Exercises for the UCSC Genome Browser Introduction

1) Find out if the mouse Brca1 gene has non-synonymous SNPs, color them blue, and get external data about a codon-changing SNP.

Skills: basic text search; Genome Viewer pulldown menus; filters; links to external resources

2) Find the protein sequence for rat leptin. BLAT this sequence vs. the human genome to find the human homolog. Look for SNPs in the coding region of this gene—are there any? Obtain the human DNA sequence for this region, and underline the SNPs.

Skills: obtaining protein sequence; BLAT; finding SNPs in exons; "get DNA" sequence with extended case/color options

3) Find the genomic region for the human NRAS [neuroblastoma RAS viral (v-ras) oncogene homolog] gene. Look for ESTs that were found in neuroblastoma tissues—by using the filter to color those blue in the genome viewer. When you are convinced that this protein is found in neuroblastoma tissue, examine the properties of this protein with the Proteome Browser (which is found as a link from Known Gene details pages). What are the pI, mw for this neuroblastoma protein?

Skills: examining ESTs; using a filter to color specific items; finding protein properties with the Proteome browser.

4) Perform an in silico PCR, to see what happens when more than 1 PCR product may arise. Determine the product sizes, and the melting temperatures of these primers.

Skills: in silico PCR of genomic sequence; finding product size and Tm.

5) Use the VisiGene interface to obtain gene expression data for a gene of interest.

Skills: find gene expression images; link back to the Genome Browser and outside sources.

Step-by-Step instructions for the UCSC Genome Browser exercises

1. Find out if the mouse Brca1 gene has non-synonymous SNPs, color them blue, and get external data about a codon-changing SNP.

Step	Action	✓
1	Go to the UCSC Genome Browser homepage, **genome.ucsc.edu**	
2	Enter the Gateway, by clicking the **Genome Browser link** from the homepage.	
3	Select **Vertebrate** as the clade, **Mouse** as the genome. Choose the most current assembly.	
4	Enter the text **Brca1** in the text box. Click **Submit**.	
5	From the results list, **click a link** that appears to be the real Brca1. I will choose Brca1 NM_009764, breast cancer 1 for this example. **Examine the gene structure, and look at the SNP track in the viewer** with default settings.	
6	From the Genome Viewer page, **scroll to the SNP pulldown menu** (way at the bottom of the page; Variation and Repeats group area). **Click the SNPs (126)** link text above the menu.	
7	On the new track settings page you will see many choices about the appearance and features of SNPs you can display.	
8	In the Color Specification area, **select "Function"** as what we want to color. Change **all the menus for types to "black" except Coding Non-Synonymous, which we will make "blue".** **Select SNP Display mode as "pack".** **Click Submit (near the top)** to make the changes back in the browser.	
9	**Examine the SNPs track now.** Your display should now show all the SNPs (in Pack mode). You should be able to quickly identify SNPs which are coding region + non-synonymous changes.	
10	Select a blue SNP from the display. For this example I will select the blue SNP in the second row: **rs28273098. Click this SNP.**	
11	On the new SNP details page, examine the SNP data. **Click the link for dbSNP.**	
12	**A new window should open with that SNP entry in the dbSNP database.** You can learn additional details about this SNP, its source, and more from dbSNP.	

2) Find the protein sequence for rat leptin. BLAT this sequence vs. the human genome to find the human homolog. Look for SNPs in the coding region of this gene—are there any? Obtain the human DNA sequence for this region, and underline the SNPs.

Step	Action	✓
1	From the Gateway page, search for the **rat leptin** gene, using the most current rat assembly. **Pick the lep gene link** from the results.	
2	From the rat leptin Genome Viewer region, get the **protein sequence** for leptin (hint: from the ***UCSC gene details*** page).	
3	**Copy the rat leptin protein sequence.** You can take everything on the page—the FASTA formatting is fine. **Use the Back button** to return to the Known Gene page.	
4	Now **access the BLAT** tool (either from the Known Gene page, or the home page).	
5	**Paste** your leptin sequence into the BLAT text box	
6	Check the BLAT options, and choose to **BLAT against the human genome**, using the **March 2006 assembly**, with a **protein** sequence. All other settings leave as default.	
7	**Submit** your BLAT search.	
8	From the BLAT results page, **click the DETAILS link for the top hit**. Examine the details page to see if the match is good.	
9	If your match is acceptable, **return to the BLAT results** page. Now **click BROWSER** to see the Genome Viewer location with this match.	
10	If you are convinced we are in the right genomic region in the viewer, we will get the DNA sequence for this region and find SNPs in exons.	
11	First, choose the **HIDE ALL** button so we can add back just the items we care about.	
12	Next, choose to see **UCSC Genes in Full** mode, and **SNPs (126) in Pack** mode. **Click Refresh** to enforce these changes. In your viewer there should only be 2 tracks now.	
13	Let's look at the SNPs in the context of the genomic sequence. **Click the link for DNA** in the blue navigation bar at the top.	
14	From the Get DNA page, **click the Extended Case/Color options** button.	
15	Choose **BOLD** for UCSC Genes, and **Underline** the SNPs. Also put **255** in one of the color boxes for SNPs.	
16	**Submit.** You should have a new page with your sequence, with the UCSC Gene exons in bold, and SNP locations underlined and in color.	
	Special note: Extended case/color options list only those tracks which are currently shown in the Genome Viewer window.	

