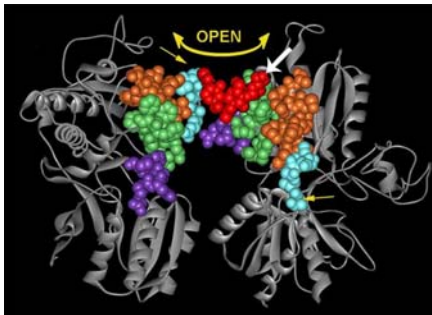


Can You Taste That?

Predicting PTC Tasting Ability among Primates Instructions



Two copies of the TAS1R3 protein (taste receptor family 1, member 3.) form a dimer and bind to sugar (red molecule in center). Like PTC tasting and the TAS2R38 protein, mice are either sweet “tasters” or “nontasters” based on their T55A and I60T genotype. Source: Genome News Network.

Researcher Background:

How and why do some people like certain foods and not others?

What we perceive as “taste” is actually a result of complex interactions among the signals that our brain receives from our nose (i.e., “smell”) and the signals that our brain receives from the **taste receptors** located in taste buds in our tongue. Taste receptors are proteins that recognize substances in food. When one of these substances (such as a sugar molecule in the structure shown in the box on the left) binds to the taste receptors, it causes a **conformational change** in the receptor that ultimately results in the brain perceiving what has been eaten.

Taste receptors are divided into two families:

Type 1 taste receptors recognize **sweet** and include receptors TAS1R1 through TAS1R3

Type 2 taste receptors recognize **bitter** and include receptors TAS2R1 through TAS2R50, and TAS2R60

These receptors mix and match with each other and, combined with distinct proteins that detect hot, cold, salty, sour, and umami, as well as smell, mouth feel and visual perceptions of food, all contribute to what we perceive as “taste.” Taste helps to identify flavors of food (i.e., nutrition). They also help us to identify **toxins**, which many perceive as being bitter.

Taste receptor 2 member 38 (TAS2R38) is a protein that in humans is encoded by the *TAS2R38* gene. TAS2R38 is a bitter taste receptor, with different genotypes associated with the ability to taste 6-n-propylthiouracil (PROP) and phenylthiocarbamide (PTC). Both are bitter tasting compounds. While scientists have long hypothesized that different taste receptor genotypes could influence tasting ability, TAS2R38 is one of the only taste receptors shown to have this function. TAS1R3, shown in the box above, is one of the others.

In humans, the dominant form of the *TAS2R38* gene, “T,” codes for the ability to taste bitter substances like PTC. The human taster “T” phenotype is associated with the sequence “PAV” – or more specifically P49, A262 and V296: human tasters have a P (Proline) at position 49, an A (Alanine) at position 262, and a V (Valine) at position 296. Non-tasters have the genotype “tt” that contains “AVI” at these same positions.

Research Question: The Basic Local Alignment Search Tool (BLAST) is a bioinformatics tool that scientists use to study and compare DNA and protein sequences. You will use BLAST to compare the TAS2R38 protein sequences from humans and non-human primates to answer the following research question:

- Are primate species other than humans “tasters” or “non-tasters”?

INSTRUCTIONS

Aims:

- To compare TAS2R38 protein sequences from various primate species to predict PTC tasting ability
- To learn how to use the bioinformatics tool BLAST (Basic Local Alignment Search Tool)

Notes:

- All protein sequences are presented using the **one letter abbreviations for amino acids**. These abbreviations can be found in the “**Codons and Amino Acid Chemistries**” handout on the “PTC Tasting Ability among Primates” webpage described in Step 1.
- Answers requiring written answers have a writing pad icon beside them. Answers can be written on your accompanying worksheet.

