Aipotu Part II:
Biochemistry & Molecular Biology

Introduction:

<u>The Biological Phenomenon Under Study</u>
In this lab, you will continue to explore the biological mechanisms behind the expression of flower color in a hypothetical plant. These flowers can be white, red, orange, yellow, green, blue, purple, or black.

<u>Scenario</u>:
You are the chief biologist for a breeder of fine flowers. Your company sells seeds that customers plant in their gardens. Since most of your customers expect that the flowers will grow each year from seeds produced the previous year, you try to produce true-breeding plants whenever you can.

You've found a new species of flower with an attractive shape. You've collected four plants from the wild: two green, one red, and one white. Your customers would really like to have purple flowers from this plant. You set out to create a true-breeding purple flower.

In Part I, you found the alleles involved in color production. You went on to describe the colors produced by different allele combinations. You then began investigations to find the biochemical mechanism behind the colors.

Today, you will complete your biochemical investigations and determine the features of the proteins' sequences that determine their colors. You will use this knowledge to engineer a purple protein. However, it is not possible to simply add a purple protein to a plant and have it produce purple offspring. You will need to do a little genetic engineering to get the pure-breeding purple plant. That is what you will do later today.

Tasks:
- Finish determining how the amino acid sequence of a pigment protein determines its color – this completes the task you began in the Aipotu I lab earlier in the semester.
- Determine the differences in DNA sequence of the alleles you defined in Part I.
- Determine how the DNA sequence of a pigment protein gene determines the color of the protein produced.
- Engineer a gene to produce a purple protein and use this to construct a pure-breeding purple flower.
- Make a mutant that has flowers of a new color.
  - Explain the color of this mutant in genetic, biochemical, and molecular biological terms.

First: Notebook Review

To be sure that you have all the information you need for this lab, with your lab partners, go over the notes you took while doing the SPOC and from lecture and be sure you have the information listed below. You should fill in any gaps in your notes so everyone in your group has all they need. You can check these items off as you go.

- o The rules for DNA and RNA from lecture
- o How transcription is controlled and how to figure out which mRNA will be made from a given DNA molecule
- o How mRNA splicing is controlled and the difference between introns and exons
- o How translation is controlled and how to figure out which protein will be made from a given mRNA molecule

You will then discuss these with your TA as a class to clarify any issues that remain.

Part I: Biochemistry II

Your TA will return your group's copy of the Aipotu Worksheet to review.

If you did not complete the worksheet, go back to the Aipotu I lab manual and your notes and complete the exercises there and put your results . You should:

- Determine how the amino acid sequence and structure of a protein controls its color
- Use this knowledge to design a purple protein

Once you have finished the Aipotu worksheet, you should review it before you proceed with the lab. You should be sure to have complete notes on the following (you should check these off as you go):

- o What are the different colors of the flowers in Aipotu?
  - o Which are dominant, co-dominant, incompletely-dominant, or recessive?
  - o How do the alleles combine to give the final flower color?
- o What are the features (structure and composition) that cause a protein to be colored and to have a particular color?

Part II: Molecular Biology

Using the tool:

Using the jsAipotu link on the On-Line Lab Manual, open Aipotu, (Do not use the Aipotu application in the Dock) and switch to this tool by clicking the "Molecular Biology" tab near the top of the window.

You will see something like this:

You use this part of the program like you used the Biochemistry section in Part II.

