



## Submitting Data and Submission Processing

Created: April 6, 2007; Updated: February 24, 2014.

### How many variations are in a small submission and how do I send a completed small submission to dbSNP?

**A small submission is defined as** a submission of 1-1000 variations/batch at a rate of 1-10 batches/day.

Once you have completed the Excel file(s) for your small submission, use the detailed instructions in the “[dbSNP VCF Submission Format Guidelines](#)” or in the Flat File “[How to Submit](#)” document to be sure you have formatted your submission correctly.

Send your completed Excel submission file to [snp-sub@ncbi.nlm.nih.gov](mailto:snp-sub@ncbi.nlm.nih.gov)

The VCF submission template ([http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf\\_template.xlsx](http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf_template.xlsx)) contains an example submission found using the “Example” tab at the bottom left corner of the template.

The three flat file submission templates are located at “[templates\\_SNPsub.xls](#)” along with flat file examples in the [submission file](#), which can be found in the /specs subdirectory of the dbSNP FTP site.

### How many variations are in a large submission and how do I send a completed large submission to dbSNP?

**A large submission is defined as** a submission of >1000 variations/batch or at a rate of > 10 batches/day.

Once you have completed the Excel file(s) for your large submission, use the detailed instructions in the “[dbSNP VCF Submission Format Guidelines](#)” or in the Flat File “[How to Submit](#)” document to be sure you have formatted your submission correctly.

Remember, **large submissions to dbSNP are accepted in Variant Call Format (VCF) only.**

To send your large submission to dbSNP, follow these steps:

1. Send an email with the subject line “dbSNP FTP account request” to: [snp-sub@ncbi.nlm.nih.gov](mailto:snp-sub@ncbi.nlm.nih.gov) with your handle information. Once we set up the account, we will send a password for the account to you.
2. Login to the FTP account we created for you and upload your dbSNP VCF formatted submission.
3. Send an email to [snp-sub@ncbi.nlm.nih.gov](mailto:snp-sub@ncbi.nlm.nih.gov), to let us know that you have uploaded to your FTP account.

The VCF submission template ([http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf\\_template.xlsx](http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf_template.xlsx)) contains an example submission found using the “Example” tab at the bottom left corner of the template.

The three flat file submission templates are located at “[templates\\_SNPsub.xls](#)” along with flat file examples in the [submission file](#), which can be found in the /specs subdirectory of the dbSNP FTP site.

### Once I've submitted my Excel submission file, how long will it take before I get a response from dbSNP?

Once you submit, you should get an email from our submission specialist confirming our receipt of your submission within 24 hours.

The VCF submission template ([http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf\\_template.xlsx](http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf_template.xlsx)) contains an example submission found using the “Example” tab at the bottom left corner of the template.

The three flat file submission templates are located at “[templates\\_SNPsub.xls](#)” along with flat file examples in the [submission file](#), which can be found in the /specs subdirectory of the dbSNP FTP site.

### **How long will it be before my submitted variations are assigned a SNP ID number?**

It will take **between one and two weeks** for your variations to be assigned a submitted SNP (ss) number, and a list of these numbers sent to you. Your submission may be delayed for a longer period if it requires correction or if it was submitted before a major holiday. **If you require the ss number by a certain date for a publication**, please let us know at the time of your submission the approximate date you need the ss number, and we will try to get your SNP loaded and the ss number to you by the date specified.

If dbSNP does not send your ss number(s) to you within two weeks after you get confirmation that dbSNP has received your submission, please email us at [snp-sub@ncbi.nlm.nih.gov](mailto:snp-sub@ncbi.nlm.nih.gov)

The VCF submission template ([http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf\\_template.xlsx](http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf_template.xlsx)) contains an example submission, found using the “Example” tab at the bottom left corner of the template.

The three flat file submission templates are located at “[templates\\_SNPsub.xls](#)” along with flat file examples in the [submission file](#), which can be found in the /specs subdirectory of the dbSNP FTP site.

### **I've sent my flat file formatted submission to dbSNP. Will the variations I submitted be assigned a refSNP (rs) number?**

No, a RefSNP number will not be assigned to a variant that was submitted using the Flat File format.

Flat File formatted submissions are now used only to submit those variants whose location could not be placed using an asserted position\*, and therefore had to be submitted with the variant positions reported on unplaced flanking sequence using the Flat File submission format. Flat File submissions will be assigned a submitted SNP (ss) number only. The ss number will be reported in the “Submitted SNP” web report, will be made available on FTP site for download, and will be available for search using the dbSNP homepage “ID search” tool.

**You can access variants submitted with flanking sequence using the Flat File format as archived data until such time as an assembly is available at NCBI that will allow mapping by BLAST** and allow us to possibly assign an rs. dbSNP cannot predict when such an assembly will be made available, or when mapping by BLAST will occur – it could be delayed by months or possibly years. SS numbers can be used in publications describing Assay Variants.

\*Note: An “asserted position” is a statement, or assertion, based on experimental evidence that a variant is located at a particular position.

