



Searching for SNPs Using Key Words, Names (Gene, Author, etc.), or Citations

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

Finding SNPs cited in PubMed

What search terms do I use in Entrez SNP to find rs numbers cited in PubMed?

Use the search term "snp pubmed cited"[Filter] in Entrez SNP; this filter is also included in as an annotation limit on the Entrez SNP "limits" page. Here is an [example](#) of the results page for such a search. Click the "PubMed" link located below the record ID to go to the PubMed listing for the article.(06/16/09)

Finding SNPs using an Organism Name

I am simply trying to find a list of soybean SNP sequences — Is this possible? I just can't figure out your web site.

You can find the soybean SNPs using the "New Batches" search:

1. Go to the dbSNP Home page, look at the left blue side bar, and click on "Search" to release a menu of search options.
2. Click on "New Batches"
3. When you get to the "New Batches" dialogue page, select "soybean_3847:Glycine max" from the "organism" drop-down menu, and select "Variation submissions (SNP)" from the "Batch Type" drop-down menu.
4. Click the "Search Batch" button at the bottom of the pane. You will be taken to the only [two submissions](#) currently in dbSNP.

(07/14/08)

I can't tell if you have horse SNPs in the db or not. If I query dbSNP for horse or txid9796, I get no results, yet there is a directory for horse on the SNP FTP site.

Search for horse using the "New Batch" search.

1. Go to the dbSNP Home page, look at the left blue side bar, and click on "Search" to release a menu of search options.
2. Click on "New Batches"
3. When you get to the "New Batches" dialogue page, select "Horse_9796: Equus caballus" from the "organism" drop-down menu, and select "Variation submissions (SNP)" from the "Batch Type" drop-down menu.
4. Click the "Search Batch" button at the bottom of the pane. You will be taken to the [submissions](#) currently in dbSNP.

You could also find horse SNPs using the dbSNP [Summary page](#). You will find "Equus caballus" toward the bottom. (07/18/08)

Finding SNPs using a Gene Name

I have a SNP that has no rs number, but does have a ID (human TNF alpha -308G>T), I've searched in dbSNP, but can't find an rs number for it.

Unless the submitter used the exact name "human TNF alpha -308G>T" in their submission, it won't show up as a result in your dbSNP search. It is also possible that the authors who mentioned "human TNF alpha -308G>T" in their publication never submitted this SNP to dbSNP.

Submission to dbSNP is voluntary, but we do try to encourage people to submit SNPs to dbSNP prior to publication so that they can cite their assigned dbSNP IDs (ss or rs) in their manuscript. Check the publication to see if the authors included a dbSNP ID for "human TNF alpha -308G>T" in the manuscript; if they have not, try BLASTing the DNA sequence of human TNF alpha -308G>T (if you have it) against dbSNP to see if there are any matching rs numbers. If there are no matches, and you wish to see this variation in dbSNP, you can submit it yourself using the [Human Variation: Annotate and Submit Batch Data](#) site (multiple submissions) or the [Human Variation: Search, Annotate, Submit](#) site (single submission). The [Human Variation Batch Submission Quick Start](#) explains how it is possible to submit published SNPs not already housed in dbSNP. (04/30/08)

I am trying to find the refSNP (rs) or submitted SNP (ss) numbers of SNPs described in the literature. The nomenclature used for these SNPs (IL1A[*gene*] -889C>T) is not supported by your search query.

Try the following search:

1. Go to the [dbSNP home page](#). Once you are there, enter IL1A[*gene*] in the search box located at the top of the page.
2. On the resulting page, you will see 94 SNPs. Click on the "Limits" tab located near the top of the page.
3. Once you are on the limits page, scroll down until you find the "Observed Allele" section, and click on the box next to IUPAC code "Y" (Meaning C or T). Now, Click on the "Go" box located next to the search box at the top of the page.
4. The resulting page shows that by limiting your search this way, you have narrowed your search result down to 27 SNPs. (3/7/05)

If I search dbSNP using a gene name, would the search results include SNPs in the UTR?

The search may return SNPs in UTR regions if a locus has been defined to contain those regions on the genome.

How do I interpret the results of a gene name query? I used the cyp2j2 gene to query dbSNP, but I am unable to determine the number of SNPs located within cyp2j2 by looking at the results page.

We have 40 human SNPs in our database associated with the cyp2j2 gene.

