dbSNP: Database of Short Genetic Variations

National Library of Medicine
National Center for Biotechnology Information

Catalog of short nucleotide changes for human genome and its annotated transcripts

Scope and Access

The NCBI Short Genetic Variation database [1], commonly known as dbSNP, catalogs short variations in human nucleotide sequences. Variations include single nucleotide variations, as well as insertions, deletions, and short tandem repeats less than 50 nucleotides in length. Short genetic variations may be common, thus representing true polymorphisms, or they may be rare. Some rare human entries have additional information including disease associations from NCBI's ClinVar database [2], genotype information and allele origin (germline / somatic).

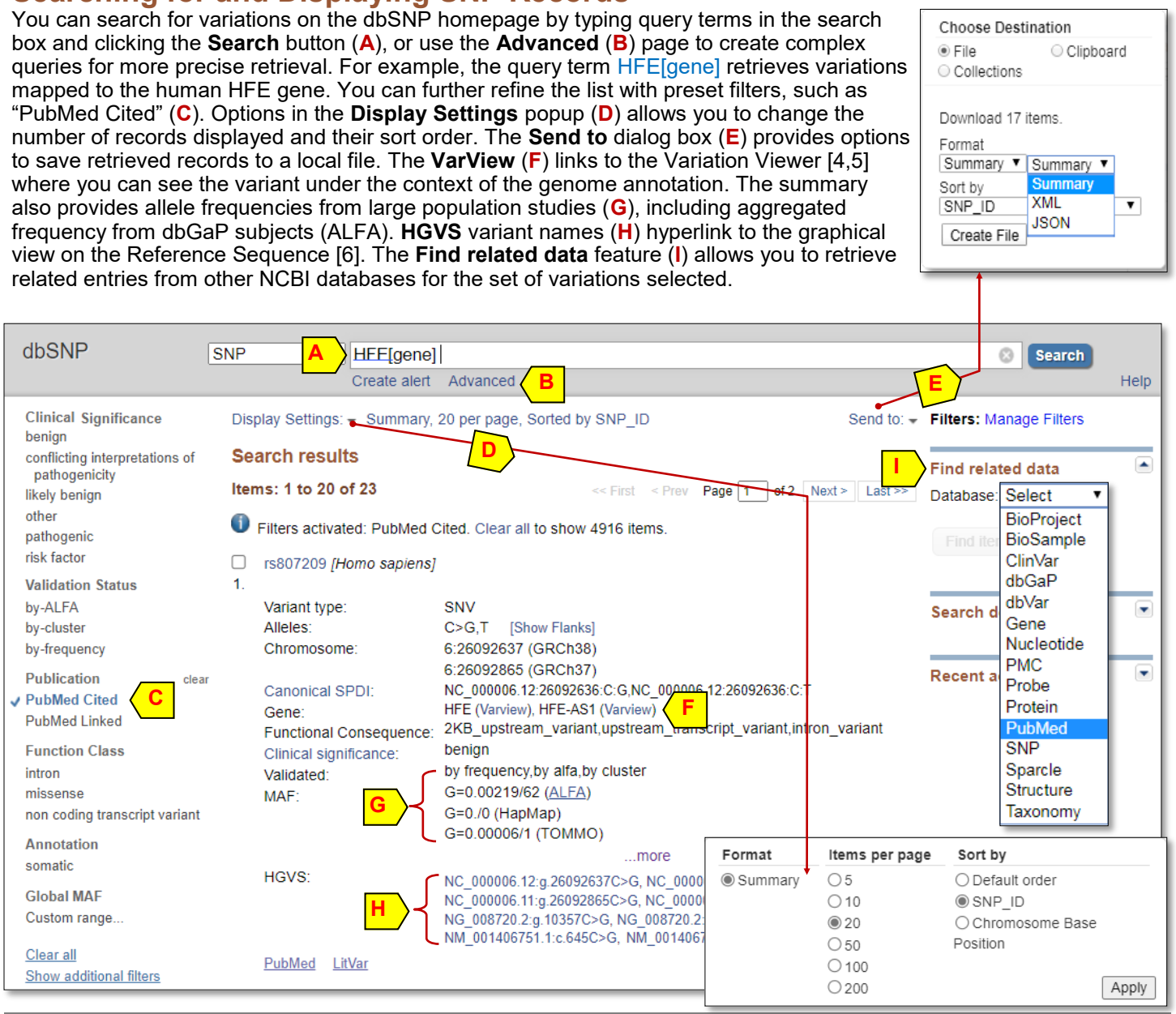
You can access short nucleotide variation data from the dbSNP homepage www.ncbi.nlm.nih.gov/snp or the Entrez Programming Utilities API www.ncbi.nlm.nih.gov/books/NBK25501.

