Prepared by the Association for Molecular Pathology Training and Education Committee
For More Educational Resources: *www.amp.org/AMPEducation*

Molecular In My Pocket™…

Bioinformatics: Interpretation Databases

Database	Description	Website
Population Databases Exclude Polymorphisms		
Genome Aggregation Database (gnomAD)	A database that aggregates and harmonizes exome and genome sequencing data from a variety of large-scale sequencing projects and makes summary data available for the wider scientific community. The database includes Exome Aggregation Consortium (ExAC) database.	https://gnomad.broadinstitute.org/
Single Nucleotide Polymorphism Database (dbSNP)	A public archive for genetic variation including SNPs and short indels from 1000 Genomes (1000g) Project and Exome Variant Server (EVS).	https://www.ncbi.nlm.nih.gov/snp/
Interpretation Databases Determine the Importance of Variants. Somatic Interpretation 🟡; Germline Interpretation 🔵.		
Catalog of Somatic Mutations in Cancer (COSMIC) 🟡	A database of somatically acquired mutations found in human cancer.	https://cancer.sanger.ac.uk/cosmic
cBioPortal for Cancer Genomics 🟡	A resource for interactive exploration of multidimensional cancer genomics data sets. The database includes Cancer Hotspots and GENIE (requires registration).	http://www.cbioportal.org/
ClinVar 🟡 🔵	A public archive that aggregates information about genomic variation and its relationship to human health.	https://www.ncbi.nlm.nih.gov/clinvar/
Human Gene Mutation Database (HGMD) 🔵	A comprehensive collection of published germline mutations in nuclear genes that underlie, or are closely associated with, human inherited disease.	http://www.hgmd.cf.ac.uk/ac/index.php
My Cancer Genome 🟡	A precision cancer medicine knowledge resource that contains information on the clinical impact of molecular biomarkers in cancer-related genes, proteins, and other biomarker types on the use of anticancer therapies in cancer.	https://www.mycancergenome.org/
OncoKB 🟡	A precision cancer database with diagnostic, prognostic and therapeutic levels of evidence for genomic alterations in cancer.	https://www.oncokb.org/
PeCan 🟡 🔵	A resource that provides pediatric cancer mutations.	https://pecan.stjude.cloud/home
Prediction Databases Predict Impact of Variants on Proteins		
Polymorphism Phenotyping v2 (PolyPhen2)	Categorical prediction: D: probably damaging; P: possibly damaging; B: benign	http://genetics.bwh.harvard.edu/pph2/
Sorting Intolerant From Tolerant (SIFT)	Categorical prediction: D: deleterious; T: tolerated	https://sift.bii.a-star.edu.sg/
SpliceAI	Delta score ranges from 0 to 1 indicating probability of variant affecting splicing. Cutoffs: 0.2 (high recall), 0.5 (recommended), and 0.8 (high precision).	https://spliceailookup.broadinstitute.org/
Notes: This is not an inclusive list, only the most common databases.		

Revised 7/2023