

# Editing the Gene Ontology

Midori A. Harris  
GO Editorial Office  
EBI, Hinxton, UK



# GO Editorial Office

People:

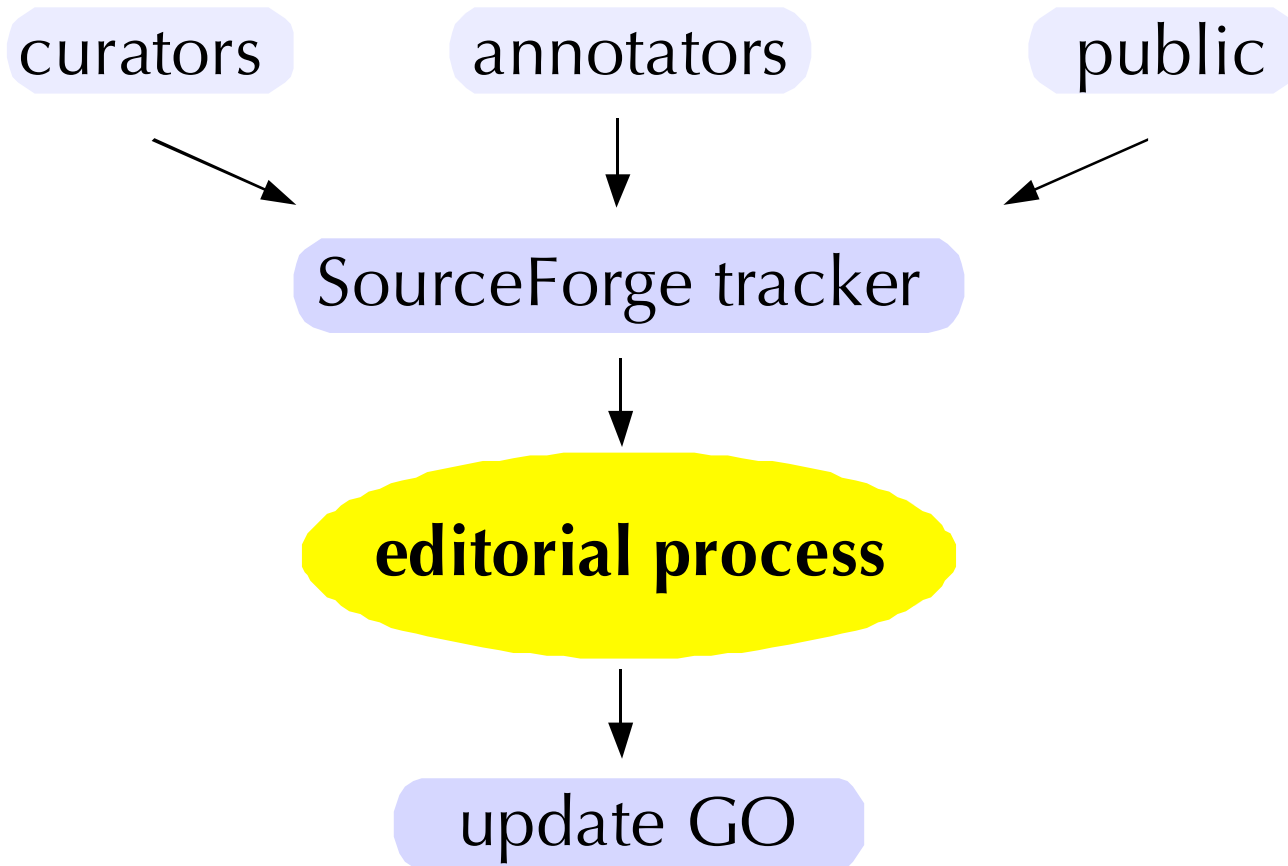
Midori Harris, Jane Lomax,  
Amelia Ireland, Jennifer Clark

Activities:

- GO content: Coordinate all changes
- Maintain web pages, including documentation
- Conferences & workshops