Download VCF and JSON files through FTP ftp.ncbi.nlm.nih.gov/snp/ or Aspera at www.ncbi.nlm.nih.gov/public/?snp/. The dbSNP API service, SPDI [3], is available at: api.ncbi.nlm.nih.gov/variation/v0/.

You can also access dbSNP and structural variation data from dbVar through the Variation Viewer www.ncbi.nlm.nih.gov/variation/view/.

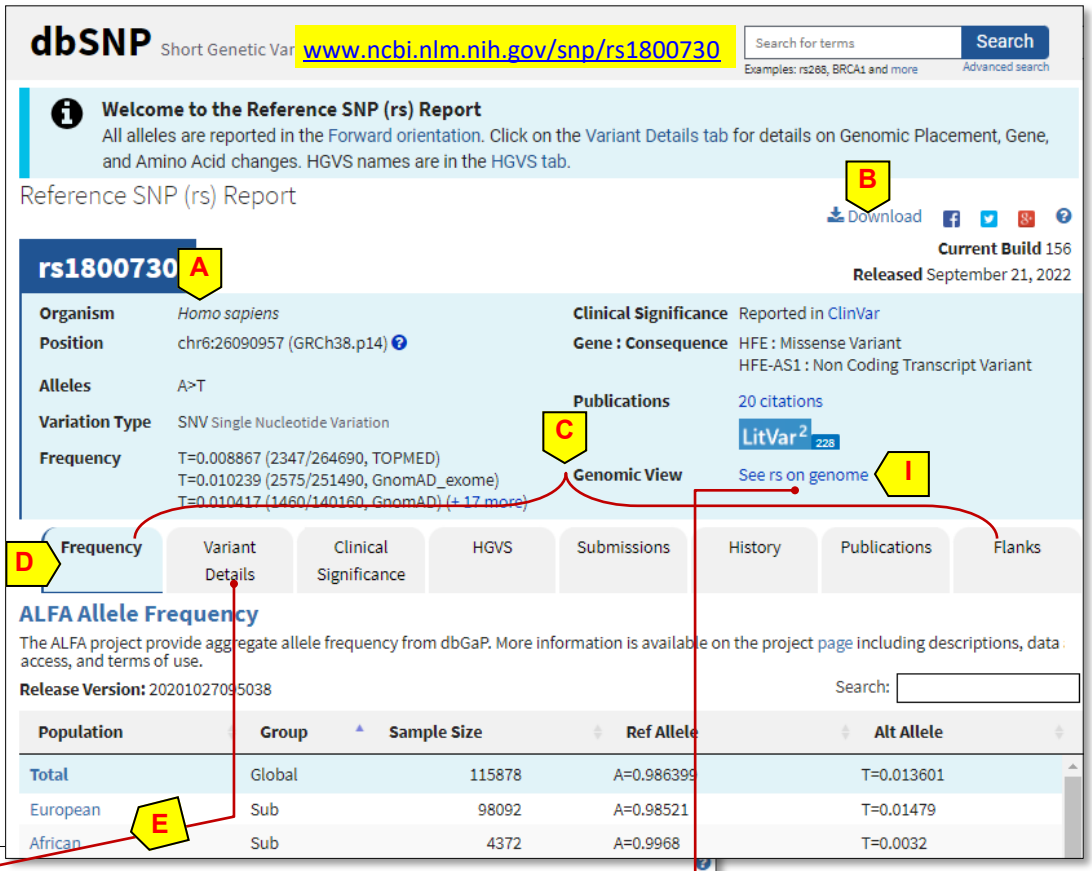
Searching for and Displaying SNP Records

