Immunology
The Journal of cells, molecules, systems and technologies

British Society for immunology

IMMUNOLOGY REVIEW ARTICLE

Access to immunology through the Gene Ontology

Ruth C. Lovering,[1] Evelyn B. Camon,[2] Judith A. Blake[3] and Alexander D. Diehl[3]

[1]*Department of Medicine, University College London, Rayne Institute, London, UK,* [2]*GOA Project, European Bioinformatics Institute, Hinxton, Cambridge, UK, and* [3]*Mouse Genome Informatics, The Jackson Laboratory, Bar Harbor, ME, USA*

doi:10.1111/j.1365-2567.2008.02940.x
Received 11 July 2008; revised 11 July 2008; accepted 29 July 2008.
Correspondence: R. C. Lovering, Department of Medicine, University College London, Rayne Institute, 5 University Street, London WC1E 6JF, UK. Email: r.lovering@ucl.ac.uk
Senior author: Alexander D. Diehl, email: adiehl@informatics.jax.org

Summary

The Gene Ontology (GO) is widely recognized as the premier tool for the organization and functional annotation of molecular aspects of cellular systems. However, for many immunologists the use of GO is a very foreign concept. Indeed, as a controlled vocabulary, GO can almost be considered a new language, and it can be difficult to appreciate the use and value of this approach for understanding the immune system. This review reflects on the application of GO to the field of immunology and explains the process of GO annotation. Finally, this review hopes to inspire immunologists to invest time and energy in improving both the content of the GO and the quality of GO annotations associated with genes of immunological interest.

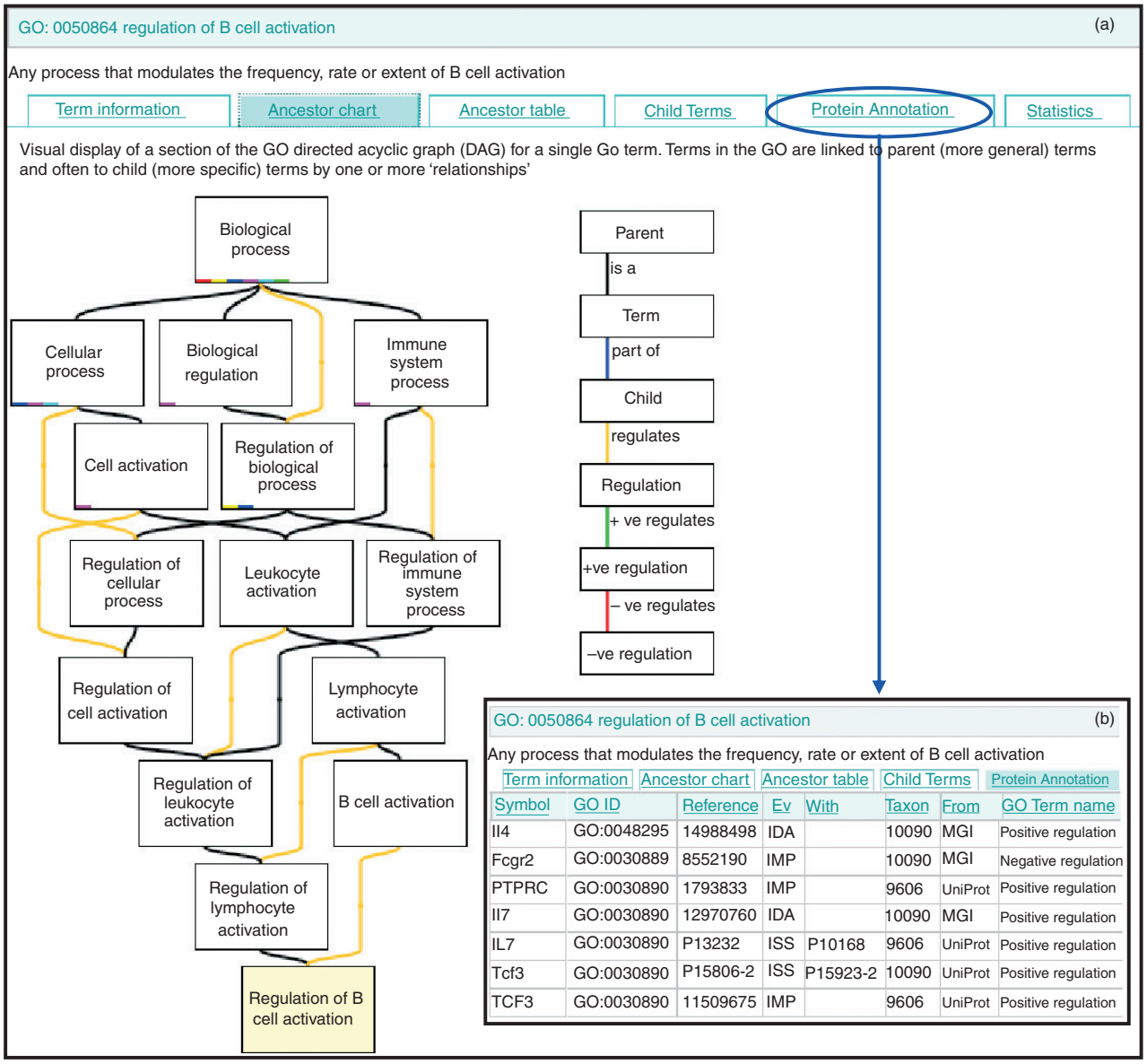
Keywords: Gene Ontology; high-throughput analysis; immunology

