

# Improving the Representation of Immunology in the Gene Ontology



Alexander D. Diehl, Jamie Lee, Richard H. Scheuermann, David P. Hill, Judith A. Blake

The Gene Ontology Consortium at MGI, The Jackson Laboratory, Bar Harbor, ME and the Immunology Database and Analysis Portal - ImmPort, UT Southwestern Medical School, Dallas, Texas

## Abstract

We have put forth a proposal to improve the representation of immunology in the biological\_process ontology of the Gene Ontology. The aims are to provide a comprehensive representation of immunological processes in the GO, improve the organization of immunological related terms in the GO to match current concepts in the field of immunology, and to revise poorly formulated and poorly defined terms already present in the GO. Whole sets of new terms have been proposed to cover areas such as mucosal immunity, tolerance induction, and B cell differentiation, which were previously absent in the GO. Discussion of the proposal at the GO Content Meeting in November 2005 led to a consensus regarding an improved high-level structure for immune system process terms, including the existing immune response hierarchy, and changes in the relationship of the immune response and inflammatory response terms to the defense response terms. We hope that with these improvements in the representation of immunology in the GO, the GO will better reflect current literature and thought in the area of immunology and thus allow the GO to be a more valuable annotation resource.

## Problems in the Current Representation of Immunology in the GO

### Incomplete Representation of Immunological Processes

Current GO lacks whole groups of terms for processes such as

*mucosal immunity*  
*tolerance induction*  
*B cell differentiation*

And individual terms for important processes such as

*granuloma formation*  
*germinal center formation*  
*affinity maturation*

### Problems in Term Organization, such as

A) Terms describing the development of immune system components are incorrectly placed under GO:0006955 immune response, when in fact many of the differentiation steps occur prior to an immune response. In particular this is true for T cell differentiation ; GO:0030217 and B cell differentiation ; GO:0030183.

B) The GO terms for cytokine metabolism and cytokine production are children of immune response, when in fact many cytokines are produced independently of immune responses and have non-immunological functions (EGF or TGF- $\beta$ , for example).

C) Separate GO terms for antigen processing and antigen presentation exist, without recognition of the relationship of these two concepts. Furthermore, the children of these terms are not well defined and important subprocesses are missing from the GO. This has led to confusion by annotators in the use of these terms, and reflects poorly on the GO.

D) GO:0006955 immune response is a child of defense response in the current GO, yet immune responses include tolerance induction and regulatory processes that are not defense responses.

E) GO:0006954 inflammatory response is a child of immune response, yet inflammatory responses include non-immunological components.

### Poorly Formulated and Poorly Defined Terms, such as

humoral immune response ; GO:0006959  
*An immune response mediated through a body fluid*

--% humoral defense mechanism (sensu Vertebrata) ; GO:0016064  
*The specific immune response mediated by antibodies, as in, but not restricted to, the vertebrates (Vertebrata, ncbi\_taxonomy\_id:7742).*

--% humoral defense mechanism (sensu Protostomia) ; GO:0016065  
*The specific immune response mediated by antibodies. As in, but not restricted to, the taxon Protostomia (Protostomia, ncbi\_taxonomy\_id:33317).*

### Two problems:

There simply are no antibodies in non-vertebrates, so sensu-terms are not needed, and the lexical string "humoral defense mechanism" has no relation to terminology in use the literature.

## History of Project

Through my work as an annotator, I (Alex) identified problems in the representation of immunology in the GO.

We decided to make immunology a focus for term development and were independently contacted by Richard Scheuermann who was interested in the same thing.

Jamie Lee, from Richard's group, and I prepared a proposal for new terms and changes to existing terms in the biological\_process ontology.

At the GO Content Meeting at TIGR in November 2005, our proposal was discussed and largely approved pending incorporation of a small number of changes.

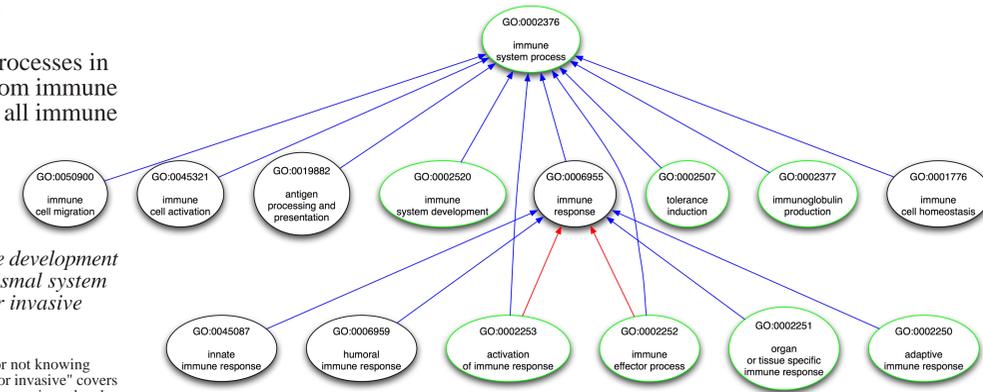
## NewTerm: Immune System Process

This term provides recognition that many processes in the immune system occur prior to or apart from immune responses, and provides a grouping term for all immune system processes.

### Definition:

*An organismal physiological process involved in the development or functioning of the immune system, an organismal system for calibrated responses to potential internal or invasive threats.*

"Calibrated" allows for tolerance and regulation, "potential" allows for not knowing whether something is really a threat or not ahead of time, "internal or invasive" covers tumors and microbes, as well as individual proteins or other immunogenic molecules of self or non-self origin that in particular contexts within the body induce immune responses or tolerance induction, centrally or peripherally.



