20130905_data_release_notes

Summary of changes in this data release: September, 2013

Content

Preferred names of short variants

In this release, the method of calculating the preferred name of each short genetic variant allele was changed from the pattern of RefSeq accession.version:cDNA change (single letter abbreviation of any protein change), *e.g.*NM_005633.3:c.1654A>G (R552G), to a pattern of gene symbol:cDNA change (3-letter code for protein change), *e.g.* SOS1:c.1654A>G (p.Arg552Gly). This change was implemented to make the affected gene more obvious.This correction was first available on the web interface August 21, 2013.

Preferred names of structural variants from dbVar

A processing error resulted in using the Variant Call ID (nssv) instead of the Variant Region ID (nsv). This correction was first available on the web interface August 21, 2013.

Alternate descriptions

Some alternate descriptions of variants were removed, based on the analysis that the terms did not provide distinctive information content.

Tab-delimited directory on FTP

A new directory for tab-delimited reports was added to the ClinVar FTP site:

ftp://ftp.ncbi.nlm.nih.gov/pub/clinvar/tab_delimited/

variant_summary.txt.gz has recently been added and will be updated monthly.

Scope of the VCF file

In this release. ClinVar began to report only variations explicitly stored in ClinVar in the VCF file. In previous releases, the VCF file included both records in ClinVar and records submitted to dbSNP that included some representation of phenotype or that had been provided via HGVS-based spreadsheets. As a result of this change, the variants reported in .vcf format and .xml format should be consistent.

ClinVar continues to contact previous submitters to dbSNP to determine if they want their submissions represented in ClinVar, so that the representation of variants of medical interest are as comprehensive as possible. There is no evidence that the variants included previously were reported in error.