Exploring Autophagy with Gene Ontology

Paul Denny, Marc Feuermann, David P. Hill, Paola Roncaglia and Ruth C. Lovering

1. Centre for Ornithological Genomics Institute of Cell and Developmental Biology, University College London, Gower Street, London WC1E 6BT
2. European Bioinformatics Institute (EMBL-EBI), The Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SD, UK
3. The Janssen Institute, 520 Main Street, Bar Harbor, ME 04609 USA
4. European Bioinformatics Institute (EMBL-EBI), European Molecular Biology Laboratory, Hinxton, Cambridge, CB10 1SD, UK

Website: https://www.ucl.ac.uk/functional-gene-annotation; Twitter: @UClgene

Introduction to GO

The Gene Ontology (GO) project is a collaborative effort to provide consistent descriptions of gene products across all kingdoms of life, and is a key resource for researchers wishing to understand the biological role of a gene product, or a list of gene products.

- GO contains three structured vocabulary (ontologies) that describe gene products in terms of their associated biological processes, cellular locations and molecular functions, in a database-driven, peer-reviewed format.
- There are now over 40,000 terms describing a wide range of concepts to differing levels of specificity.

What is Autophagy?

Autophagy is the process in which cells dispose parts of their own cytoplasm and organelles, including recycling membrane structures under conditions of cellular stress and remodeling of the intracellular structure during cell differentiation. Key features include:

- Well-conserved across species
- Involves in several pathophysiological events relevant to human health: cancer, metabolic disorders, cardiovascular and pulmonary diseases, neurodegenerative processes, such as Parkinson’s disease

Autophagy Ontology Development

This project is a collaboration between members of the Gene Ontology Consortium at University College London (UCL), the European Bioinformatics Institute (EMBL-EBI) and The Janssen Research Laboratory at the Janssen Institute. Our aim is to modify and extend the Gene Ontology for the Biological Process of autophagy and provide high-quality GO annotations to the products of genes involved in the process.

In this project, we have focused on characterizing mammalian autophagy (CMA), microautophagy and macropathological forms of autophagy, such as mitophagy.

Mitophagy is well characterized in yeast, but literature searches revealed conserved mammalian macropathological forms. Details of a related process, called ‘late endosomal autophagy’ in mammals and a suitable GO term created (GO: 0006738) (see Figure 1). It remains unclear whether use of this term should be limited to some taxonomic groups only.

Chaperone-mediated autophagy was, prior to the start of this work, erroneously represented in the ontology as a synonym of autophagy; this has been corrected. Terms to describe cellular components and subpathways involved in CMA have also been created.

In addition to the revision of existing terms, the project has so far led to creation of 28 new GO terms, including ‘proteasome activity involved in autophagy assembly’ (GO:0006739) and ‘paranucleae mitochondria’ (GO:0006743). We have also created logical definitions for some of the new terms.

In summary, improvements to an ontology domain benefit both those new to a topic and scientists who wish to analyze their existing data more accurately.

- A specific class of autophagy, known as macroautophagy, has been the focus of a related project (see prior abstract #46).

Methods

Announcement: We extract data from primary papers, and some reviews, to associate GO terms with proteins. Our primary focus is human and mouse but we also capture information from other model organisms including rat.

Functional analysis: Transcriptome microarrays expressing in left ventricular myocardium versus right and a myocardium were analyzed using the functional analysis tool (gOstat) (http://biit.cs.ut.ee/gostat/index), to identify over-representation of GO terms. The analysis included the GO annotation database available on 21 January 2009 and GO annotation released on 8 December 2015. 1


Summary of Progress and Functional Analysis

- So far we have created 417 annotations to 347 distinct proteins (299 human proteins) from 75 papers (March 10th 2016).
- To improve annotation coverage, we compared the proteins associated with the GO terms ‘autophagy’ and ‘macroautophagy’ with the proteins annotated with similar processes in two complementary resources, Reactome (http://www.reactome.org/pathwaybrowser/) and PD-map (http://golm.mdc-berlin.de/PDmap).
- 67 proteins in macroautophagy/Reactive pathway.
- 59 also annotated to GO term macroautophagy.
- 714 annotations – resolved by creating new GO annotations:
- 135 proteins associated with autophagy in PD-map.
- 54 also annotated to GO term autophagy.
- 814 annotations – resolved by creating new GO annotations, 127 outstanding.

- To examine the impact of over- and under-expression in left ventricular myocardium by comparison with the right atrial myocardium: Ventricle has a higher demand for oxygen than atrium and this is reflected in the enrichment of mitochondrial processes (Table 1).

- Our functional analysis with the current GO annotation dataset also demonstrates that both ‘autophagy’ and ‘mitophagy’ are processes that are significantly enriched in the ventricle trans-membrane, due to the need to remove damaged mitochondria.

Repeating the analysis using the 2015 GO annotation dataset, only mitophagy is significantly enriched and only just achieving significance. Autophagy is not enriched in this analysis.

This project demonstrates the need to continue to improve both the GO terms and annotations so that the GO dataset can provide an effective resource for researchers, from those analyzing high-throughput data to those wishing to identify gene variants associated with specific diseases.

Table 1. Comparison of functional analysis of adult ventricular myocardium data using Gene Ontology. A selection of GO terms identified by two analyses using GO annotation datasets in 2015 and 2019.