This part of the program simulates the expression of eukaryotic genes using the same algorithm as in the **Gene Explorer.**

Eukaryotic genes, like those found in our simulated organisms, 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 tool are shown below:

<table>
<tr><td>a promoter:</td><td>a terminator:</td></tr>
<tr><td>

```
5'-TATAA-3'
   |||||
3'-ATATT-5'
```
</td><td>

```
5'-GGGGG-3'
   |||||
3'-CCCCC-5'
```
</td></tr>
<tr><td>*Transcription begins with the first base-pair to the right of this sequence and proceeds 5'=>3' to the right.</td><td>*Transcription ends with the first base-pair to the left of this sequence.</td></tr>
</table>

therefore, a gene would look like this:

```
5'-TATAAXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXGGGGG-3'
   ||||||||||||||||||||||||||||||||||||||||
3'-ATATTXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXCCCCC-5'
```

*the region shown as X's would be transcribed into a <u>single strand</u> of 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, 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". 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 tool 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. Now that you understand how these processes work, the tool will do all the tedious work of:
- finding the promoter and terminator
- reading the DNA sequence to produce the pre-mRNA
- finding the splice sites
- splicing and tailing the mRNA
- finding the start codon
- translating the mRNA

The tool will then allow you to make specific mutations in a gene sequence and it will then calculate and display their effects on the mRNA and protein. You do not have to deal with all the details listed above; the software will take care of it all. Researchers use tools like this to analyze the genes they are studying.

There are three kinds of experiments you can perform with this tool. The following sections use examples to show you how to do each; you will need to devise your own experiments to carry out the tasks above.

I) Examine the Pigment Protein Genes Present in an Organism from the **Greenhouse**. This simulates sequencing the DNA from the two pigment proteins genes that an organism possesses, expressing them, and displaying their colors.

> 1) Click on the **Green-2** organism in the **Greenhouse**. You should see this:

The **Green-2** organism contains two alleles of the pigment protein gene. Each of these alleles has a distinct DNA sequence. One of these sequences is shown in the **Upper Gene Window**; it is a blue-colored protein as shown by the blue square next to the "Color:" label. The other protein is shown in the **Lower Gene Window**; this is a yellow-colored protein. The combined color of the two proteins is green as shown by the **Combined Color** in between the two **Gene Window**s.

<u>IMPORTANT NOTE</u>:
This software is under development. Please treat it gently and be patient. Please report any bugs to your TA.

The **Gene Window**s also show the key features of the genes:

- Clicking on a base in the upper DNA strand highlights the corresponding bases in the mRNA and the corresponding amino acid in the protein (as appropriate).
- In the DNA sequence:
  - Promoters are shown in green
  - Terminators are shown in red
- In the pre-mRNA sequence:
  - Exons are shown in color
  - Introns are not colored
- In the mature mRNA sequence:
  - Exons are colored as in the pre-mRNA
  - Start and stop codons are underlined

The **Gene Windows** work like the **Gene Explorer** with the addition of colors.

II) Compare the DNA sequences of two pigment protein genes. You can compare the DNA sequence of the two genes by clicking on the "Compare" menu and choosing "Upper vs. Lower". A window will appear showing the differences between the two sequences.

III) Edit a DNA Sequence or Create a New DNA Sequence and Determine the Resulting Protein's Amino Acid Sequence and Color. You can edit the sequence in either of the **Gene Windows** and click the "Fold" button to predict the color of the protein. The tool will also give the color that results from the combination of the colors in the Upper and Lower windows.

You can edit the DNA in several ways:

- To <u>delete</u> the selected base, use the "*delete*" key.
- To <u>replace</u> the selected base with another base, type a *lower-case* letter (a, g, c, or t).
- To <u>insert</u> bases *to the left of the selected base*, type an *upper-case* letter (A, G, C, or T).
- You can enter an entirely new (or copied) DNA sequence by clicking the "Enter New DNA Sequence" button.

For example:

1) Double-click the **Red** organism in the **Greenhouse**.

2) In the **Upper Gene Window**, click on the top DNA strand near base 20. The base you have clicked on will be highlighted in blue. In addition:
   - the corresponding bases in the mRNAs and the corresponding amino acid in the protein will be selected.
   - at the bottom of the **Upper Gene Window**, you will see "Selected Base = 20" or some similar number corresponding to the base you selected.

   If you have not selected base 20, use the arrow keys to select it.