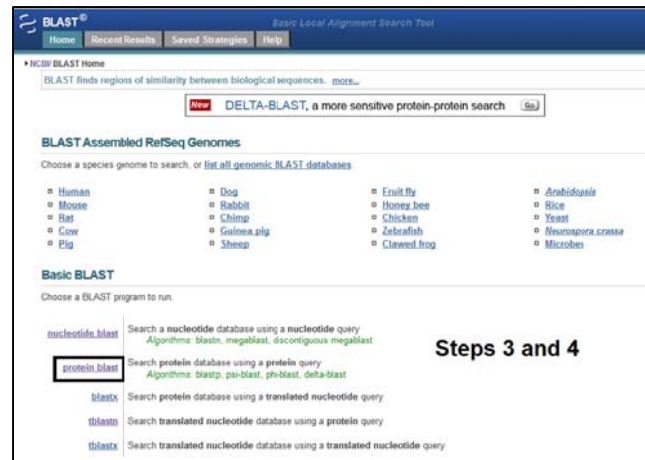


PART I: Aligning DNA Sequences

1. Complete **Section 1** on your “Predicting PTC Tasting Ability among Primates” worksheet.

2. To complete **Section 2** of your workshop, in a new internet web browser window, go to the “PTC Tasting Ability among Primates” webpage at Digital World Biology:
<http://digitalworldbiology.com/dwb/abe/ptc-tasting-ability-among-primates>

3. In another new web browser window, go to the BLAST homepage at the National Center for Biotechnology Information (NCBI):
<http://blast.ncbi.nlm.nih.gov/Blast.cgi>



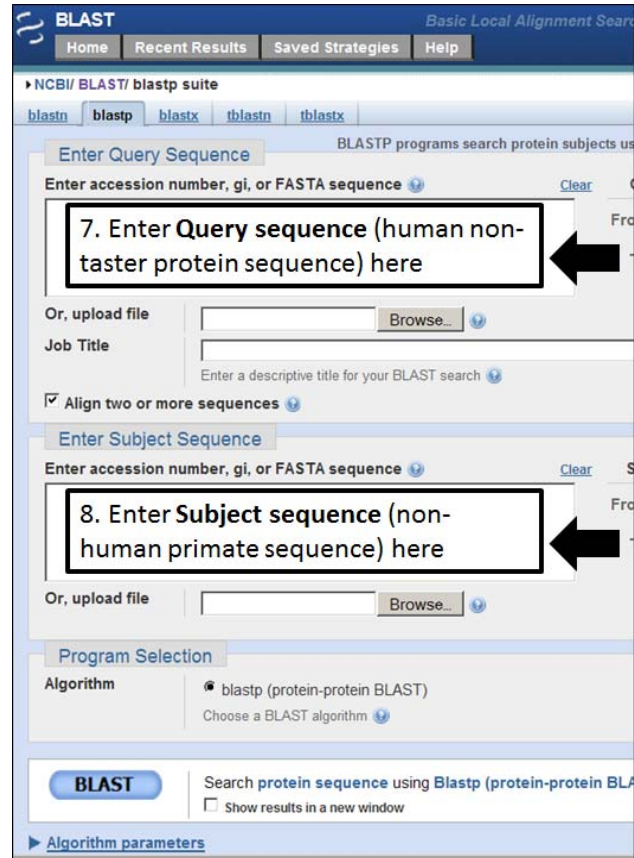
4. Select “**protein blast**,” as shown in the image to the right. We will use a protein BLAST (or “**blastp**”) because we will be comparing a protein sequence (sequence of amino acids) to another protein sequence. Note that there are options for other types of BLAST, such as for comparing DNA sequences to DNA sequences.

5. From the protein blast page (image, right), click the box to choose the option to “**Align two or more sequences.**”

6. A second text box will appear.



7. From the “PTC Tasting Ability among Primates” webpage, click on the **human non-taster protein sequence** link to open the sequence file. Copy and paste the “**Human Non-Taster Protein Sequence**” into the top BLAST box (the “**Query Sequence Box**”). Be sure to include the caret (“>”) symbol and the sequence name. This is the TAS2R38 protein sequence from a human that cannot taste PTC (“non-taster”). **Note:**
- To highlight all use Control + A [PCs] or Command + A [Macs].
 - To copy use Command + C [PCs] or Command + C [Macs].
 - To paste into BLAST use Control + V [PCs] or Command + V [Macs].

