NIH U.S. National Library of Medicine
National Center for Biotechnology Information

SNP FAQ Archive
NCBI Help Manual

National Center for Biotechnology Information
U.S. National Library of Medicine

Requirements for dbSNP Submissions

Created: June 6, 2005; Updated: June 15, 2010.

Do my SNPs need to be published before I can submit them to dbSNP?

You don't have to publish a SNP before submitting it to dbSNP. Since some journals require prior submission to dbSNP before they will accept a manuscript, however, it's actually better to submit to dbSNP prior to publication so you'll have the dbSNP assigned accession numbers for your SNPs to cite in your manuscript. (**6/8/06**)

Can I still submit a SNP if I don't have genotype or frequency data?

Go ahead and submit; just skip the SNPINDUSE and SNPPOPUSE sections if you don't have genotype or frequency data. (**5/16/05**)

When I submit a SNP to dbSNP, must the submission contain an allele frequency for every SNP submitted?

Allele frequency data are not required data for SNP submissions. (**2/13/07**)

I was wondering if NCBI has some universal recommendations concerning the reporting of new SNPs for our publication. Should we use HGVS standards?

I don't know specific NCBI policy on offering "universal recommendations" to publishers about their citation policies, but I will point out that dbSNP identifiers [submitted SNP (ss) number; refSNP (rs) number] are stable and unique identifiers within the NCBI dbSNP database, and that they provide flanking sequence context and alleles for a specific variation.

The issue of requiring use of HGVS nomenclature as you have asked is useful when all the information pieces are available (i.e. accessioned sequence records, gene annotation, functional analysis). Many polymorphism discovery projects, however, occur before such organized knowledge is available for a new species genome, and the rs number/ss number paradigm has been sufficient to describe sequence variations at this minimal level of detail.

I would say use HGVS standards when available, but use the rs numbers and ss numbers as well. This will encourage submission of the variation data to a public repository since using HGVS alone does not guarantee that the variation is in a public database. (**8/14/07**)

What are the requirements for dbSNP submissions?

The minimal SNP submission consists of the observed alleles, the 5′ and 3′ flanking sequences, the gene name, and the NCBI GenBank accession of the genomic sequence in which the SNP is located (sample shown below).

```
SNP SAMPLE:

SNP:SELPLG-000083  <-- user-defined SNP ID
ACCESSION:AY331789 <-- NCBI GenBank accession
```

```
SAMPLESIZE:42        <-- number of chromosomes assayed
GENENAME:SELPLG
LENGTH:338           <-- total length of 5' and 3' flanking + SNP
5'_FLANK: tcacctcag atatcattgc ttcaatcttg gaaggcaac ttggagtat ttttcttgg
          gcttcattc ctagaagca GAAACATA
OBSERVED: C/T
3'_FLANK: TCACTTTGA GTCTTTGACT CTCACAGATG TCAGGTGCG GTGAAGTGT GGTTTTATG
          GTCTTCACA CAGACGACAT ACACCAGATT TATCTTAGC AATATTTGC TCCTCTGAG
          AAGGAAATT TCACTGCTCA AGTCTCACCT TGAAATAGA TTACTTGGA GCTGGGCGT
          GGTGGTGTG CTCCTGTAAC AGCTACTGGG GAGGCTGAG GCAGGAGGA TTGCTTGCT
          TGAGACCAG GAGTTTGAGG CTGCAGTGAG CT
||
```

An example of a submitted SNP in submission format is available.

You can also submit additional SNP annotation fields as well as the genotype of each strain (if it is known). Additional submission instructions are available on the web.

If we have a SNP that is less than 25 bases from the end of our sequence, should we submit it, even though the instructions say that the minimum flanking sequence either 5' or 3' must be at least 25 bases?

You can submit it if you want; we will load it, but I'm not sure it will map to the correct position. (**7/25/07**)

GenBank Accession Numbers

I don't have enough information to get a GenBank number, but want to submit a SNP to dbSNP. Your site says one is required, so what do I do?

GenBank accession numbers or dbSTS reference numbers are optional, not required. Look in the "Using the Excel submission Templates…" section located in the "Submission Quick Start", to find information showing what data is required and what is optional in a dbSNP submission. (**01/22/08**)

I would like to submit to dbSNP, but I don't have a GenBank reference number to include in my application. What do I do?

In a dbSNP submission, a GenBank reference number is optional information – all you really need is the flanking sequence for your variations. If you are able to find the GenBank reference number and include it in your submission, however, it would make mapping and your submission easier. (**04/10/08**)

Do I submit to GenBank first and wait to submit to dbSNP until the accession numbers are assigned?

You can submit both the GenBank and the SNP data simultaneously since dbSNP data submission does not require a GenBank ID. You need only submit the variations and their flanking sequences to dbSNP. (**04/29/08**)