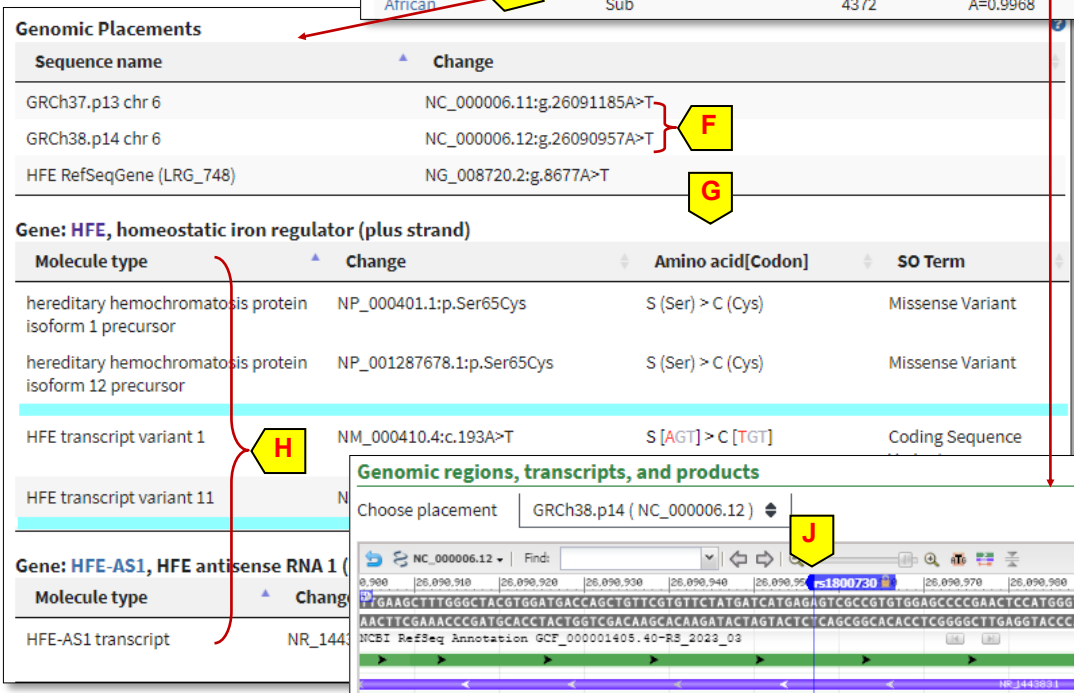
You can search for variations on the dbSNP homepage by typing query terms in the search box and clicking the **Search** button (**A**), or use the **Advanced** (**B**) page to create complex queries for more precise retrieval. For example, the query term HFE[gene] retrieves variations mapped to the human HFE gene. You can further refine the list with preset filters, such as "PubMed Cited" (**C**). Options in the **Display Settings** popup (**D**) allows you to change the number of records displayed and their sort order. The **Send to** dialog box (**E**) provides options to save retrieved records to a local file. The **VarView** (**F**) links to the Variation Viewer [4,5] where you can see the variant under the context of the genome annotation. The summary also provides allele frequencies from large population studies (**G**), including aggregated frequency from dbGaP subjects (ALFA). **HGVS** variant names (**H**) hyperlink to the graphical view on the Reference Sequence [6]. The **Find related data** feature (**I**) allows you to retrieve related entries from other NCBI databases for the set of variations selected.