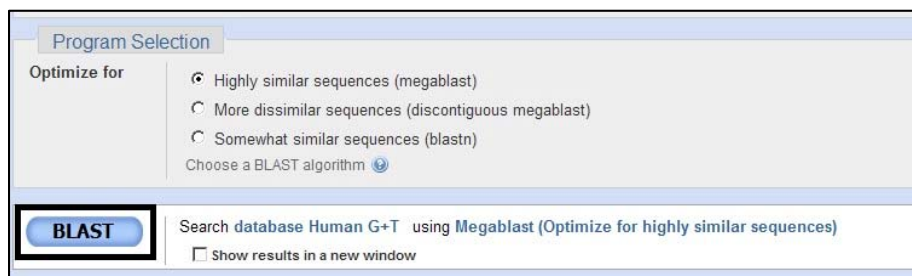


A **query sequence** (sometimes called a “reference sequence”) is the sequence to which other sequences are compared when performing a BLAST alignment. One can think of this as the independent or control variable. In this experiment, we are comparing a non-human primate sequences **to** the human sequence, so the **human sequence is our query sequence**.

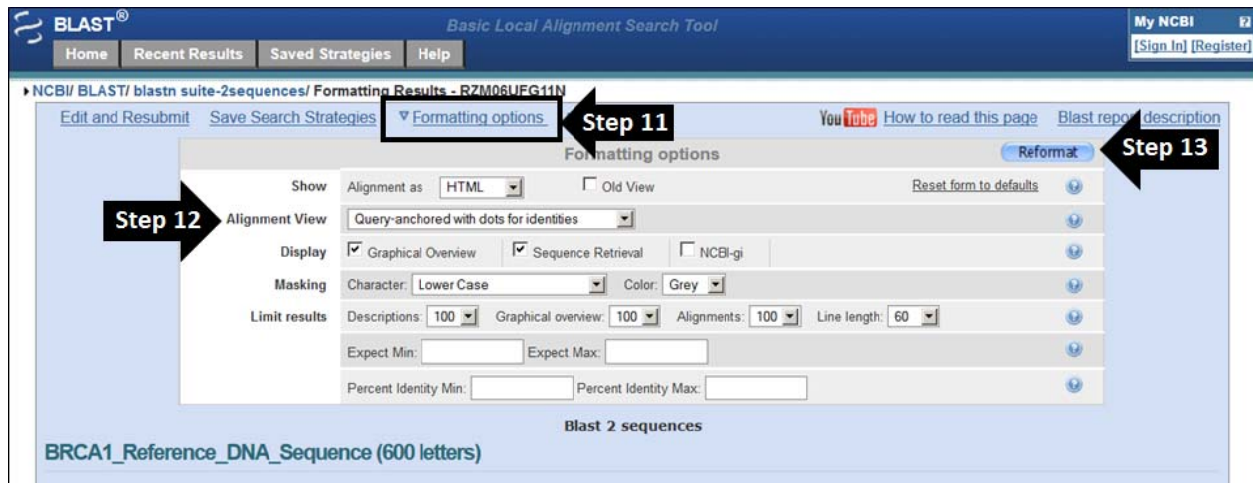
8. From the “PTC Tasting Ability among Primates” webpage, click on the sequence link for the non-human primate that you are studying, then copy and paste the non-human primate PTC protein into the bottom BLAST box (the “**Subject Sequence Box**”). Be sure to include the caret (“>”) symbol and the sequence name. This is the TAS2R38 protein sequence from a non-human primate.

A **subject sequence** is the sequence being compared when performing a BLAST alignment. One can think of this as the dependent or experimental variable. In this experiment, a non-human primate sequence is being compared to our human sequence, so the **non-human primate sequence is our subject sequence**.

9. Click “**BLAST.**”



- When your search is complete, you will see a window with the BLAST results showing an alignment of the two protein sequences that you entered in steps 6 and 7.



