

Increased expressivity of mouse protein function representation using the gene ontology

Harold J Drabkin, David P Hill, Mary E Dolan and Judith A Blake

The Jackson Laboratory, USA

The Gene Ontology (GO) is a structured controlled vocabulary used by numerous model organism databases for the functional annotation of gene products. GO terms define molecular functions, biological processes, and cellular components that cover all life forms. Until recently, the GO annotation framework was limited in its ability to adequately capture the biological context of each annotation such as when and where a protein was active. Now the GO Consortium has created a methodology that supports the addition of contextual detail to GO annotations extracted from biomedical literature. These extensions include several “effector-target” relationships, such as cellular and anatomical localization dependencies, enzyme substrates, and regulation targets of signally pathways and transcription factors, in addition to the spatial and temporal aspects of processes. These details are captured using relationships to a variety of external ontologies such as the Cell Type Ontology. The Mouse Genome Informatics (MGI) database system supported capturing such details privately for many years. Here we describe the migration of this information into the structured method now available. The enhanced annotations will support more specific queries and allow computational reasoning to augment discovery by inference.

Biography

Harold Drabkin received his Ph.D. from Wesleyan University and postdoctoral training at Roche Institute of Molecular Biology (Nutley NJ) and Massachusetts Institute of Technology (Cambridge, MA). He was a research scientist and lecturer at MIT, and is currently a Senior Scientific Curator for the Mouse Genome Informatics system at the Jackson Laboratory (Bar Harbor, ME), working on the Gene Ontology and Protein Ontology projects.

harold.drabkin@jax.org