



## General Information about dbSNP as a Database Resource

### What is dbSNP?

dbSNP is world's largest database for nucleotide variations, and is part of the National Center for Biotechnology Information (NCBI), an internationally respected resource for molecular biology information. As of this date, dbSNP is comprised of a large cluster of species-specific databases that contain over 12 million non-redundant sequence variations (single nucleotide polymorphisms, insertion/deletions, and short tandem repeats) and over 1 billion individual genotypes from HapMap and other large-scale genotyping activities — over 200GB of data and growing daily.

dbSNP, along with other NCBI resources, is anticipated to house data from the Genome-Wide Association Study Database (GWAS), a project designed to produce 10 terabytes of variation and health/disease association data. It is hoped that the ability to discover genetic variants associated with health and disease will facilitate our understanding of the genetic contribution to the pathogenesis and progression of disease. The genes and pathways discovered through genome-wide association studies may lead to new treatment strategies for preventing disease. (8/11/06)

### What type of database is dbSNP?

dbSNP is a relational database implemented on an SQL server. [Schema and table data](#) can be found on the dbSNP FTP site.

### Can you tell me more about the back-end database technology dbSNP utilizes?

- Here are a few links relevant to your question:
- [Database schema and datafile](#)
- [ERD](#)
- [Schema change from build to build](#)
- [Data Dictionary](#)

The ERD and data dictionary are current with the exception of a few minor changes listed in schemaChange doc.

### When was dbSNP created and released to the public?

dbSNP was first created and released to public in September of 1998. (11/10/05)

### I am doing a report on the SNP database. Can you provide me with information about the usefulness and completeness of dbSNP?

Here are links to our [Coffee break page](#), the [SNP handbook](#) and the [dbSNP primer](#) that you may find helpful. (ud:03/13/08)

### Is there a published manuscript that describes the dbSNP database and website?

In addition to a [NAR article](#) published in 2001, you might want to look at the [dbSNP Handbook](#) as well as the “[General Information](#)” and “[Data Content Information](#)” sections the dbSNP FAQ Archive. (11/15/07)

**How may I contribute my DNA for participation, anonymously or not, in dbSNP, human genome projects, and/or in other genetic projects?**

Thank you for your interest, but dbSNP does not engage in direct research with, or the collection of, DNA samples. dbSNP is a public database that stores genetic variation data submitted from research laboratories located around the world, and provides tools for searching and retrieving these data. For more information about dbSNP, consult the [General information section of the Archive](#).

You may wish to contact the [NIH Clinical Center](#) or another biomedical institution looking for research volunteers. (11/23/07)