2015-12-03_Web_Release

Release notes, Dec 1, 2015

Re-arrangement of location of reporting the review status and clinical significance

In early November 2015, the ClinVar variation page was updated so that the sections reporting the review status and clinical significance are now located prominently in the upper left. Review status is in the boxed section under the title; the clinical significance is now reported in the section titled "Interpretation".

Two identifiers were also added to the display. The Variation ID is displayed near the review status; this identifier represents the set of variants that were interpreted (for most ClinVar records, this is a single variant). The Allele ID is displayed in each Allele(s) section; each individual variant in the set is represented by its own Allele ID. For example on this record for a haplotype of six SNPs:

http://www.ncbi.nlm.nih.gov/clinvar/variation/208474/

the set of six SNPs is represented by Variation ID 208474. Each allele in the haplotype is represented by its own Allele ID, e.g. Allele ID 198596 represents NM_016124.4:c.602C>G.

Alleles reported in studies in dbGaP

There is a new section on the right of the page that reports whether an allele has been observed in studies submitted to dbGaP and what the frequency was. This section is entitled **Variant frequency in dbGaP**. The info icon provides more details, as copied here:

This table lets you know if the ClinVar variant is observed in dbGaP data. It provides both the count of submissions to dbGaP with the variant and the frequency of the variant in the database, not in the population. It reports two measures: called variants when the variant is observed in the set formally submitted with a dbGaP study, and potential variants when NCBI found the variant in aligned NGS with the dbGaP study, but we can take reads that were submitted with the study, align them to the genome, and see if that variant is observed in the reads.

Formally, called variants are submitted to dbGaP as subject genotypes and counted when a sample (subject genotype from VCF file) has at least one allele present; potential variants are computed by examination of any aligned next generation sequence for the sample that covers the position. A sample (SRA run accession) is counted when the allele is observed in at least 30% of the reads covering the position, and 10 or more clean reads (not flagged as artifact or PCR duplicate) cover the position.

Variant counts are updated on a monthly basis.

Separation of the report of overall statistics from the list of submitters

ClinVar now provides two pages specific to reporting statistics:

- http://www.ncbi.nlm.nih.gov/clinvar/docs/submitter_list/, which provides a listing of all submitters and counts specific to each submitter
- http://www.ncbi.nlm.nih.gov/clinvar/submitters/, which provides aggregate counts of submitters and what they submitted

Helpful hint

Did you know you can query ClinVar by identifiers from OMIM?

ClinVar supports retrieval of data by searching on MIM numbers for genes or phenotypes (*e.g.* http://www.ncbi.nlm.nih.gov/clinvar/?term=312750[mim]) as well as allelic variant identifiers (*e.g.* http://www.ncbi.nlm.nih.gov/clinvar/?term="191315.0003"[Variant ID])