



Using SNP IDs to Find Genbank Accessions, and Vice Versa

Created: July 7, 2005; Updated: February 25, 2014.

Is SNP information annotated in GenBank entries? If so, how do I access this annotation?

dbSNP annotates GenBank reference sequences and mgc mRNAs based on BLAST analysis, rather than on submitter accessions. Reference Sequences, or RefSeqs (rs), are a curated, non-redundant set of records for mRNAs, proteins, contigs, and gene regions constructed from a GenBank exemplar for that protein or sequence.

I will use rs11542065, a non-synonymous variation in the LPL gene, [as an example](#). If you look under GeneView, you'll see that the SNP maps to protein NP_000228. If you click the link for this protein, it will take you to the GenBank protein record. From here, click the features button at the top of the page, and then check SNP in the checkbox. Once this is done, scroll down, and find the variation blurb for this record:

```
variation 71
/replace="h"
/replace="q"
/db_xref="dbSNP:11542065"
```

If you repeat this exercise for the mRNA NM_000237, you will find the variation annotation:

```
Variation complement(387)
/gene="LPL"
/replace="g"
/replace="c"
/db_xref="dbSNP:11542065"
```

Always remember to set the SNP feature checkbox whenever you load a GenBank record.

I want to retrieve GeneBank accession numbers associated with a list of rs ID and ssID numbers. Do you have a SNP text file containing this information that I can parse?

There is an eLink utility that will allow you to retrieve GenBank records associated with an rs number. Please see the online instructions and examples for [Elink](#). (10/27/05)

Does dbSNP have a search that will take a sequence accession (or gi number) as input and will return the SNP information for that sequence?

I'm sorry that we don't have the exact service you requested. We try to provide resources (search service and reports) that are broad and generic enough to suit a broad range of users. The information that you requested is available, but it will require parsing a dbSNP XML or ASN report that you can retrieve automatically using [Entrez Eutils](#).

Please read the online document and examples for automatic search/retrieval using the Eutils web API.

[Here's an example](#) of using eLink to get all the SNP using the protein LPL gi(4557727).

[Here's an example](#) of retrieving an XML report for a snp id (rs).

Additional [Entrez Tools](#) are available.

I see in “Submitter-Referenced Accessions” that Genbank accession numbers are linked to the submitted SNP (ss) ID number, but I can’t find these Genbank accessions in the XML formatted reports.

The accession numbers you see under “Submitter-Referenced Accessions” is annotation that is included with a submitted SNP (ss) when it is submitted to dbSNP. Submitter-Referenced Accessions are currently not included in the XML dump, but you can look up the accession using the ss number in the [SubSNPacc.bcp table](#). (1/5/06)