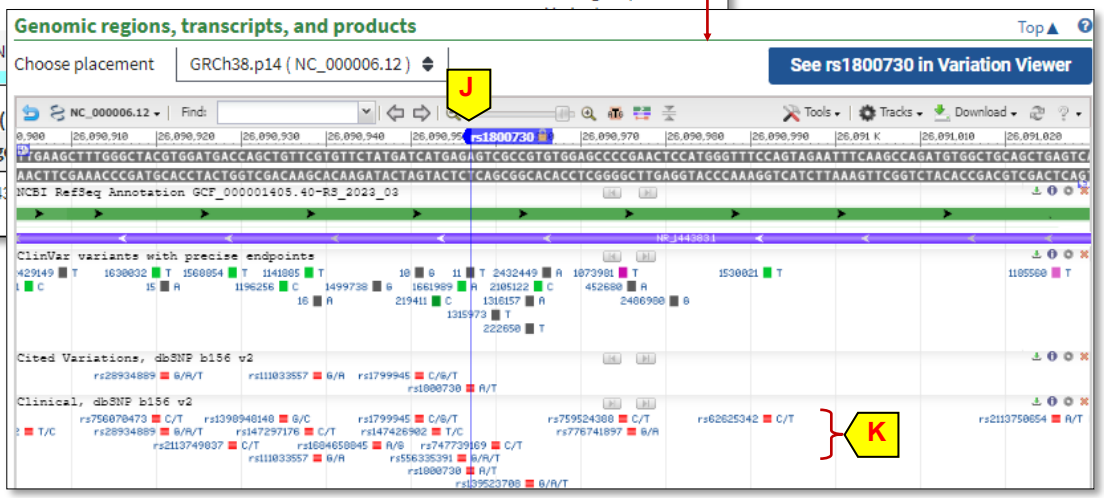
The Reference SNP Report

The Reference SNP Report linked from a rsID (right & below) shows details of a variant record. The summary at the top (**A**) provides an overview of the variant. The allele change is relative to the plus strand of the chromosome. The report is also available in JSON format through the **Download** link (**B**). The Reference SNP report groups available details into categories (**C**) under separate tabs. The default **Frequency** tab (**D**) lists allele frequencies from major studies, such as ALFA from dbGaP samples and 1000 Genomes, broken down by subpopulation if available.





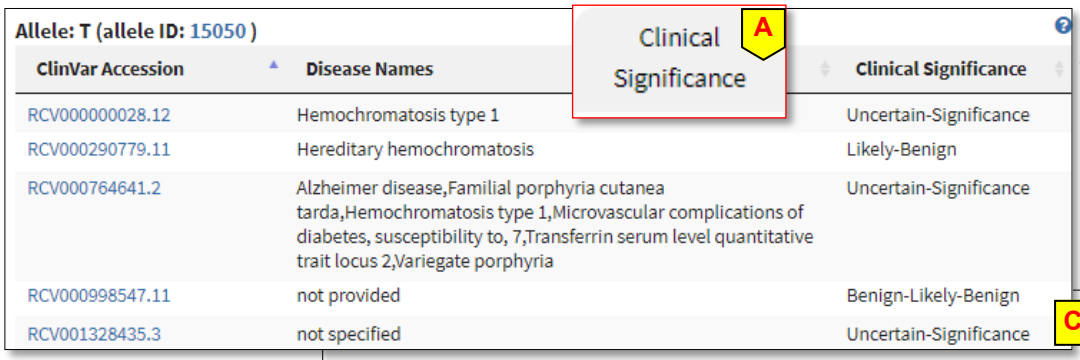
The **Variant Details** tab (**E**) displays the genomic placement (**F**) of the variant on the previous and current genome assemblies. The gene table (**G**) presents the gene and transcript mapping plus their relevant molecular consequences. Variants mapping to multiple genes, there will be multiple gene tables (**H**).



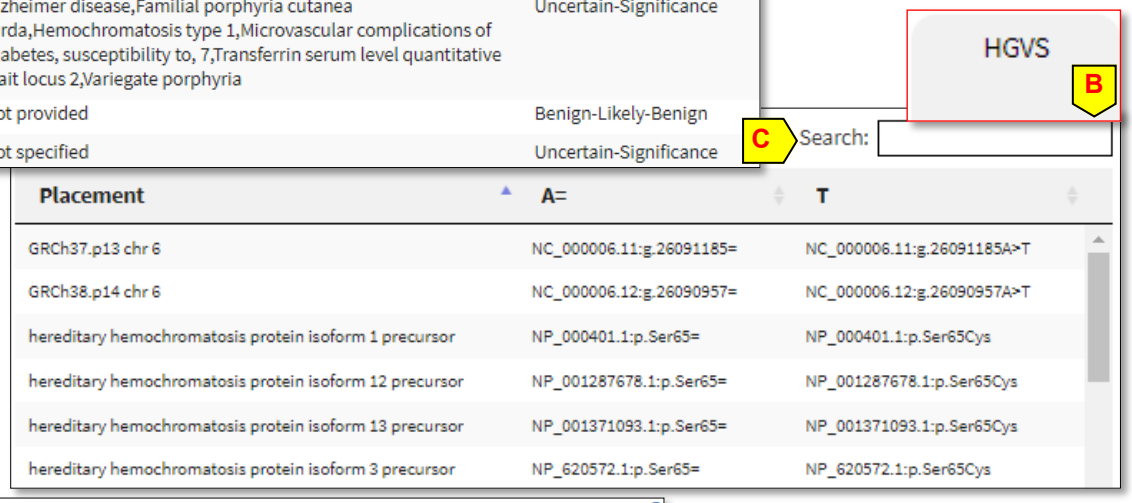
The **See rs on genome** (**I**) links to the graphical panel of the display (right) that shows the variant in the context of genome and other neighboring variants in a graphical display (**J**). Different tracks show variants with specific attributes, such with literature citations in the literature or with clinical significance from ClinVar (**K**).

