Schema Changes and Updates

Created: August 8, 2005; Updated: June 15, 2010.

How is information on schema changes and data updates passed to users?

Schema changes/updates are listed in the Schema Change Document. You can access this document from the dbSNP web page sidebar.

Does each build represent schema changes or data updates?

Each build represents both data updates and schema changes.

Are there any plans to modify the XML Schema for dbSNP?

We will make future modifications to the XML schema as necessary to fit evolving SNP data model. We'll announce any schema changes on the dbSNP homepage and the dbSNP maillist, should you care to subscribe.**(6/8/06)**

Table Updates

The dbSNP Data Dictionary states that RsMergeArch.bcp has 8 columns; however, when I looked at RsMergeArch.bcp I found 9 columns.

A new "comment" field was added to RsMergeArch to store information for future internal reference when we were conducting special processing to handle PHARMGKB's request to withdraw and update their submission.

The Data Dictionary description of the RsMergeArchtable will be updated to include this information as time allows. **(08/12/08)**

Discrepancies between Tables

Your web site directed me to the SNP.bcp and SnpValidationCode.bcp tables to find the validation status for each SNP, but the data dictionary and the database schema don't match up with the fields in SNP.bcp.

The updated SNP table description can be found online, and the introduction to our Schema Dictionary states:

"Table/column descriptions should be updated whenever there are database changes. This task takes time since dbSNP has well over a hundred tables. Instead of waiting for all table/column descriptions to be completed before releasing the data dictionary, we are letting a table/column description go public as soon as it is finished to allow faster user access. Please click on "Contact Us" on the sidebar if the table of your interest still lacks the updated description." **(10/25/05)**

I noticed that the validation_status field for build 123 has a much wider range of validation_status codes, and no longer corresponds to the SnpValidationCode table in the same build.

SNP.validation_status still ranges from 0-31, but the SNP table has had column changes. Perhaps you are reading a different column into the validation_status field. Please review the build 123 Schema change documentation for any column changes that occurred during build 123. **(1/31/05)**

Replaced or Retired Tables and fields

I see that the b130_ContigExon_36_3 table is empty. Is this intentional? Where can I get exon coordinates?

dbSNP is using a new pipeline to annotate SNPs onto genes and as a result, we no longer use the ContigExon table. You can get exon coordinates from contigs and chromosomes using the Map Viewer FTP site. To get to the Map Viewer FTP site:

1. Go to the Map Viewer site
2. Scroll down and click on the "Map Viewer FTP" link located under "Related Resources" on the left side bar.
3. Click on the "Homo_Sapiens"
4. Click on the "Sequence" file
5. Click on BUILD.36.3
6. Click on initial_release
7. Click on seq_gene.md.gz

(06/11/09)

What happened to the SNPSeqPick table?

The SNPSeqPick table was replaced with a view starting with dbSNP build 110 as a space saving measure and so that the SNPSeqPick table will always be in sync with refSNP cluster change in the SNPSubSNPLink table.

This and other schema changes are reported in the Schema Change Document, which you can access from dbSNP sidebar. Search for SNPSeqPick in the document.

What happened to the locus_id table?

LocusLink was replaced by Entrez Genes as the official resource for gene information. "locus_id" is being replaced by "gene_id" though it is my understanding that the values will be the same for existing data.

What happened to a field called physMapStr that was part of the MapLoc struct Table?

This field was deprecated (made invalid or obsolete) starting with build 125 due to a location type change. Please see the (old) Column Description for SNPContigLoc and look at the information for "phys_pos". You may also wish to look at the new location type description.

(6/28/06)