3) Type "C". This will change base 20 from G to C, thus removing the start codon for the gene. You will immediately see that the protein sequence shown under the mRNA sequence, now reads "none" indicating that no protein will be made from this gene. Furthermore, the border of the **Upper Gene Window** will turn pink to indicate that the new protein has not been folded.

4) Click the "Fold Protein" button in the **Upper Gene Window** (or hit the return key) to fold the resulting protein. You will see the following:
- the border of the **Upper Gene Window** will turn back to gray
- the color chip in the **Upper Gene Window** will show white (the color of the new protein)
- a blank entry with a white background will be generated in the **History List** to show the new protein. You can bring it back to one of the **Gene Window**s by selecting it in the **History List** and clicking the ">Upper" or ">Lower" button as appropriate.

You can save this new organism to the **Greenhouse** by clicking the "Add…" button above the **Greenhouse**. You will be prompted to give this organism a name. You might want to switch over to the Genetics tool and self-cross this organism to see what happens.

You can also click the "Load Sample DNA" button. This will load a sample amino acid sequence that folds to a white-colored protein with a shape that is similar to many colored proteins.

If you click an entry in the **History List**, you can then examine it in the upper or lower windows by using the "=>Upper Window" and "=>Lower Window" buttons:
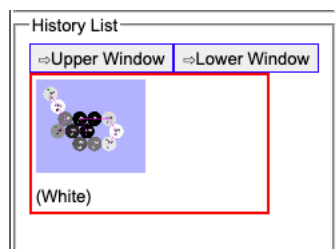
Specific Tasks for this section
a) Give the genotype of each starting strain.
b) What are the differences in the DNA sequences of the alleles you defined in Part I.
c) Do all the white alleles have the same DNA sequence? Hint: use the **Compare** menu to compare the sequences.
d) Using this knowledge, construct a pure-breeding purple organism.

You should put your data in the tables below:

a) **Genotypes**

- Green 1

- Green 2

- Red

- White

b) **DNA changes in the** alleles

The best way to do this is to use the "Sample" sequence as the "reference" sequence and express changes as we have in class. For example, if, in the protein, Ala 12 is changed to Val by a mutation, you can write "Ala12Val" and if, in the DNA, the base at position 22 is changed from A to T, you can write "A22T".

allele	**color**	**change(s) in amino acid sequence**	**change(s) in DNA sequence**

Are all the white alleles the same?

(d) Show your TA that you have made a pure-breeding purple organism. For full credit, you need to explain to your TA *why* it is purple and pure-breeding.

Part III: Genetics - Mutation

Another very important way scientists study a process is to alter it by mutation and see what happens. Analyzing mutants is a very powerful technique for understanding biological processes – finding out how something can be "broken" or changed can give insight into how it works normally.

Tasks:
- Make a mutant version of one of the white alleles that has a different color.
- Examine the changes to its color, protein sequence, and DNA sequence.
- Your analysis of the mutant will be a very important part of your answer to Exam 7.

Using the Tool
Mutate one of the white organisms. This simulates treating the cells of one organism with a strong dose of a mutagen and then growing each of those mutant cells into a separate plant. The flowers of the resulting mutant plants are displayed in a new **Tray**. This process can take a few minutes.

1) Go to the **Genetics Window**.

2) Select the White organism from the **Greenhouse**. When you have only one organism selected, the "Cross Two Organisms" buttons will be grayed out and the "Self-Cross One Organism" and "Mutate One Organism" buttons will be activated.

3) Click the "Mutate One Organism" button in the **Upper Work Panel**. A pop-up will appear telling you that the program is busy making mutants. You can cancel this process if you like.

You will then see a new **Tray** with the mutants and a new entry in the **History List.**

Procedure

(1) Select the White organism from the **Greenhouse** and click the **Mutate One Organism** button. Wait while it makes the mutants.

(2) With luck, you will see an organism of a different color. If not, try mutating the White strain again until you see another color. Do not mutate an organism that was already mutated.

(3) Save your mutant organism to the **Greenhouse** with a descriptive name (like "Green from White").

