



**GENEONTOLOGY**  
Unifying Biology

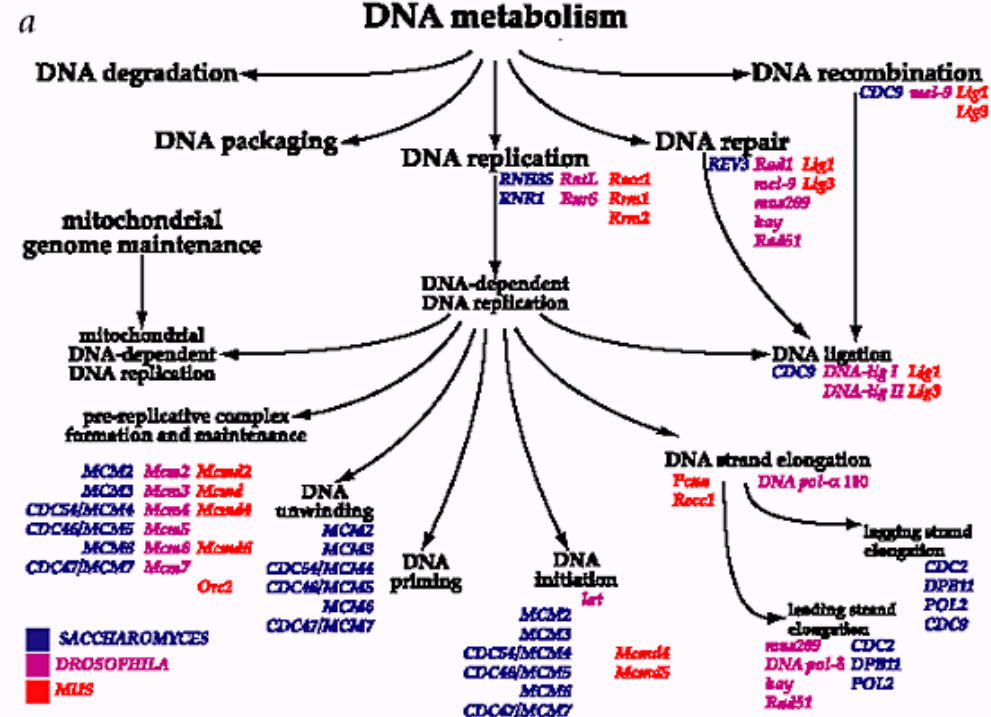
# **Gene Ontology – анализ обогащения**

Борисенко Евгений Александрович

# Что такое GO?

Онтология генов (GO) - это контролируемый словарь, связывающий каждый ген с одной или несколькими функциями. Как оказалось, название «онтология генов» вводит в заблуждение. Онтология предназначена для классификации продуктов генов, а не самих генов.

Онтология GO структурирована как ориентированный ациклический граф, где каждый термин имеет определенные связи с одним или несколькими другими терминами в той же области, а иногда и с другими областями. Словарь GO разработан таким образом, чтобы быть независимым от вида и включает термины, применимые к прокариотам и эукариотам, а также к одноклеточным и многоклеточным организмам.



Commentary | Published: May 2000

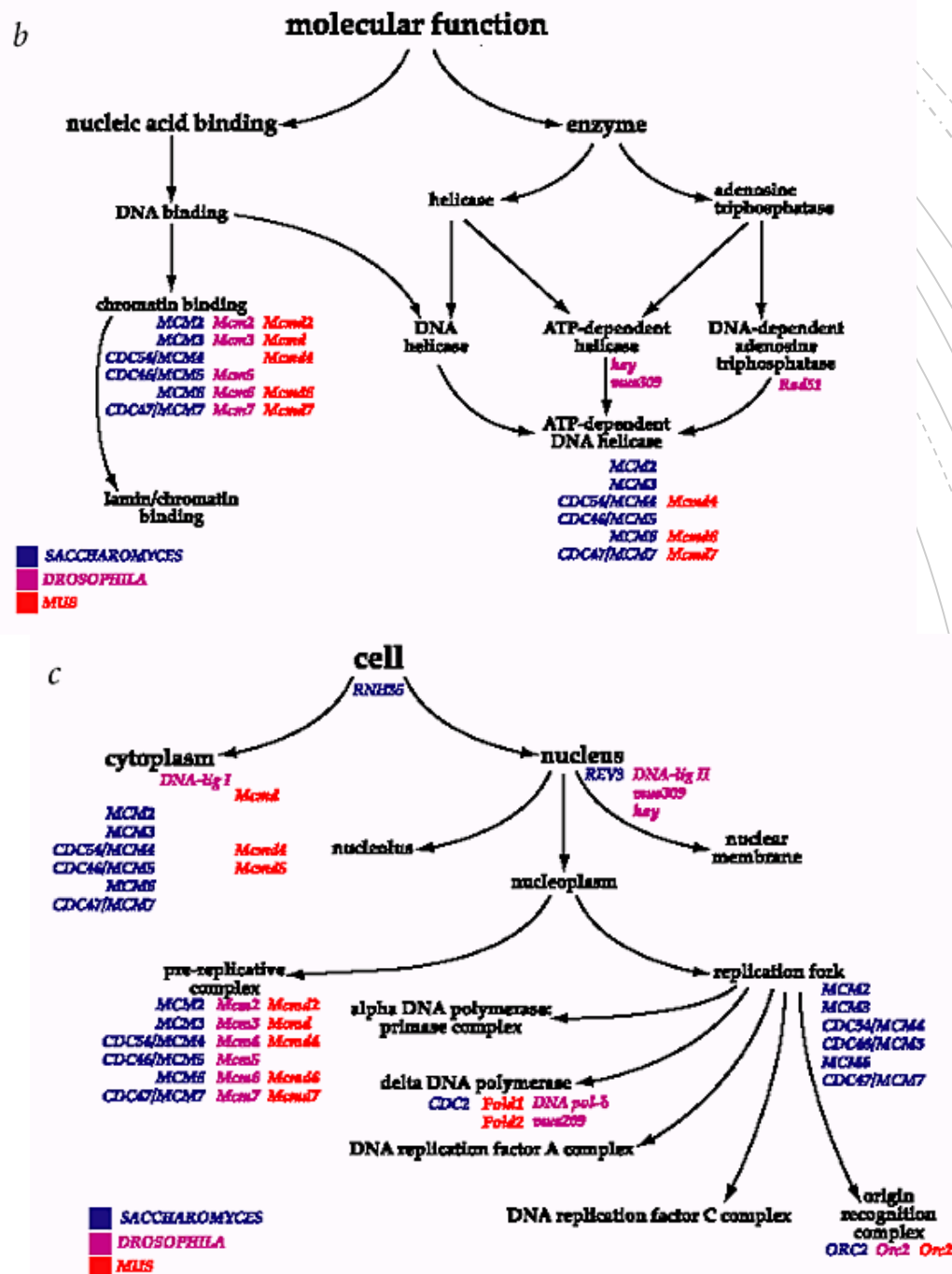
## Gene Ontology: tool for the unification of biology

