

Gene Explorer Lab

Purpose:

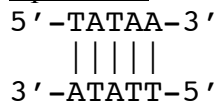
To work with a computer simulation of eukaryotic gene expression in order to help you to understand:

- more about transcription and translation
- eukaryotic gene structure
- exons and introns
- the effects of mutations on gene expression

Introduction:

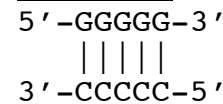
Up until now, you have worked with prokaryotic genes of increasing complexity. We will now move on to examine the structure and function of the genes of eukaryotes (creatures like us). Eukaryotic genes have promoters and terminators for controlling transcription as well as start and stop codons for controlling translation. Although the promoter and terminator sequences are different in different organisms, the genetic code, including the start and stop codons is **identical in all organisms**. The promoter and terminator sequences used in the hypothetical organism simulated by the gene explorer are shown below:

a promoter:



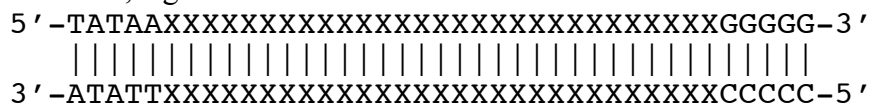
*Transcription begins with the first base-pair to the *right* of this sequence and proceeds to the *right*.

a terminator:



*Transcription ends with the first base-pair to the *left* of this sequence.

therefore, a gene would look like this:



*the region shown as X's would be transcribed into pre-mRNA.

In addition, eukaryotic genes have a few features that prokaryotic genes do not have. These are:

- transcription produces an mRNA called a pre-mRNA which is not yet ready for translation.
- this pre-mRNA is then processed in several steps:
 - The introns are removed and the exons are joined; this is called mRNA “splicing”. This is controlled by splice signal sequences. In real organisms, these sequences are not well known. In general, introns start with 5'-GU-3' and end with 5'-AG-3'. In the hypothetical organism simulated by the Gene Explorer, introns start with 5'-GUGCG-3' and end with 5'-CAAAG-3'.
 - a modified G nucleotide is added to the 5' end of the mRNA; this is called the “cap”. In the Gene Explorer, this is not shown.
 - many A's are added to the 3' end of the mRNA; this is called the tail. In real organisms, as many as 400 A's can be added at a specific signal sequence; the Gene Explorer adds 13 A's as a tail to the 3' end of any mRNA. Note that these A's do not correspond to T's in the DNA.

In previous labs, you did the work by hand; this was necessary in order to make the processes clear to you. Today, you will use the Gene Explorer to do the tedious work of:

- finding the promoter and terminator
- reading the DNA sequence to produce the pre-mRNA
- finding the splice sites
- splicing, capping, and tailing the mRNA
- finding the start codon
- translating the mRNA

The Gene Explorer will then allow you to make specific mutations in a gene sequence and it will then calculate their effects on the mRNA and protein. Researchers use tools like this to analyze the genes they are studying.

Procedure

Part I: Start Gene Explorer and familiarize yourself with the program

Note: This program will work on both Macs and PCs with internet access. You can also download the Gene Explorer as an application from the course website. These instructions refer to the Macs in the lab.

1) From the Dock, start Safari (on PCs, Internet Explorer will work fine).

2) From the Bio 111 home page, go to the OLLM and to the Gene Explorer link for this lab. Be sure you use the “new Gene Explorer” link.

3) It will ask you for permission to print. You must click “YES” or the program won’t run right. You are not doing anything risky since you have to tell it when to print anyway.

4) You will see something like this on the web page:

The screenshot displays the Gene Explorer interface with the following components and annotations:

- DNA strands:** The top section shows the DNA sequence with a green-shaded promoter region (TATAA) and a pink-shaded terminator region (TAAAGG). A box labeled "DNA strands" points to this section.
- pre-mRNA:** The middle section shows the pre-mRNA sequence with a yellow-shaded exon and a green-shaded intron. A box labeled "RNA strands" points to this section.
- mature-mRNA and Protein:** The bottom section shows the mature-mRNA sequence with colored exons (yellow, green, blue, red) and the corresponding protein sequence: N-MetProLeuSerAspValGluArgGlyPro-C. A box labeled "protein" points to this section.
- Annotations:**
 - A box on the right states: "• The promoter is shaded in green. • The terminator is shaded in pink."
 - Another box on the right states: "The exons are shaded in colors in both the pre-mRNA and mRNA."
- Interface Elements:** The interface includes buttons for "Enter a Caption", "Page Set", "Reset DNA Sequence", and "Enter New".