3) Find the genomic region for the human NRAS [neuroblastoma RAS viral (v-ras) oncogene homolog] gene. Look for ESTs that were found in neuroblastoma tissues—by using the filter to color those blue in the genome viewer. When you are convinced that this protein is found in neuroblastoma tissue, examine the properties of this protein with the Proteome Browser (which is found as a link from Known Gene details pages). What are the pI, mw for this neuroblastoma protein?

This exercise demonstrates the use of a FILTER for ESTs.

Step	Action	✓
1	From the Gateway page, **search for the human NRAS** gene in the current assembly. **Choose the NRAS gene link for uc001efg.1** from the search. *Click the default tracks button to restore default view if needed.*	
2	From the NRAS Genome Viewer region page, **click the link for the Human ESTs track**—above the pulldown menu in the track controls section.	
3	On the Human ESTs track information and filter page, **make 3 changes**: a) change the Display to **Pack**; b) type **neuroblastoma** in the tissue box; c) **select the Filter color radio button** for what color you want these ESTs to be.	
4	Now **click Submit,** and you will return to the genome viewer.	
5	The Genome Viewer will now display neuroblastoma ESTs in the color you selected.	
6	If you are convinced that this protein is likely to be expressed in neuroblastomas, learn more about the protein properties with the Proteome Browser.	
7	**Click the track for the NRAS we selected in the UCSC Genes section**.	
8	On the UCSC Genes detail page **click the link for the UCSC Proteome Browser** (Hint: in the **Quick Links** box)	
9	**Examine the protein properties** that are available in the UCSC Proteome Browser. **Find the pI** for the human NRAS protein:_____ **Find the molecular weight** of the NRAS protein:_____ How many **KEGG pathways** does this protein participate in? _____	
	Special note: Not all tracks have filters; but you can find out the filtering options for a track using the links to the track information page above the pulldown menus, OR by clicking the gray or blue boxes on the left side of the genome viewer image section. Also note—filters will remain in force until you reset everything!!	

Answers: pI = 5.0, molecular weight = 21.2 kDa, KEGG pathways = 10.

4) Perform an in silico PCR, to see what happens when more than 1 PCR product may arise. Determine the product sizes, and the melting temperatures of these primers.

Step	Action	✓
1	Go to the UCSC Genome Browser homepage, **genome.ucsc.edu**	
2	Enter the PCR tool by clicking either of the **PCR or In Silico PCR links** from the homepage.	
3	Select **human** as the species, and the current assembly.	
4	Enter this as the FORWARD primer (with or without spaces): **TTC AAG GAG GCC TTC TCC CT**	
5	Enter this as the REVERSE primer: **CTG GGG GAG AAG CTG A**	
6	**Click 'flip reverse primer' checkbox** if it isn't selected.	
7	**Click Submit.**	
8	The results page will show that these particular primers would amplify **2 different genomic regions**—one on chr19 and one on chr10. The product size would vary and be detectible. Product size on chr19:_____ Product size on chr10:_____ *This set of primers is clearly not specific for one region, if that is the goal.*	
9	What is the melting temperature for the primers? Forward primer:_____ Reverse primer:_____	
	Note: this primer pair was specifically selected to demonstrate what would be seen if there was more than one match for the primer sequences.	

Answers:
Chr19 = 714 bp, Chr10 =316 bp;
Forward TM = 62.5 C, Reverse TM = 55.4 C

5) Use the VisiGene interface to obtain gene expression data for a gene of interest.

Step	Action	✓
1	Go to the UCSC Genome Browser homepage, **genome.ucsc.edu**	
2	From the blue navigation links on the left side of the page, **click the link for VisiGene.**	
3	From the VisiGene Image Browser interface, type **drd2** in the text box. **Click search.**	
4	Examine the result. In the left frame, **use the scroll bar to move down** through all the images.	
5	At the bottom of the left images panel, **click the 2** to move to the next page of results.	
6	**On the second page, click the top image.** (At the time of this writing, it is a green spotted image). That image should replace the first one in the larger frame on the right.	
7	**Move to the large right frame. Click the "zoom in" button** to get a better look at the image in the center.	
8	Use the scroll bar on the far right to access the details about this image. **Click the link for the gene Drd2** in the information area. This should bring you to the gene details page in the genome browser.	
9	**Click other links in the details area** to understand the data available. Clone information, reference information, and links to original source databases should be available.	