# GO Editorial Procedure Overview





# SourceForge Tracker

- Item for every proposed change
- Submissions from curators, annotators, others
- Each item assigned to a curator
- Submitters and others can comment on items
- Records discussions and progress
- Archive of all entries

[https://sourceforge.net/tracker/?func=add&group\\_id=36855&atid=440764](https://sourceforge.net/tracker/?func=add&group_id=36855&atid=440764)



# SourceForge Tracker

The screenshot shows the SourceForge Tracker interface for the Gene Ontology project. The browser address bar displays <https://sourceforge.net/tracker/index.php>. The page title is "Project: Gene Ontology: Browse Curator requests".

Navigation links include: [Summary](#), [Admin](#), [Home Page](#), [Forums](#), [Tracker](#), [Bugs](#), [Support](#), [Patches](#), [RFE](#), [Lists](#), [Tasks](#), [Docs](#), [News](#), [CVS](#), [Files](#).

Filtering options:

- Assignee: Any
- Status: Open
- Category: Any
- Group: Any
- Sort By: Submitter, Descending

Table of curator requests:

Request ID	Summary	Date	Assigned To	Submitted By
<input type="checkbox"/> 852977	additional germ layer specification terms	2003-12-03 00:04	nobody	vanaukenk
<input type="checkbox"/> 854677	dolichyl monophosphate biosynthesis	2003-12-05 12:00	nobody	val_wood
<input type="checkbox"/> 854029	Golgi vesicle budding/ Golgi vesicle targeting	2003-12-04 13:38	nobody	val_wood
<input type="checkbox"/> 837124	3' RNA processing	2003-11-06 11:27	kchris	val_wood
<input type="checkbox"/> 848923	cohesin compnew cohesin complexes	2003-11-25 13:10	gamidori	val_wood
		2003-12-03		

[https://sourceforge.net/tracker/?func=add&group\\_id=36855&atid=440764](https://sourceforge.net/tracker/?func=add&group_id=36855&atid=440764)



# SourceForge Request

The screenshot shows a web browser window displaying a SourceForge request page. The browser's address bar shows the URL: [https://sourceforge.net/tracker/index.php?func=detail&aid=668448&group\\_id=...](https://sourceforge.net/tracker/index.php?func=detail&aid=668448&group_id=...). The page title is "SourceForge.net: Modify 668448 - mannosome".

**Navigation:** Submit New | Browse | Reporting | Admin

## [ 668448 ] mannosome

**Monitor** (?)

**Submitted By:** Marion Grenee (brendamg)

**Date Submitted:** 2003-01-15 12:58

**Date Closed:** 2003-01-20 13:18

**Date Last Updated:** 2003-01-20 13:18

**Number of Attachments:** 0

**Submit Changes**

**Last Updated By:** j1242 - Comment added

**Number of Comments:** 1

**Data Type:** (?)  
Curation request

**Category:** (?)  
New term request (admin)

**Assigned To:** (?)  
j1242 (admin)

**Status:** (?)  
Closed

**Group:** (?)  
GO (admin)

**Priority:** (?)  
5 - Medium

**Resolution:** (?)  
Accepted

**Summary:** (?)  
mannosome

Could we have the term 'mannosome' for the localization of the enzyme 1.1.3.40?

<pmannosome ; GO:0005777  
<mannosome ; GO:000nnnn

<http://www.chem.qmul.ac.uk/lubmb/enzyme/EC1/1/3/40.html>  
"The enzyme from the snail *Helix aspersa* and *Arion ater* is found in a specialised tubular organelle that has been termed the mannosome."

<http://asp.wlv.ac.uk/Level7.asp?UserType=6&Level7=40>  
"The Group has also demonstrated the existence in Gastropods of a novel membranous subcellular organelle, the mannosome of *Strophodonta*...

