

# Aipotu III: Molecular Biology

## Introduction:

### The Biological Phenomenon Under Study

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.

### Scenario:

Recall from the Aipotu I&II lab exercise two weeks ago: you are the chief biologist for a breeder of fine flowers. Your company sells seeds to customers to 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. In Part II, you found the biochemical mechanism behind the colors. You also made 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 in this part.

## Tasks:

- 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.
- Explain, in terms of the genes present, the interactions between the alleles you found in part I.

Why is the color phenotype of some pigment proteins dominant while others are recessive?

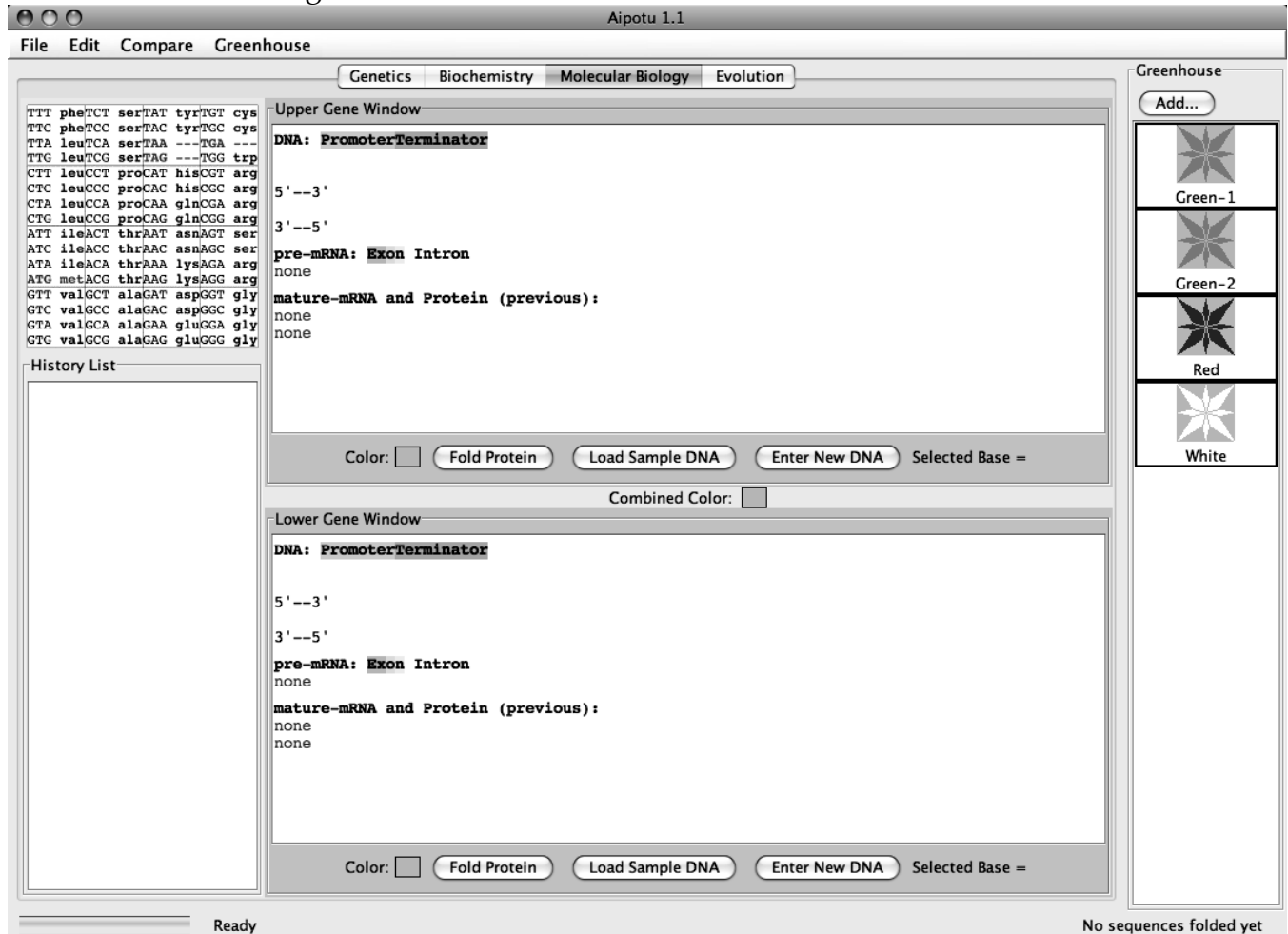
How do the pigment proteins combine to produce the overall color of the plant?

- Construct a purple protein to demonstrate your understanding of this process.

## Using the tool:

Launch Aipotu 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 software **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:

a promoter:

5' -TATAA-3'

|||||

3' -ATATT-5'

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

a terminator:

5' -GGGGG-3'

|||||

3' -CCCCC-5'

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

therefore, a gene would look like this:

5' -TATAAXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXGGGGG-3'

|||||

3' -ATATTXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXCCCCC-5'

\*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, 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.