The Reference SNP Report (cont.)

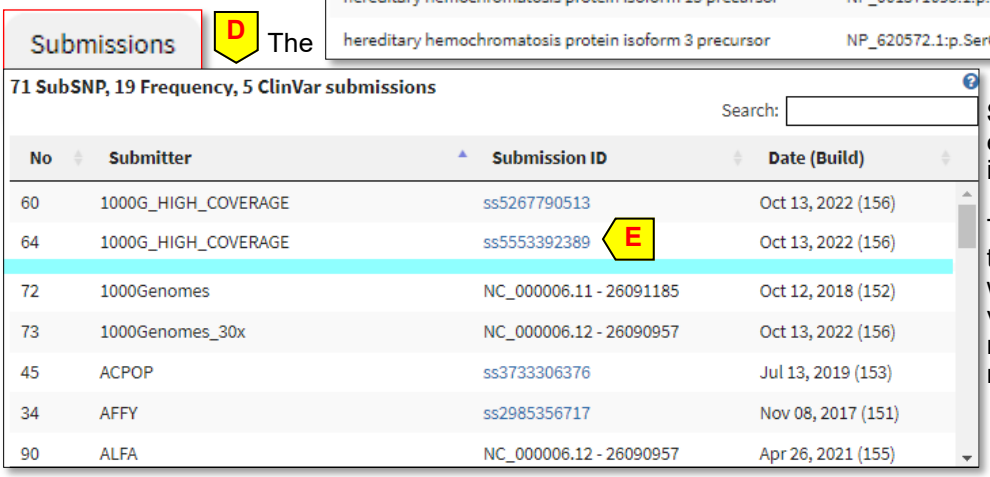
Other tabs in the new Reference SNP Report provide additional information.



The **Clinical Significance** tab (**A**) lists related clinical assertions for the variant from ClinVar, with links to ClinVar records.