- Click the “**Formatting Options**” link located near the top of the page.
- Find the “**Alignment View**” and use the drop-down menu to choose “**Query-anchored with dots for identities.**”

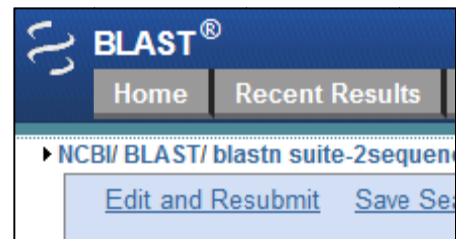
The **query is the human sequence**. The query-anchored view shows the human sequence at the top with the subject sequence (your non-human primate sequence) aligned below it. Dots are used to show amino acids that are identical in both sequences and letters are used to show amino acids that differ.

- Click the “**Reformat**” button.
- Scroll down the page and look at your alignment between the human and non-human primate sequences. The numbers on either side of the alignments indicate amino acid positions in the alignment. The human taster phenotype is associated with the sequence “PAV” – or more specifically P49, A262 and V296. In other words, human tasters have a P (Proline) at position 49, an A (Alanine) at position 262, and a V (Valine) at position 296. Non-tasters are associated with “AVI” at these same positions. This is also summarized in the table below.

	Amino Acid Position		
	49	262	296
Taster:	P	A	V
Non-Taster:	A	V	I



- Based on your analysis of the TAS2R38 protein sequence from a non-PTC tasting human and your non-human primate species, do you predict that your non-human primate species is a “taster” or a “non-taster”? **Explain or justify your answer on your worksheet.**



16. To further support your prediction in step 15, repeat steps 3 through 14 using the human taster protein sequence. You can quickly return to the BLAST homepage by clicking the “Home” button in the upper left corner of your screen.



17. Does your analysis of the TAS2R38 protein sequence from a PTC tasting human and your non-human primate species support the prediction that you made in step 15? **Explain or justify your answer on your worksheet.**

18. BLAST can also be used to compare many sequences at the same time. Return to the BLAST homepage and select “Protein BLAST.”

19. Click the box to “Align two or more sequences.”

20. From the “PTC Tasting Ability among Primates” webpage, click on the link to and open the file “**All TAS2R38 Protein Sequences.**” You may need to save this file on your desktop to open it.

21. Copy and paste the **human non-taster protein sequence** into the top **Query Sequence** box. Be sure to include the caret (“>”) symbol and the sequence name.

22. Copy and paste **all** of the remaining TAS2R38 protein sequences (listed under the heading “Subject Sequences”) into the bottom **Subject Sequence** box. Be sure to include the caret (“>”) symbols and the sequence names.

23. Click **BLAST**.

24. Reformat your BLAST results as you did in steps 10 through 12 above.

25. Scroll through your sequence alignments as you did in step 14 above and **answer the following questions on your worksheet:**



a. Which sequence is your reference sequence [Step 21]?



b. Fill in the **Results Summary Table** using the information in your BLAST results.



c. Based on the protein sequence data in your BLAST alignment, do you predict that the majority of the non-human primates studied here are PTC tasters or non-tasters? Explain or justify your answer on your worksheet.



d. Are there any non-human primate species that do not fit into either the “PAV = taster” or “AVI = non-taster” phenotype categories? If so, what sequence do they have at these three key sites? What do you predict their phenotype to be?



e. Find at least three other amino acid sites where the human sequences differ from the majority of the other non-human primate sequences and list them on your worksheet.

26. In addition to comparing two or more sequences to each other, **BLAST scores** help us *quantify* BLAST results, explained in detail on the next page.