Introduction

Traditionally, the field of immunology has embraced any emerging technology that enables accelerated progress in the understanding of the intricate network of molecular and cellular interactions associated with immunological processes and disease. Until recently, the study of the specific pathways or individual molecules has been the major approach taken; however, during the past decade genome-sequencing projects have led to the identification of thousands of novel genes in higher vertebrates. Experimental investigation to understand how, or if, these genes are involved in immune-related processes is an ongoing process. High-throughput methodologies, such as expression arrays or proteomics, provide substantial information about the properties of these newly identified genes, through the detailed characterization of the molecular composition of entire tissues, cells or organelles at both specific developmental and disease states or through protein-binding studies or cellular-location studies. Consequently, such methodologies provide researchers with the potential to increase rapidly our understanding of the complex interactions and biological functions within the immune system. Integrating high-throughput data with the accrued knowledge about individual genes, gained through focused experimental approaches, is an essential step to ensure that data derived from all experimental approaches informs current research projects.

The Gene Ontology Consortium (GOC) has been developing the terms necessary to describe each gene product to enable the integration of all this data and to facilitate bridging the gap between data collation and data analysis[1] (http://www.geneontology.org), including a full set of terms for immunological processes.[2] However, translating the wealth of knowledge available in the field of immunology into comprehensive functional annotations using the Gene Ontology (GO) is a substantial undertaking. The depth of knowledge in this area is immense and, as an outsider, it can be difficult to appreciate fully the complexity of the system and the range of processes a single gene can be involved in. Consequently, the annotation with GO terms of many of the genes involved in immunological processes has yet to reflect the volume of literature available in this field. This is where the support and input of immunologists is sought.

GO provides three detailed, structured vocabularies of terms (ontologies) which describe the molecular functions that gene products normally carry out, the biological processes (Fig. 1) that gene products are involved in



Symbol	GO ID	Reference	Ev	With	Taxon	From	GO Term name
Il4	GO:0048295	14988498	IDA		10090	MGI	Positive regulation
Fcgr2	GO:0030889	8552190	IMP		10090	MGI	Negative regulation
PTPRC	GO:0030890	1793833	IMP		9606	UniProt	Positive regulation
Il7	GO:0030890	12970760	IDA		10090	MGI	Positive regulation
IL7	GO:0030890	P13232	ISS	P10168	9606	UniProt	Positive regulation
Tcf3	GO:0030890	P15806-2	ISS	P15923-2	10090	UniProt	Positive regulation
TCF3	GO:0030890	11509675	IMP		9606	UniProt	Positive regulation