**Left Sidebar:**

- sf.net Resources
  - Site Docs
  - Site Status
  - Site Map
  - SF.net Supporters
  - Compile Farm
  - Foundries
  - Project Help Wanted
  - New Releases
  - Get Support
- Site Sponsors
- Most Active
  - 1 Gaim
  - 2 phpMyAdmin
  - 3 gulvekl
  - 4 Azureus - BitTorrent Client
  - 5 phpCedView
  - 6 fma
  - 7 Compiere ERP + CRM Business Solution
  - 8 FileZilla
  - 9 WinMerge
  - 10 moregroupware
- More Activity >>
- Top Downloads
  - 1 eMule
  - 2 BitTorrent
  - 3 DC++
  - 4 VirtualDub
  - 5 Azureus - BitTorrent



# GO Editorial Procedure

- Claim SourceForge item
- Consider:
  - does the term belong in GO?
  - term name & definition
    - standard wording?
  - relationships to other terms
- Consult literature, interest group, other curators & researchers as needed



# GO Editorial Results 1

Easy items:

- Unambiguous request
- Clear term name(s) and meaning(s)
- Clear relationships to other terms

Curators make the change(s) at once

Example: cobalt ion binding

(SourceForge entry [896544](#))





# Cobalt ion binding

- ⊕ ⓘ GO:0003674 : molecular\_function ( 97507 ) 🌐
  - ⊕ ⓘ GO:0005488 : binding ( 26483 ) 🌐
    - ⊕ ⓘ GO:0046872 : metal ion binding ( 2138 ) 🌐
      - ☰ ⓘ **GO:0046914 : transition metal ion binding ( 1299 )** 🌐
        - ⓘ GO:0046870 : cadmium ion binding ( 6 )
        - · ⓘ **GO:0050897 : cobalt ion binding ( 1 )**
        - ⓘ GO:0005507 : copper ion binding ( 204 )
        - ⊕ ⓘ GO:0005506 : iron ion binding ( 104 )
        - ⓘ GO:0030145 : manganese ion binding ( 23 )
        - ⓘ GO:0045340 : mercury ion binding ( 1 )
        - ⓘ GO:0030151 : molybdenum ion binding ( 10 )
        - ⓘ GO:0016151 : nickel ion binding ( 6 )
        - ⓘ GO:0008270 : zinc ion binding ( 966 )

AmiGO tree view



# GO Editorial Results 2

Moderately difficult items:

- Some aspect of request requires clarification
- Different proposals to be resolved

Discussion via SourceForge comments and email before curators make change(s)

Example: fore-, mid-, hindbrain development  
(SourceForge entry [854736](#))



# Brain development

- Original request: add forebrain development
- Additional requests: midbrain development  
hindbrain development
- Question: How complicated must the DAG structure be? Do we need *sensu* terms?

## Existing DAG:

```
(i) organogenesis ; GO:0009887
  (i) neurogenesis ; GO:0007399
    (p) central nervous system development ; GO:0007417
      (p) brain development ; GO:0007420
        (p) central complex development ; GO:0048036
          (p) mushroom body development ; GO:0016319
```



# Brain development

Simple structure option (no *sensu* terms)

```
(i) organogenesis ; GO:0009887
  (i) neurogenesis ; GO:0007399
    (p) central nervous system development ; GO:0007417
      (p) brain development ; GO:0007420
        (p) forebrain development ; GO:new
        (p) midbrain development ; GO:new
        (p) hindbrain development ; GO:new
        (p) central complex development ; GO:0048036
        (p) mushroom body development ; GO:0016319
```



# Brain development

More complex structure, with  
'*sensu* Chordata' and '*sensu* Insecta'

```
...
(i)neurogenesis
  (p)central nervous system development ; GO:0007417
    (p)brain development ; GO:0007420
      (i)brain development (sensu Chordata); GO:new
        (p)forebrain development ; GO:new
        (p)midbrain development ; GO:new
        (p)hindbrain development ; GO:new
      (i)brain development (sensu Insecta) ; GO:new
        (p)central complex development ; GO:0048036
        (p)mushroom body development ; GO:0016319
```