Sequences producing significant alignments:

Select: All None Selected:0

Alignments Download Graphics Distance tree of results Multiple alignment

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/> Human bitter taste receptor TAS2R38	663	663	100%	0.0	99%	Query_32253
<input type="checkbox"/> Chimpanzee taste receptor TAS2R38	657	657	100%	0.0	98%	Query_32254
<input type="checkbox"/> Bonobo taste receptor TAS2R38	654	654	100%	0.0	98%	Query_32255
<input type="checkbox"/> Gorilla taste receptor T2R38	649	649	100%	0.0	97%	Query_32256
<input type="checkbox"/> Bornean Orangutan taste receptor T2R38	639	639	100%	0.0	95%	Query_32257
<input type="checkbox"/> Sumatran Orangutan taste receptor TAS2R38	627	627	100%	0.0	94%	Query_32258
<input type="checkbox"/> Siamang Money taste receptor TAS2R38	640	640	100%	0.0	95%	Query_32259
<input type="checkbox"/> Rhesus Macaque taste receptor T2R38	637	637	100%	0.0	95%	Query_32260
<input type="checkbox"/> Japanese Macaque taste receptor TAS2R38	636	636	100%	0.0	95%	Query_32261
<input type="checkbox"/> Olive Baboon taste receptor TAS2R38	633	633	100%	0.0	95%	Query_32262
<input type="checkbox"/> Hamadryas Baboon taste receptor TAS2R38	635	635	100%	0.0	95%	Query_32263

Ordered List of Subject Sequences and corresponding Accession ID in the far right column

In the summary data table (above), BLAST lists all of the Subject Sequences you entered in the order that you entered them, and then assigns each sequence a temporary “**Accession**” or ID number, which are used in the BLAST alignment.

The **Max score** and **Total score** are related to the length of the sequences compared. Generally, the higher the score, the better the two sequences match each other. These scores are particularly helpful when comparing multiple sequences to each other.

Query coverage & Percent identity quantify how much of the sequences match, and how well they match. For example, a small portion of the sequences (for example, *25% query coverage*) may match well (*100% identity*). Alternatively, 100% of the sequences may align to one another, but might share only 50% of the same nucleotides (*50% identity*).

The **e-value** or **expect value** is an indication of how likely these results are based purely on chance. Similar to a p-value, low e-values mean you can be more confident in your results.

EXAMPLES:

30% Query Coverage, 100% Identity
3/10 bases (30%) match perfectly (100%)

```
ATGGATACGT
TGAGATGATC
```

100% Query Coverage, 50% Identity
All 10 bases (100%) align, but only 50% match

```
ATGCCGATTG
AGGGCAACAG
```

The “Query-anchored with dots for identities” BLAST alignment would look like this, with a dot in the subject sequence at each position where it matches the query sequence:

ATGGATACGT
TGA•••GATC

ATGCCGATTG
•G•G•A•CA•


27. Click the “Ident” link at the top of the list of BLAST results to sort your BLAST results by Percent Identity. This will result in a list BLAST results in which the sequence most similar to the human non-taster sequence will be at the top of the list and the sequence most dissimilar to the human non-taster sequence at the bottom of the list.


Sequences producing significant alignments:

Select: [All](#) [None](#) Selected: 0

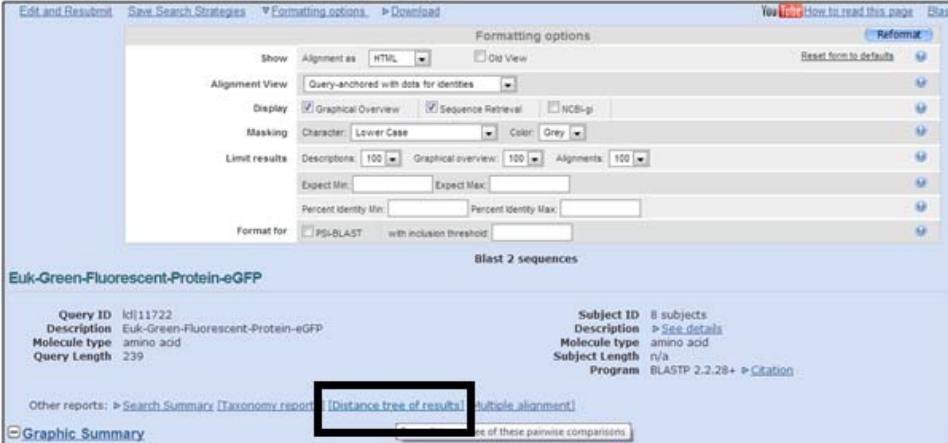
[Alignments](#) [Download](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#)