Figure 1. QuickGO browser view of the Gene Ontology (GO) biological process term 'regulation of B-cell activation' (http://www.ebi.ac.uk/ego/GTerm?id=GO:0050864). The QuickGO record displays the GO term, identifier and definition along with ancestry information (a), child terms, proteins annotated to the GO term (b) and statistical data. The blue arrow and circle indicate the 'protein annotation' tab view option. (a) In the QuickGO ancestor graph, the term 'regulation of B-cell activation' is shown to be a child of the broader parent terms 'regulation of lymphocyte activation' and 'B-cell activation'. (b) Only a few of the 125 proteins associated with 'regulation of B-cell activation' and its child terms are shown in this example.

and lastly the localization (cellular component), relative to the cell, where gene products are active.[2–5] For example, the annotations for tumor necrosis factor (*TNF*) include the molecular function term 'tumor necrosis factor receptor binding', the biological process term 'leukocyte tethering or rolling' and the cellular component term 'extracellular space'; whereas the annotations for caspase recruitment domain family, member 11 (*CARD11*) include the molecular function term 'guanylate kinase activity', the biological process term 'positive regulation of cytokine production' and the cellular component term colocalizes_with 'T-cell-receptor complex'.

The GOC also provides data sets of GO terms associated with the appropriate genes and their products, for many different species[1] (http://www.geneontology.org/GO.current.annotations.shtml). Depending on the amount of published data available, gene and protein identifiers can be annotated with multiple GO terms from any, or all, of the three gene ontologies (Fig. 2). Annotations can be produced either by a professional curator (typically a

VCAM1 Homosapiens P19320

Accession	P19320
Gene	VCAM1
Taxonomy	Homosapiens
Description	Vascular cell adhesion protein 1 precursor (V-CAM1) (INCAM-100) (CD106 antigen)

GO ID	GO Term name	Reference	Ev	With	From
GO:0016337	Cell-cell adhesion	interpro	IEA	IPR003987	UniProt
GO:0007155	Cell adhesion	spkw	IEA	KW-0130	UniProt
GO:0044419	Interspecies interaction between organisms	spkw	IEA	KW-0945	UniProt
GO:0007159	Leukocyte adhesion	12082081	IEP		BHF-UCL
GO:0030183	B cell differentiation	1715889	IC	GO:0007159	BHF-UCL
GO:0007159	Leukocyte adhesion	1715889	IDA		BHF-UCL
GO:0050901	Leukocyte tethering or rolling	2688898	IEP		BHF-UCL
GO:0007159	Leukocyte adhesion	2688898	IDA		BHF-UCL
GO:0007159	Leukocyte adhesion	1381355	IDA		BHF-UCL
GO:0042102	Positive regulation of T cell proliferation	1381355	IDA		BHF-UCL
GO:0007157	Heterophilic cell adhesion	1715889	IDA		BHF-UCL
	Function				
GO:0005515	Protein binding	12082081	IPI	P15311	BHF-UCL
GO:0005178	Integrin binding	1377228	IDA		BHF-UCL
	Component				
GO:0016021	Integral to membrane	spkw	IEA	KW-0812	UniProt
GO:0016020	Membrane	spkw	IEA	KW-0472	UniProt
GO:0009986	Cell surface	1715889	IDA		BHF-UCL
GO:0009897	External side of plasma membrane	1377228	IDA		BHF-UCL
GO:0005615	Extracellular space	9290466	IDA		BHF-UCL

Figure 2. Protein record page of the QuickGO browser, showing all annotations for the human vascular cell adhesion molecule 1 (VCAM-1) protein (http://www.ebi.ac.uk/ego/GProtein?ac=P19320). The Gene Ontology (GO) terms associated with the VCAM-1 protein record were derived from only six of the thousands of *VCAM1* publications. It is unlikely that the annotation of any highly studied gene will ever be completed, as new publications are available each week. However, for this protein the addition of annotations extracted from a few more key publications would enable the GO curators to increase substantially the number and quality of annotations. The QuickGO protein record page displays a short summary of the chosen UniProtKB protein accession number, together with its associated GO annotations. The evidence column (*Ev*) displays the acronym of the GO evidence code used to categorize the type of evidence that has been found to support the association of the GO term with the protein. Seventeen different evidence codes can be used to categorize the type of evidence available to support the annotation, one of which (IEA) is use to indicate an 'electronic' source for the annotation; the remaining 16 are all used for manually curated annotations. The 'from' column lists an acronym for the database accredited for the annotation. Additional supporting data for the annotation, such as a GO:ID or InterPro domain is listed in the 'with' column. Hyperlinks on this page enable users to retrieve the paper supporting the annotation and to view the information available for other database identifiers.