5) You can select a base in the DNA strand and Gene Explorer will show the corresponding bases in the mRNA and the corresponding amino acid in the protein, if applicable. To select a base, here are some notes:

- You can only select bases in the top strand of DNA. If you click anywhere else on the gene panel, nothing will happen.
- It can be tricky to select a particular base on the first click; here are some tips:
 - You can use the “+” key to move one base to the right or the “-“ key to move one base to the left.
 - To hit a particular base, click on the space between it and the next base. For example, to select 25, click on the space between 25 and 26.
- The number of the selected base is shown at the bottom right of the Gene Explorer.

An example is shown below. Here, base 63 has been selected.

The screenshot shows the Gene Explorer interface with the following elements and callouts:

- Top DNA strand:** GATGCCTTGTGCGATAAGGTGTGCCCCCCCAAAGTGTCCGGATTCGAGTGC GCGTG. A blue vertical bar highlights the base 'A' at position 63. Callout: "This DNA base was clicked on."
- Bottom DNA strand:** CTACGGAACACGCTATTCCACACAGGGGGGGTTTCACAGCCTA...CTCAGCGCAC. Callout: "This is the corresponding base in the other strand of DNA."
- pre-mRNA:** GAUGCCUUGUGCGAUAAGGUGUGUCCCCCCCAAAGUGUCGGAUUCGAGUGCGCGUG. Callout: "This is the corresponding base in the pre-mRNA; this lines up with the DNA"
- mature mRNA:** CCGAUCGAGCGAGGACCUAAGAAGGUGUGAAAAAAAAAAAAA-3'. Callout: "This is the corresponding base in the mature mRNA."
- Protein sequence:** erAspValGluArgGlyPro-C. Callout: "This is the corresponding amino acid in the protein; this lines up with the mature mRNA."
- Control bar:** Reset DNA Sequence | Enter New DNA Sequence | Selected Base = 63. Callout: "This shows the number of the selected base."

6) You can edit the DNA in several ways:

- To delete the selected base, use the “delete” key.
- To replace the selected base with another base, type a *lower-case* letter (a, g, c, or t).
- To insert bases *to the left of the selected base*, type an *upper-case* letter (A, G, C, or T).

7) There are several useful buttons on the Gene Explorer:

- Enter A Caption – click this to add a few words to the bottom of the screen. You can use this to uniquely identify your gene when you print it out from a common printer.
- Page Setup – click this to set up the page for printing. **You should use landscape mode.**
- Print – click this to print the gene (the DNA, RNA, and protein).
- Reset DNA Sequence – click this to reset the DNA to the starting gene sequence.
- Enter New DNA Sequence – click this to enter a new DNA sequence.

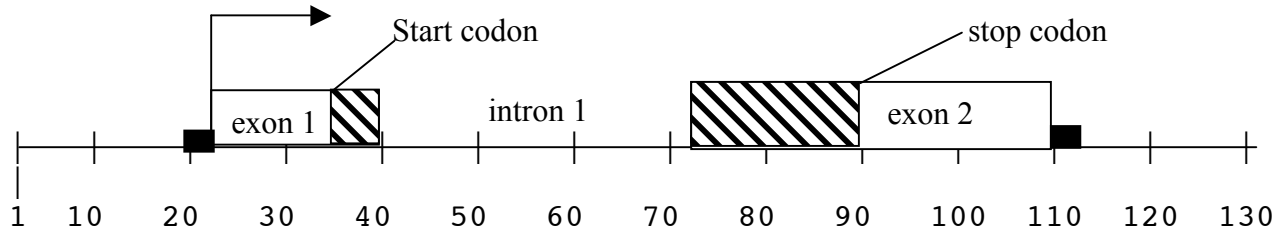
8) If it starts acting strange, quit and re-launch Safari.