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/> Human bitter taste receptor TAS2R38	663	663	100%	0.0	99%	Query_168808
<input type="checkbox"/> Chimpanzee taste receptor TAS2R38	657	657	100%	0.0	98%	Query_168809

28.  Based on your BLAST scores, which three non-human primate sequences are most closely related to the human query sequence? **Justify your answer on your worksheet using max score, query coverage, and percent identity.**

29.  Based on your BLAST scores, which three non-human primate sequences are most distantly related to the human query sequence? **Justify your answer on your worksheet using max score, query coverage, and percent identity.**

30. Another way to visualize your results is by using a **distance tree**. Similar to a phylogenetic tree, a distance tree is a graphical representation of relationships – in this case, which sequences are most similar to one another. BLAST can perform this analysis for you using the comparisons that you have already made. To view your tree, click the “**Distance Tree of Results**” link located near the top of the page just below the formatting options menu.



Format for: PSI-BLAST with inclusion threshold:

Euk-Green-Fluorescent-Protein-eGFP

Query ID	kl 11722	Subject ID	8 subjects
Description	Euk-Green-Fluorescent-Protein-eGFP	Description	See details
Molecule type	amino acid	Molecule type	amino acid
Query Length	239	Subject Length	n/a
		Program	BLASTP 2.2.28+ Citation

Other reports: [Search Summary](#) [Taxonomy report](#) **[Distance tree of results](#)** [Multiple alignment](#)

[Graphic Summary](#) [View all pairwise comparisons](#)

31. Distance trees are **models**, or a specific way to visualize and analyze data. As scientists, you must consider different tree formats different tree variables or **parameters**:

- **Rectangle, slanted, radial and force:** These are different ways to draw the same tree
- **Tree Method:** Scientists choose which statistical method (or **algorithm**) is used to compare protein sequences to one another and create the tree.
- **Maximum Sequence Difference:** Scientists choose how dissimilar subject sequences can be from the query sequence and still be included in the tree. Larger numbers allow inclusion of more dissimilar sequences in the same tree.
- **Distance:** Scientists choose which method to use to predict evolutionary relationships by considering things like: How many codons are used for a specific amino acid? How likely (or **probable**) is a given nucleotide change in the corresponding codons? For example, how often in nature does a guanine (G) change to a cytosine (C), resulting in a codon change from CCA=Proline to GCA=Alanine?

The screenshot shows the 'Blast Tree View' interface. At the top, there are links for 'Reset Tree' and 'Download Tree'. The 'BLAST RID' is 'GPM0GV3M11N' and the 'Query ID' is 'Icl|Q'. The 'Tree method' is set to 'Fast Minimum Evolution'. The 'Max Seq Difference' is set to '0.85'. The 'Distance' is set to 'Grishin (protein)'. Below these settings are four buttons for tree visualization: 'rectangle', 'slanted', 'radial', and 'force'. A checkbox labeled 'Show distance' is checked.

32. Experiment with various distance tree parameters until you obtain a tree that you believe is the best visualization of your results. Then, **either download the tree as an image** file by clicking the “**Download Tree**” link **or draw a sketch of your tree on your worksheet**. You do not need to include all of the branches in your tree, but your tree should include the positions of your non-human primate species, the species closest neighbors to your sequence, and how far these sequences are from the reference (query) sequence, which is highlighted in yellow.



33. Based on what you have learned in this activity, which do you think is the more ancient or **ancestral** phenotype: being a taster or a non-taster? In other words, do you think that the common ancestor of all of these primates was a taster or a non-taster? **Explain or justify your answer on your worksheet.**