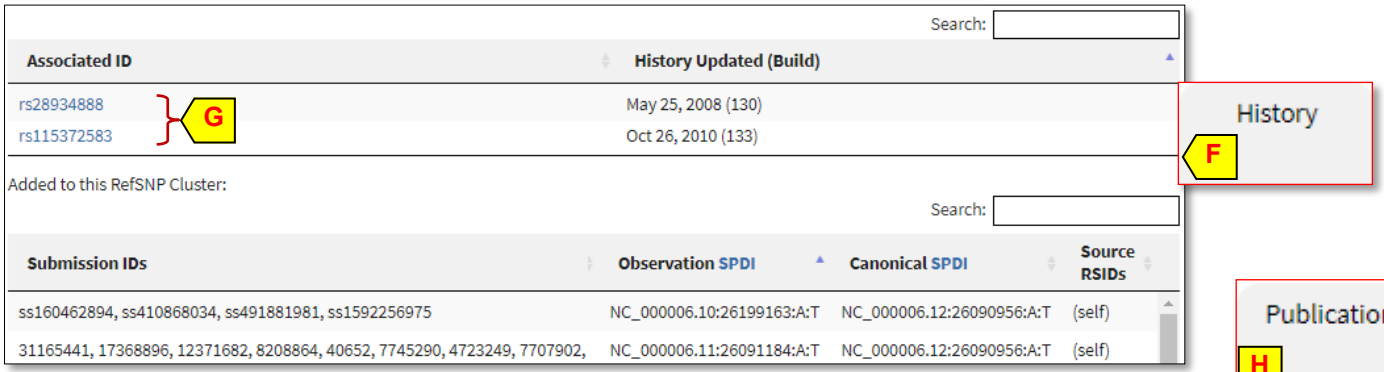


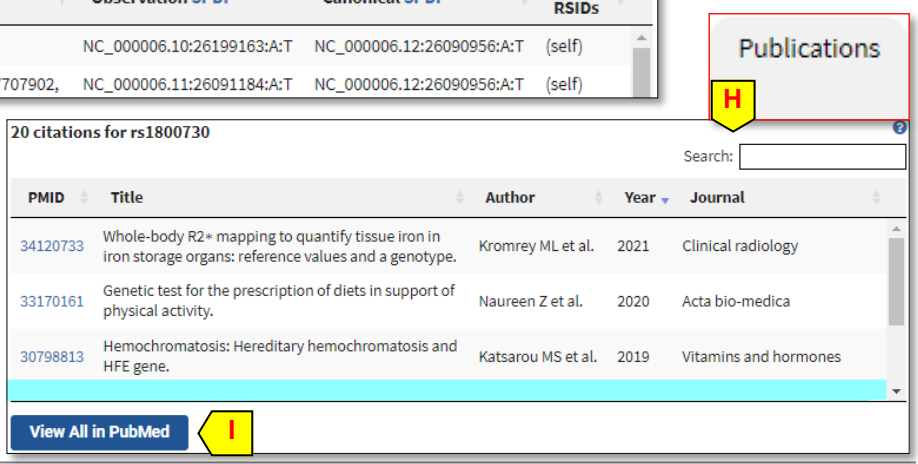
The **HGVS** tab (**B**) lists HGVS names describing the variant using different reference accessions. Use the table header to sort names or the search box (**C**) to filter the list.



The **Submission** tab (**D**) lists submitted entries from large projects or individual submitters (**E**).

The **History** tab (**F**) tracks changes of the cluster and lists other rsIDs that were merged with the variant. Here, variants rs115372583 and rs28934888 (**G**) were merged into rs1800730.

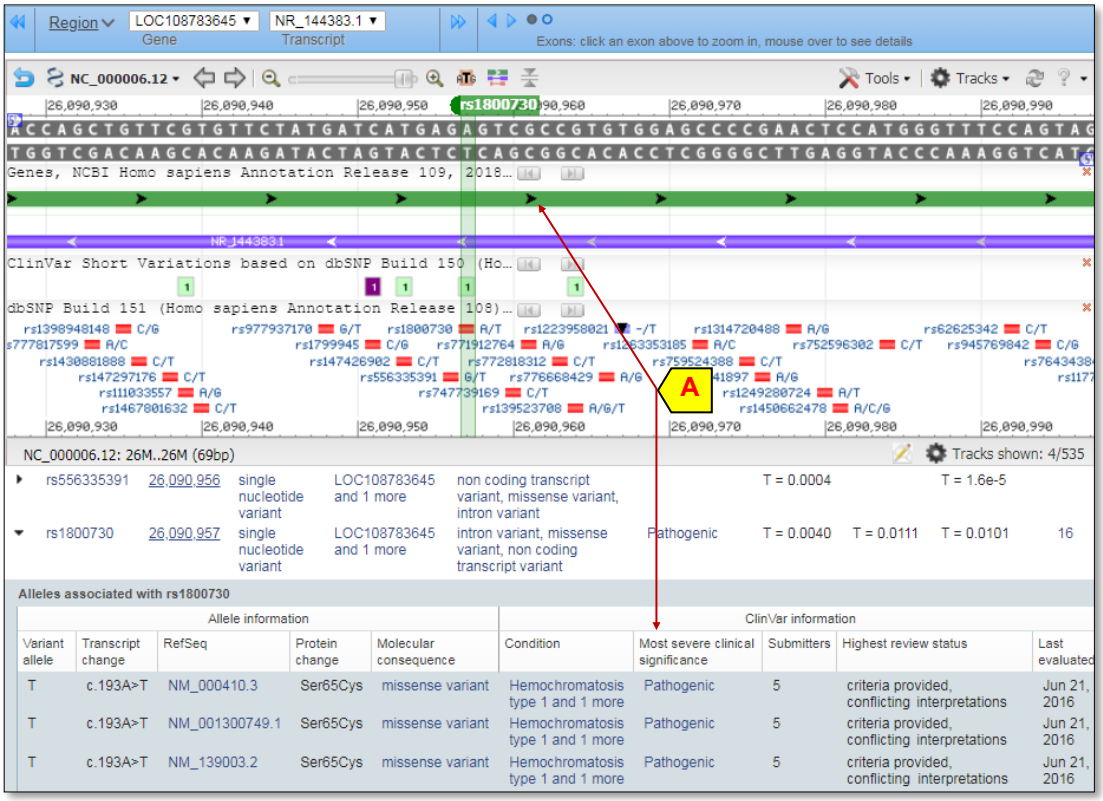




dbSNP connects Reference SNP variants and biomedical literature citations through text-mining. The Reference SNP Report displays relevant citations under the **Publications** tab (**H**). Use the "View All in PubMed" button (**I**) to retrieve the list of citations in PubMed and examine their abstracts for more information. Some of these citations may also have free full-text available from PubMed Central (PMC).

