



The 1000 Genomes Web based Tutorial Exercises.

These exercises should present a set of questions whose answers can be found in our tutorial slides that are found on: <http://www.1000genomes.org/using-1000-genomes-data>

All these questions can be answered using our web based tools and browser and the software Haploview.

Haploview is available from:
<http://www.broadinstitute.org/scientific-community/science/programs/medical-and-population-genetics/haploview/downloads>

Finding Data

1a. Find what Omni VCF files we have on our ftp site using the website ftp search. (Omni is a high throughput genotyping platform from Illumina on which all 1000 genomes samples are being genotyped)

1b. Find the most recent Omni VCF file on build 37 from the 31st January 2012

2. Use the Website search box found in the top right hand corner of all pages to find the FAQ question about getting subsections of VCF files.

3. Use the Data slicer to get this section of the Omni VCF file 6:31830969-31846823:

http://browser.1000genomes.org/Homo_sapiens/UserData/SelectSlice

Browsing Data

4. Find the variant rs45562238 using <http://browser.1000genomes.org>.

5. In what 1000 Genomes Super Population is this variant detected?

6. What are its global allele frequencies in the 1000 Genomes Data set?

7. In which gene is the variant found?

Using the 1000 Genomes Tools

8. Use the browser to find the SLC44A4 gene.

9. Use the get VCF button in the left hand menu on the gene page to get a slice of a vcf file for this Gene.

10. Unzip this VCF file using a tool link winzip or Archive Utility.

11a. Upload this VCF file to the Variant Effect Predictor.

http://browser.1000genomes.org/Homo_sapiens/UserData/UploadVariations

11b. Do any of the variants have negative Sift or Polyphen predictions?

12. Using the example URLs on the Variation Pattern Finder tool menu look at the patterns of inheritance for this region: 6:31830700-31840700

http://browser.1000genomes.org/Homo_sapiens/UserData/VariationsMapVCF

13a. For the same region use the VCF to PED tool to produce a ped and info file for the CEU population.

13b. Look at these files in haploview.

13c. How many haplotype blocks does haploview think there are in this section?