post-PhD biologist) reading published scientific papers and creating each association or by a computational biologist applying bioinformatic techniques to predict associations.[6–9] These two broad categories of techniques have their own advantages and disadvantages, but both require skilled scientists to ensure that conservative, high-quality annotations are created. The annotation of each gene can therefore be a laborious process, which for a highly-studied gene such as B-cell CLL/lymphoma 10 (*BCL10*) or vascular cell adhesion molecule 1 (*VCAM1*) (Fig. 2) could take several days or, for a more recently described gene such as melanoma inhibitory activity family, member 3 (*MIA3*), may only take a few hours. GO annotation does not attempt to replace the need for researchers to read scientific literature; it simply provides a computable, yet comprehensive, description of a gene or protein drawn from relevant and traceable publications.

A large range of applications have been developed specifically for the visualization of GO and its associated annotation data, and for the computational and statistical analysis of large data sets using GO. Currently there are 48 tools for gene expression and microarray analysis and 20 GO browsers, all of which are listed on the GOC tools web page (http://www.geneontology.org/GO.tools.shtml). Additionally, GOC annotation data sets are imported into many of the top biological databases, including UniProtKB, Ensembl, EntrezGene, GeneCards and InnateDB.

The application of GO to immunological research

GO is being used to identify gene groupings within data sets derived from a wide range of sources relevant to the field of immunology. In addition to the classical differential expression of genes identified by microarray analysis,[10] other high-throughput technologies use GO to identify statistically enriched functional, pathway or component-associated gene groups.[5] Such technologies may identify gene data sets as mRNA targets,[11] soluble proteins,[12] membrane proteins,[13] transcription factor targets[14] or protein-binding networks.[15] Furthermore, GO

is also being used to ensure the accuracy of new experimental methods.[16–19]

GO annotation data are often used to guide development of a hypotheses to explain proteome-wide alterations in response to certain diseases, such as arthritis,[13] or stress states, such as hypoxia.[20] In such studies, an indication of underlying cellular mechanisms, which may account for an observed phenotype, can be obtained using GO to cluster subsets of proteins that share related GO annotation and are found to be similarly over-expressed or under-expressed in the disease or stress state. Ishikawa *et al.*[13] applied GO to a data set derived from comparing the expression of genes in the peripheral blood cells of healthy children to those of patients with systemic juvenile idiopathic arthritis (sJIA). As expected, the analysis indicated an involvement of the defense-response system in sJIA. Unexpectedly, the analysis also suggested that a mitochondrial disorder may play a role in this disease.[13]

The ability to include GO within the analysis of large data sets has also been found to be useful for the identification of new sets of biomarkers for a certain disease. This approach has enabled investigators of graft-versus-host disease,[21] osteoarthritis[12] and chronic kidney disease[22] to identify new diagnostic biomarkers and to indicate disease-associated deregulated processes. Differentially expressed gene clusters, associated with specific diseases, identified with the use of GO, may also provide targets for developing new disease therapies and/or candidates for genetic research.[23] Furthermore, the broad gene categories available through the use of GO are leading to unexpected links between immune-associated genes and non-immune-associated tissues and disease.[23]

GO can also be used to provide a link between the protein-binding network and the activities and locations of the participant proteins. Dyer *et al.*[15] used GO data to investigate interactions of human proteins with viral pathogens and found that many different pathogens target the same processes in the human cell, such as regulation of apoptosis, even though they may interact with different proteins. The use of GO enabled Emmonds *et al.*[11] to identify that in dendritic cells 25% of the proteins encoded by the mRNA targets bound by the mRNA stabilizing protein tristetraprolin (ZFP36) were associated with protein synthesis. This suggested that tristetraprolin has a broader role in regulating the immune response than previously suspected.[11]

