

Gene Ontology annotation of the Blood-Brain Barrier

B. Kramarz¹, M. Rodríguez-López², K. Thurlow¹, S. De Miranda Pinheiro¹, S.C.C. Saverimuttu¹, A. Ignatchenko², M.J. Martin², S. Orchard², R. Bandopadhyay³, D. Brough⁴, N.M. Hooper⁴, H. Attrill⁵, R.C. Lovering¹

1. Functional Gene Annotation, Preclinical and Fundamental Science, Institute of Cardiovascular Science, University College London (UCL); 2. European Bioinformatics Institute (EMBL-EBI), European Molecular Biology Laboratory, Wellcome Genome Campus, Cambridge; 3. UCL Institute of Neurology and Reta Lila Weston Institute of Neurological Studies, London; 4. Division of Neuroscience and Experimental Psychology, School of Biological Sciences, University of Manchester, Manchester; 5. FlyBase, Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge.



What is Gene Ontology?

- The Gene Ontology (GO) Consortium aims to provide a comprehensive, computational model of biological systems, ranging from the molecular to the organism level
- The GO knowledgebase is the world's largest source of functional information about gene products, such as proteins or non-coding RNAs, which are all referred to as 'genes' throughout this poster
- The biological knowledge is captured in both human-readable and machine-readable formats
- The freely-available descriptions of genes provided by GO form a foundation for computational analysis of large-scale cellular and genetic biomedical research data
- The association of a descriptive GO term with a gene results in a gene 'annotation'
- The GO terms are standardised 'phrases' that describe either the function of a gene product, its biological role or its location in the cell

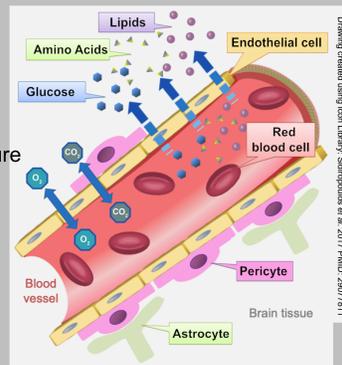
Annotation of the Blood-Brain Barrier

The role of the blood-brain barrier (BBB) in Alzheimer's disease and other neurodegenerative disorders is now being investigated by many research groups. However, the BBB is poorly represented in the majority of online biomedical resources, which is likely to be having an impact on a variety of neurological functional and big data analyses.

To address this problem we have prioritised capture of experimental data describing this tissue.

In order to appropriately capture the knowledge about the BBB we have:

- 1) Revised the GO
 - Improving the 'cell junction' ontology structure
 - Improving the 'cell junction organisation' ontology structure
- 2) Prioritised annotation of 105 proteins with a role in:
 - Maintaining the structural integrity of the BBB
 - Transport across the BBB
- 3) Prioritised annotation of microRNAs with a role in:
 - regulating the expression of BBB priority proteins

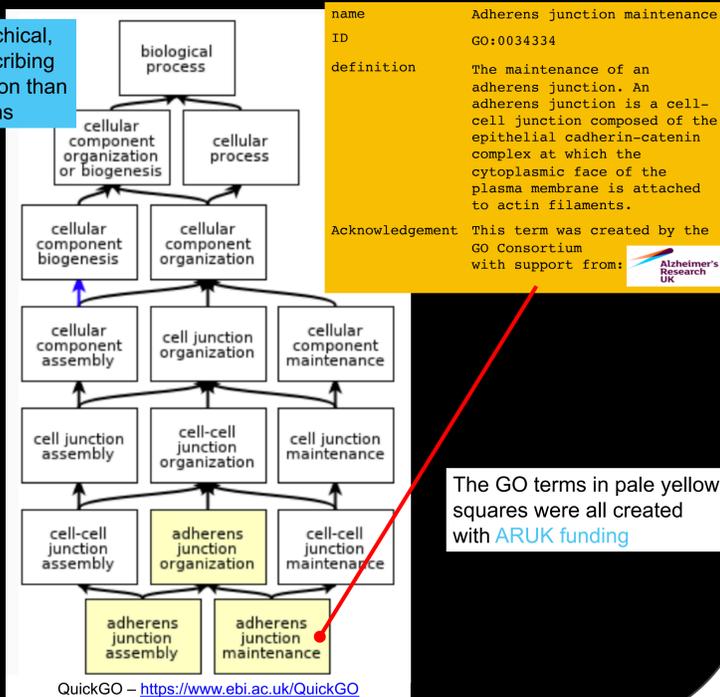


Visit <https://tinyurl.com/BBB105> for the full list of prioritised proteins

What is an ontology?

- There are currently 45,000 GO terms describing a wide range of biological concepts to differing levels of specificity.
- The terms are organised as a unidirectional network, each term is a node and each edge a relation

An ontology is hierarchical, with 'child' terms describing more specific information than their 'parent' terms



name Adherens junction maintenance
ID GO:0034334
definition The maintenance of an adherens junction. An adherens junction is a cell-cell junction composed of the epithelial cadherin-catenin complex at which the cytoplasmic face of the plasma membrane is attached to actin filaments.
Acknowledgement This term was created by the GO Consortium with support from:

The GO terms in pale yellow squares were all created with ARUK funding

Further reading

- Kramarz B, Roncaglia P, Meldal BHM, Huntley RP, Martin MJ, Orchard S, Parkinson H, Brough D, Bandopadhyay R, Hooper NM, Lovering RC. **Improving the Gene Ontology Resource to Facilitate More Informative Analysis and Interpretation of Alzheimer's Disease Data.** *Genes(Basel)*. 2018 9(12). pii: E593. PMID:30501127.
- Kramarz B, Lovering RC. **Gene Ontology: A Resource for Analysis and Interpretation of Alzheimer's Disease Data.** In: Wisniewski T, editor. *Alzheimer's Disease*. Brisbane (AU): Codon Publications; 2019 Dec 20. Chapter 2. PMID:31895510.
- The Gene Ontology Consortium. **The Gene Ontology Resource: 20 years and still GOing strong.** *Nucleic Acids Res*. 2019 47(D1):D330-D338. PMID: 30395331.

