



## Downloading From dbSNP's FTP Site

Created: July 7, 2005; Updated: February 18, 2014.

### **I want to download flat files that list all SNPs for which allele frequencies exist across the major human populations (European, East Asian, and African). Where can I find such data?**

We currently do not have a flat file report of allele frequencies across major populations for all human SNPs. If you can work with XML files, you can find FTP files of genotype data (including genotype and allele frequency data) in the dbSNP FTP site. They are organized by chromosomes:

[ftp://ftp.ncbi.nlm.nih.gov/snp/organisms/human\\_9606/genotype/](ftp://ftp.ncbi.nlm.nih.gov/snp/organisms/human_9606/genotype/)

[ftp://ftp.ncbi.nlm.nih.gov/snp/organisms/human\\_9606/genotype\\_by\\_gene/](ftp://ftp.ncbi.nlm.nih.gov/snp/organisms/human_9606/genotype_by_gene/)

If you are working with a set of SNPs, you could use [eutils](#) to fetch the allele frequency information.

In the meantime, we will look at making a new FTP file format for reporting SNP frequencies. (7/24/07)

### **How do I download the genotypes for Batch ID:AFFY\_071103? The file is too large and cannot be handled by a normal PC program.**

To download the genotypes for Batch ID:AFFY\_071103, please follow these instructions:

1. From the dbSNP homepage, select the new batches link located under the submission information section.
2. Fill out the form using the following: organism="Homo Sapiens-9606", Batch type="Variations with individual genotype data (IND)", Submitter Handle "starts with" ="AFFY", and click on search batch.
3. Select the link AFFY\_071103 under Local Batch ID.
4. Click on the download, all sort by button and click on Go. This will allow you to download the genotype information to your PC in a zipped file (~7 MB). Use Winzip to unzip it. Unzipped, this file will be ~231 MB.

### **I would like to use a script to fetch average allele frequency data for each human SNP from every web page, but I'm afraid that my IP will be blocked by the server for the heavy usage.**

In general, large amounts of data can be obtained using our ftp, efetch or batch query services. Specifically, if you only need SNP allele frequency data, then use SNPAlleleFreq.bcp.gz, which is located on our [ftp site](#). The Allele.bcp.gz file, also available on the ftp site, has Allele\_id to allele\_string mapping. (3/25/05)

### **I need to identify SNPs that are different between C57BL/J and A/J mice from a list of 50,000 SNPs. Downloading the data by chromosome is very time consuming. Is there a better way to do this?**

All the mouse genotype data submitted to dbSNP are available on the FTP site. Go to the mouse organism database, and select "[genotype](#)" from the list of available files. You can then download all of the data and do your own analysis.(6/19/06)

**I want to download the current set of sequences in dbSNP to format for a local BLAST search. Could I use a "wget --mirror" to pull down the contents of /snp/organisms/\*/ss\_fasta/\*/\*", and then concatenate the files?**

We already have the files you need on the dbSNP FTP site as FASTA files. For example, to get [human ss FASTA data](#), go to the human organism directory, and select ss\_fasta, and then select the year in which the data you need was submitted.

You can also get [rs FASTA organized by chromosome](#) in the dbSNP FTP site, or you could also blast dbSNP rs sequences directly [online](#).

For more information, please check the [dbSNP handbook](#).(9/13/06)

## Problems Downloading

**I have been trying to download snpBatch\_BCMHGSC\_JDW\_59658.gz from the human viewBatch FTP directory, but I get a "no such file or directory" error message.**

Thank you for reporting this. snpBatch\_BCMHGSC\_JDW\_59658.gz has been generated. It is fairly large, and contains data 300K+ SNPs, so opening it directly in web browser is too slow for such large files.

If you use Windows Explorer, please cut and paste this path: [ftp://ftp.ncbi.nlm.nih.gov/snp/organisms/human\\_9606/viewBatch/snpBatch\\_BCMHGSC\\_JDW\\_59658.gz](ftp://ftp.ncbi.nlm.nih.gov/snp/organisms/human_9606/viewBatch/snpBatch_BCMHGSC_JDW_59658.gz) to Explorer and save the file to your local drive.

If you use Linux/Unix, use the ftp command to save the file.(04/17/08)

**I have a problem downloading SNP XML file from Entrez. When I use "db=SNP&uid=100&cmd=file&dopt=XML" in the URL, Entrez returns with a file not in XML format.**

Please use the XML request example located on the [utils for EntrezSNP](#) page. (6/14/07)

**I've tried downloading FTP files to create a local copy of dbSNP, but get an error message that reads: "Passive mode refused. Turning off passive mode. No control connection for command. No such file or directory". What do I do?**

This message indicates that your firewall is blocking passive FTP — at least some of the time. Try a "smart" FTP client like NCFTP (often available on most UNIX machines). It is better at auto-negotiating active/passive FTP connections than some of the older FTP clients (e.g. SUN SOLARIS FTP). (4/25/06)