



Missing FTP Files

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

I was able to retrieve submitted SNPs (ss) mentioned in two recent zebrafish publications, but could not find them in the zebrafish xml file, or by using Entrez SNP.

Zebrafish SNPs were not mapped in the last build, but will be mapped in the next build (b128). Once they're mapped and clustered, the SNPs will have rs numbers as well as chromosome positions, and will therefore be searchable in Entrez. We expect to have build 128 ready in a few weeks. (7/27/07)

The ftp download site contains the data file b127_SNPChrPosOnRef_36_2.bcp.gz. Do you have a corresponding data file for build 35?

b127_SNPChrPosOnRef_36_2 is a new table starting b127. We do not have the same data file for genome build 35. (5/17/07)

I searched db SNP several months ago for SNPs in SLC22A2 and its splicing variations, and found non-synonymous SNPs. I searched again today, and found that many of these SNPs have been deleted from dbSNP. Why?

The SNPs mentioned in your question were originally submitted by PHARMGKB (to see the submitter of a ss#, click on the ss#, which is located in the "NCBI Assay ID" column of the "Submitter Records" section of the refSNP report page). PHARMGKB has withdrawn 11740 of their submitted SNPs (ss#) for error corrections and will resubmit the data early next year (2007). PHARMGKB did not indicate the percentage of corrected SNPs that will have flanking sequence changes. (12/19/06)

I have just downloaded the b125 XML files from the dbSNP ftp site and can't find the population/frequency information. Why did you take it out?

Population and frequency information is now located in the genotype files for the organism of interest. You can find this file by going to the dbSNP ftp [organism directory](#), select an organism of interest (in this case I will choose "human_9606"), and then select "genotype" from the list of the organism's subdirectories. (11/4/05)

I see that SNP has updated to build 117, but I cannot find the subdirectory within the MSSQL directory that contains the SQL server backup files.

We have discontinued this format. It was pointed out to us that providing the database in a single proprietary format could create legal problems for NCBI because we do not provide the database in all vendor formats. We are committed, however, to work with users like you in making the creation of your own copy of dbSNP as painless as possible.

A [sample Unix C shell script](#), cmd.create_local_dbSNP, has been provided in the FTP site that shows how to use the dbSNP_table.sql and *.bcp to create a local sql database.

First, `cmd.create_local_dbSNP` creates all tables without indexes to allow for a fast data load. Then, it loads table data and compares the lines in the data file with table row counts to check for errors. Users of other database platforms can use the same method with a few modifications of platform-specific commands.

What happened to SNPContigLoc_b34_3? I can't find it anymore.

SNPContigLoc_b34_3 is an old file, and has been renamed: b125_SNPContigLoc_b34_3.

Beginning with b125, all mapping related tables (e.g. SNPContigLoc, SNPMapInfo, etc) will contain the SNP build number as a prefix, and the NCBI build number as a suffix.

So the file in question, b125_SNPContigLoc_b34_3, contains the mapping data for b125 SNPs that map to NCBI genome build 34, version 3. (12/9/05)

The SNPFlankStatus table disappeared from dbSNP_124; is there any other table that contains the same information?

dbSNP no longer independently repeat masks the submitted SNP flank. Masking is now completed during the BLAST portion of the build process. SNPFlankStatus and SubSNPRepmask were removed from the schema as a consequence of this change. Submitted SNP flanks from submissions later than build 122 are no longer mix-cased masked in dbSNP. (2/1/05)

Although the XML brief format has been dropped, I was wondering if I could still get it via esummary/efetch.

The XML brief format has been dropped and is not available via esummary/efetch on Entrez.