Part II: An eukaryotic gene

1) Look at the Gene Explorer Display; it shows the un-modified gene.

2) Make a map of this gene's structure using the printout of the gene's DNA, mRNA and protein. You can select each DNA base one-by-one to do this; use the "+" key to advance through the sequence.

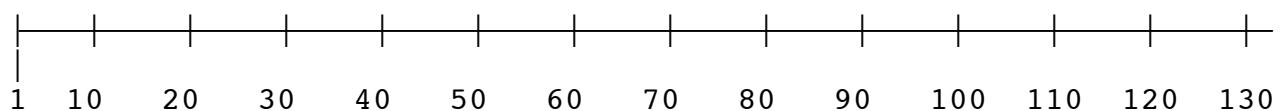
Maps have the following format; the numbers correspond to the numbers on the DNA strands. The map shown below is **only for illustration purposes** and does not correspond to the gene we are studying.



Here is what this map shows (and a list of all the features that a map must contain); note that all the map coordinates are approximate.

- The promoter is indicated by the small black rectangle at position 20.
- The start of transcription is indicated by a bent arrow; in this gene it is at position 22 (roughly).
- Exon 1 is indicated by a labeled box; it starts at 22 and ends at 40.
- Intron 1 is indicated by a labeled blank space; it starts at 41 and ends at 72.
- Exon 2 is indicated by a labeled box; it starts at 73 and ends at 110.
- The end of transcription is indicated by the end of the last exon; here it is at 110.
- The terminator is indicated by a small black rectangle around position 110.
- The start codon is indicated by the start of the hashed region in exon 1; it is at position 35.
- The stop codon is indicated by the end of the hashed region in exon 2; it is at position 90.
- The coding region - the region that encodes the protein - is indicated by the hashed parts of the exons; it extends from 35 to 90. Note that it **does not** include the intron.

Make your map of the gene shown by the Gene Explorer on the line below; be sure to include all the features listed above. If there are more than 2 exons and one intron, be sure to include them as well.



- 3) There are several important things to notice about this gene; you will discuss these with your TA.
- a) Note that the poly A tail, the string of A's at the 3' end of the mature mRNA, do not have any corresponding bases in the DNA or the pre-mRNA. How did they get there?

 - b) Students often think that introns do not split inside codons. This is not true, as you can see if you look at intron 1 and 2. Why doesn't it matter if an intron occurs in the middle of a codon?

 - c) Students also often think that an intron has to be a multiple of three nucleotides long. This is also not true, as you can see if you measure the length of introns 1 and 2 or if you try inserting some bases in an intron to see if it has an effect. Why is it that introns can be any number of nucleotides long?

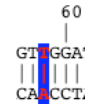
Part III: Mutations

- You will now make several mutations in the gene and explore their effects.
- For each mutation, draw a map of the mutated gene; this is a **mutant allele**.
- Indicate the location of the mutation on your map with an asterisk “*”.
- After each mutation, be sure to hit the “Reset DNA Sequence” button so that each mutation is made on an unaltered copy of the gene.

1) Click “Reset DNA Sequence”

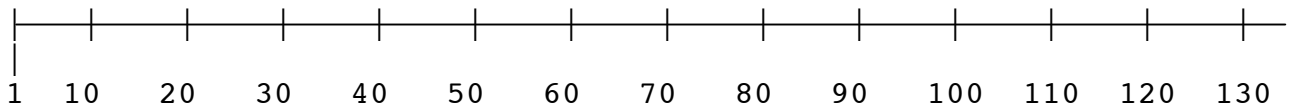
2) You will now do mutation #1: changing the C/G base-pair at position 58 to a T/A base-pair.

- Select base number 58; be sure that you have selected the right base by looking at the number in the lower right of the screen.
- Type ‘t’ (be sure it is lower-case).
- Look at the sequence of the DNA strands; it should look like this:
- Note that base-pair 58 is now T in the top strand and A in the bottom strand.
- If it does not, go back to step 1 and try again.



3) What is the amino acid sequence of the mutant protein? How does it differ from the original sequence (this should be shown in blue in the line below the new protein sequence)? What kind of mutation is this?

