

Name That Gene, Disease, and Protein Lesson Plan

By Sharlene Denos (UIUC) & Kathryn Hafner (Danville High)

Lesson Goals

1. To improve students' understanding of mutations and how a single point mutation in a gene can result in a large phenotypic change like Cystic Fibrosis or Alzheimer's.
2. To reinforce students' learning of the Central Dogma of molecular biology, and that genes act by expressing proteins.
3. To reinforce students' understanding that some mutations are more harmful than others, and the most harmful ones will be in a sensitive area of the protein's structure. It will either disrupt the protein's ability to fold (stability) or its function by changing its binding or active site.

Prior Knowledge Required

This lesson is designed as a follow up to the genetics unit covered in weeks 2-4 of the second semester at DHS. Students should understand the central dogma of molecular biology and have a working knowledge of the terms gene, nucleotide, base, base-pair, centromere, chromosome, genome, mutation, genetic disease, genotype, phenotype, bioinformatics, sequence alignment, BLAST and query. In addition, students should understand that a mutation can be a small change in just one nucleotide or the change can involve many nucleotides and that the normal nucleotides can be substituted for something else, deleted altogether, or an extra nucleotide can be added.

Definitions:

- **Gene:** A sequence of DNA that codes for a protein.
- **Nucleotide** (or base): The monomer unit of a DNA polymer, often referred to as a “base”. The size of a gene is often given in base pairs (bp) or kilobase pairs (kbp). An average human gene is about 75-100kbp.
- **Gene sequence:** The type & order of nucleotides in a gene.
- **Protein:** A polymer of amino acids which folds up into a compact structure in order complete tasks such as catalyzing reactions.
- **Protein sequence:** The exact type and order of amino acids in a protein polymer. An average protein contains about 300 amino acids.
- **Protein structure:** The 3-dimensional form a protein takes. This shape is entirely dependent on the sequence of amino acids in the protein.
- **Enzyme:** A protein catalyst. It will speed up a chemical reaction without changing the ratio of products to reactants.
- **Active site:** The place on the enzyme where catalysis takes place. The active site is usually in an interior pocket of the enzyme, protected from the environment so that a special kind of chemistry can take place there.
- **Binding site:** Some proteins must bind to another molecule in order to perform their function. The place where the other molecule binds is the binding site. There are usually some special amino acids there which make this site stick to the target molecule.
- **Mutation:** A change in the gene sequence. This can be a small change in just

one nucleotide or the change can involve many nucleotides. The normal nucleotides can be substituted for something else, deleted altogether, or an extra nucleotide can be added.

- **Amino acid:** The 20 basic building blocks of proteins.
- **Bioinformatics:** The study of biology using computer and information science. One of the most lively areas of study is the attempt to store & organize the large amount of DNA sequences that have been obtained from organisms from E. coli to humans.
- **Queue:** This is a line. In computer science, it is the line of jobs waiting to be performed by a computer.
- **Query:** A query is a question. In bioinformatics, though, it is the DNA or protein sequence you input into the computer in order to compare it to the online database of DNA & protein sequences.
- **Hits:** When you submit a query sequence, matches from the database are returned and these are called “hits.”
- **BLAST:** “Basic Local Alignment Search Tool”. This is a search tool that is used to quickly & accurately search online DNA databases for matches against any input sequence.
- **Bit score:** With any BLAST search, hits will have a “bit score”. This is a normalized score that allows you to compare results from multiple searches. The score tells you how well your hit matches your query.
- **Alignment:** This is the process of lining up two or more sequences in order to compare (in a quantitative way) how similar they are.

- **Multiple sequence alignment:** An alignment with more than 2 sequences.
- **Alignment score:** A number that represents how good the alignment is. The score is calculated by adding up every match and subtracting any gap or substitution (non-match) between your query sequence and the hits returned by the BLAST search.
- **Expectation, Expect, or E-Value:** This is the number of different alignments that are expected to occur in a database search by chance, which have the same or better score than your hit. The lower the E value for a hit, the more meaningful the match is likely to be.
- **Chromosome:** The structures in which DNA of higher eukaryotes are found.
- **Autosome:** Any chromosome other than a sex chromosome.
- **Genetic Disease:** A disease that is inherited. It is usually the result of a mutation which causes the protein for which the gene codes to function improperly or not at all.
- **Genome:** The full set of an organisms DNA. For humans, this would include all 21 autosomes and the two sex chromosomes.
- **Centromere:** The structure occurring roughly in the middle of a chromosome that holds individual chromatids together.
- **Protein Data Bank (PDB):** A database of 3-dimensional structures for proteins and nucleic acids.
- **GenBank:** A database of all known nucleotide sequences and their protein translations.