Start [here](#):

1. Enter: cyp2j2[GENE] AND human[ORGN] in the Search box at the top of the form and click the Preview button.
2. The SNP count for the gene name queried will display as an active link that will take you to a list of available SNPs associated with the gene queried.
3. Try to explore the Help links on the left sidebar to learn more about Entrez SNP and the Preview/Index feature.

Finding SNPs using a Mutation Name

How do I find the rs number for the Thr790Met mutation in the human EGFR gene?

dbSNP allows searches for mutations using Human Genome Variation Society (HGVS) nomenclature, so you will first have to convert the mutation name into HGVS nomenclature before you can search for it:

Visit the [HGVS site](#) for details on how to convert a named mutation into HGVS nomenclature.

The NCBI [Sequence Viewer](#) might be useful to you in finding the sequence details required for determining HGVS nomenclature.

Once you have determined the HGVS name of your variant, you can search for it by going either to the [Human Variation: Search Annotate, Submit site](#), or to [Entrez SNP \(06/26/09\)](#)

Finding SNPs Using a Key Word

I tried a query using the keyword “mtDNA” and found 60 records, but many of them show SNP localization on chromosomes rather than in the mitochondrial genome. Why?

The SNPs that you retrieved using the mtDNA query are not true mitochondrial moltypes; they are incidentally related in some way with the keyword mtDNA. We only have mitochondrial moltype data for one organism, *Cooperia oncophora*.

When I search the dbSNP database for “microsatellite” with Limits of coding nonsynonymous and Homo sapiens, none of the records I retrieve contains the term “microsatellite”. Why?

Unless you specify the correct field tag, every field is searched, and every SNP record that has the word “microsatellite” anywhere in it will be returned. If you search using "microsat"[SNPCLASS] without limits set, you should retrieve 4954 records. You can use the Preview/Index option to get the correct field and value.

Finding SNPs using a Submitter’s Name

Although I can search dbSNP for CSHL-HAPMAP, the output is too large and lists SNPs submitted by CSHL-HAPMAP and others, while I need just those submitted originally by HAPMAP.

CSHL-HAPMAP genotyped about 4 million of the 10 million refSNPs in dbSNP, so I’m assuming you’d like to find those SNPs that have genotype data from HapMap as opposed to novel SNP assays submitted by HapMap,

When you said the EntrezSNP gives you too much output, what display format did you choose? If you select the "brief" format, you will get a list of just the search’s resulting refSNP numbers. Or, you can take a look at [eUtils](#) as a means of retrieving data in batches.

Another alternative would be to use the *.bcp file approach:

A *.bcp file is a text file of table column values separated by tabs. If you can work with *.bcp files either by using scripts or databases, then you can achieve what you want by using two tables: SNP_bitfield and SNPChrPosOnRef, both of which are located in the [organism_data](#) directory of the dbSNP FTP site. First, take a look at page 2 (F6-HapMpa properties) of the [bitfield pdf](#). Here, you’ll see that when the SNP_bitfield.hapmap_prop equals 2, it means that the SNP has had a HapMap genotype (phase2) submitted. By using this information in conjunction with the data in SNPChrPosOnRef, which lists the chromosome and position for each SNP on reference assembly, you can get the list of refSNP numbers you wanted.

For future queries of this sort, you can use the *.bcp file approach or check back with us for a SNP Genome Workbench plug in which will allow flexible queries and reporting using the data in SNP_bitfield within [Genome Workbench](#). This system is currently under development, and should be available soon. (11/01/07)

The current Nature Genetics advance online papers indicates that David Page's group submitted 95 new SNPs to dbSNP. How do I find these SNPs?

1. From the [SNP home page](#), scroll down to the "Submission Information" section, and click on the words "By Submitter". This will take you to the [Search/View Submitter Detail page](#).
2. In the "Search By" section of the Search/View Submitter Detail page, choose "Submitter name" and then choose "contains". Then type the name, "Page" (without quotation marks) in the text box, and click on the "Search" button.
3. Click on the name "Page" in the Handle column of the response page to go to the [Contact Detail page](#).
4. Scroll down the Contact Detail page until you see a list of data batches submitted by the Page lab. To see a list of the SNPs you are interested in, click on the text "2006.01.23" in the "Submitter batch id" column to go to the [View SNP Submission Batch](#) page. The Submitted SNP (ss) ID numbers displayed on this page are links to the submitted SNP data. (3/9/05)