Lab 1: Genome Browsers

During the Human Genome Project, researchers realized that a long string of A,T,C, and G was totally unintelligible to people. Therefore, Jim Kent (a biology PhD student) and David Haussler (the a CS professor, now in Biomolecular Engineering) at UCSC developed a way to view, search, and visually interpret the Human Genome. This tool is the UCSC Genome Browser. Since then many other browsers have been developed. Most are essentially a direct descendent of this first browser.

In this lab you will become accustomed to looking at genes in this genome browser, and look at a few others too. We'll also start to look ad differences across human populations.

Section 1: Genetics of Taste

We are going to start by exploring the genetic variation of members of this class.

Experimental Protocol

Each of you has 4 strips (1-4).

1. Rinse your mouth with water prior to starting the experiment. Taste the strip labeled 1 and classify your bitter perception as taster, mild taster or nontaster. Record your result.
2. Rinse your mouth with water prior to starting the experiment. Taste the strip labeled 2 and classify your bitter perception as taster, mild taster or nontaster. Record your result.
3. Rinse your mouth with water prior to starting the experiment. Taste the strip labeled 3 and classify your bitter perception as taster, mild taster or nontaster. Record your result.
4. Rinse your mouth with water prior to starting the experiment. Taste the strip labeled 4 and record whether you perceive bitter, salty and/or sweet (taster). You may also taste nothing (nontaster). Record whether you are a taster or nontaster; if you are a taster describe the flavor. Keep in mind that you may taste one or all three of the sensations.

Individual Results:

Strip 1: _____

Strip 2:_____

Strip 3:_____

Strip 4:_____

If taster, do you taste bitter?_____

If taster, do you taste salty?_____

If taster, do you taste sweet?_____

Class Results:

What percentage of the class tasted strip 1?

What percentage of the class tasted strip 2?

What percentage of the class tasted strip 3?

What percentage of the class tasted strip 4?

Of the tasters, what percentage tasted sweet?

Of the tasters, what percentage tasted bitter?

Of the tasters, what percentage tasted salty?

Analysis:

1. What do you think strip 1 was?
2. Of those students who tasted strip 2, were they more likely to taste strip 3?
3. Of those students who tasted strip 2, were they more likely to think strip 4 tasted bitter?
4. Do you think the same protein that recognizes and binds to the chemical on strip 2 can recognize and bind to the chemical on strip 3 or strip 4? Why or why not?

Genetics

Strip 2 is imbued with the chemical phenylthiocarbamide, also known as phenylthiourea. The structure can be seen in Figure 1A. This chemical is perceived as bitter, sometimes very strongly so, by individuals that have particular genetic changes in the gene *TAS2R38*[1]. This gene encodes for a G-protein coupled receptor associated with taste. Let's look at it in the genome browser.

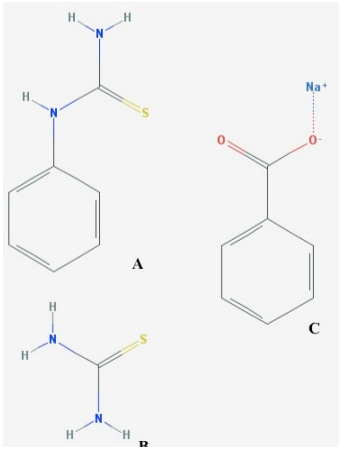

Figure 1: Structures of chemicals: A) phenylthiocarbimide, B) thiourea, and C) sodium benzoate. All structures from Pubchem (pubchem.ncbi.nlm.nih.gov)

1. Google for UCSC Genome Browser[2], the UCSC Genome Browser should come up first. The url is: genome.ucsc.edu

2. In the genomes tab, select the human build with the largest number; that will always be the most recent one.

3. Type the name of the gene of interest (TAS2R38) into the text box.

4. Click on the gene under known genes.

5. Let's configure the tracks to be useful for this exercise.

a. Change OMIM alleles to pack
b. Change Human mRNAs to squish
c. Change Conservation to squish
d. Change Common SNPs to pack
e. Change RepeatMasker to dense
f. Push any of the "Refresh" buttons
g. Move SNP track up to below OMIM track

6. By zooming in and out on the gene, determine the following:
    a. What chromosome is this gene on
    b. Which strand the coding sequence is on (5'→3' toward the centromere or toward the telomere)
    c. How many exons and introns are in this gene
    d. How many common SNPs there are in this gene
    e. If these SNPs are correlated with a human phenotype
    f. How many other genes are nearby (within 10kb)

7. By clicking on the SNPs, we can get information about them. For each of the common SNPs, determine the following:
    a. If the chimp allele different than the human allele
    b. If the SNP causes a change in the protein coding
    c. What the minor allele is and how common it is (%)
    d. If this SNP is associated with a human phenotype

Experiments in large populations to determine the genetics of PTC tasting have determined that three positions in the protein determine ones ability to taste, residues 49,262, and 296. Those with a haplotype AVI/AVI are non-tasters and those with PAV/AVI or PAV/PAV are[3]. Given your tasting profile, what would you predict your haplotype to be?