See the “[VCF: Reporting Variant Positions](#)” section of this document to see various submission scenarios based on the ID associated with the sequence used to report variant location.

### **I've sent my VCF formatted submission to dbSNP. Will the variations I submitted be assigned a refSNP (rs) number?**

**All of your variations will be assigned a submitted SNP (ss) number**, but the assignment of a refSNP (rs) number depends on whether or not the variation is reported on a sequence housed in the [NCBI Assembly Resource](#):

- 1 **If you reported your variant position(s) on a sequence that is part of an assembly housed in the NCBI Assembly Resource**, your variations will be assigned a submitted SNP (ss) number, and then a refSNP number during the next dbSNP build that follows your submission if we can validate the submitted variants and can place them on an assembly.

**Note:** Use dbSNP's "Search by ID" feature on the dbSNP home page to find the corresponding refSNP (rs) number for the ss number you cite in your publication. Again, the refSNP numbers will be assigned to your variants during the next dbSNP build that follows your submission if we can validate the submitted variant and can place it on an assembly.

**2. If you reported a variant position on a sequence that does not yet align to an assembly in NCBI's Assembly Resource** either because there is not yet an assembly to which the sequence aligns, or because the submitted sequence aligns to a gap in an existing assembly, the variant will be assigned an ss number only.

If, however, at some future date a new assembly is created or an old assembly is updated such that the reported variant sequence aligns to an assembly in the NCBI Assembly Resource, the reported variant will be assigned an rs number at that time.

**Note:** If a reported variation is not assigned a refSNP number, it will not appear on maps or graphic representations of the assembly, and will not be integrated with NCBI's other resources. The ss number will, however, be reported on the 'Submitted SNP' web report, will be available for search using dbSNP homepage's 'ID search' tool, and will be made available on FTP site for download.

If a refSNP number is assigned to a variant at a later date, it will appear on maps or graphic representations of the assembly, and will be integrated with NCBI's other resources at that time.

**3. If you reported your variant position on a sequence known only through an assay that provides just the variant and flanking sequence, it will be assigned an ss number only.**

**You can access Assay Variants only as archived data until such time as an assembly is available at NCBI that will allow mapping by BLAST** and allow us to possibly assign an rs. dbSNP cannot predict when such an assembly will be made available, or when mapping by BLAST will occur – it could be delayed by months or possibly years. SS numbers can be used in publications describing Assay Variants.

Note: ss numbers assigned to Assay Variations will be reported In the 'Submitted SNP' web report, will be available for search using the dbSNP homepage 'ID search' tool, and will be made available on FTP site for download.

dbSNP will send you a list of your ss numbers when they have been assigned. If you cite your variations in a publication, you should cite the submitted SNP (ss) numbers since the ss number is unique to each variant you submit.

Since refSNP numbers can change between builds if the genome changes, citing the refSNP number in your publication could possibly mislead users that want to access your SNPs.

The VCF submission template ([http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf\\_template.xlsx](http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf_template.xlsx)) contains an example submission, found using the "Example" tab at the bottom left corner of the template.

The three flat file submission templates are located at "[templates\\_SNPsub.xls](#)" along with flat file example s in the [submission file](#), which can be found in the /specs subdirectory of the dbSNP FTP site.

## Publication Issues

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

You don't have to publish a variation before submitting it to dbSNP. Since some journals require prior submission to dbSNP before they will accept a manuscript, it's actually better to submit to dbSNP prior to publication so you'll have the dbSNP issued ID number (submitted SNP[ss] number) for the variation to cite in your manuscript.

The VCF submission template ([http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf\\_template.xlsx](http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf_template.xlsx)) contains an example submission, found using the "Example" tab at the bottom left corner of the template.

The three flat file submission templates are located at “[templates\\_SNPsub.xls](#)” along with flat file example in the [submission file](#), which can be found in the /specs subdirectory of the dbSNP FTP site.

### **What is dbSNP’s Hold Until Published (HUP) policy?**

dbSNP doesn't have a "hold until published" policy, and will not release data on a specific date or for a specific dbSNP build. If your manuscript, however, requires a dbSNP ID number (submitted SNP[ss] number) for the review process, we can hold the submitted data until the publication is accepted and the submitter has given us notification that dbSNP may release the data. dbSNP will then attempt to release it during the next build release cycle after notification.

The normal dbSNP build cycle is between 6-8 weeks, but can be longer depending on how long it takes to complete the dbSNP pipeline.

Note: The VCF submission template ([http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf\\_template.xlsx](http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf_template.xlsx)) contains an example submission, found using the “Example” tab at the bottom left corner of the template. The three flat file submission templates are located at “[templates\\_SNPsub.xls](#)” along with flat file example in the [submission file](#), which can be found in the /specs subdirectory of the dbSNP FTP site.

### **I need to get dbSNP IDs to put in my paper. Can I get dbSNP IDs assigned by a specific date so I won’t miss my journal’s deadline?**

If you require an ss number for a variation by a certain date for a publication, please let us know the approximate date you need the ss number at the time of your submission, and we will try to get your SNP loaded and the ss number to you by the date specified.

