

Quantitative Trait Loci Mapping Part IV: Identifying Candidate Genes using Microarrays

Based on a module developed by William Grisham and Natalie A. Schottler at the University of California, Los Angeles (mdcune.psych.ucla.edu). Creative Commons Copyright 2009

- *Pierce, pp 548-549 explains the difference between a Genetic Map and a Physical Map*
- *Pierce, pp 560-562 has a basic overview of microarrays*
- *I also provided a link on our site to an NIH page that also has a good overview of microarrays.*

Now that we have achieved linkage to a particular chromosomal region, we will turn to examining the physical map of the mouse genome to pick some candidate loci that may be responsible for the phenotypic variation we have quantified. Up until this point we have talked only of genetic maps, which are maps that show the organization of genes on a chromosome based on recombination rates (see the fly example on page 180 of our text). Today we will explore the physical map of the mouse which is based on direct DNA sequencing. This map is a more accurate representation of a genome and shows all of the genes and the distance between them in base pairs. When you examine your region of the mouse genome today you can zoom in to the base pair sequence in the region if you would like. The UCSC genome browser will show you all of the genes in your chosen region. You can search through this list by clicking on the names of the genes, which will lead to a page where you can find a wealth of information about that particular gene and the protein that it encodes.

Our ultimate goal today is to choose two candidate genes that are expressed in the olfactory bulb. A simple criteria for picking a candidate gene in the region identified by our QTL mapping is whether or not it is expressed at high levels in the olfactory bulbs. It would make sense that if a gene is important for the variation we have observed in olfactory bulb volume that it would be expressed in that tissue. We will use microarray data that is summarized for each gene in the UCSC genome browser to determine each

gene's olfactory bulb expression levels. The data collected in the database is from a 2002 paper (Su *et al.*, 2002 *PNAS*) that examined the expression levels of all the genes in the human and mouse genome in a particular subset of tissues. This collection of expression data is called the GNF Expression Atlas Data and it was acquired with a U74a Mouse Microarray Chip. The data shows you the level of expression of a particular gene in 45 different mouse tissues. Each of these tissue types were isolated from C57/Bl6 mice and then used to make tissue specific pools of mRNA. This mRNA was then used to probe individual mouse microarrays to quantify the expression level of every single mouse gene in each of these tissues. They first determined the median expression level of each gene in all tissue types combined. Then they calculated the ratio of a gene's expression level in a particular tissue to the median level. A red colored box for a particular tissue type indicates that the gene of interest is expressed higher than the median level in that cell type. The more intense the red color the higher the gene's expression. A black colored box indicates that the gene of interest is expressed at a level equal to the median in that particular tissue. Finally, a green colored box for a particular tissue type indicates that the gene of interest is expressed lower than the median level in that cell type. The more intense the green color the lower the gene's expression.

To pick a candidate gene, you will need to examine its expression levels in the olfactory bulb. Many of the genes in the region will have expression in the olfactory bulbs. As long as the gene shows above average expression levels—any kind of red box—it is acceptable for you to pursue. Admittedly, a gene that has influence on the size of the olfactory bulb may be expressed in early development and not at the ages represented in the UCSC Genome Browser, but this would still be a starting point for further analysis of the QTL. On a more technical note, we search for genes that show relatively high expression in olfactory bulb because you are going to go on and explore the *in situ* expression pattern of a particular gene in the mouse brain using the Allen Brain Atlas. Genes that are expressed at high levels in this microarray analysis are more likely to show an interesting pattern in the olfactory bulb in our next analysis.

Finding Candidate Genes Using the UCSC Genome Browser

Take Home Points Part IV (Testable Material)

- What is a physical map of the genome? How is it different from a genetic map?
- How does a gene expression microarray work? What do you hybridize to the chip?
- How was the GNF Expression Atlas Data generated?
- Why did we choose genes that were highly expressed in the olfactory bulb?

Open a word document for your group. Name it **Candidate Gene Analysis** and save it to the desktop. This will be the document that you will use to compile your data from this section. Lightning bolts will highlight each place where data must be recorded. Email it to your group members at the end of this class.

1. Open the UCSC Genome browser

It can be found at: genome.ucsc.edu. There is also a link on our course website under the date for this activity.

When you get to the home page, click on the “**Genomes**” link on the menu bar that stretches across the top of the page. This will take you to the Genome search page. The default setting should be the Mouse Genome Browser Gateway that is shown below. If it is not, just set the clade, genome, and assembly as shown below and the mouse gateway will open. When you have the mouse gateway open, type in the positional information that you saved from the previous assignment. It has the form: **chr6: xxx,xxx,xxx-xxx,xxx,xxx**. Enter that data and press submit, you will be taken to the physical map of the chromosome in the region that you have specified.

Mouse (Mus musculus) Genome Browser Gateway - Mozilla Firefox

http://genome.ucsc.edu/cgi-bin/hgGateway

Home Genomes Blat Tables Gene Sorter PCR Session FAQ Help

Mouse (Mus musculus) Genome Browser Gateway

The UCSC Genome Browser was created by the [Genome Bioinformatics Group of UC Santa Cruz](#).
Software Copyright (c) The Regents of the University of California. All rights reserved.

clade	genome	assembly	position or search term	image width	
Mammal	Mouse	July 2007	chr6:112,000,000-136,000,000	800	submit

[Click here to reset](#) the browser user interface settings to their defaults.


[add custom tracks](#) [configure tracks and display](#) [clear position](#)

About the Mouse July 2007 (mm9) assembly ([sequences](#))

The July 2007 mouse (*Mus musculus*) genome data were obtained from the Build 37 assembly by [NCBI](#) and the [Mouse Genome Sequencing Consortium](#).

Sample position queries

A genome position can be specified by the accession number of a sequenced genomic region, an mRNA or EST, a



2. Looking at the Physical Map



The physical map interface shows you where on the chromosome you are and highlights the locus you are currently examining (arrow and red box). **Write down the locus on your Candidate Gene sheet.**

UCSC Genome Browser on Mouse Feb 2006 Assembly

position/search chr6:130,166,666-140,166,666 jump clear size 10,000,001 bp. configure

chr6 (qF3-qG2) qG1 qG2 qG3 qG4 qG5 qG6 qG7 qG8 qG9 qG10 qG11 qG12 qG13 qG14 qG15 qG16 qG17 qG18 qG19 qG20 qG21 qG22 qG23 qG24 qG25 qG26 qG27 qG28 qG29 qG30 qG31 qG32 qG33 qG34 qG35 qG36 qG37 qG38 qG39 qG40 qG41 qG42 qG43 qG44 qG45 qG46 qG47 qG48 qG49 qG50 qG51 qG52 qG53 qG54 qG55 qG56 qG57 qG58 qG59 qG60 qG61 qG62 qG63 qG64 qG65 qG66 qG67 qG68 qG69 qG70 qG71 qG72 qG73 qG74 qG75 qG76 qG77 qG78 qG79 qG80 qG81 qG82 qG83 qG84 qG85 qG86 qG87 qG88 qG89 qG90 qG91 qG92 qG93 qG94 qG95 qG96 qG97 qG98 qG99 qG100

Clickable genes!

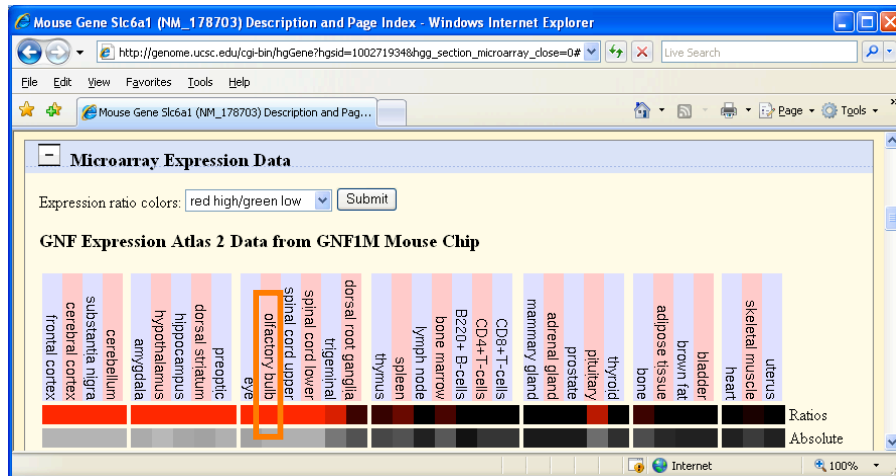
Mouse Chromosomal Locus

The genome browser allows you to see the genes in the region and clicking on an individual gene name will take you to a Gene Index page that has a collection of functional information about that particular gene in the mouse. On this page you will find a library of information about the gene, the mRNA that is made from it, and the encoded protein.

3. Examining Gene Expression

Scroll down the page to find the **Microarray Expression Data** section. High expression in a particular tissue is shown as a shade of red. While lower than average expression is given in shades of green.

If your candidate gene displays high expression as seen below (any shade of red will do), then you may use this gene.



If the olfactory bulbs do not display high expression (there is a black or green box under the olfactory bulb label), then return to the physical map (from the previous page) and select another gene.



On your Candidate Gene sheet write down the name of your first candidate gene. Take a screen shot of the microarray data for that gene and paste it into your Candidate Gene sheet. This is candidate gene number one.

Return to the gene list on the physical map by going back one page on the browser. Find another candidate gene that has high expression in the olfactory bulbs.



On your Candidate Gene sheet write down the name of the second candidate gene. Take a screen shot of the microarray data for that gene and paste it into your Candidate Gene sheet. This is candidate gene number two.

We are picking two candidate genes to make sure that at least one of them has an interesting expression pattern in the Allen Brain Atlas.

To summarize what should be on your Candidate Gene document:

1. The chromosomal locus you examined
2. The exact name and microarray data for your first candidate gene
3. The exact name and microarray data for your second candidate gene

Email this document to each group member. You will each have to turn this in at the end of the project.

You will need the gene names for the next assignment.