# DAG-Edit

tree view

DAG view

editing

The screenshot displays the DAG-Edit version 1.311 interface. On the left is a tree view of biological terms, with 'nuclear membrane' selected and highlighted in red. The main area shows the details for the selected term, including its ID (CO:0005435), name, definition, synonyms, and general Dlxrefs. A search bar at the top right contains the text 'nuclear membrane'. A yellow oval labeled 'tree view' points to the left-hand list, another yellow oval labeled 'DAG view' points to the hierarchical structure on the right, and a third yellow oval labeled 'editing' points to the 'Add' and 'Delete' buttons for synonyms and Dlxrefs.



# Brain development

Present solution: simpler structure

- GO:0008150 : biological\_process ( 96312 )
- └─ GO:0007275 : development ( 14496 )
- └─ GO:0009653 : morphogenesis ( 7513 )
- └─ GO:0009887 : organogenesis ( 5601 )
- └─ GO:0007399 : neurogenesis ( 2206 )
- └─ GO:0007417 : central nervous system development ( 676 )
- └─ **GO:0007420 : brain development ( 323 )**
- ├─ GO:0048036 : central complex development ( 4 )
- ├─ **GO:0030900 : forebrain development ( 10 )**
- ├─ GO:0030902 : hindbrain development ( 4 )
- ├─ GO:0030901 : midbrain development ( 5 )
- ├─ GO:0030917 : midbrain-hindbrain boundary development ( 1 )
- └─ GO:0016319 : mushroom body development ( 100 )



AmiGO tree view



# GO Editorial Results 3

Very difficult items:

- Challenging biology to model
- Extensive change in ontology structure or interpretation

Protracted discussion via SourceForge comments and email; resolve at face-to-face meeting

Example: cell killing & pathogenesis

(SourceForge entry [900600](#) & >30 emails)





# Pathogenesis & cell killing

Discussion if time permits ...



# SourceForge Statistics

May 5, 2004

- 143 items open
- 1715 total

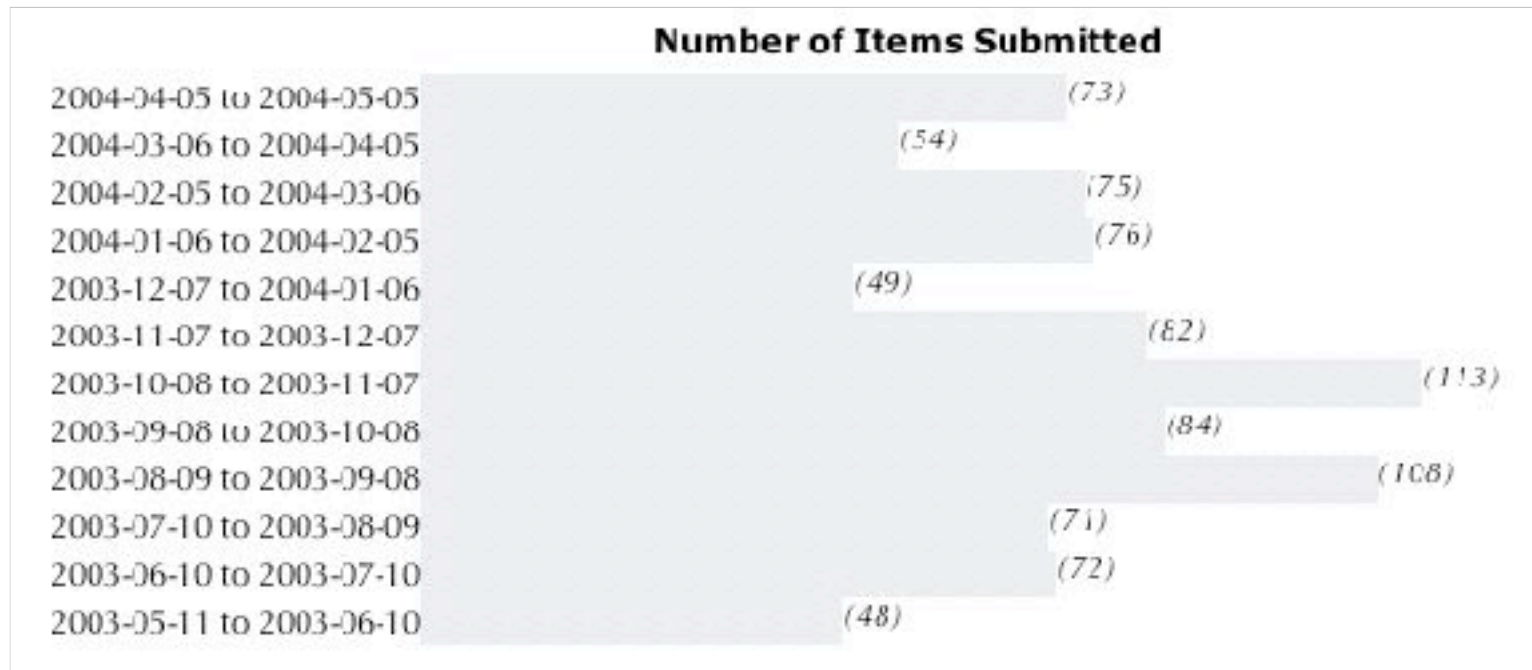
Reporting available online:

