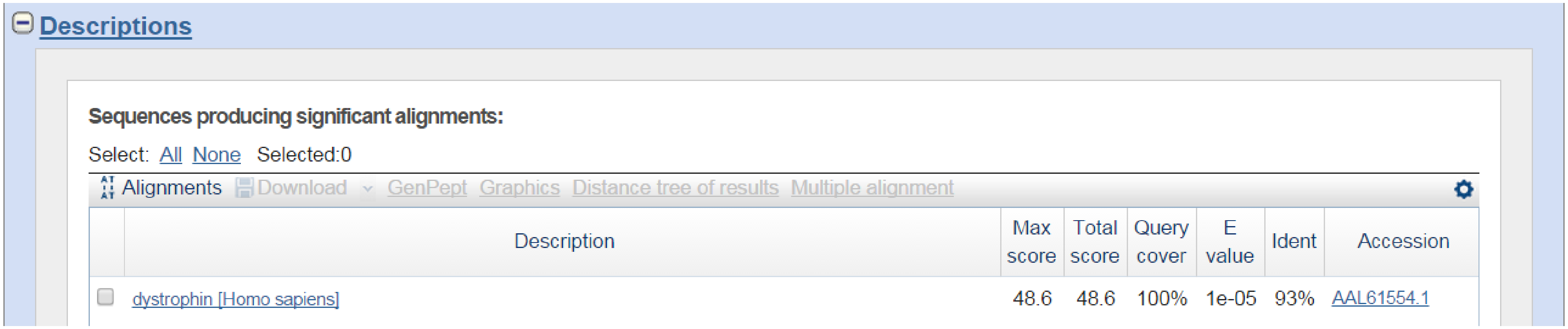
7. Using the table on the last page, enter the one-letter abbreviations for your amino acid sequence in the SEARCH box labeled: “Enter accession number(s), gi(s), or FASTA sequence(s)” – be sure to enter them in the correct order!



8. Scroll down and 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 “Description” of proteins matched your sequence. Choose one that matches one on the list of possible proteins on the last page.



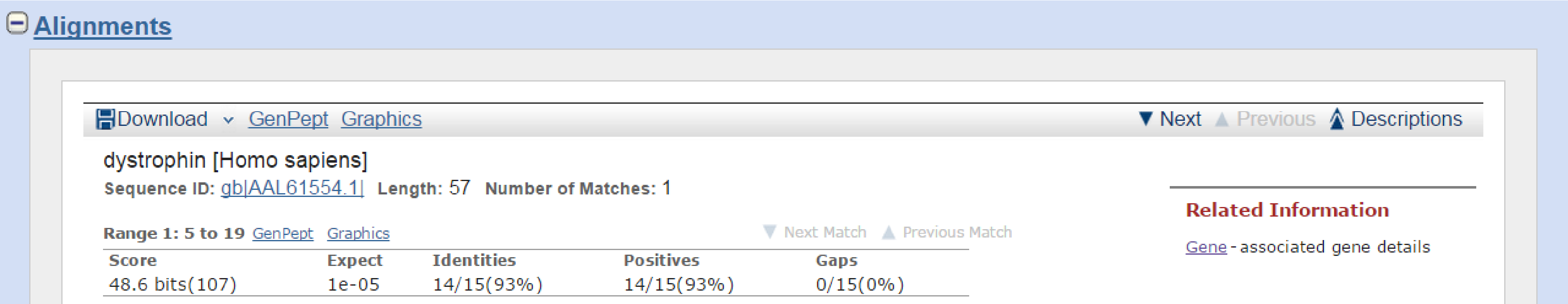
Sequences producing significant alignments:

Select: All None Selected:0

	Description	Max score	Total score	Query cover	E value	Ident	Accession
☐	dystrophin [Homo sapiens]	48.6	48.6	100%	1e-05	93%	AAL61554.1

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

12. Scroll down to the “Alignments” section. Locate the alignment for Homo sapiens (Humans). Click on the Gene link under related information to the right. This link should provide information regarding the function or associated disease of the gene and protein. Additionally, conduct an internet search for the name of the protein. Look for information regarding the protein and associated disease.



dystrophin [Homo sapiens]

Sequence ID: gb|AAL61554.1| **Length:** 57 **Number of Matches:** 1

Range 1: 5 to 19 GenPept Graphics

Score	Expect	Identities	Positives	Gaps
48.6 bits(107)	1e-05	14/15(93%)	14/15(93%)	0/15(0%)

Related Information

Gene - associated gene details

13. This protein is involved in the following disease: ___________________________________

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

AMINO ACID CHARTS AND PROTEIN NAMES

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

FIRST NUCLEOTIDE	SECOND NUCLEOTIDE								THIRD NUCLEOTIDE
	U		C		A		G		
U	UUU	Phenylalanine (Phe)	UCU	Serine (Ser)	UAU	Tyrosine (Tyr)	UGU	Cysteine (Cys)	U
	UUC		UCC		UAC		UGC		C
	UUA	Leucine (Leu)	UCA		UAA	STOP	UGA	STOP	A
	UUG		UCG		UAG		UGG	Tryptophan (Trp)	G
C	CUU	Leucine (Leu)	CCU	Proline (Pro)	CAU	Histidine (His)	CGU	Arginine (Arg)	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	Glutamine (Gln)	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	Isoleucine (Ile)	ACU	Threonine (Thr)	AAU	Asparagine (Asn)	AGU	Serine (Ser)	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	Lysine (Lys)	AGA	Arginine (Arg)	A
	AUG	Methionine (Met) START	ACG		AAG		AGG		G
G	GUU	Valine (Val)	GCU	Alanine (Ala)	GAU	Aspartic Acid (Asp)	GGU	Glycine (Gly)	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	Glutamic Acid (Glu)	GGA		A
	GUG		GCG		GAG		GGG		G

BLAST Amino Acid Abbreviations	
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**