4) Using the line below, draw a map of the mutant gene. Indicate the location of the mutation in the DNA sequence with an asterisk “*”. How does the structure of the mutant gene differ from the original gene?



5) Explain how the mutation caused the change you diagrammed in part (4).

6) Click “Reset DNA Sequence”.

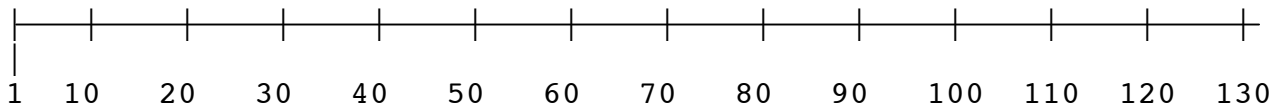
7) You will now do mutation #2: deleting the T/A base-pair at position 26.

- Select base number 26; be sure that you have selected the right base by looking at the number in the lower right of the screen.
- Type the delete key once.
- Look at the sequence of the DNA strands; it should look like this:
- If it does not, go back to step 6 and try again. Note that base-pair 26 is now missing.



8) What is the amino acid sequence of the mutant protein? How does it differ from the original sequence? What kind of mutation is this?

9) Using the line below, draw a map of the mutant gene. Indicate the location of the mutation in the DNA sequence with an asterisk “*”. How does the structure of the mutant gene differ from the original gene?



10) Explain how the mutation caused the change you diagrammed in part (9).

11) Click “Reset DNA Sequence”.

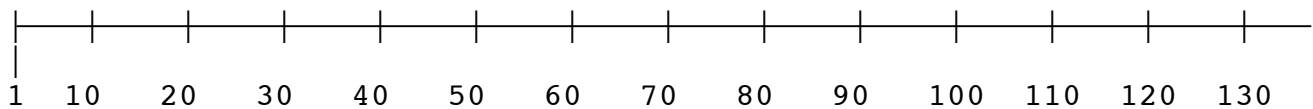
12) You will now do mutation #3: changing the A/T base-pair at position 51 to a T/A base-pair.

- Select base number 51; be sure that you have selected the right base by looking at the number in the lower right of the screen.
- Type ‘t’ (be sure it is lower-case).
- Look at the sequence of the DNA strands; it should look like this:
- Note that base-pair 51 is now T in the top strand and A in the bottom strand.
- If it does not, go back to step 11 and try again.



13) What is the amino acid sequence of the mutant protein? How does it differ from the original sequence?

14) Using the line below, draw a map of the mutant gene. Indicate the location of the mutation in the DNA sequence with an asterisk “*”. How does the structure of the mutant gene differ from the original gene?

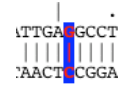


15) Explain how the mutation caused the change you diagrammed in part (14).

16) Click “Reset DNA Sequence”.

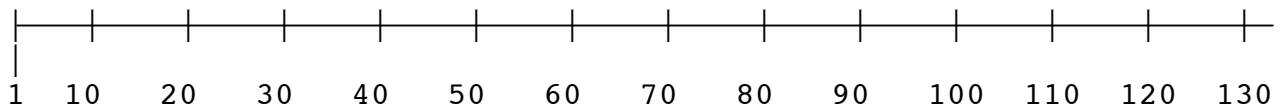
17) You will now do mutation #4: changing the T/A base-pair at position 21 to a G/C base-pair.

- Select base number 21; be sure that you have selected the right base by looking at the number in the lower right of the screen.
- Type ‘g’ (be sure it is lower-case).
- Look at the sequence of the DNA strands; it should look like this:
- Note that base-pair 21 is now G in the top strand and C in the bottom strand.
- If it does not, go back to step 16 and try again.



18) What is the amino acid sequence of the mutant protein? How does it differ from the original sequence?

19) Using the line below, draw a map of the mutant gene. Indicate the location of the mutation in the DNA sequence with an asterisk “*”. How does the structure of the mutant gene differ from the original gene?