## Improvements to the GO

Created separate term hierarchies for processes involved in activation of the immune system and effector mechanisms of the immune system.

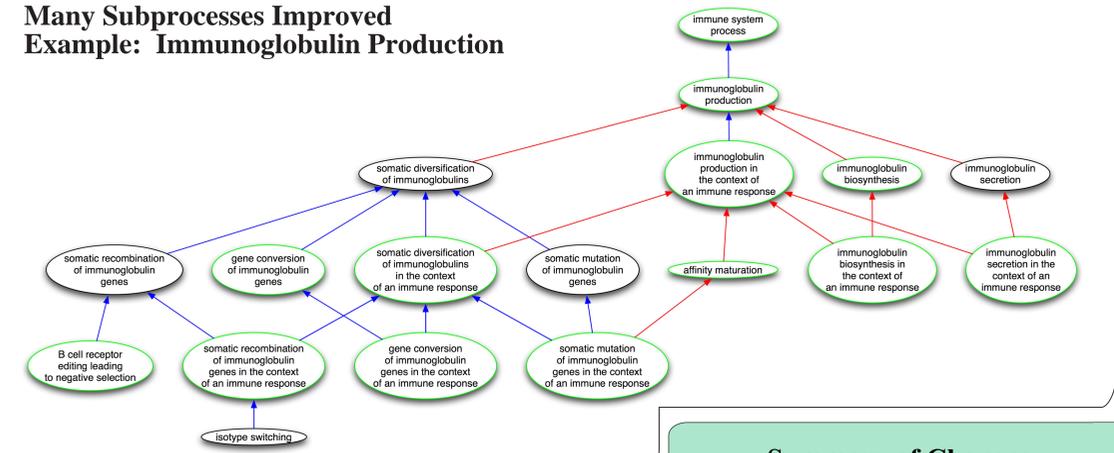
Improved the representation of innate immunity and incorporated plant innate immunity terms.

Improved the representation of immune cell differentiation, particularly for T and B cells.

Improved the inflammatory response hierarchy.

Incorporated long planned changes in the so-called response and detection terms.

## Many Subprocesses Improved Example: Immunoglobulin Production



## Summary of Changes

345 new GO biological\_process terms created for immunological processes.

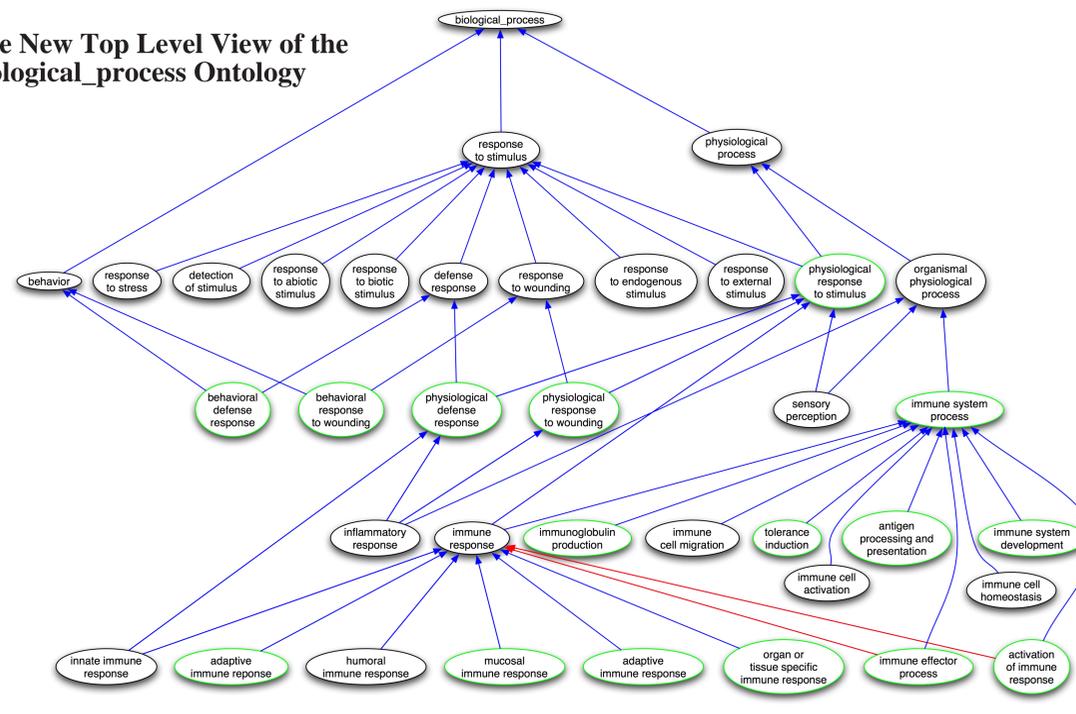
Large scale rearrangements of existing terms performed to correct true path violations.

Regulation term hierarchies also simplified to avoid true path violations.

Structure should allow addition of new terms in a consistent fashion.

Please see the printout of the full term hierarchy and selected subhierarchies on view nearby.

## The New Top Level View of the biological\_process Ontology



## Implementation

Final term hierarchy is in place.

Definitions are still being written for the new terms.

A short review period will be given to allow for additional comments.

Expect to incorporate changes into the official GO by the end of April 2006.

## Thanks to

All the participants in the GO Content Meeting, November 2005 at TIGR, and especially Michelle Gwinn and Linda Hannick.

The GO Editorial Office: Amelia Ireland, Jane Lomax, Jen Clark, and Midori Harris

The members of PAMGO, particularly Candace Collmer, and also Suparna Mundodi and the Plant Pathogen Experts for help with development of the response and detection hierarchies.

The GO Consortium.