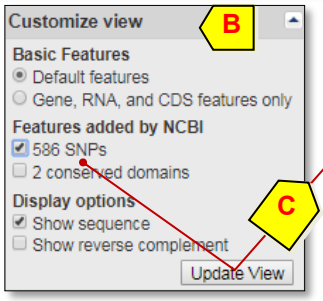
Variation Viewer

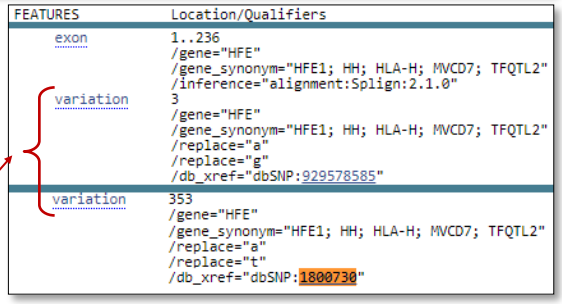
The Variation Viewer provides an interactive display of the variant under the context of annotation of the selected genome assembly. It also provides a table of variants in the display with molecular and clinical consequences (**A**). Facet filters in the left column (not shown) allows you to filter for variants with particular characteristics. More information on this tool is available online [4, 5]



Other Ways to Access dbSNP Data

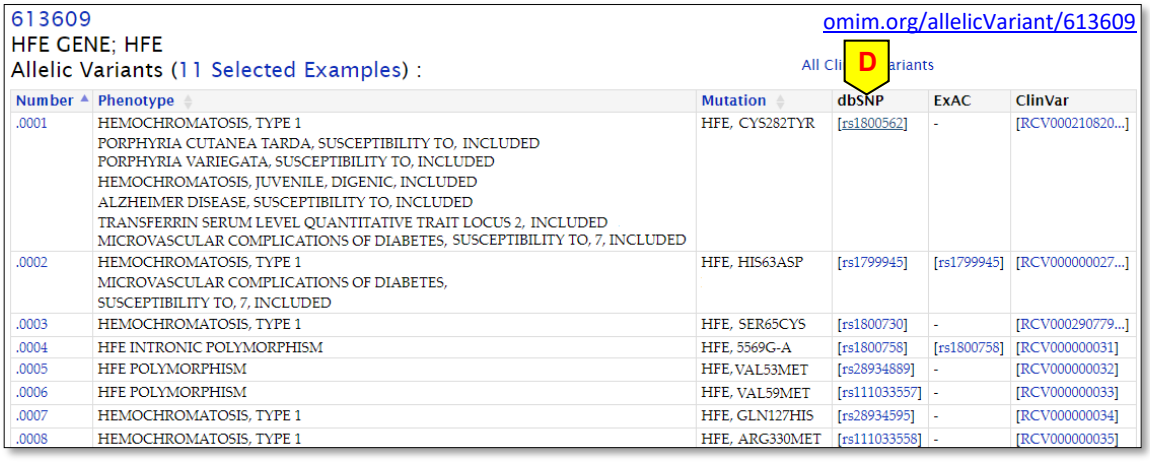
The dbSNP database is fully integrated with the Entrez system that provides links to SNP records from other NCBI databases. For example, you can show variations mapped to a RefSeq mRNA record (with NM_ accessions) by using the **Customize view** (**B**) menu in the upper right hand corner of the sequence record, simply check the SNPs checkbox and click **Update View** (**C**).





www.ncbi.nlm.nih.gov/nuccore/NM_000410.3

dbSNP also includes variations reported in literature, collected by OMIM or submitted to ClinVar. For example, the OMIM Allelic Variants table links to dbSNP entries (**D**).



References

1. The Database of Short Genetic Variation (dbSNP). Kitts A, et. al. In The NCBI Handbook [Internet], 2nd ed.
2. ClinVar: improving access to variant interpretations and supporting evidence. Landrum MJ, et al. Nucleic Acids Res. 2018 Jan 4;46(D1):D1062-D1067.
3. New Web Services for Comparing and Grouping Sequence Variants. go.usa.gov/xUeKT.
4. Variation Viewer factsheet. ftp.ncbi.nih.gov/pub/factsheets/Factsheet_Variation_Viewer.pdf
5. Variation Viewer video tutorial. www.youtube.com/watch?v=rnWZ9MFBwUM
6. Graphical Sequence Viewer factsheet. ftp.ncbi.nih.gov/pub/factsheets/Factsheet_Graphical_SV.pdf