**Verifying Bases of Interest**

1. Open two files using **ApE: BlockA\_Close\_Primer.ab1** and **BlockA\_Far\_Primer.ab1**. Adjust the Y and X sliders on the far right side until the tallest peaks are just touching the letters (Y slider) and the peaks are nicely separated as shown below (X slider). These are two different sequencing reactions of the exact same clone of DNA.



1. When you get your sequences back, you might see some N bases. The software inserts an N whenever it cannot tell for sure which base should go there. The software that “calls” each base uses two criteria: phred score of quality and spacing between bases. Phred scores can be seen as the light bar graphs behind the colored lines. Good scores result in a bar graph about half way up the sequence box. Each base is associated with a different color line.
2. Under the edit menu, choose “Find”. Check the box next to “Also find rev-com of string” so that you search this strand and the complementary strand in the reverse (5’ to 3’) orientation. Also check the “Wrap” box so that you can search upstream and downstream of wherever your courser is currently located.
3. Search for this sequence “actaggggcccaagttca” in both of the files you have open. ApE will highlight the sequence if it exists in your file.
4. In both files, look immediately upstream of the highlighted sequence for the five red (T) peaks in a row. Write down the bases the sequencing software assigned to these 5 peaks and the next few peaks upstream of the highlighted sequence.
5. Which sequence is correct and how do you know? Be able to explain why the software was unable to correctly assign the bases in the file that you determined was wrong.
6. Now you can close these two files and prepare for the second half of your guided tour.

**ApE TAS2R38 Guided Tour**

 By now you should have read about the function of TAS2R38 protein and the two most common protein variants (PAV and AVI; see Wikipedia for a summary). Based on the phenotype and genotypes of many people, one allele is often called a taster allele while the other is called a non-taster. However, taste perception is not completely determined by one protein and therefore the real world phenotypes can be more complex than simply yes or no. There are super tasters and there are mild tasters. In this lab, you will determine your genotype and phenotype and see if you conform to typical outcomes or not. This Guided Tour will help you use ApE (A plasmid Editor; [biologylabs.utah.edu/jorgensen/wayned/ape/](http://biologylabs.utah.edu/jorgensen/wayned/ape/) ) to evaluate DNA sequences coming from homozygotes for the four SNPs listed in Table 1 below. Tables 1 and 2 will help you understand how a single nucleotide polymorphism can affect your ability to collect information from your environment. But first, visit the web site (<http://www.bio.davidson.edu/113/TAS2R38_ORF.html>) and answer the two questions on that page. You will need Tables 1 and 2 to help you answer these questions.

**Table 1.** SNPs and corresponding amino acids in four common alleles for TAS2R38. Dots between letters denote the letters are not located next to each other.

|  |  |  |
| --- | --- | --- |
| **DNA SNPs** | **Protein Amino Acids** | **Taster Status** |
| C.C.G | P.A.V | “yes” |
| G.T.A | A.V.I | “no” |
| C.T.G | P.V.V | ? |
| G.C.A | A.A.I | ? |

**Table 2.** Codons and the SNPs (blank spot) they contain.

|  |  |  |
| --- | --- | --- |
| **Codon Number** | **Codon Sequence** | **SNPs** |
| 49 | \_ C A | C or G |
| 262 | G \_ T | C or T |
| 296 | \_ T C | G or A |

1) Launch ApE by opening the file called TAS2R38\_CCG.ape. If your computer does not have ApE, you will need to download it (free for Mac and Windows computers). Next, open the file TAS2R38\_GTA.ape. Use the “Align Two Sequences…” command under the “Tools” menu at the top. Choose the two different sequences to align and click OK. You will see the two sequences aligned and a red hash tag # everywhere there is a SNP. Locate the three critical SNPs and verify that one allele contains the **C C G** SNPs and the other contains **G T A** SNPs.

2) Click on the CCG sequence and make sure your cursor is in front of the first letter. Under the “ORF” menu at the top, choose “Find Next” to locate the largest open reading frame in this sequence. You should see a large portion of the sequence is now highlighted. What are the first three letters? Which amino acid is encoded by this first codon of the ORF?

3) Under the “ORF” menu, use the “Translate” command to produce the protein encoded by the TAS2R38 mRNA. You will get an options window. Leave the default settings except choose to show line number on both sides and display the DNA below the amino acid sequence (see screen shot below).

4) Find amino acids 49, 262 and 296 to verify that CCG encodes PAV. You should repeat this ORF translation for the GTA allele as well.

5) Now open these three additional sequence files in this order: TAS2R38\_CTG.ape, TAS2R38\_GCA.ape and TAS2R38\_8green.ape (sequence from an unknown person). Edit the TAS2R38\_8green.ape sequence by deleting the first 24 bases (mostly Ns) to the left of the TAG sequence.

6) Under the “Tools” menu, choose “Align Sequences…”. Choose TAS2R38\_CCG.ape as the reference gene using the drop down menu. In the “Align to Windows” field, select all 5 sequence files by clicking on the top one (should be TAS2R38\_CCG.ape), hold down the shift key, and click on the bottom one (should be TAS2R38\_8green.ape). Click on the button “Show Alignment Parameters and set them as shown below. This will cause your sequence to not be split into small fragments. Click OK.



7) In the resulting window, you should be able to find the position of all three SNPs. Determine the genotype of the unknown 8green sample at all three positions.

8) Now you want to open the chromat file called TAS2R38\_8green\_chromat.ape. Adjust the X and Y sliders so that the peak heights nearly reach the letters and the spacing between peaks is big enough to clearly see each peek.



9) Under the “Edit” menu, choose “Find…” and type in the bases that appear on your multiple sequence alignment immediately after the first SNP (CACTGAG). Hit “find next” and the chromat should highlight the bases you asked the software to find. Look immediately to the left of the highlighted sequence. Compare the peaks at this SNP site to the nearby peaks. You should notice that this SNP peak looks different. Explain what you are looking at and what this indicates about the person’s genotype for TAS2R38 at codon number 49. Go find the other two SNP locations using the Find function and the appropriate DNA sequences by looking at the non-chromat ApE files. Is this person heterozygous at all three SNPs? Explain what you found.

10) Now you are able to analyze your personal DNA sequence when it arrives. Your goal is to determine your genotype for all three SNPs, and the corresponding three amino acids. Predict whether you a taster, a non-taster, or something else. Are you homozygous at all three SNPs, or do you have at least one heterozygous SNP as you found in TAS2R38\_8green\_chromat.ape.

11) Once you have a prediction, get a test strip of PTC and record your response. Emergency chocolate is on hand should you be a super taster.