ClinVar: Clinically Relevant Sequence Variations

An archive of medically relevant variants and their clinical interpretations

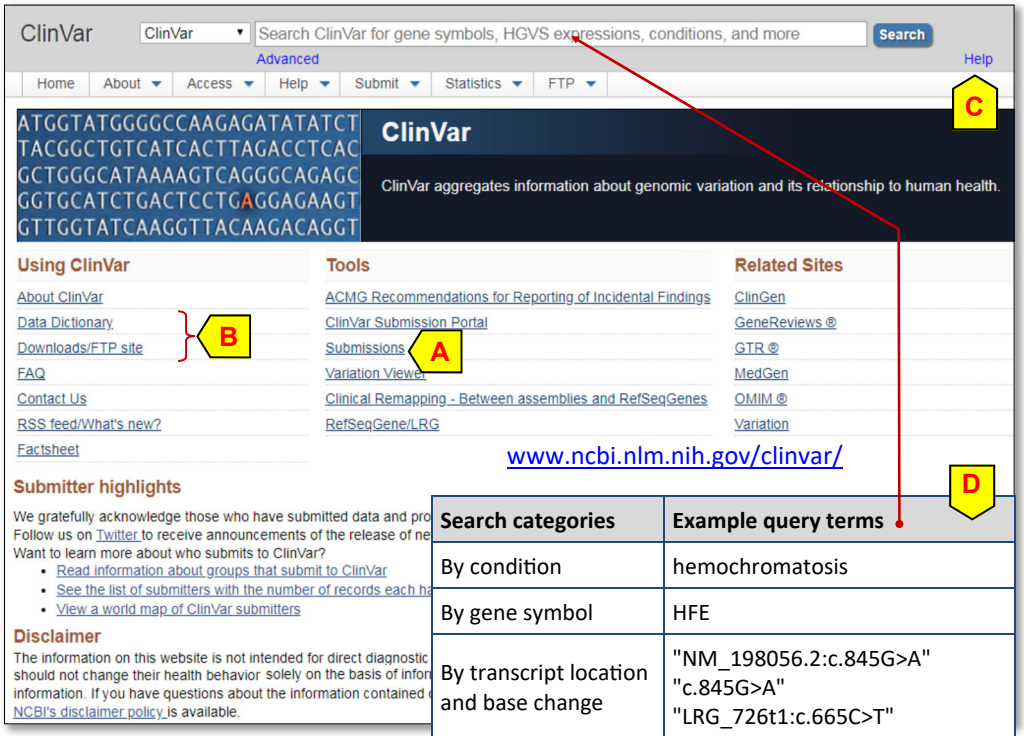
https://www.ncbi.nlm.nih.gov/clinvar/

National Center for Biotechnology Information • National Library of Medicine • National Institutes of Health • Department of Health and Human Services

Overview

Accurate and timely interpretation of genetic testing results is critical to translating genomics to clinical care. ClinVar supports the medical genetics community as a freely available, public archive of the relationships between medically important variants and phenotypes. It allows testing laboratories access to a broader set of clinical interpretations than they may have collected on their own, and the ClinVar data can be incorporated into their daily workflow. ClinVar is also available to individual users and organizations that want to incorporate it into their own applications. Data providers submit observed variants and make an assertion about the clinical significance of each variant with respect to a phenotype. The interpretation may be based on clinical testing, research, or the literature. Various types of evidence may also be provided to support the assertion. We continue to work closely with several genetic testing labs and other end users to refine the submission process and display of the data to maximize its utility for the clinical genetics community.

ClinVar adds value to submissions in several ways. Submissions for the same variant and phenotype pair from different submitters are aggregated, so that agreement or conflict in clinical significance is clear and evidence from different submitters can be viewed together. Accession numbers are assigned to individual submissions (SCV) and to aggregate records (RCV) to facilitate retrieval; version numbers allow tracking of updates to each record as submitters refine clinical interpretations over time. Both SCV and RCV records are given a review status which allows the user to evaluate the validity of each interpretation. ClinVar supports standardized descriptions of both variant and phenotype, by providing HGVS expressions at the genomic, cDNA, and protein level and phenotype terms reported in MedGen. The molecular consequence is predicted for variants within a coding region based on the effect of the sequence change on translation, and for others variants in a gene by reporting their location (UTR, splice site). Curation by NCBI staff may also add published allele names, citations, and links to the same variant in other databases. Although ClinVar provides aggregation, standardization, and a central repository, the database is driven by submission of data from the clinical genetics community. ClinVar provides limited curation of variant and phenotype terms, but **clinical interpretations are provided by submitters.**



D

Search categories	Example query terms
By condition	hemochromatosis
By gene symbol	HFE
By transcript location and base change	"NM_198056.2:c.845G>A" "c.845G>A" "LRG_726t1:c.665C>T"
By protein location and residue change	Ile105Thr
By genomic location and base change	"NC_000001.10:g.11856378G>A" "g.11856378G>A"

Data Submission

ClinVar welcomes submissions from clinical testing labs, research labs, locus-specific databases, clinicians, patient registries, expert panels and professional societies. Two Excel spreadsheet templates are available from the **Submissions** link (**A**), one for submissions with minimal data and one for all types of submissions. For more detailed submissions by XML, the xsd is available on the ftp site and a Data Dictionary (**B**), which defines data elements in ClinVar, is available from the home page. The data required for submission includes a valid variant description (by HGVS, genomic location, or cytogenetic description), the disease or phenotype for which the variant was interpreted, and the interpretation. Consider submitting supporting evidence, such as the number of observations of the variant, mode of inheritance, presence of family history or segregation, since they greatly enhance the utility of the submitted interpretation. On NCBI's ClinVar Submission Portal (submit.ncbi.nlm.nih.gov/clinvar/), use the Submission Wizard for guided entry of a single interpretation, or upload a submission file directly. Refer to the help document (**C**) in the upper right hand corner for additional details.

The table (**D**) lists a few sample query terms that can be used in ClinVar searches.

Data Access

For bulk download and analysis, the data is available on ClinVar's ftp site (ftp.ncbi.nlm.nih.gov/pub/clinvar/) as VCF, XML, and tab-delimited summary files.

The ClinVar website provides access to variation records. Users can search (**A**) for variants, phenotypes, genes, proteins, MIM numbers, dbSNP RSIDs, and other data fields. Filters (**B**) can be used to restrict the results by clinical significance, variation type, molecular consequence, review status, and other fields (as in **C**, by pathogenic & multiple submitters).

Search results list variations reported to ClinVar and link to each variation page, which provides general information about the allele(s), clinical significance (**D**), conditions reported for the variant, and a link to ClinVar's variant-disease (RCV) record (**E**), as well as affected genes (**F**).

The variation page displays tabs with the details of clinical assertions (**G**), a summary of the evidence provided per submitter, and the details of each observation made by each submitter (**H**).