Many proteomic investigations have used GO data to verify the success of subcellular enrichment strategies or large-scale confocal microscopy analyses.[16–19] Crockett *et al.*[16] applied the GOC data set to confirm that their subcellular fractionation protocol efficiently isolated the appropriate subcellular compartments. For example, of the 553 proteins detected in the cytoplasmic fraction, over half of the proteins were annotated solely to the GO component term 'cytoplasm', whereas nearly half of the membrane fraction proteins were annotated to the GO terms 'membrane' or 'extracellular region'.[16]

Although it may appear that only scientists using high-throughput methodologies will benefit from GO, scientists looking for new disease-associated genes are also expected to benefit. For example, genome-wide association scans for Crohn's disease[24] are identifying novel genes with previously unknown connections to immunological processes, and the comprehensive annotation of these novel genes might provide clues to their link with the immune system.

The manual annotation approach

In general, scientific papers are written as interesting text rather than as a list of results and conclusions; a single function for one gene may be discussed using a series of similar, yet non-overlapping, descriptions. Consequently, although current text mining tools are able to locate useful papers for curation, they cannot provide the correct, detailed descriptions of the functions and processes in which a gene is involved.[25] Curated gene annotation typically results in high-quality annotation, but is labour intensive[6] (http://www.ebi.ac.uk/GOA/annotationexample.html). GO curators are required to annotate each gene product efficiently. Consequently, they may not have the time to read every paper published about a gene, and may rely on the literature cited in a review to identify key references.

Association of GO terms with specific gene products

During the annotation process, curators read in full the appropriate publications to gather detailed experimental data. The curator then uses a GO browser to identify an appropriate GO term for the process, function or cellular location with which the gene product has been experimentally shown to be associated.[26] At this stage, the curator will confirm that the GO term is appropriate by reading the GO term definition, viewing the location of the term within the GO hierarchy and ensuring that none of its child terms are more appropriate. If the GO term is not appropriate, a new term or change in definition is easily requested from the GO editorial office.[27] In addition to associating a GO term with a gene product (via database identifiers), one of 17 evidence codes (http://www.geneontology.org/GO.evidence.shtml) is included in the annotation to indicate the type of experimental evidence supporting the annotation. For example, anti-VCAM1 immunoglobulin IgG1 was shown to block adhesion of normal B-cell precursors to bone marrow-derived fibroblasts,[28] and therefore the biological process term 'heterophilic cell adhesion' was associated with the VCAM1 record and given the evidence code 'IDA', an acronym for 'inferred from direct assay' (Fig. 2). Reading a single publication can lead to the association of several

GO terms with one protein, or to the association of a single GO term with multiple proteins. For instance, use of an enzyme-linked immunosorbent assay (ELISA) detected the secretion of four cytokines, interleukin (IL)-10, IL-5, IL-6 and IL-13, into the culture supernatant by murine splenic dendritic cells,[29] and using this information a GO curator associated the cellular component term 'extracellular space' with these four proteins, with the evidence code 'IDA'.

Call for community contributions

Clearly, scientists working in specific research areas could provide the most comprehensive knowledge about specific genes and gene products. However, the majority of bench scientists do not have the time to learn the complexities of GO annotation, or the time to edit databases. Several GO annotation groups are now looking for ways in which scientists can submit relevant information to GO curators, thereby reducing the time each curator takes to annotate genes, without unduly burdening the research scientists.[30,31]

Although the GOC database provides sufficient GO terms to enable interpretation of high-throughput analysis, the current quality and quantity of the annotations will, and currently does, limit the level of interpretation that can be achieved. Lee *et al.*[10] identified an 'intracellular transport-related' gene cluster following microarray analysis of B cells stimulated with a variety of ligands. However, by adding their own annotations to the 38 genes in this cluster, Lee *et al.*[10] were able to identify specific transport processes up-regulated following B-cell stimulation. Rather than individual groups annotating their own genes of interest, it would be more cost effective if these resources were pooled and the annotations stored within the GOC database. The scientific community could help to improve the quality of GO annotation, and potentially increase the impact value of their own publications, by investing only a small amount of time in GO. One of the most time-consuming aspects of GO annotation is finding the appropriate paper to use for annotation. This time could be reduced if scientists were willing to spend a small amount of time sending details of key publications to