Creation of GO annotations at UCL

GO biocurators search the literature for experimental data and capture information about biological roles using the universal GO terms. The resulting GO annotations are regularly submitted to GO browsers (QuickGO and AmiGO 2), and exported to other biological databases, such as UniProt, Ensembl and NCBI Gene, as well as to Wikipedia.

Symbol	Qualifier	GO Term	Evidence	Reference	Assigned By
JAM3	involved_in	GO:0098609 cell-cell adhesion	ECO:0000314 IDA	PMID:11823489	ARUK-UCL
JAM3	enables	GO:0098632 cell-cell adhesion mediator activity	ECO:0000314 IDA	PMID:11823489	ARUK-UCL
JAM3	part_of	GO:0070160 tight junction	ECO:0000314 IDA	PMID:19060272	ARUK-UCL
JAM3	involved_in	GO:0034113 heterotypic cell-cell adhesion	ECO:0000314 IDA	PMID:12208882	ARUK-UCL
JAM3	involved_in	GO:0034333 adherens junction assembly	ECO:0000315 IMP	PMID:19060272	ARUK-UCL

A subset of the 84 GO annotations associated with the human protein JAM3 (junctional adhesion molecule 3). 22 of these annotations were created during this focused annotation project. Adapted from the EMBL-EBI GO browser QuickGO (www.ebi.ac.uk/QuickGO).

Biocurators at UCL have submitted 10% of all manual GO annotations for human gene products. Funding from ARUK has enabled the UCL team to create 1500 GO annotations describing the role of 227 human gene products at the BBB. Of these 49 proteins are associated with the GO term 'cell junction organization', or one of its child terms, and 121 are associated with a transport child term.

Curated	Total all species	Total human	Priority proteins human
Articles	227	194	181
Gene products	452	204	72 (out of 105)**
GO annotations	2058	1490	1101

GO annotations are used to interpret 'big data'

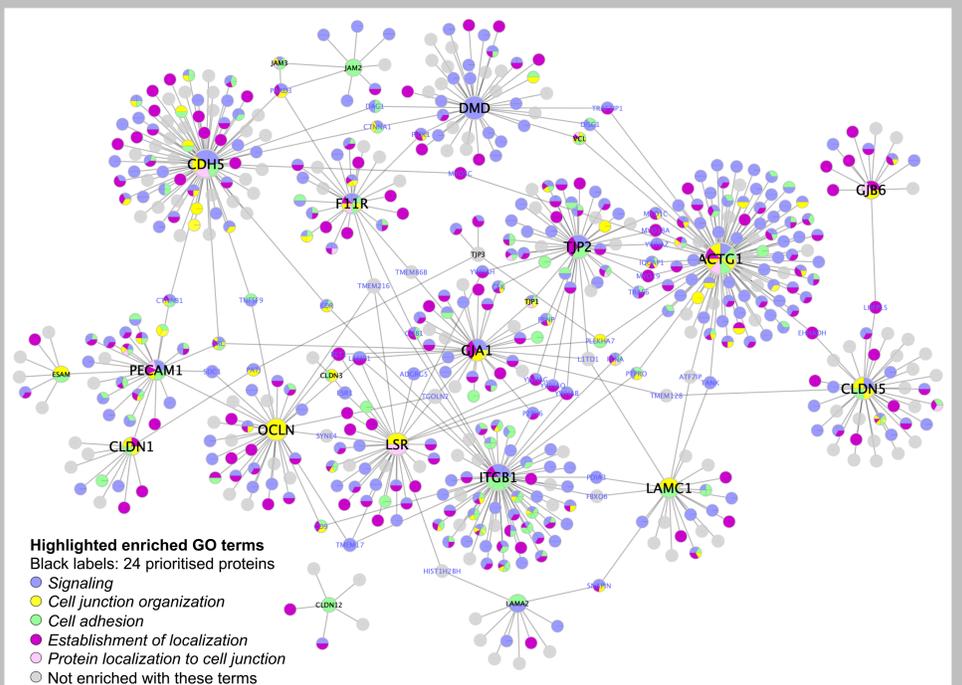
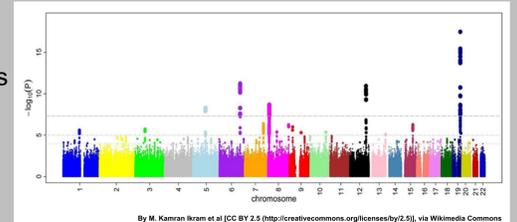
- GO annotations group genes based on their function, biological role or cellular location
- GO annotations are incorporated into > 50 functional analysis tools
- Researchers around the world use GO annotations to analyse high-throughput datasets

Genome-wide association studies

- Inform choice of candidate genes
- Detect new risk genes

Transcriptomics and Proteomics

- Identify dysregulated processes
- Validate results
- Summarise the role of the expressed genes



Network of a selection of BBB-associated proteins. 24 prioritised proteins involved in maintaining the structural integrity of the BBB used to seed an interaction network using Cytoscape and analysed using Golorize, BiNGO and GO ontology and annotation files (downloaded 5th March 2020). The proteins associated with a selection of the enriched GO terms (or one of their child terms, including regulation child terms) are shown in the network.



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

