DO Buzz Newsletter April 2020

Latest News:

Check out our new Youtube video, by James Munro - "How is the Human Disease Ontology FAIR?"

In this video tutorial, we look at the FAIR Guiding Principles - these being Findable, Accessible, Interoperable, Reusable - for scientific data management and stewardship of digital assets.

How is the Human Disease Ontology FAIR?

DO Spotlight of the Quarter: Disease Ontology's application in Gene Co-expression Network Analysis

In a recent paper published by Rhead et al. in PLOS ONE in March 2020, genes in the top tumor necrosis factor-alpha associated module were associated with 136 Disease Ontology terms, including autoimmune/inflammatory, infectious and cardiovascular diseases, and cancers. The researchers used weighted gene expression network analysis (WGCNA) to construct a gene expression network to find groups of genes that have both high correlation of expression and high topological overlap. Each module was tested with TNFa treatment, and differentially methylated regions were identified. Disease Ontology enrichment analysis was performed on genes in each WGCNA module to find known associations of genes to disease.

"The Disease Ontology enrichment analysis was performed on genes in each module using the XGR R package which utilizes the Disease and Gene Annotations database to map genes to diseases...To be considered as an enriched disease term, at least 10 and at most 2,000 genes were required to be annotated for that term, and at least 5 genes were required to overlap with the input gene list. Fisher's exact test was used to determine significance, and parent-child relations were accounted form using the "lea" algorithm. Disease terms enriched at an FDR-adjusted $p<0.05$ were reported."

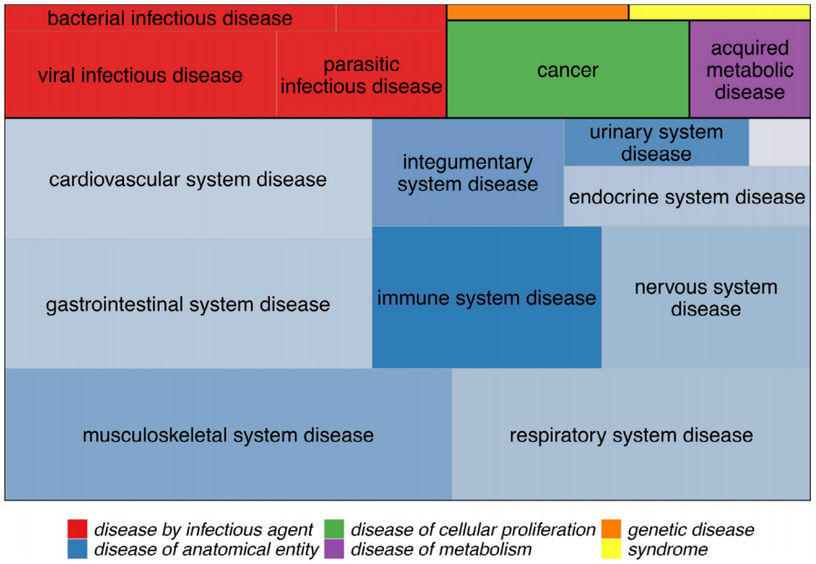
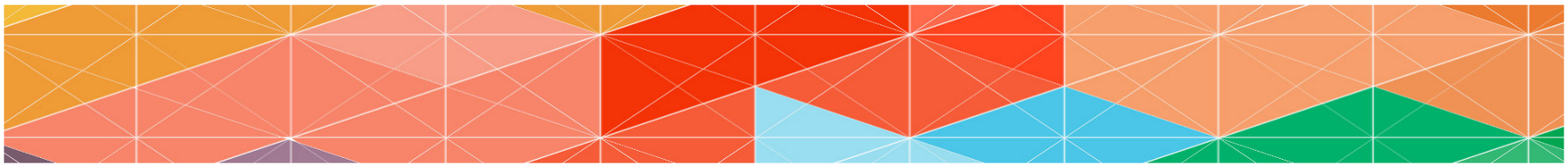


Fig 3. Disease Ontology categories for the 136 diseases that showed overrepresentation of green module genes. Boxes are colored according to the top-level disease categories and labels show the second-level categories. The size of each box is proportional to the number of disease terms in that category with significant overrepresentation of green module genes. Some disease terms belong to more than one category (e.g., multiple sclerosis is both a "nervous system disease" and an "immune system disease"), but each term is only represented once. For "disease of anatomical entity" terms, squares are shaded by the proportion of terms that represent autoimmune/inflammatory diseases (e.g., 3 of 16 "gastrointestinal system disease"terms are autoimmune/inflammatory, while 9 of 11 "immune system disease" terms are). The full list of disease terms, with (manually curated) autoimmune/inflammatory terms highlighted, is given in S3 File.



Latest Release Notes:

DO Data Release: Available in DO's GitHub repository: (previous release notes)

Release # 80: April 20, 2020 Release Notes

In the DO April 2020 release: Includes 10,041 terms with 71% defined. Additions include the updated COVID-19 parent term to Coronavirus infectious disease, updated NCBI cross-references, and the update of the benign neoplasm nomenclature. Since the last quarterly newsletter, over 400 Symptom logical definitions have been added, and ~290 OMIM additions.
We have also incorporated British synonyms into the ontology, in addition to the existing US spellings.

Disease Ontology Citations:

The DO team has identified a body of 594 DO project citations (as of April 2020).
This set of citations has been compiled as a public PubMed MyNCBI collection (DO citing papers). This MYNCBI collection represents the growing number of instances of integration of DO in databases, research studies, and bioinformatics tools. The DO Citations are identified through PubMed data mining (direct DO paper citations, inclusions of 'Disease Ontology', DO URL or DOID).

Publications:

Human Disease Ontology 2018 update: classification, content and workflow expansion.
Schriml LM, Mitraka E, Munro J, Tauber B, Schor M, Nickel L, Felix V, Jeng L, Bearer C, Lichenstein R, Bisordi K, Campion N, Hyman B, Kurland D, Oates CP, Kibbey S, Sreekumar P, Le C, Giglio M, Greene C. Nucleic Acids Res. 2019 Jan 8;47(D1):D955-D962. doi: 10.1093/nar/gky1032.
PDF: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6323977/pdf/gky1032.pdf

ECO, the Evidence & Conclusion Ontology: community standard for evidence information.
Giglio M, Tauber R, Nadendla S, Munro J, Olley D, Ball S, Mitraka E, Schriml LM, Gaudet P, Hobbs ET, Erill I, Siegele DA, Hu JC, Mungall C, Chibucos MC.
Nucleic Acids Res. 2019 Jan 8;47(D1):D1186-D1194. doi: 10.1093/nar/gky1036.
PDF: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6323956/pdf/gky1036.pdf