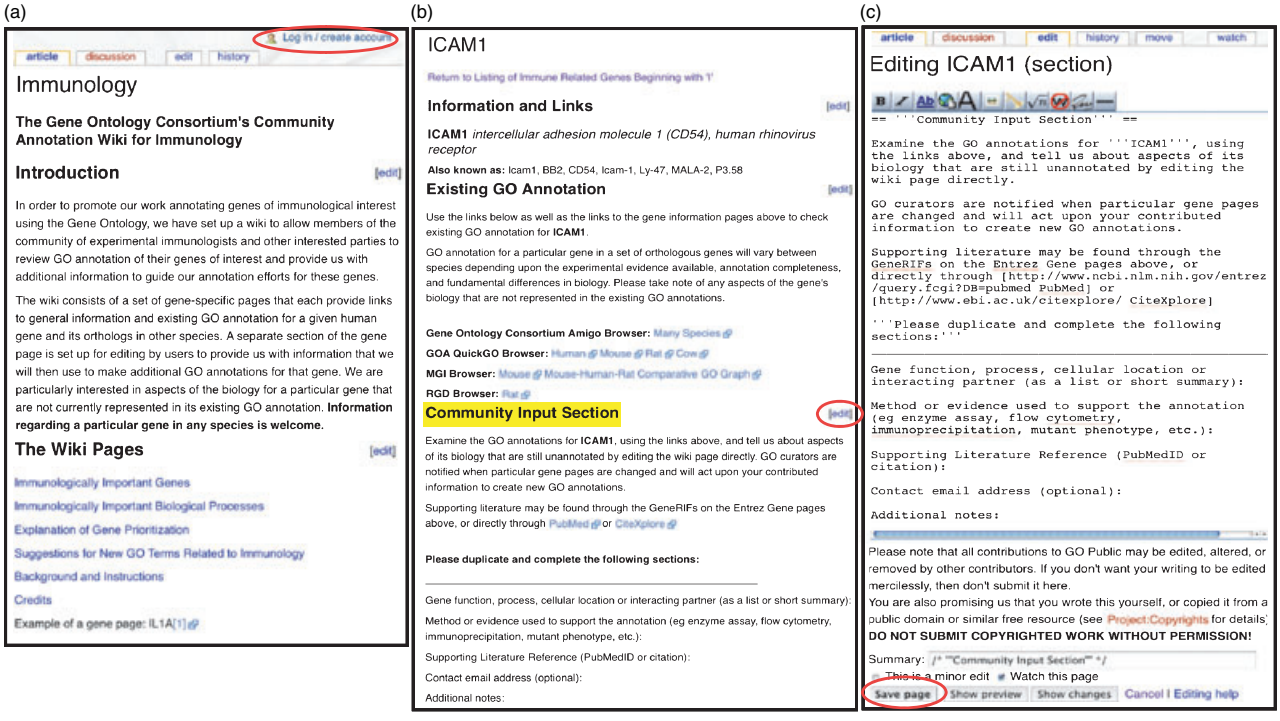


Figure 3. Wiki pages available at the Gene Ontology Consortium (GOC) immunology Wiki site. (a) Immunology Wiki index page (http://wiki.geneontology.org/index.php/Immunology), an index for the 1326 high-priority genes associated with the immune system, is available at the bottom of this page. While there is full public access to this site, scientists interested in commenting on the annotation of a gene will need to create an account by completing a short form (link to account page is circled in red). (b) Each of the immune-associated genes has an individual Wiki page with hyperlinks (in blue) to a variety of external sites (http://wiki.geneontology.org/index.php/ICAM1). The 'Community Input Section' is highlighted in yellow, with the link to the editing page (c) circled in red. (c) Information can be submitted to the Gene Ontology (GO) curators by adding comments to the 'community input section' in a gene record and using the 'save page' option (circled in red, http://wiki.geneontology.org/index.php?title=ICAM1&action=edit&section=3).

GO curators (mgi-go@informatics.jax.org, goa@ebi.ac.uk, GOAnnotation@ucl.ac.uk).