[https://sourceforge.net/tracker/reporting/?group\\_id=36855&atid=440764](https://sourceforge.net/tracker/reporting/?group_id=36855&atid=440764)



# SourceForge Statistics

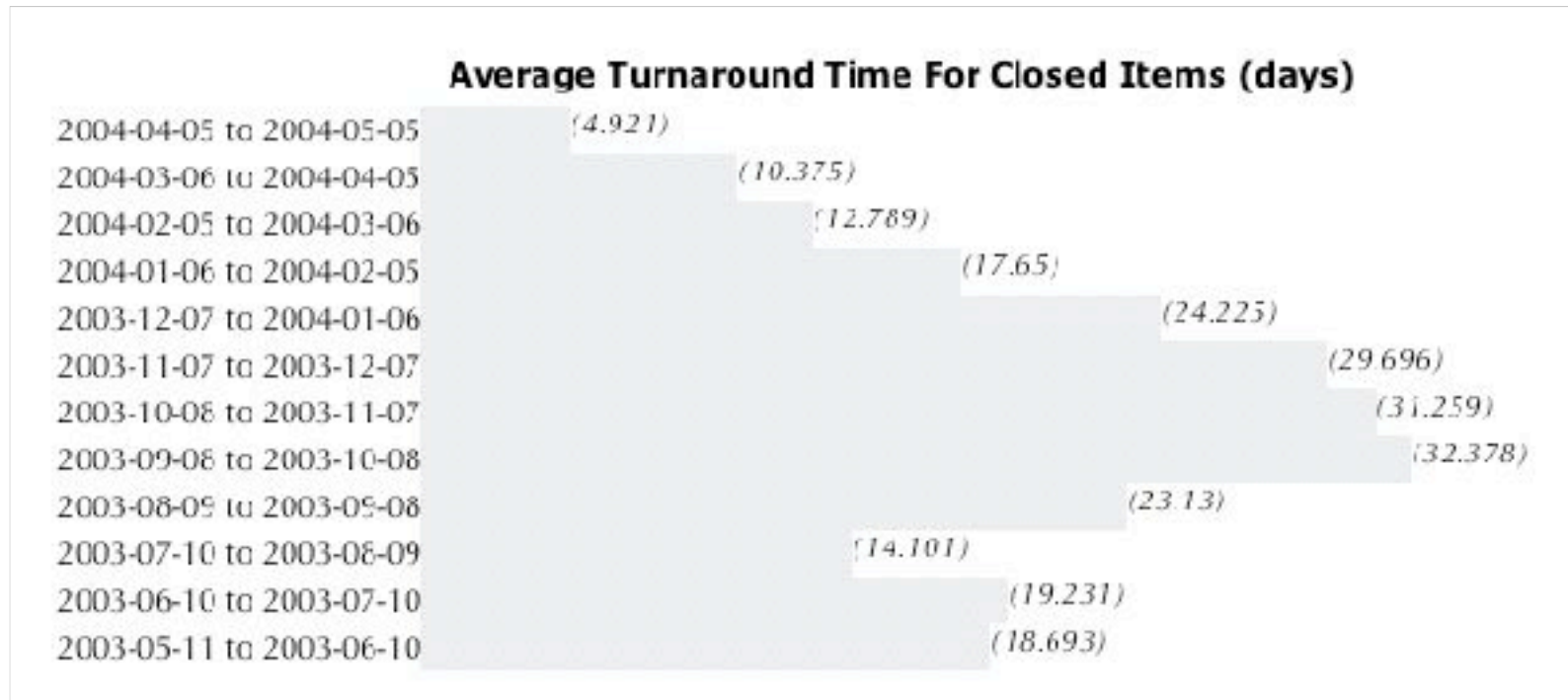
Aging report: items submitted per month