(4) From the **Greenhouse** menu, choose **Save Greenhouse**. You should e-mail this file to yourself for safe-keeping. You can load it into any browser by choosing "Load Greenhouse File…" from the Greenhouse menu.
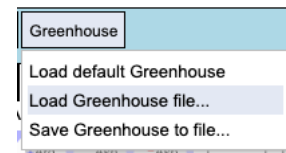
(5) Find out how the mutation has changed the *Protein Sequence*.
   a. Click on the **Biochemistry** tab at the top of **Aipotu**.
   b. Click on the organism you just saved in the **Greenhouse**.
   c. Aipotu then displays the shapes of the proteins produced by the two alleles that this organism is carrying. One of these will be unchanged – it will be white; the other will be mutated, resulting in an altered color.
   d. From the **Compare** menu, choose **Upper vs Sample** or **Lower vs Sample** as appropriate. A window will pop up showing the two amino acid sequences with differences highlighted in red. Make a note of the difference in the space below.

Note that, sometimes, mutations cause a dramatic change in protein sequence due to frame-shifts, splicing changes, etc. In this case, it is not possible to describe the mutation as a simple amino acid change. In this case, you should describe mutations that involve many changes something like this: "the first 5 amino acids are unchanged, but the 16 amino acids that follow are all completely different from the starting sequence" as appropriate.

In addition, because the Mutate button simulates a very high mutation rate, your new strain may show more than one amino acid change. If this is the case, you will need to find out which amino acid change(s) are the ones that caused the color change. You can do this by going to the **Biochemistry** tab, loading the sample protein (it has the same sequence as the starting white allele), and making individual changes to the sequence to see which one(s) are necessary for color change. You will need to have this information for your exam answer.

(6) Find out how the mutation has changed the *DNA Sequence*.
   a. Click on the **Molecular Biology** tab at the top of **Aipotu**.
   b. Double-click on the organism you just saved in the **Greenhouse**.
   c. Aipotu then displays the genes that encode the two alleles that this organism is carrying. One of these will be unchanged – it will be white; the other will be mutated, resulting in an altered color.
   d. From the **Compare** menu, choose **Upper vs Sample** or **Lower vs Sample** as appropriate. A window will pop up showing the two DNA sequences with differences highlighted in red. Make a note of the difference in the space below.

If your mutant has more than one DNA sequence change, you will need to figure out which of the mutations is responsible for the color change. You can use a procedure like the one you used to find the consequential amino acid change(s) to find the consequential DNA sequence changes. This information is also essential for your take home exam answer.

Preparing for the Take Home Exam
    Take Home Exam 7 will be, in part, based on this lab. You should look at the exam on Blackboard before you leave lab today. You may want to use some of the remaining time in lab to prepare for the exam.

The Genetic Code in mRNA

	U	C	A	G	
U	UUU phe	UCU ser	UAU tyr	UGU cys	U
	UUC phe	UCC ser	UAC tyr	UGC cys	C
	UUA leu	UCA ser	UAA STOP	UGA STOP	A
	UUG leu	UCG ser	UAG STOP	UGG trp	G
C	CUU leu	CCU pro	CAU his	CGU arg	U
	CUC leu	CCC pro	CAC his	CGC arg	C
	CUA leu	CCA pro	CAA gln	CGA arg	A
	CUG leu	CCG pro	CAG gln	CGG arg	G
A	AUU ile	ACU thr	AAU asn	AGU ser	U
	AUC ile	ACC thr	AAC asn	AGC ser	C
	AUA ile	ACA thr	AAA lys	AGA arg	A
	AUG met	ACG thr	AAG lys	AGG arg	G
G	GUU val	GCU ala	GAU asp	GGU gly	U
	GUC val	GCC ala	GAC asp	GGC gly	C
	GUA val	GCA ala	GAA glu	GGA gly	A
	GUG val	GCG ala	GAG glu	GGG gly	G