Additionally, the quality of GO annotation for each gene is currently dependent on the knowledge that a curator has in each field. If scientists reviewed the GO annotations available for their favourite gene and sent comments about missing or incorrect annotations to the GO curators, completeness of the annotations would be achieved. To facilitate this approach an editable Wiki system has been set up by the GOC (Fig. 3, http://wiki.geneontology.org/index.php/Immunology). From the immunology Wiki index page it is possible to access gene-specific pages, which have hyperlinks to the GO annotation available for each gene and an area to suggest improvements to these annotations. Immunologists can view the gene record and current annotations associated with their favourite gene and then use the edit page to add comments about the annotation and to suggest additional references or missing terms. Contributing scientists and GO curators can use the page recognition facility (by checking the 'Watch this page' option) to receive notification when any edits are made to a gene-specific Wiki page (Fig. 3).

Consulting experts from the immunological community will ensure that the current accumulated knowledge in this field has been comprehensively reviewed and correctly summarized. A comprehensive representation of immunological knowledge within the GOC database will ensure that the analysis of future experiments will lead to well-supported hypotheses.

Acknowledgements

We thank Emily Dimmer for her editorial contributions to this manuscript. The Cardiovascular GO Annotation Initiative is funded by the British Heart Foundation (SP/07/007/23671). The GO work at MGI is funded by a P41 grant from the National Human Genome Research Institute (NHGRI grant HG002273).

References

1 The Gene Ontology Consortium. Creating the gene ontology resource: design and implementation. *Genome Res* 2001; **11**:1425–33.

2 Diehl AD, Lee JA, Scheuermann RH, Blake JA. Ontology development for biological systems: immunology. *Bioinformatics* 2007; **23**:913–5.

3 Lomax J. Get ready to GO! A biologist's guide to the Gene Ontology. *Brief Bioinform* 2005; **6**:298–304.

4 Ashburner M, Ball CA, Blake JA *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000; **25**:25–9.

5 Dimmer EC, Huntley RP, Barrell DG *et al.* The Gene Ontology – providing a functional role in proteomic studies. *Practical Proteomics* 2008; July [Epub ahead of print].

6 Camon E, Magrane M, Barrell D *et al.* The Gene Ontology Annotation (GOA) Database: sharing knowledge in Uniprot with Gene Ontology. *Nucleic Acids Res* 2004; **32** (Database issue):D262–6.

7 Camon E, Magrane M, Barrell D *et al.* The Gene Ontology Annotation (GOA) project: implementation of GO in SWISS-PROT, TrEMBL, and InterPro. *Genome Res* 2003; **13**:662–72.

8 Hill DP, Begley DA, Finger JH *et al.* The mouse Gene Expression Database (GXD): updates and enhancements. *Nucleic Acids Res* 2004; **32** (Database issue):D568–71.

9 Mulder NJ, Apweiler R, Attwood TK *et al.* InterPro, progress and status in 2005. *Nucleic Acids Res* 2005; **33** (Database issue):D201–5.

10 Lee JA, Sinkovits RS, Mock D *et al.* Components of the antigen processing and presentation pathway revealed by gene expression microarray analysis following B cell antigen receptor (BCR) stimulation. *BMC Bioinformatics* 2006; **7**:237.

11 Emmons J, Townley-Tilson WH, Deleault KM, Skinner SJ, Gross RH, Whitfield ML, Brooks SA. Identification of TTP mRNA targets in human dendritic cells reveals TTP as a critical regulator of dendritic cell maturation. *RNA* 2008; **14**:888–902.

12 Wu J, Liu W, Bemis A, Wang E, Qiu Y, Morris EA, Flannery CR, Yang Z. Comparative proteomic characterization of articular cartilage tissue from normal donors and patients with osteoarthritis. *Arthritis Rheum* 2007; **56**:3675–84.

13 Ishikawa S, Mima T, Aoki C *et al.* Abnormal expression of the genes involved in cytokine networks and mitochondrial function in systemic juvenile idiopathic arthritis identified by DNA maicroarray analysis. *Ann Rheum Dis* 2008; April 3 [Epub ahead of print].

