CLOSE

Exploring Autophagy with Gene Ontology

Paul Denny[1], Marc Feuermann[2], David P. Hill[3,5], Paola Roncaglia[4,5] and Ruth C. Lovering[1]

1. Centre for Cardiovascular Genetics, Institute of Cardiovascular Science, University College London, Rayne Building, 5 University Street, London, WC1E 6JF
2. SIB, Swiss Institute of Bioinformatics, Geneva, Switzerland
3. The Jackson Laboratory, 600 Main Street, Bar Harbor, Maine 04609 USA
4. European Bioinformatics Institute (EMBL-EBI), European Molecular Biology Laboratory, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, UK
5. The Gene Ontology Consortium http://geneontology.org/

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

UCL

Introduction to GO

• The Gene Ontology (GO) project[1] 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 controlled vocabularies (ontologies) that describe gene products in terms of their associated **biological processes**, **cellular locations** and **molecular functions**, in a species-independent manner.

• There are now over 40,000 terms describing a wide range of concepts to differing levels of specificity.

Figure 1 'Mitophagy' & selected parent terms in the Gene Ontology. Black lines indicate is_a relations to the parent terms. Asterisks (*) indicate terms revised, whereas tilde (~) indicates terms created by this project.

What is Autophagy?

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

• Well-conserved across species

• Involved in several pathophysiological events relevant to human health
  - cancer, metabolic disorders, cardiovascular and pulmonary diseases
  - neurodegenerative processes, such as Parkinson's disease

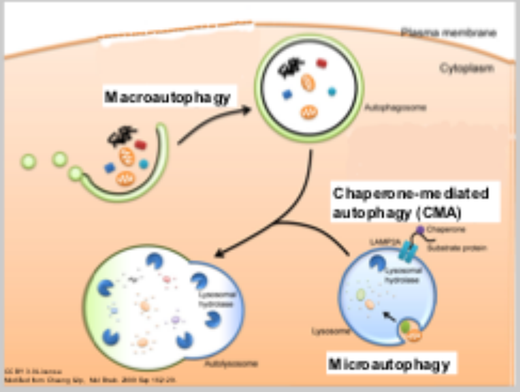

Figure 2. Main types of autophagy: macroautophagy, microautophagy and chaperone-mediated autophagy[4].

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), The Jackson Laboratory and the Swiss Institute of Bioinformatics (SIB).

• 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 chaperone-mediated autophagy (CMA), microautophagy and targeted forms of autophagy, such as mitophagy.

• Microautophagy is well characterised in yeast, but literature searches failed to reveal homologous mammalian molecular participants. Details of a related process, called 'late endosomal microautophagy', were found in mammals and a suitable GO term created (GO:0061738) (see Figure 1). It remains unclear whether use of this process 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 sub-processes 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 'protein lipidation involved in autophagosome assembly' (GO:0061739) and 'parkin-mediated mitophagy in response to mitochondrial depolarization' (GO:0061734). We have also created logical definitions for many of the new terms.

• Improvements to an ontology domain benefit both those new to a field, and scientists who wish to analyse their existing datasets accurately.

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

Methods

Annotation: 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: Transcripts over-expressed in left ventricular myocardium versus right atrial myocardium[2] were analysed using the functional analysis tool g:GOSt[3] (http://biit.cs.ut.ee/gprofiler/index.cgi), to identify over-representation of GO terms. The analysis included the GO annotation dataset available on 21 January 2016 or March 2015 and the GO ontology released on 8 December 2015.

1. Gene Ontology Consortium: going forward. The Gene Ontology Consortium (2015). Nucleic Acids Res. 43, D1049-56.
2. Barth, A.S., et al., Functional profiling of human atrial and ventricular gene expression (2005). Pflugers Arch. 450, 201-208.
3. Reimand, J., Arak, T., Vilo, J., g:Profiler-a web server for functional interpretation of gene lists (2011). Nucleic Acids Res 39, W307-315.
4. Cheung, Z.H. & Ip, N.Y. (2009). Mol Brain. Sep 16;2:28.
5. Fabregat A, et al., The Reactome pathway Knowledgebase (2016). Nucleic Acids Res. Jan 4;44(D1):D481-7.
6. Fujita, K.A., et al., Integrating pathways of Parkinson's disease in a molecular interaction map (2014). Mol Neurobiol. Feb;49(1):88-102.

Summary of Progress and Functional Analysis

• So far we have created 477 annotations to 347 distinct proteins (299 human proteins) from 75 papers (March 10th 2016).

• To improve our 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[5] (http://www.reactome.org/PathwayBrowser/) and PD-map[6] (http://minerva.uni.lu/MapViewer/).
  - 67 proteins in the macroautophagy Reactome pathway
    - 59 also annotated to GO term macroautophagy
    - 8 discrepancies – resolved by creating new GO annotations.
  - 135 proteins associated with autophagy in PD-map
    - 53 also annotated to GO term autophagy
    - 81 discrepancies – 54 resolved by creating new GO annotations, 28 outstanding.

• To examine the impact of our annotation on data analysis we analysed transcripts over-expressed in left ventricular myocardium by comparison with the right atrial myocardium[2].

• Ventricles have a higher demand for respiration than atria and this is reflected in the enrichment of mitochondria-relevant terms (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 transcriptome, 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 analysing high-throughput data to those wishing to identify gene variants associated with specific diseases.

GO term		Number of proteins annotated with GO term		p-value	
Identifier	**Name**	March 2015	January 2016	March 2015	January 2016
Cellular Component					
GO:0005739	mitochondrion	354	618	8.36E-87	2.06E-146
GO:0031982	vesicle	418	594	4.49E-26	1.21E-06
Biological Process					
GO:0007005	mitochondrion organization	158	303	3.74E-41	1.69E-76
GO:0006914	**autophagy**	**#N/A**	**110**	**#N/A**	**3.80E-04**
GO:0006936	muscle contraction	57	72	3.46E-08	1.21E-02
GO:0000422	**mitophagy**	**29**	**56**	**4.34E-02**	**9.70E-04**

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 data available in 2015 and 2016

MGI · GENEONTOLOGY Unifying Biology · SIB Swiss Institute of Bioinformatics · EMBL-EBI · PARKINSON'S UK CHANGE ATTITUDES. FIND A CURE. JOIN US.

Funded by: Parkinson's UK, grant G-1307 (PI: RC Lovering); NIH NHGRI HG002273 Gene Ontology Consortium (PIs: JA Blake, JM Cherry, S Lewis, PW Sternberg and P Thomas).