Connecting the Gene Ontology (GO) to Parkinson's disease

Rebecca E. Foulger[1], Paul Denny[1], Maria J. Martin[3], John Hardy[2] 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. Department of Molecular Neuroscience & Reta Lila Weston Laboratories, Institute of Neurology, University College London, Queen Square, London, WC1N 3BG
3. European Bioinformatics Institute (EMBL-EBI), European Molecular Biology Laboratory, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD

The UK Parkinson's GO annotation project: aims, priorities and progress

- Our project is a Parkinson's UK-funded collaboration between University College London (UCL) and the European Bioinformatics Institute (EMBL-EBI). We are extending GO annotation into neurological areas to provide high-quality GO annotations to the products of genes relevant to Parkinson's disease.
- This is the first annotation effort to focus on Parkinson's, and we have established collaborations with local and international neurological researchers to guide our priorities.
- We extract data from primary papers and reviews to attach GO terms to Parkinson's-relevant proteins. Our primary focus is human, but we also capture information from model organisms including fly, rat and mouse.

Our annotation priorities include:

Cellular pathways known to be dysregulated in Parkinson's

Including: **mitophagy**, **autophagy**, oxidative stress response, unfolded protein response, regulation of neuron death, **Wnt-regulated dopaminergic neuron differentiation**, and synaptic vesicle transport

Our 2014-2015 focus

Products of genes associated with Parkinson's, including causative genes and risk genes

Includes 48 high priority genes, identified through GWAS and genetic analyses

Our current focus

Parkinson's UK-funded research

Within each topic, and for each priority protein, we prioritise publications funded by Parkinson's UK

Our ongoing focus

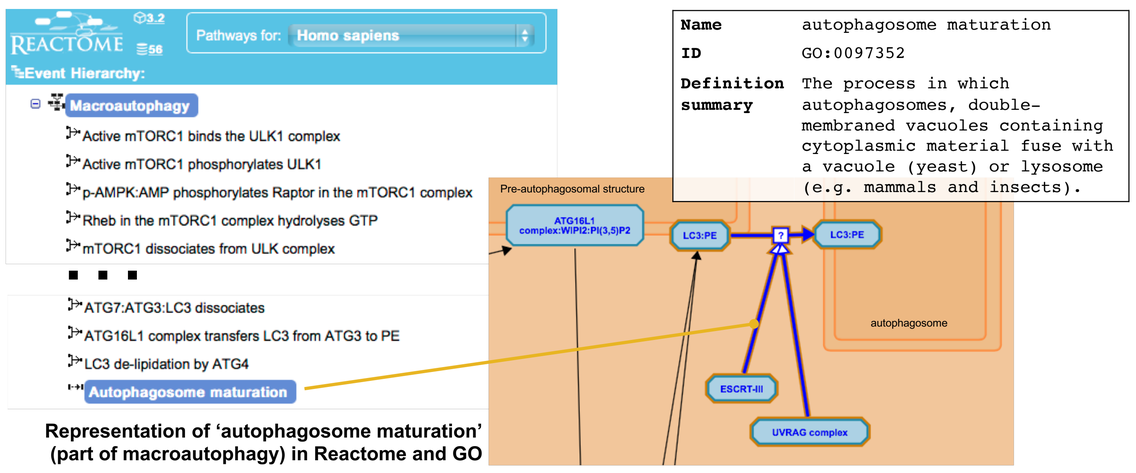
- We have so far created over 6000 annotations to more than 1400 distinct proteins (including over 800 human proteins) from over 500 papers (statistics correct as of June 2nd 2016).

A short summary of Parkinson's disease

- Parkinson's disease is a progressive, neurological condition resulting from loss of dopamine-producing neurons in the substantia nigra, a region of the brain controlling balance and movement.
- One in every 500 people develop Parkinson's disease, and this complex condition can affect people in different ways:
  - Motor symptoms of Parkinson's disease include tremor (shaking), slowed movement (bradykinesia) or loss of movement (akinesia) and rigidity/stiffness.
  - Non-motor symptoms also affect the day-to-day life of sufferers and include problems with the bladder, bowel and eyes, sleep disruption, loss of cognition, depression, dementia and other mental health effects.
- There's currently no cure, however some drugs are available that do help manage the symptoms of Parkinson's.
- The role of genetics in Parkinson's is becoming more understood with the ongoing discovery of genes involved in disease onset, progression and risk.