14 Long F, Liu H, Hahn C, Sumazin P, Zhang MQ, Zilberstein A. Genome-wide prediction and analysis of function-specific transcription factor binding sites. *In Silico Biol* 2004; **4**:395–410.

15 Dyer MD, Murali TM, Sobral BW. The landscape of human proteins interacting with viruses and other pathogens. *PLoS Pathog* 2008; **4**:e32.

16 Crockett DK, Seiler CE 3rd, Elenitoba-Johnson KS, Lim MS. Annotated proteome of a human T-cell lymphoma. *J Biomol Tech* 2005; **16**:341–6.

17 Cao R, He Q, Zhou J, Liu Z, Wang X, Chen P, Xie J, Liang S. High-throughput analysis of rat liver plasma membrane proteome by a nonelectrophoretic in-gel tryptic digestion coupled with mass spectrometry identification. *J Proteome Res* 2008; **7**:535–45.

18 Stevens SM Jr, Duncan RS, Koulen P, Prokai L. Proteomic analysis of mouse brain microsomes: identification and bioinformatic characterization of endoplasmic reticulum proteins in the mammalian central nervous system. *J Proteome Res* 2008; **7**:1046–54.

19 Barbe L, Lundberg E, Oksvold P *et al.* Toward a confocal subcellular atlas of the human proteome. *Mol Cell Proteomics* 2008; **7**:499–508.

20 Boraldi F, Annovi G, Carraro F, Naldini A, Tiozzo R, Sommer P, Quaglino D. Hypoxia influences the cellular cross-talk of human dermal fibroblasts. A proteomic approach. *Biochim Biophys Acta* 2007; **1774**:1402–13.

21 Oh SJ, Cho SB, Park SH *et al.* Cell cycle and immune-related processes are significantly altered in chronic GVHD. *Bone Marrow Transplant* 2008; **41**:1047–57.

22 Perco P, Wilflingseder J, Bernthaler A, Wiesinger M, Rudnicki M, Wimmer B, Mayer B, Oberbauer R. Biomarker candidates

for cardiovascular disease and bone metabolism disorders in chronic kidney disease: a systems biology perspective. *J Cell Mol Med* 2008; **12**:1177–87.

23 van Gassen KL, de Wit M, Koerkamp MJ, Rensen MG, van Rijen PC, Holstege FC, Lindhout D, de Graan PN. Possible role of the innate immunity in temporal lobe epilepsy. *Epilepsia* 2008; **49**:1055–65.

24 Mathew CG. New links to the pathogenesis of Crohn disease provided by genome-wide association scans. *Nat Rev Genet* 2008; **9**:9–14.

25 Camon EB, Barrell DG, Dimmer EC, Lee V, Magrane M, Maslen J, Binns D, Apweiler R. An evaluation of GO annotation retrieval for BioCreAtIvE and GOA. *BMC Bioinformatics* 2005; **6**(Suppl. 1):S17.

26 Hill DP, Smith B, McAndrews-Hill MS, Blake JA. Gene Ontology annotations: what they mean and where they come from. *BMC Bioinformatics* 2008; **9**(Suppl. 5):S2.

27 The Gene Ontology Consortium. The Gene Ontology (GO) project in 2006. *Nucleic Acids Res* 2006; **34**(Database issue):D322–6.

28 Ryan DH, Nuccie BL, Abboud CN, Winslow JM. Vascular cell adhesion molecule-1 and the integrin VLA-4 mediate adhesion of human B cell precursors to cultured bone marrow adherent cells. *J Clin Invest* 1991; **88**:995–1004.

29 McKee AS, Pearce EJ. CD25 + CD4 + cells contribute to Th2 polarization during helminth infection by suppressing Th1 response development. *J Immunol* 2004; **173**:1224–31.

30 Ort DR, Grennan AK. Plant Physiology and TAIR partnership. *Plant Physiol* 2008; **146**:1022–3.

31 Lovering RC, Dimmer E, Khodiyar VK, Barrell DG, Scambler P, Hubank M, Apweiler R, Talmud PJ. Cardiovascular GO annotation initiative year 1 report: why cardiovascular GO? *Proteomics* 2008; **8**:1950–3.