Issue No. 1

The Gene Ontology Newsletter

May 2006

Welcome to the first issue of the GO Newsletter

This is the first issue of the Gene Ontology newsletter, which highlights significant improvements and changes to the ontologies, annotations, and software that are developed by members of the GO Consortium. This newsletter will be sent to those subscribed to the GO Friends mailing list on a quarterly basis.

Representing Blood Pressure Regulation in the GO

One of the main goals of the Gene Ontology Consortium is to improve the biological structure and content of the ontologies, as well as their use in the annotation of genes. A primary method to achieve this aim involves a curator reviewing the existing terms in the ontologies that cover a particular area; consulting primary literature, reviews, and textbooks, as appropriate; and if required, developing new terms that describe the area accurately. Following this model, curators at Mouse Genome Informatics (MGI; http://www.informatics.jax.org/) at The Jackson Laboratory (http://www.jax.org/) have targeted the comprehensive annotation of mammalian genes involved in the regulation of blood pressure.

MGI's curator for this project, David Hill, began reading papers in this area, and realized that the existing terms in the GO were insufficient to annotate genes involved in the various processes that regulate blood pressure. After consulting a medical textbook and review articles regarding the physiology of blood pressure regulation, David proposed 43 new GO terms, which were discussed and refined with other GO curators through the GO Consortium's SourceForge discussion forum. Using these new terms, David then began annotating genes using primary literature dealing with blood pressure regulation. Again, he quickly found that additional terms were needed, and another cycle of annotation and term development ensued that has now resulted in 23 additional terms.

By developing the ontology during the active annotation of genes, the GO has gained 66 new terms addressing blood pressure regulation, and now represents this area of biology far more comprehensively than before. David's annotation efforts ultimately yielded 49 annotations for 23 mouse genes (http://www.informatics.jax.org/searches/GO.cgi?id=GO:0008217) directly involved in the process of blood pressure regulation, bringing the total to 264 GO annotations for those genes.

The process of annotation-driven ontology development has greatly improved the utility of the GO for biologists studying the regulation of blood pressure. The depth of these annotations will certainly aid analysis of large-scale datasets resulting from blood pressure studies. In addition, these annotations may be further mapped to human and rat orthologs through the GO evidence code ISS, Inferred from Sequence or structural Similarity.
(contributed by Alex Diehl and David Hill, MGI)

Tutorial: AmiGO

What is AmiGO?
AmiGO is a web application that allows browsing and visualization of the ontologies and annotations from multiple species at the same time.

How do I access AmiGO?
www.godatabase.org/cgi-bin/amigo/go.cgi

Where can I get the data in AmiGO?
The ontology data in AmiGO can be found in the gene_ontology.obo file. The annotations can be found in the gene_association files that are submitted by the various database groups to the GO project. Both files can be downloaded from the main GO website (www.geneontology.org/GO.downloads.shtml).

I would like to retrieve a list of 'helicases' from different organisms (Fly, Yeast, Arabidopsis, Mouse). How do I do that?
Go to the AmiGO URL and enter your query (helicase, for example) in the query box located at the top left hand corner. You will see a list of terms on the results page that match your query string. Click on the term name that you think is a close match to your query to see a page that provides details on the term itself (definition, GO ID, synonyms, placement in the ontology). On the bottom half of the page are annotations for gene_products from various organisms. An annotation consists of the following core pieces of information: gene name, information about the gene including the source organism, database name, evidence (abbreviated as IDA, IMP etc.) and the reference used to make the annotation.

What else can I learn about helicases from the Term Details page?
The Term Lineage section of the Term Details page shows the relationships (is_a and part_of) between the term of interest (e.g. helicase activity) and other terms. Using the Term Context filters located above the Term Lineage section, one can see the siblings of the term. The Term Ancestor filter can be used to see the parents of a given term all the way up to the root node.
(contributed by Rama Balakrishnan, SGD)

Cellular Component to be is_a Complete

The Cellular Component ontology will soon be "is_a complete", meaning that each term in the hierarchy will have at least one is_a relationship path to the top node. This is important for making the Cellular Component ontology ontologically correct, allowing for more accurate queries and reasoning. In addition, the Cellular Component ontology will become compatible for use with other ontologies and ontology tools developed by other scientific communities. Once the Cellular Component Ontology has been made is_a complete, the GO Consortium is committed to maintaining the completeness. Any new terms entered must not be is_a orphans. This transition will be implemented no sooner than May 29, 2006.

Converting to OBO Flat File Format, version 1.2

The Gene Ontology Consortium is converting its gene_ontology.obo file to OBO Format version 1.2, which supports a more complete representation of the terms and relationships in the ontology as well as meta-data associated with terms (such as synonyms and comments). Details are available in the specifications for OBO Flat File Format, version 1.2 (http://www.godatabase.org/dev/doc/obo_format_spec.html). To be implemented no sooner than June 15, 2006.

Upcoming Meetings

- GO Neurobiology Ontology Workshop
  June 14-16, 2006, Bar Harbor, ME
  contact jblake@informatics.jax.org *for more info*
- GO Annotation Camp
  July 12-14, 2006, Stanford University, Stanford, CA
  www.geneontology.org/meeting/AnnotCamp2006info.shtml
- GO Consortium User Meeting (in conjunction with MGED)
  September 10, 2006, Seattle, USA
  www.geneontology.org/meeting/mged-2006-meeting.shtml
  www.mgedmeeting.org/

Participants of the GO Consortium meeting, March 2006

Contact GO

To receive this newsletter and other announcements from the GO Consortium, subscribe to the GO Friends mailing list (gofriends@genome.stanford.edu).

Please contact the Gene Ontology Consortium with any comments or suggestions at: gohelp@geneontology.org. Frequently asked questions may appear as tutorials or tips in upcoming newsletters.