Aligning with Reactome

- Reactome (http://www.reactome.org) is a peer-reviewed, pathway database, and curates many cellular events that overlap with our priority processes, including macroautophagy:



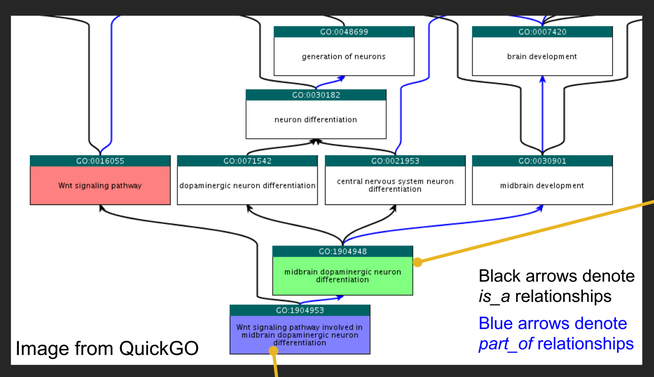
Name	autophagosome maturation
ID	GO:0097352
Definition summary	The process in which autophagosomes, double-membraned vacuoles containing cytoplasmic material fuse with a vacuole (yeast) or lysosome (e.g. mammals and insects).

Representation of 'autophagosome maturation' (part of macroautophagy) in Reactome and GO

- All proteins in the Reactome **Macroautophagy** pathway were downloaded and cross-checked against our GO-annotated proteins, and additional GO annotations added where required.
- We have suggested new mappings between Reactome and GO, and for Wnt signaling this has created more than 250 new Reactome GO annotations to > 150 human proteins.

Introduction to GO and GO annotation

- The Gene Ontology (GO) project is a collaborative effort to provide freely-available, consistent descriptions of gene products across all species.
- GO contains three structured controlled vocabularies (ontologies) that describe gene products in terms of their associated **biological processes**, **cellular locations** and **molecular functions.**
- Originally developed in 1998, the ontologies have grown to include over 42,000 terms describing a wide range of concepts to differing levels of specificity.
- GO is an essential resource for high-throughput data analysis, facilitating the grouping of genes to common pathways, functions and cellular locations.



name	midbrain dopaminergic neuron differentiation
ID	GO:1904948
definition	The process in which a relatively unspecialized cell acquires the specialized features of a midbrain dopaminergic neuron.
synonym	midbrain DA neurogenesis

This term was created by the GO Consortium with support from PARKINSON'S UK CHANGE ATTITUDES. FIND A CURE. JOIN US.

GO terms are associated with a gene product in a GO annotation

Evidence codes give an indication of the underlying assay supporting the annotation
IDA = inferred from direct assay
ISS = inferred from sequence similarity
NAS/TAS = author statements

Each annotation is attached to a reference for traceability

Symbol	GO Term Name	Evidence	With	Reference	Assigned By
LRP6	Wnt signaling pathway involved in midbrain dopaminergic neuron differentiation	TAS		PMID:24431302	ParkinsonsUK-UCL
WNT1	canonical Wnt signaling pathway involved in midbrain dopaminergic neuron differentiation	TAS		PMID:22988876	ParkinsonsUK-UCL
WNT2	canonical Wnt signaling pathway involved in midbrain dopaminergic neuron differentiation	IDA		PMID:20018874	ParkinsonsUK-UCL
RYK	Wnt signaling pathway involved in midbrain dopaminergic neuron differentiation	ISS	Q6BC88	GO_REF:0000024	ParkinsonsUK-UCL
WNT3A	Wnt signaling pathway involved in midbrain dopaminergic neuron differentiation	TAS		PMID:22988876	ParkinsonsUK-UCL
WNT5A	planar cell polarity pathway involved in midbrain dopaminergic neuron differentiation	TAS		PMID:22988876	ParkinsonsUK-UCL
FZD1	Wnt signaling pathway involved in midbrain dopaminergic neuron differentiation	NAS		PMID:23461676	ParkinsonsUK-UCL
LMX1A	midbrain dopaminergic neuron differentiation	ISS	Q9JKU8	GO_REF:0000024	ParkinsonsUK-UCL
NR4A2	midbrain dopaminergic neuron differentiation	TAS		PMID:24431302	ParkinsonsUK-UCL

A subset of human proteins annotated with 'midbrain dopaminergic neuron differentiation' and descendants during our project. Adapted from in the EBI GO browser QuickGO (www.ebi.ac.uk/QuickGO).