# SourceForge Statistics

Aging report: turnaround time





# SourceForge Statistics

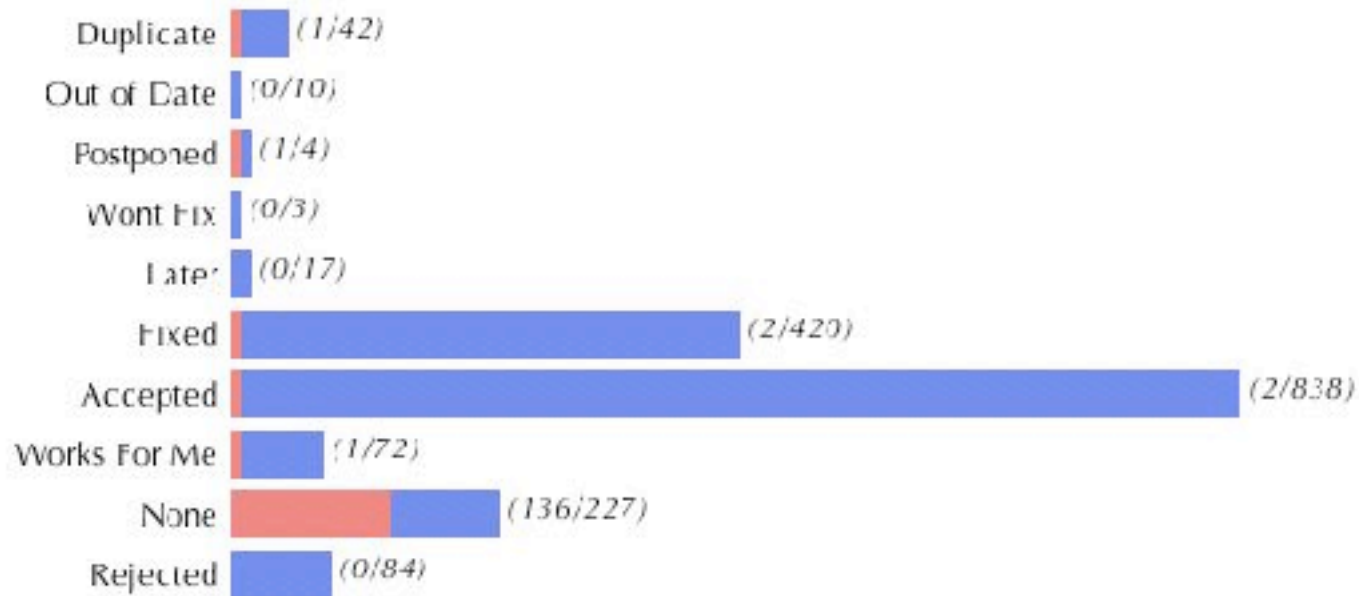
Aging report: open items



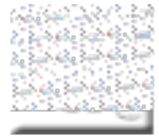


# SourceForge Statistics

## Distribution By Resolution



Key: ( Open / All )



