

PSEUDOGENE SUITE
Lesson C: Exploring Primate Pseudogenes with Biology Workbench

Mary Ball and Steve Karr
Carson-Newman College

STUDENT HANDOUT

Introduction:

In the lessons on Vitamin C and Pseudogenes, you examined only a small section of several sequences that are available online in sequence databases. To analyze the huge volumes of sequences that are now available, special software is used to automatically align sequences and to color-code the alignments. Biology Workbench is a website that allows you to access sequences and use online software to compare them. In this activity you will use some of the tools available at the Workbench website to look at the GULO and Psi Beta data.

Information on the Sequences:

The following sequences will be available to you for exploration

Human GULOP (Exon10)
Macaque GULOP (Exon10)
Chimpanzee GULOP (Exon10)
Orangutan GULOP (Exon10)
Guinea pig GULOP (Exon10)
Rat GULO (Exon10)

Human Beta Globin
Human Psi Beta
Chimpanzee Psi Beta
Gorilla Psi Beta

GULOP is the notation for a GULO pseudogene, and **Psi Beta** is the notation for a beta globin pseudogene. The GULO sequence we will use is that for just one part of the protein-coding sequence, the part in Exon 10. The globin data is the coding sequence for the entire beta globin protein. (The species happen to be ones for which these sequences are available.)

Deciding on One or More Questions to Explore:

The guinea pig is another mammal that cannot synthesize Vitamin C even though it has a DNA sequence similar to that of the rat's functional GULO gene. You may wonder, "**How does the guinea pig pseudogene compare to the human pseudogene?**" A good question, and one you can answer with Biology Workbench. What other comparisons would you like to make?

Decide on two or more sequences to compare and list them on the **Report Form**. (You can do other comparisons later, so just pick something to start.)

PROCEDURE FOR USING BIOLOGY WORKBENCH TO COMPARE SEQUENCES:**A. Log Into the Student Interface to the Biology Workbench:**

1. Go to the **Student Interface to Biology Workbench** at:
<http://bsw.ncsa.uiuc.edu/cgi-bin/sib.py> (be sure to save this as a "favorite" or "bookmark" for future easy access).
2. To enter, **type your username and password**, then click on **Submit**.
(If you do not already have a username and password, please click on "REGISTRATION". This is a free account. Use your first initial, second initial, last name as your username, for example, **btsmith**, and enter the password **biology**, so the instructor can examine your account at a later date.)
This brings up the **Student Interface to the Biology Workbench (SIBW)** page, subtitled "Preferences Tools with Current Session: Default Session"

B. Upload the Primate Pseudogenes Data to Your Account:

3. Create a new session by clicking on **New**. (Wait for a response.) This takes you to the **NEW** screen.
4. In the **Session Description** box, type in **Primate Pseudogenes** and click on **Start New Session**. (Wait for a response.) This returns you to the SIBW page, but indicating Primate Pseudogenes as the Current Session.
5. Under your browser (e.g., Netscape, or Explorer) **File** pull-down menu, select **Open (Web) Location**. (A dialog box will open.)
6. In the dialogue box, enter: <http://deltas.animal.uiuc.edu/sibdoc/sequences/ensi.html> then click on **Open**. This opens the SIBW page for Importing assigned sequences into your Biology Workbench account. You will see the DNA sequences for all the primates in the collection.
7. Input your **Username** and **Password** and then click on **Upload Sequences**. (Wait for a response.) This will display a somewhat condensed version, showing all sequences in the set.
8. Click in the circle in front of **Primate Pseudogenes** and then click on **Transfer Sequences**. (Wait for a response.) This will bring up the **SIBW Nucleic Tools** page for the set of Primate Pseudogenes.

C. Aligning Two or More of the Sequences:

9. Scroll down to the bottom of the screen, where you will see a list of the organisms whose nucleic acid sequences were transferred. The names are not very descriptive, so they are listed here (below) in the **same sequence**, with details important to know about each:

User Human (GULOP) (Exon 10)	User Human (Beta Globin)
User Macaque (GULOP) (Exon 10)	User Chimpanzee (Psi Beta)
User Chimpanzee (GULOP) (Exon 10)	User Human (Psi Beta)
User Orangutan (GULOP) (Exon 10)	User Gorilla (Psi Beta)
User Guinea [-pig] (GULOP) (Exon 10)	
User Rat (GULO) (Exon 10)	

On the screen, **click** in the **box on the left of each of the entries** that you decided you want to compare, then **scroll back up** and click on **CLUSTALW**. (Be prepared to justify your choices.) This takes you to the CLUSTALW page (Multiple Sequence Alignment), on which the selected names will be listed. The actual selected sequences will be shown further down (scroll down), with the full identification given for each. The alignments are clearly indicated, and the different contiguous portions of the sequence are shown in separate sets below.

10. Scroll back up and **click** on **Import Alignment**. (Wait for a response.) This takes you to the SIBW Alignment Tools for your selected sequences.

D. Create a Shaded Alignment Chart:

11. Scroll down and **click** in the box on the left of the list of species whose DNA sequences are aligned (CLUSTALW - Nucleic), then scroll back up and click on **BOXSHADE**. (Wait for a response.) This produces a BOXSHADE page, on which portions of the aligned sequences are color-coded.

E. Examine the Results:

12. Those sites that are **identical** in all the species aligned are color-coded **green**. Where not all, but more than half, of the species match, the predominant amino acid is color-coded **yellow**. **Blue** means a mismatch, but one in which the blue-coded amino acid is similar to the more common, yellow one. **No Color** indicates a mismatch in which the amino acids are **not** chemically similar or a position at which there is a gap.

Thus, a green column indicates absolute identity across the species being compared and a column with only yellow and blue indicates strong similarity across species, but one with any uncolored letters indicates greater variability at that position.

13. How similar are the sequences you compared? Were you surprised? What other comparisons would you like to make? Can you figure out a way to quantify the differences? Complete the **Report Form**.

Note: Since **BOXSHADE** plots cannot be saved in your **Workbench** account, you may want to print out the plot on a color printer or save it to a disc or your hard drive. **To print the plot, “right click” (or double-click) in the plot window and select View Image. When the image appears, select Print. To save the image as a GIF image, select Save Image As, [give it a name], and click on Save.**

F. Additional Explorations:

Clicking the **Return** button in the **BOXSHADE** window takes you back to the **Alignment Tools** window. (Wait for a response.)

Try using the other **Alignment Tools**, and discuss what they seem to suggest about your selected sequences and the species selected: **DRAWTREE**, **DRAWGRAM**, and **CLUSTALDIST**.

To get back to the list of nucleic acid sequences, click on the **Nucleic Tools** button. (Wait for a response.)

Select and compare a different set of sequences. Complete a different **Report Form** for each set of sequences that you explore.

To Delete a Session (or a particular set of sequences):, go back to the **Preferences Tools** page, click in the circle of the session to delete (be careful! It's permanent), then click on **DELETE**. Then you can begin again, as described from step one.

Report Form
(Attach a copy of the BOXSHADE plot if possible.)

Researchers:

Sequences Compared:

Alignment Tools Used:

Reasons for Comparing These Sequences (and using selected tools):

Observations:

Conclusions: