Mapping Homozygous Regions using HomozygosityMapper

Part 1 – Analyzing SNP data using HomozygosityMapper

A consanguineous pedigree (Figure 1) segregates autosomal recessive trait P, that is genetically heterogeneous but with Mendelian inheritance and complete penetrance. DNA from family members (marked by an *asterisk*) was submitted to a genome-wide linkage scan using an Affymetrix SNP chip. We would like to scan the genome for homozygous regions to identify region(s) which contain the causal variant. These regions should also be followed up by performing linkage analysis. One additional benefit of homozygosity mapping is that it is robust to intermarker linkage disequilibrium unlike multipoint linkage analysis. For this exercise, we will be using the program Homozygosity Mapper (Seelow et al. 2009).

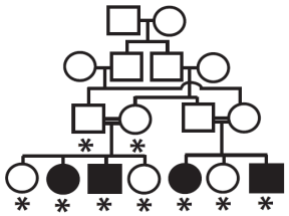

Figure 1

Using Firefox or Chrome, go to http://www.homozygositymapper.org/ (Tip: Running the program with Internet Explorer can take a longer time.) Under *options*, one can create a personal profile and create projects with private data by clicking on *upload genotypes*. Having a personal profile allows one to have limited access to one's own data and also to delete the data after analysis. When creating an account name and file names, only alpha-numeric characters and the underscore symbol are allowed, otherwise data upload will not commence. Both Affymetrix and Illumina data are accepted, with the genotype file having the following format:

SNP_ID	Sample1	Sample2	Sample3	Sample4
SNP_A-1234567	AB	AB	AB	BB
SNP_A-5678910	AA	AA	AA	AA
SNP_A-9101234	BB	NoCall	AB	AB
Etc.				

For this exercise, the data has been uploaded to the server so you do not have to create a new personal profile or upload the data. Under the tab "**add data**" click "**(re)analyse a project**". From the pull-down menu for **project**, select **SML_exercise**. Assign your own analysis name.
For the **cases** box, type in the following to identify which individuals should have homozygous genotypes: S1, S2, S3, S4

Then within the **controls** box, type in the following: P1, P2, H1, H2, H3

Make sure to *check the tickbox* for the field "**require genetic homogeneity**". To "**exclude homozygous stretches in controls**", put in 2000 as value. For **allele frequencies** select from the pull-down menu the option **HapMap: CEPH (European origin)**. For this exercise you can leave the rest of the fields blank. Click **Submit.**

It will take a few minutes to analyze the uploaded genotypes. (The website recommends taking your coffee instead of staring at the screen while you wait.) When the analysis is done, click **Show results**.

The website will present a genome-wide bar graph with regions attaining homozygosity scores >80% in **red**. You can click on the chromosome number (in *green font*) to view each chromosomal region in greater detail.

Question 1: In which chromosome(s) were regions of homozygosity detected? _____

In the middle of the webpage under the listed regions there is a drop-down menu to ***change threshold***; choose **0.5 x max** then click **change.**

Below the graph there are two lists of homozygous regions. Go to the list marked **narrow**. Based on the physical positions for the markers that are listed, the markers cluster within a 10Mb region. For each line you can click on region or genotypes for more details. Find the most distal region within the list then click on genotypes. Here you can see the SNPs bounding the region in *red* and the homozygous SNPs for the selected region within a *black rectangle*. There is also an option to zoom in/out or to sort the haplotype blocks by size. Click Zoom out until you see the full length of the haplotype (this might take several clicks). Occasionally the haplotype would seem to be broken up by uninformative genotypes. After viewing the full-length haplotype, click Click here for the genotypes table to inspect the actual genotypes.

Question 2: Which markers flank the homozygous region? _____

After performing linkage analysis on selected SNP markers from the genome scan data, a significant LOD score of 3.63 was obtained at Θ=0. The 3-unit support interval is bounded by SNP markers rs6550024 and rs11544593 and therefore contains 26.5 Mb of sequence within chromosomal region 3p24.1-p14.3.

Question 3: From the HM results is the homozygous region smaller than the 3-unit support interval? (Note: You have to check the updated physical position from the UCSC Genome Browser (http://genome.ucsc.edu/)) _____

Reference:

Seelow D, Schuelke M, Hildebrandt F, Nürnberg P. HomozygosityMapper – an interactive approach to homozygosity mapping. Nucleic Acids Res 2009; 37(Web Server issue):W593-599.

Answers

Question 1: In which chromosome(s) were regions of homozygosity detected? __3__

Question 2: Which markers flank the homozygous region?
On chromosome 3, a relatively contiguous region in which the affected individuals are homozygous for the trait is flanked by markers rs6550024 to rs17052830.

Question 3: Is the homozygous region smaller than the 3-unit support interval?
Yes, the homozygous region is smaller by 4.1 Mb.