**SUPPLEMENTARY DATA**

**GENEASE: Real time bioinformatics tool for multi-omics and disease ontology exploration, analysis and visualization**

Sudhir Ghandikota1,2, Gurjit K. Khurana Hershey2 and Tesfaye B. Mersha2

1,2Computer Science and Engineering, University of Cincinnati, Cincinnati, OH, 45221, USA

2Department of Pediatrics, Cincinnati Children’s Hospital Medical Center, Cincinnati, OH, 45229, USA

**Supplementary Table S1: SNP explore databases:** List of the annotation databases used in GENEASE exploration module for “SNP” Input category.

|  |  |  |
| --- | --- | --- |
| **Database/Tool Name** | **Group in Tool** | **URL** |
| **Exome Variant Server** | Genome Variation | <http://evs.gs.washington.edu/EVS/>  |
| **GWAS Catalog** | Genome Variation | <https://www.ebi.ac.uk/gwas/home> |
| **Genome Variant Server** | Genome Variation | <http://gvs.gs.washington.edu/GVS144/>  |
| **UCSC** | Genome Variation | <https://genome.ucsc.edu/>  |
| **Ensembl** | Genome Variation | <http://useast.ensembl.org/index.html> |
| **dbSNP** | Genome Variation | <https://www.ncbi.nlm.nih.gov/projects/SNP/index.html>  |
| **Geno2MP** | Genome Variation | <http://geno2mp.gs.washington.edu/Geno2MP/#/> |
| **DGV(Database of Genomic Variants)** | Genome Variation | <http://dgv.tcag.ca/dgv/app/home>  |
| **VarSome** | Genome Variation | <https://varsome.com/>  |
| **RegulomeDB** | Variant Annotation | <http://www.regulomedb.org/index>  |
| **HaploReg** | Variant Annotation | <https://archive.broadinstitute.org/mammals/haploreg/haploreg.php>  |
| **GWAVA** | Variant Annotation | <https://www.sanger.ac.uk/sanger/StatGen_Gwava>  |
| **GTEx** | Variant Annotation | <https://gtexportal.org/home/> |
| **SNPeffect** | Variant Annotation | <http://snpeffect.switchlab.org/menu>  |
| **Spliceman 2** | Variant Annotation | <http://fairbrother.biomed.brown.edu/spliceman2/upload>  |
| **Varian Effect Predictor** | Variant Annotation | <http://www.ensembl.org/Tools/VEP/>  |
| **PubMed** | Other Information | <https://www.ncbi.nlm.nih.gov/pubmed>  |
| **SNPedia** | Other Information | <https://www.snpedia.com/index.php/SNPedia>  |

**Supplementary Table S2: GENE explore databases:** List of databases used in GENEASE exploration module for “gene” input category.

|  |  |  |
| --- | --- | --- |
| **Database/Tool Name** | **Group in Tool** | **URL** |
| **UCSC** | Gene Information | <https://genome.ucsc.edu/>  |
| **Ensembl** | Gene Information | <http://useast.ensembl.org/index.html>  |
| **HGNC** | Gene Information | <http://www.genenames.org/>  |
| **OMIM** | Gene Information | <https://omim.org/>  |
| **HAVANA** | Gene Information | <http://vega.archive.ensembl.org/index.html>  |
| **Illumina** | CpG Information | <https://www.illumina.com/>  |
| **GWAS Catalog** | Genome & Variation | <https://www.ebi.ac.uk/gwas/home>  |
| **ENCODE** | Genome & Variation | <https://www.encodeproject.org/>  |
| **Exome Variant Server** | Genome & Variation | <http://evs.gs.washington.edu/EVS/>  |
| **Genome Variant Server** | Genome & Variation | <http://gvs.gs.washington.edu/GVS144/>  |
| **SNPeffect** | Genome & variation | <http://snpeffect.switchlab.org/menu>  |
| **dbVar** | Genome & Variation | <https://www.ncbi.nlm.nih.gov/dbvar>  |
| **ClinVar** | Genome & Variation | <https://www.ncbi.nlm.nih.gov/clinvar/>  |
| **Geno2MP** | Genome & Variation | <http://geno2mp.gs.washington.edu/Geno2MP/#/>  |
| **DGV(Database of Genomic Variants)** | Genome & Variation | <http://dgv.tcag.ca/dgv/app/home>  |
| **GEO Profiles** | Gene Expression | <https://www.ncbi.nlm.nih.gov/geoprofiles>  |
| **ArrayExpress** | Gene Expression | <https://www.ebi.ac.uk/arrayexpress/>  |
| **GTEx** | Gene Expression | <https://gtexportal.org/home/> |
| **TargetScan** | Gene Expression | <http://www.targetscan.org/vert_71/> |
| **Encode Promoter** | Gene Expression | <http://promoter.bx.psu.edu/ENCODE/index.html> |
| **KEGG** | Pathways | <http://www.kegg.jp/kegg/kegg2.html> |
| **Reactome** | Pathways | <http://www.reactome.org/> |
| **BioCarta** | Pathways | <https://cgap.nci.nih.gov/> |
| **Protein Atlas** | Protein Information | <http://www.proteinatlas.org/> |
| **UniProt** | Protein Information | <http://www.uniprot.org/> |
| **Pfam** | Protein Information | <http://pfam.xfam.org/> |
| **SMART** | Protein Information | <http://smart.embl.de/> |
| **MGI (Mouse genome)** | Ortholog Information | <http://www.informatics.jax.org/>  |
| **RGD (Rat genome)** | Ortholog Information | <http://rgd.mcw.edu/wg/home>  |
| **ZFIN (zebrafish information)** | Ortholog Information | <http://zfin.org/>  |
| **FlyBase** | Ortholog Information | <http://flybase.org/>  |
| **WormBase** | Ortholog Information | <http://www.wormbase.org/#012-34-5>  |
| **PomBase** | Ortholog Information | <https://www.pombase.org/>  |
| **SGD (Yeast genome)** | Ortholog Information | <https://www.yeastgenome.org/>  |
| **NCBI Map** | Other Information | <https://www.ncbi.nlm.nih.gov/> |
| **GeneCards** | Other Information | <http://www.genecards.org/> |
| **Vega** | Other Information | <http://vega.archive.ensembl.org/index.html> |
| **DECIPHER** | Other Information | <https://decipher.sanger.ac.uk/>  |
| **PubMed** | Other Information | <https://www.ncbi.nlm.nih.gov/pubmed>  |
| **NCBI Gene** | Other Information | <https://www.ncbi.nlm.nih.gov/gene/>  |