1) *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) Double-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 Windows**.

### IMPORTANT NOTE:

This software is under development. Please treat it gently and be patient. Please report any bugs to your TA. **You should save your Greenhouse regularly, especially if you save a large number of organisms.**

The **Gene Windows** 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 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).
- 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.

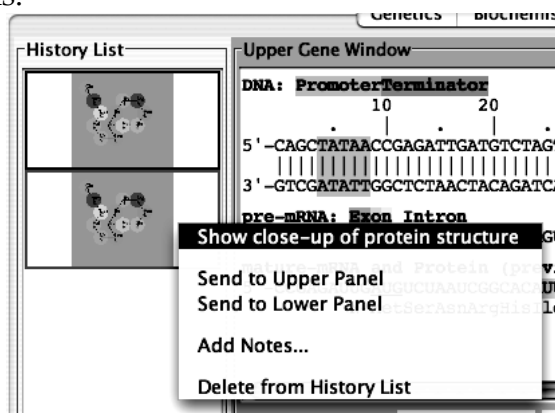
4) Click the “Fold Protein” button in the **Upper Gene Window** (or hit the return key) to fold the resulting protein. You will see the following:

- 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 Windows** 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 double-click an entry in the **History List**, you will you get a pop-up menu with a list of useful options:



- **Show close-up of protein structure:** this will pop up a window that shows a larger picture of the protein's 2-dimensional structure than is shown in the **History List**.
- **Send to Upper Panel:** Sends this **Tray** to the **Upper Panel** so you can cross those organisms.
- **Send to Lower Panel:** Sends this **Tray** to the **Lower Panel** so you can cross those organisms.
- **Add Notes...:** Allows you to add notes to the **Tray** in the **History List**. These notes will appear if you leave the cursor over the **Tray** for a few seconds.
- **Delete from History List:** deletes the **Tray** from the **History List**; this is cannot be undone.

⇒ You can also save snapshots of the **Gene Windows** to the clipboard or a file as you did in the Aipotu II lab.

### Specific Tasks for this section

- a) What are the differences in the DNA sequences of the alleles you defined in Part I.
- b) Do all the white alleles have the same DNA sequence? Hint: use the **Compare** menu to compare the sequences.
- c) Which DNA sequences are found in each of the four starting organisms?
- d) Using this knowledge, construct a pure-breeding purple organism.
- e) *Advanced tasks:* How does the DNA sequence of the different alleles explain the effects of mutations you found in part I?
- f) *Try making this protein:* MLVKEIAMYRFATHER ("M LVKE I AM YR FATHER" thanks to Grier Belter and Griffin Hancock from the Nova Classical Academy)

**You should put your data in the tables below:**

(a) and (b)

allele color

change(s) in amino acid sequence

change(s) in DNA sequence

Are all the white alleles the same?



(c) **Green-1**

**Green-2**

**Red**

**White**

(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.

## Follow-Up Assignment

Although you worked together as a class, each student must submit an individual assignment written in their own words.

Using the tools in the **Genetics Workbench**, make random mutants of one of the starting organisms in the **Greenhouse**. Select one with a color not found in the original **Greenhouse** organisms and save it to the **Greenhouse**.

You will use what you have learned in the three Aipotu labs to provide a complete explanation of why your mutant has the color it does. Specifically, you must answer the following questions as they apply to the mutant you created. *An example is given in italics; your answer should be similar, but more thoroughly explained.*

**1) Starting Organism.** Which organism from the **Greenhouse** did you start with? *For example, "We started with Green-1"*

**2) Resulting Color.** What was the color of your mutant organism? *For example, "We found a black mutant of Green-1"*

**3) Molecular Biology.** Your mutant strain contains two alleles of the pigment protein gene. For each allele, describe the change(s) in DNA when compared to the un-mutated starting sequences found in the organism you listed in (1). Then describe how each of these changes to the DNA sequence led to the resulting changes in protein sequence. *For example, "In the top allele, C47 had been mutated to G; this changed Arg17 to Glutamic acid. In the bottom allele, T12 had been deleted; this inactivated the promoter, resulting in no mRNA or protein being made."*

**4) Biochemistry.** Your mutant strain produces one or more copies of the pigment protein. For each, describe how the changes in amino acid sequence you described in (3) result in the color of each allele. *For example, "The top allele produced a black protein because..... Since the bottom allele produced no protein, it's contribution was colorless."*

**5) Genetics.** Explain why the two alleles you have described in (3) and (4) combine to form the color they do. *For example, "From our previous work, color is always dominant to colorless. From above, the genotype of our strain is black/white. Therefore, it will be black."*

## The Genetic Code in mRNA

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