Are there differences in this gene across ethic populations? To find out, we'll use a different genome browser – ENSEMBL[4].

1. Copy the SNP name of the allele you want to look at.
2. Go to ensemble.org
3. In the search tool, input the SNP name and select human as the species
4. Using the Population Genetics tool, identify if there is a difference in ethnic groups.
5. Hypothesize about how these differences might have arisen.

Little is known about the genetics of the tasting of thiourea or sodium benzoate (strips 3 and 4). Given the data from the class, do you think *TAS2R38* is likely to play a role? What about the data supports that conclusion?

<u>Section 2: Other genetic characteristics</u>

Let's look at a few genes that are related to diseases or interesting phenotypes.

LCT:
This gene is involved in digesting lactose in milk in our body. There are known mutations upstream of the *LCT* gene that promote continued mRNA production into adulthood. This gives a person the ability to digest lactose even as an adult, called lactose persistence phenotype. Without this mutation, you will have to drink lactose–free milk! (You may need to change the zoom to find these mutations.)

BRCA1:
Unlike cystic fibrosis, breast cancer susceptibility can be influenced by mutations in many different genes. *BRCA1* is one of the best–known breast cancer susceptibility genes.

IL4R:
Malfunctioning of the immune response causes many diseases, like asthma. *IL4R* is involved in controlling the immune system, and mutations in this gene can cause asthma.

For each of these genes, identify the following:

a) The chromosome it is on, and how many introns and exons are in the gene
b) The coordinates of the beginning and ending of the gene
c) If the gene is on the 5'→3' into the centromere or out toward the telomere
d) Identify at least 1 SNP that is related to a known phenotype or disease characteristic. For this SNP, determine which allele is the minor allele and how common the minor allele is.
e) Determine if there are population differences in the occurrence of this SNP.

<u>Section 3: Explore gene of your choice</u>
Identify a gene you want to know more about. Answer the flowing questions:

1. How did you pick this gene? Why are you interested in learning more about this gene?
2. Where is this gene in the human genome?
3. Are there SNPs that correlate with human phenotypes?
4. How prevalent is the phenotype causing allele? Does this vary across populations?

Instructions for Analysis report

The report should have 3 sections, one for each section of the lab assignment. All sections should be in paragraph form with answers to the questions posed contained within. This report is worth 64 points.

Section 1 should have 3 subsections. 1) A results section that presents, visually the results from the class taste tests. 2) An analysis section should include information on the correlations (if any) between the different tastes as well as information on the gene *TAS2R38.* 3) A genetics section that describes the position of the gene on

the chromosome, the number of exons, the number of common SNPs and the phenotypes associated with those. It should also include information on the commonness of these alleles, the population genetics, and a hypothesis about what your haplotype is.

Section 2 should contain a summary of the three genes explored. This can be represented visually if you would like. It should contain answers to all the posed questions.

Section 3 should contain an explanation of the gene you chose to explore as well as information about mutations in the gene and population differences.

This report should be no longer than 6 pages. **It is due on Fri. April 7 at 11:59 PM**. It should be submitted as a pdf through Canvas.

References:

1. Conte, C., Ebeling, M., Marcuz, A., Nef, P. & Andres-Barquin, P. J. Identification and characterization of human taste receptor genes belonging to the TAS2R family. *Cytogenet. Genome Res.* (2002). doi:10.1159/000068546
2. Kent, W. J., Sugnet, C. W., Furey, T. S. & Roskin, K. M. The Human Genome Browser at UCSC W. *J. Med. Chem.* **19,** 1228–31 (1976).
3. Kim, U. *et al.* Positional Cloning of the Human Quantitative Trait Locus Phenylthiocarbamide. **299,** 1221–1226 (2003).
4. Yates, A. *et al.* Ensembl 2016. *Nucleic Acids Res.* **44,** D710 (2016).