

APPENDIX A: USEFUL WEBSITES FOR CLINGEN GENE CURATORS

The following websites are free and publically available. While this list is not exhaustive, it includes websites that are often used during the ClinGen gene curation process. A brief description for each website is given below; please go to the websites for more information.

If there are additional websites that you think curators should be aware of, please contact Jenny Goldstein (jennifer.goldstein@unc.edu)

LITERATURE SEARCHES

- **PubMed**
 - <https://www.ncbi.nlm.nih.gov/pubmed>

REVIEWS/DISEASE ENTITIES

- **Online Mendelian Inheritance in Man (OMIM)**
 - <http://www.ncbi.nlm.nih.gov/omim>
A comprehensive compendium of human genes and phenotypes that is updated regularly. Summaries of gene-disease associations and references to primary literature can be found here.
- **GeneReviews**
 - <http://www.ncbi.nlm.nih.gov/books/NBK1116/>
Provides clinically relevant information for hundreds of different inherited conditions. The “Molecular Genetics” section of each entry may be useful for information on common variants for a gene. The “Establishing the Diagnosis” section typically contains a summary of the genetic testing options, including the different genes involved and proportion of cases caused by variants in each gene.
- **Monarch Disease Ontology (MonDO)**
 - <https://www.ebi.ac.uk/ols/ontologies/mondo>
Human disease ontology merging information from multiple disease resources.
- **ORPHANET**
 - <http://www.orpha.net>
Online inventory of human diseases.

PHENOTYPES

- **Human Phenotype Ontology (HPO) Browser**
 - <http://www.human-phenotype-ontology.github.io/>
Standardized vocabulary and codes for human phenotypic abnormalities.

GENES AND GENE PRODUCTS

- **HUGO Gene Nomenclature Committee (HGNC)**
 - <http://www.genenames.org>
An online repository of approved gene nomenclature.
- **National Center for Biotechnology Information (NCBI) gene**
 - <http://www.ncbi.nlm.nih.gov/gene>
Integrates information from a wide range of species. Includes gene nomenclature, reference sequences, maps, expression, protein interactions, pathways, variations, phenotypes, functional evidence (in **GeneRIFs**) links to locus-specific resources.

GENES AND GENE PRODUCTS (continued)

- **Ensembl**
 - <http://www.ensembl.org/index.html>
Nomenclature, splice variants, references sequences, maps, variants, expression, comparative genomics, ontologies, and function.
- **UCSC Genome Browser**
 - <https://genome.ucsc.edu/>
Genome browser with access to genome sequence data from a range of species.
- **UniProt**
 - www.uniprot.org
Comprehensive resource for protein sequence and functional information.

VARIANT DATABASES

- **ClinVar**
 - <http://www.ncbi.nlm.nih.gov/clinvar/>
Public archive of human gene variants and phenotypes submitted by clinical and research laboratories, genetics clinics, locus specific databases, expert groups, and OMIM.
- **Leiden Open Variation Database (LOVD)**
 - <http://www.lovd.nl/3.0/home>
Listings of variants within human genes and associated phenotypes; includes links to locus-specific databases.

ALLELE FREQUENCIES

- **Exome Aggregation Consortium (ExAC)**
 - <http://www.exac.broadinstitute.org>
 - Database with aggregated and harmonized data from over 60,000 human exomes from unrelated individuals. Provides allele frequencies in different major racial and ethnic groups.
- **Genome Aggregation Database (gnomAD)**
 - <http://gnomad.broadinstitute.org/>
Database with aggregated and harmonized data from over 123,000 human exomes and 15,000 human genomes from unrelated individuals. Provides allele frequencies in different major racial and ethnic groups.

GENE EXPRESSION

- See data on individual gene pages on **NCBI Gene** and **Ensembl**
 - <https://www.ncbi.nlm.nih.gov/gene>
 - <http://www.ensembl.org/index.html>
- **The Human Protein Atlas**
 - <http://www.proteinatlas.org/>
- **Genotype-Tissue Expression (GTEx) project**
 - <https://gtexportal.org/home/>
- **BioGPS**
 - <http://biogps.org/#goto=welcome>

PROTEIN INTERACTION

- See data on individual gene pages on **NCBI Gene** and **Ensembl**
 - <https://www.ncbi.nlm.nih.gov/gene>
- **Biological General Repository for Interaction Datasets (BioGRID)**
 - <https://thebiogrid.org/>
Compilation of genetic and protein interaction data from model organisms and humans.
- **Agile Protein Interactomes DataServer (APID)**
 - <http://cicblade.dep.usal.es:8080/APID/init.action#tabr2>
Comprehensive collection of protein interactions from over 400 organisms.
- **STRING database**
 - <http://string-db.org/>
Database of known and predicted protein interactions.

MOUSE MODELS

- **Mouse Genome Informatics**
 - <https://www.jax.org/jax-mice-and-services>
Database of laboratory mice, providing integrated genetic, genomic, and biological data.
- **Knockout Mouse Project (KOMP)**
 - <https://www.komp.org/>
Initiative to generate a public resource of **mouse** embryonic stem cells containing a null mutation in every gene in the **mouse** genome.