If no time constraint is reported to us, submission processing and assignment of ss numbers will take one to two weeks.

Note: The VCF submission template ([http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf\\_template.xlsx](http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf_template.xlsx)) contains an example submission, found using the “Example” tab at the bottom left corner of the template. The three flat file submission templates are located at “[templates\\_SNPsub.xls](#)” along with flat file example s in the [submission file](#), which can be found in the /specs subdirectory of the dbSNP FTP site.

## **Data Types that can be Submitted to dbSNP**

### **Can I submit a variation if I don’t have genotype or frequency data?**

Yes, you can submit a variation without genotype or frequency data.

If you are submitting using VCF, submitting genotype and frequency data is optional, so all you have to do is fill out the VCF submission template without using the optional INFO tags for frequency and genotype.

If you are submitting using the flat file format, just use the SNPASSAY template alone to create a variation submission — don’t use the SNPINDUSE and SNPPOPUSE templates if you don't have genotype or frequency data.

Note: The VCF submission template ([http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf\\_template.xlsx](http://www.ncbi.nlm.nih.gov/projects/SNP/docs/vcf_template.xlsx)) contains an example submission, found using the “Example” tab at the bottom left corner of the template. The three flat file submission templates are located at “[templates\\_SNPsub.xls](#)” along with flat file example s in the [submission file](#), which can be found in the /specs subdirectory of the dbSNP FTP site.

### **Can I submit just genotypic data for existing variants in dbSNP?**

Yes, you can submit genotypic data for existing refSNP (rs) numbers using either the VCF or flat file format:

**VCF:** Submit genotype data for existing variations, you can submit it using the dbSNP's [VCF Submission template](#). Follow the [genotype submission example](#) provided by 1000 Genome Project in their description of VCF version 4.1.

**Flat File:** Submit genotypic data for existing SNPs by submitting a completed SNPINDUSE worksheet, which you can create by using the SNPINDUSE template, located at "[templates\\_SNPsub.xls](#)", in the [submission file](#), which can be found in the /specs subdirectory of the dbSNP FTP site.

### Can I submit genotype or allele frequency information for existing variants in dbSNP?

Yes, you can submit genotypic or allele frequency data for existing refSNP (rs) numbers using either the VCF or flat file format:

**VCF:** Submit genotype data for existing variations, you can submit it using the dbSNP's [VCF Submission template](#). Follow the instructions for frequency data submissions found in the [dbSNP VCF Submission Format Guidelines](#) and follow the [genotype submission example](#) provided by 1000 Genome Project in their description of VCF version 4.1.

**Flat File:** Submit genotype frequency and allele frequency information for existing variants by submitting a completed SNPPOPUSE worksheet, created by using the SNPPOPUSE template, located at "[templates\\_SNPsub.xls](#)", in the [submission file](#), found in the /specs subdirectory of the dbSNP FTP site.

### What is the smallest data submission dbSNP will accept?

The smallest SNP submission consists of a variation's observed alleles, its asserted position\* (or 5' and 3' flanking sequences in flat file submissions), the name of the gene in which the SNP is located, and the NCBI GenBank accession number of the genomic sequence in which the SNP is located:

**VCF:** Fill in the [VCF Submission template](#) using your data and follow the instructions provided in the grey portions of the template and in the [dbSNP VCF Submission Format Guidelines](#). The VCF submission template contains an example submission you can access by clicking the "example" tab at the bottom left corner of the Excel Spreadsheet.

**Flat File Format:** Submit this data using a SNPASSAY submission worksheet, created using the SNPASSAY template, located at "[templates\\_SNPsub.xls](#)", in the [submission file](#), found in the /specs subdirectory of the dbSNP FTP site.

\*NOTE: An "asserted position" is a statement, or assertion, based on experimental evidence that a variant is located at a particular position.

## Using your Handle to Find your Assigned ss Numbers in dbSNP

**I submitted variations some time ago, but I'm unable to find the list of my submitted SNPs (ss) numbers. How do I find them this using my Handle?**

You can find the submitted SNP (ss) numbers assigned to your submissions by following these steps:

1. Go to the old [dbSNP Home page](#)
2. Go to the "Submission Information" section and click on the first link in the section: "By Submitter", which will take you to the "[Search/View Submitter Detail](#)" page
3. Type your Handle (I used KWOK) into the text box, make sure the radio buttons for "Submitter handle" and "Starts with" have been selected, and click the "Search" button.
4. The "Search/View Submitter Detail" page will [refresh](#), showing a table of handles at the bottom.
5. Click on your handle (located on the far right in the handle column (for this example I will click on KWOK), which will take you to the "[Submitter Contact Detail](#)" page.

6. On the the “Submitter Contact Detail” page, click on the “Submitter Batch ID” of interest (in this case I chose Assay 10.23.98), to go to the “[View SNP Submission Batch](#)” page, where all the submitted SNP (ss) numbers assigned to the SNPs submitted in that particular batch are listed.