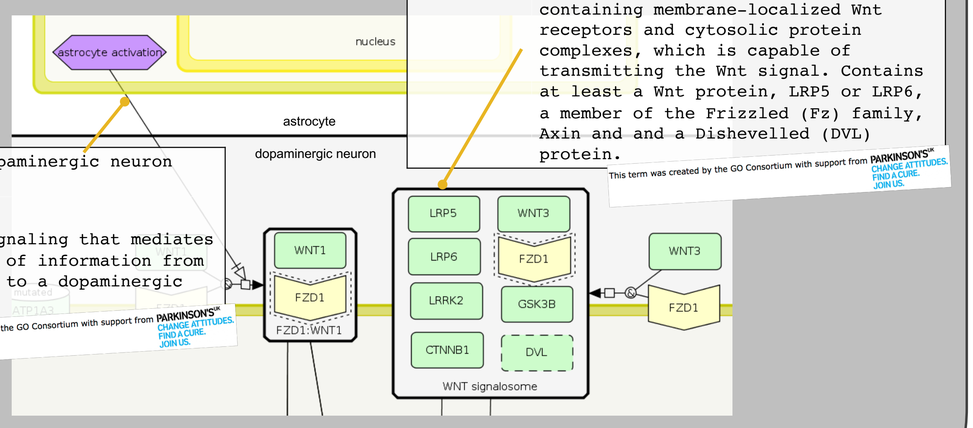
Annotation transferred from mouse Lmx1a

Denotes GO terms assigned by our project

Aligning with PD-map

- Parkinson's disease map (PD-map; http://minerva.uni.lu/MapViewer/) offers a pathway-view of Parkinson's disease. We are continuing to work with PD-map to align our projects so users can easily access all relevant data, and to ensure both resources represent the cellular processes underlying Parkinson's consistently.
- Our curation feeds back into the development of the Gene Ontology itself as we expand and improve areas of the ontology relevant to Parkinson's disease. We have worked with GO editors to create >360 new GO terms relevant to Parkinson's, and many of our new terms are based on PD-map; for Wnt signaling, we created new terms including '**Wnt signalosome**' and '**astrocyte-dopaminergic neuron signaling**':

A subsection of Wnt signaling adapted from PD-map



Name	astrocyte-dopaminergic neuron signaling
ID	GO:0036520
Definition	Cell-cell signaling that mediates the transfer of information from an astrocyte to a dopaminergic neuron.

This term was created by the GO Consortium with support from PARKINSON'S UK CHANGE ATTITUDES. FIND A CURE. JOIN US.

Name	Wnt signalosome
ID	GO:1990909
Definition	A multiprotein protein complex containing membrane-localized Wnt receptors and cytosolic protein complexes, which is capable of transmitting the Wnt signal. Contains at least a Wnt protein, LRP5 or LRP6, a member of the Frizzled (Fz) family, Axin and and a Dishevelled (DVL) protein.

This term was created by the GO Consortium with support from PARKINSON'S UK CHANGE ATTITUDES. FIND A CURE. JOIN US.

- We also aligned our annotations with PD-map; for both **Wnt signaling** and **autophagy** all proteins and processes represented in PD-map were curated with GO terms, to ensure that we had annotated all key areas.

Further reading

- **Using the Gene Ontology to annotate key players in Parkinson's disease**. Foulger RE, Denny P, Hardy J, Martin MJ, Sawford T, Lovering RC. Neuroinformatics 2016, Jul;14(3):297-304. *PMID 26825309*
- **Computational analysis of the LRRK2 interactome.** Manzoni C, Denny P, Lovering RC, Lewis PA. PeerJ. 2015, Feb 19;3:e778. doi: 10.7717/peerj.778. *PMID 25737818*
- **Ten quick tips for using the Gene Ontology**. Blake J.A. PLoS Comput Biol. 2013, Nov;9(11):e1003343. *PMID 24244145*
- **The impact of focused Gene Ontology curation of specific mammalian systems**. Alam-Faruque *et al*. PLoS One, 2011, 6(12):e27541. *PMID 22174742*

Email us for more information or to receive our quarterly newsletter, or follow us on Twitter:

goannotation@ucl.ac.uk

@UCLgene

www.ucl.ac.uk/functional-gene-annotation/neurological

www.geneontology.org

The Parkinson's UK annotation project is funded by Parkinson's UK, grant G-1307. Project members are part of the GO Consortium.