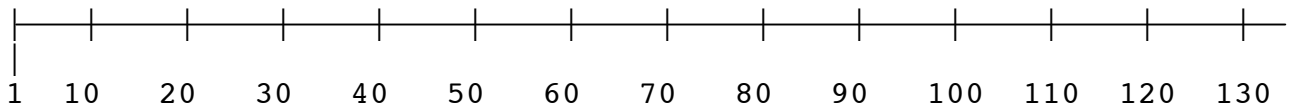


20) Explain how the mutation caused the change you diagrammed in part (19).

21) Click “Reset DNA Sequence”.

22) Invent a nonsense mutation of your own.

- Which base-pair did you change? _____
- What did you change it to? _____
- Draw a map of the resulting gene:



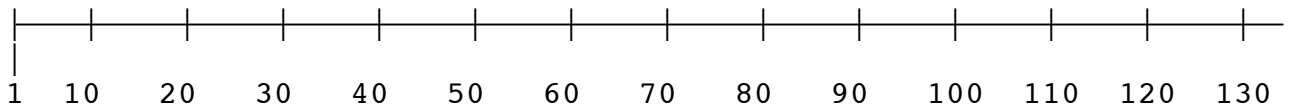
- Have your TA check off your mutation and map.

23) Explain how the mutation caused the change you diagrammed in part (22).

24) Click “Reset DNA Sequence”.

25) Invent a mutation that changes the mature mRNA sequence but does not change the protein sequence.

- Which base-pair did you change? _____
- What did you change it to? _____
- Draw a map of the resulting gene:



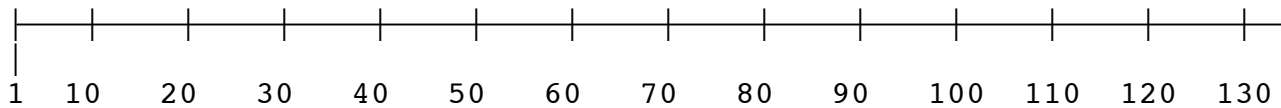
- Have your TA check off your mutation and map.

26) Explain how the mutation caused the change you diagrammed in part (25).

27) Click “Reset DNA Sequence”.

28) Invent a mutation that results in no mRNA or protein being made at all.

- Which base-pair did you change? _____
- What did you change it to? _____
- Draw a map of the resulting gene:



- Have your TA check off your mutation and map.

29) Explain how the mutation caused the change you diagrammed in part (28).

This is a very challenging problem, you can try it if you have time.

30) Click “Reset DNA Sequence”.

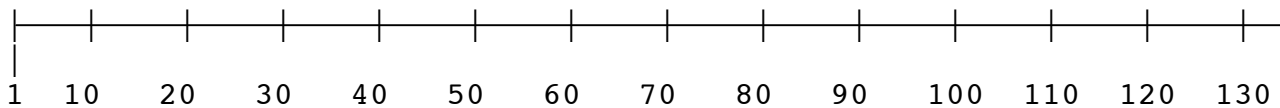
31) A single base mutation (one base inserted, deleted, or changed) in the starting gene results in the following protein sequence:

N-MetProLeuSerAspValAspAlaArgAlaLysLysAsnLysGlyGluAspLeuLysLysVal-C

Note that the first 6 amino acids are the same as the normal protein (underlined).

What was the mutation that led to this? (There may be more than one right answer; give only one)

- Which base-pair was changed? _____
- What was the change? _____
- Draw a map of the resulting gene:



- Have your TA check off your mutation and map.

32) Explain how the mutation caused the change you diagrammed in part (31).

Lab report

- Must be typed; handwritten reports will not be accepted.
- Due next week at the start of the lab session you are currently in. This is a firm deadline.
- Although you have worked as a group, each member of the group must turn in an individual lab report.

Your lab report must include:

An entirely new gene that you have invented. Using the new Gene Explorer, this gene must:

- produce a protein of at least 5 amino acids (including the N-terminal met)
- contain at least one intron

Your report can consist entirely of a printout from the Gene Explorer showing your gene (with your name on it, of course...).

Tips

- Use the “Enter New DNA Sequence” button and delete the starting sequence from the entry blank.
- Type in a promoter, a little DNA and a terminator and be sure your RNA is made.
- Click on your gene and add start codon, coding region, and stop codon; be sure your protein is made. Type slowly so the program can keep up.
- Similarly, add an intron in the coding region and be sure your gene works.