# GO Editorial Procedure

- Claim SourceForge item
- Consider:
  - does the term belong in GO?
  - term name & definition
  - relationships to other terms
- Consult literature, interest group, other curators & researchers as needed
- More details in curator documentation

---

<http://www.geneontology.org/GO.contents.curator.guides.html>



# Documentation for Curators

Guidelines for:

- using SourceForge
- making changes in the ontologies
- molecular function terms  
(biological process and cellular component guidelines planned)





# Documentation for Curators

GO Curator Guides

http://www.geneontology.org/GO.contents.curator.guides.html

## A Guide to Modifying the Gene Ontologies

These guides explain how to add to or alter the gene ontologies. They are intended for those members of the consortium whose job it is to make such modifications to the GO. However we make them available here for anyone who is interested to know how this work is carried out.

<a href="#">The Beginner's Guide to Modifying the Ontologies</a>	A guide for new GO curators
<a href="#">Submitting suggestions to GO using SourceForge.net</a>	How to suggest possible changes to the ontologies
<a href="#">Guide to Addressing a SourceForge Request</a>	How to make changes to the ontologies

**Downloads**

- .. [Current Ontologies](#)
- .. [Current Annotations](#)
- .. [GO Database](#)

**Documentation**

- .. [FAQ](#)
- .. [Introduction](#)
- .. [Editorial Style Guide](#)
- .. [File Format Guide](#)
- .. [Function Ontology](#)
- .. [Synonym Guide](#)
- .. [Annotation Guide](#)
- .. [Evidence Codes](#)
- .. [SourceForge links](#)
- .. [Modifying GO](#)
  - ... [For Beginners](#)
  - ... [Suggest changes](#)