Materials

- A computer with VMD installed and an LCD projector for the instructor.
- A computer with VMD installed for each student group.
- The VMD tutorial (“Getting to Know RNA Polymerase”).
- Student handout.
- Disease Gene sequences.
- Data storage is required for saving VMD images.
- Students will need a pen and paper or word processing program to take notes.

Protocol

(See Wefer et. al⁹ or the student handout for a more detailed protocol)

1. Print out the DNA sequences in the quantity required for your class. Each group will need a sequence, but the same sequence can be given to more than one student/group if they are not sitting at adjacent stations. Make sure to cut out the DNA sequence without its corresponding disease name.
2. Lead students to the NCBI website and then to the BLAST page. It is helpful if everyone performs their BLAST searches together, despite the different sequences.
3. Once students find their match, help them navigate to the “Genes and Diseases” online book. Some genes will contain the name of the disease, while others will not. Lead students to the “search book” option until they find their disease.
4. Students should also explore the chromosomes and determine where their gene is located. They will need to take extensive notes during this portion of the activity.

5. Under the main page for each genetic disease is a link to Entrez, which will provide detailed information about the protein and mutations leading to the disease. Students should copy the protein name down, as well as any information about the location of the mutation in the protein sequence and the final folded structure.
6. Help students navigate to the PDB website and search for their disease protein there. They will need to copy down the PDB ID.
7. Help students load their protein into VMD and locate the site of the mutations. In some cases, a mutant form of the protein will be available and that structure can be compared with the normal version. Otherwise, the student should look to see if the mutation site is near a binding or active site. It could also disrupt stability by interfering with disulfide bonds, hydrophobic interactions, salt bridges or hydrogen bonds.
8. Students will need to render an image and save or print it for their report, which should be turned in a 2-3 days later.

Note to the teacher:

On this assignment it is important to warn students about plagiarism. They typically have difficulty with rephrasing pathology descriptions and may need extra help with this.

Grading Rubric (continued next page)

Criteria	Poor	Satisfactory	Excellent
----------	------	--------------	-----------

Name and location of gene with explanation of how student found it (10 points)	Name or location is wrong. Methods explanation is weak or not detailed enough.	Name and location are correct. Methods explanation is weak or not detailed enough.	Name and location are correct. Methods explains exactly how student performed the BLAST, how student knew which hit was correct and how student determined for which disease gene was responsible.
Name and description of disease (20 points)	Name is correct but description contains more than 2 sentences which are direct quotes from the “Genes and Disease” online book or another resource are not properly referenced.	Name is correct but description contains one or two sentences which are direct quotes from the “Genes and Disease” online book or another resource are not properly referenced.	Name is correct and description is detailed without copying the “Genes and Disease” online book or any other resource. Any direct quotes are properly referenced.

Grading Rubric, continued

Criteria	Poor	Satisfactory	Excellent
Name and normal function of protein (20 points)	Name and function are correct, but insufficiently detailed. Contains one or more direct quotes that are not properly referenced.	Name and function are correct, but insufficiently detailed. OR Contains one or more direct quotes that are not properly referenced.	Name and function are correct and sufficiently detailed. Any direct quotes are properly referenced.
Where, in the protein sequence and folded structure, does the mutation occur? (20 points)	Mutation positions are incorrect or missing.	Mutation positions in the protein sequence OR folded structure are incorrect or absent.	Mutation positions are both correct.
Explanation of how the mutation disrupts the protein’s function. (15 points)	Explanation is incorrect and not properly explained or infeasible.	Explanation is incorrect, but detailed and plausible. Answer shows the student	Explanation is incorrect, but detailed and plausible. Answer shows the student

		properly understands how mutations disrupt can protein function.	properly understands how mutations can disrupt protein function.
VMD image and description. (15 points)	No image or no description is given.	Image or description does not support the explanation of how the mutation disrupts the protein's function.	Image and description support the explanation of how the mutation disrupts the protein's function.