**Supplementary Table S3: Disease explore databases:** List of databases used in GENEASE exploration module for “disease” input category.

|  |  |  |
| --- | --- | --- |
| **Database/Tool Name** | **Group in Tool** | **URL** |
| **GWAS Catalog** | Genome-wide Association/Gene Expression | <https://www.ebi.ac.uk/gwas/home> |
| **dbGaP** | Genome-wide Association/Gene Expression | <https://www.ncbi.nlm.nih.gov/gap>  |
| **Geo DataSets** | Genome-wide Association/Gene Expression | <https://www.ncbi.nlm.nih.gov/gds>  |
| **PhenX Toolkit** | Phenotype Genotype | <https://www.phenxtoolkit.org/index.php>  |
| **DECIPHER** | Phenotype Genotype | <https://decipher.sanger.ac.uk/>  |
| **ENCODE** | Variation Information | <https://www.encodeproject.org/> |
| **ClinVar** | Variation Information | <https://www.ncbi.nlm.nih.gov/clinvar/>  |
| **OMIM** | Variation Information | <https://omim.org/>  |
| **GeneCards** | Variation Information | <http://www.genecards.org/> |
| **Swissvar** | Variation Information | <http://www.expasy.org/>  |

**Supplementary Table S4:** Execution times of bulk input requests of both SNP and gene lists (in cache and not in cache scenarios).

|  |  |  |
| --- | --- | --- |
| **Gene** |  | **SNP** |
| **Number of Genes** | **Runtime(not in cache)** | **Runtime(in cache)** |  | **Number of SNPs** | **Runtime(not in cache)** | **Runtime(in cache)** |
| 100 | 28 secs | 2.5 secs |  | 100 | 67 secs | 1.7 secs |
| 500 | 144 secs | 14 secs |  | 200 | 123 secs | 2.5 secs |
| 1000 | 290 secs | 30 secs |  | 500 | 300 secs | 9.7 secs |

**Supplementary Table S5:** List of annotation databases used in “Enrichment Analysis” module. Annotation terms download and parsed in real-time for each request.

|  |  |  |
| --- | --- | --- |
| **Annotation Type** | **Database/Project Name** | **URL(s)/File(s)** |
| **Biological Pathways** | KEGG(Kyoto Encyclopaedia of Genes and Genomes) | <http://rest.kegg.jp/list/pathway/hsa> <http://rest.kegg.jp/link/hsa/pathway>  |
| **Gene Ontologies** | GOC(Gene Ontology Consortium) | <http://geneontology.org/ontology/go.obo> <http://geneontology.org/gene-associations/goa_human.gaf.gz>  |
| **Phenotype Ontologies** | HPO(Human Phenotype Ontology) | <http://compbio.charite.de/jenkins/job/hpo.annotations.monthly/lastStableBuild/artifact/annotation/ALL_SOURCES_ALL_FREQUENCIES_genes_to_phenotype.txt>  |

**Supplementary Table S6:** Gene coverage in the ontology terms used in GENEASE in comparison with other tools.

|  |  |
| --- | --- |
|   | **Gene Counts** |
|  | **KEGG Pathways** | **GO Biological Process** | **GO Cellular Component** | **GO Molecular Function** | **Human Phenotype Ontologies** |
| **Actual Counts** | 7320 | 17783 | 18932 | 17758 | 3699 |
| **DAVID** | 6910 | 16792 | 18224 | 16881 | Unsupported |
| **Enrichr** | 7010 | 13822 | 10427 | 10601 | 3096 |
| **GENEASE** | 7320 | 17257 | 18631 | 17505 | 3699 |

**Supplementary Figure S1**: Exploring proxy variants: SNP explore result for **rs9272346** with LD option to find its proxy variants in CEU (European) population (r2>0.8). A particular proxy variant can be further explored by selecting the particular row in the result table.