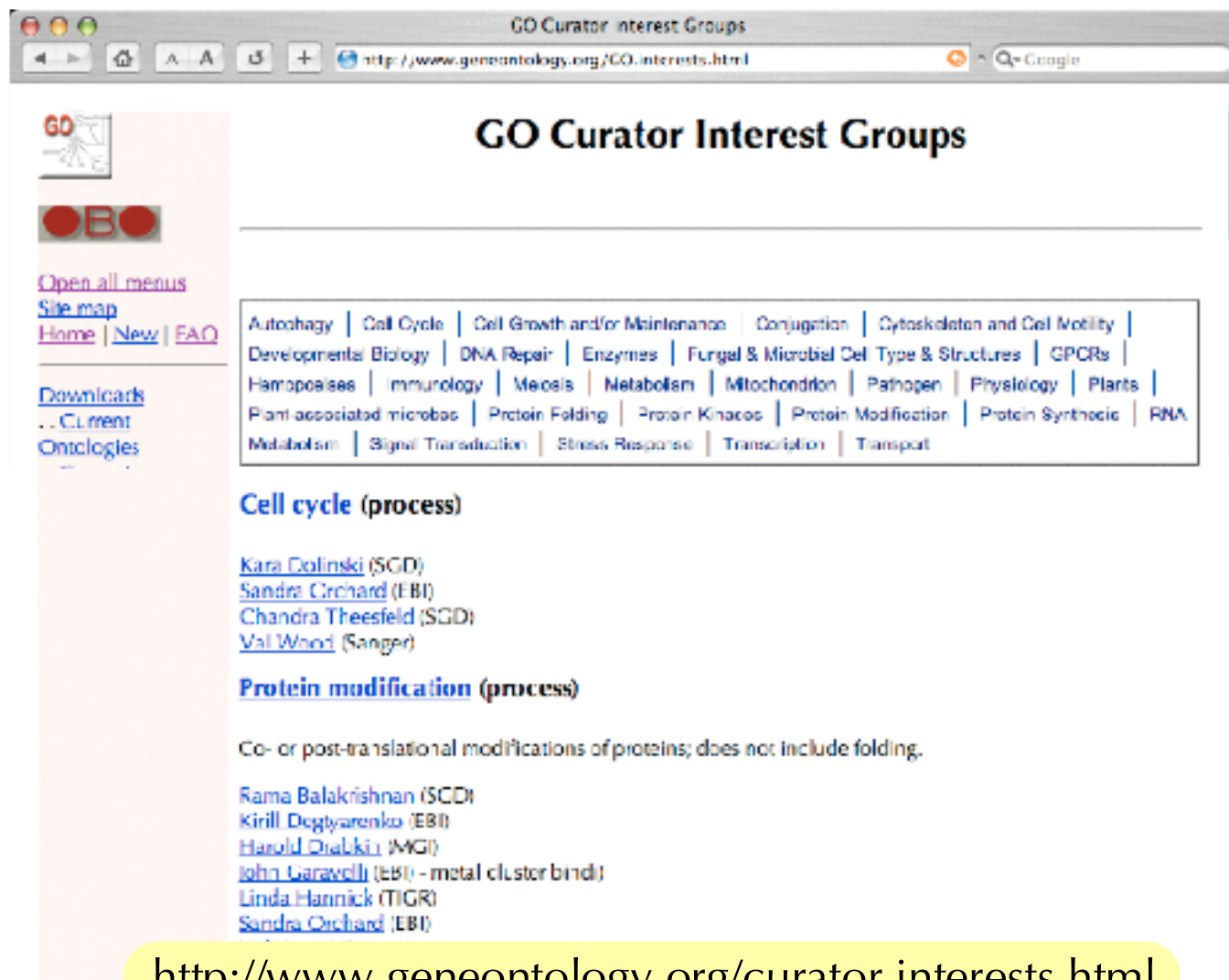


# GO Curator Interest Groups

- Cover specific topics within GO
- Include GO curators and outside experts
- Group members develop portions of ontology, then report to Consortium
- Sample topics:
  - developmental biology *e.g. limbs & fins*
  - plant biology *e.g. pollination, flower development*
  - protein modification
  - transcription



# GO Curator Interest Groups



The screenshot shows a web browser window with the title "GO Curator Interest Groups" and the URL "http://www.geneontology.org/GO.interests.html". The page content includes a navigation menu on the left with links for "Open all menus", "Site map", "Home", "New", and "FAQ". Below the menu are sections for "Downloads" and "Ontologies". The main content area features a grid of interest group categories: Autophagy, Cell Cycle, Cell Growth and/or Maintenance, Conjugation, Cytoskeleton and Cell Motility, Developmental Biology, DNA Repair, Enzymes, Fungal & Microbial Cell Type & Structures, GPCRs, Hemopoiesis, Immunology, Meiosis, Metabolism, Mitochondrion, Pathogen, Physiology, Plants, Plant-associated microbes, Protein Folding, Protein Kinases, Protein Modification, Protein Synthetic, RNA Metabolism, Signal Transduction, Stress Response, Transcription, and Transport. Two specific interest groups are highlighted: "Cell cycle (process)" with members Kara Dolinski (SGD), Sandra Orchard (EBI), Chandra Theesfield (SGD), and Val Wood (Sanger); and "Protein modification (process)" with members Rama Balakrishnan (SCD), Kirill Degtyarenko (EBI), Harold Oubkhi (MGI), John Garavito (EBI) - metal cluster bind, Linda Hannick (TIGR), and Sandra Orchard (EBI).

<http://www.geneontology.org/curator.interests.html>



# Current Priorities

Ontology topics:

- cellular vs. organismal processes
- developmental processes
  - cell fate specification & determination
  - plants: flower development
- cell cycle

Other activities:

- correct part\_of relationships
- guidelines for process & component
- GO slims