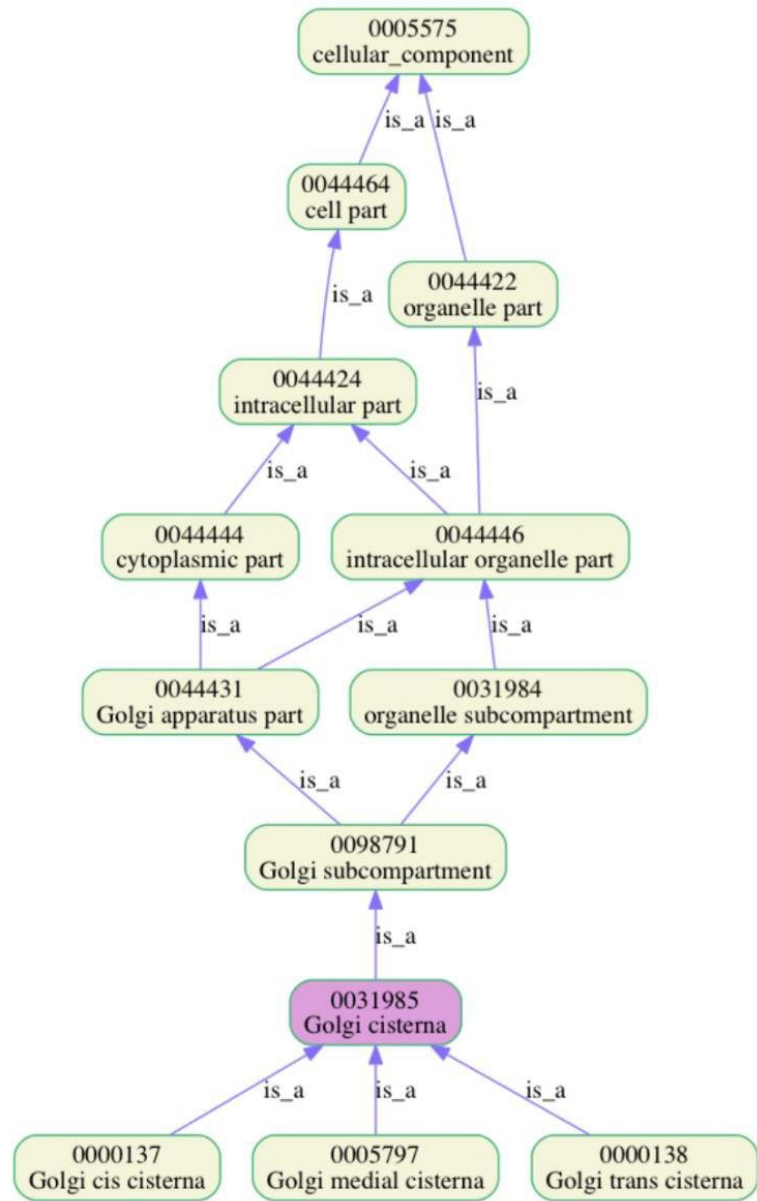
Michael Ashburner, Catherine A. Ball, Judith A. Blake, David Botstein, Heather Butler, J. Michael Cherry, Allan P. Davis, Kara Dolinski, Selina S. Dwight, Janan T. Eppig, Midori A. Harris, David P. Hill, Laurie Issel-Tarver, Andrew Kasarskis, Suzanna Lewis, John C. Matese, Joel E. Richardson, Martin Ringwald, Gerald M. Rubin & Gavin Sherlock

Nature Genetics 25, 25–29 (2000) | Cite this article

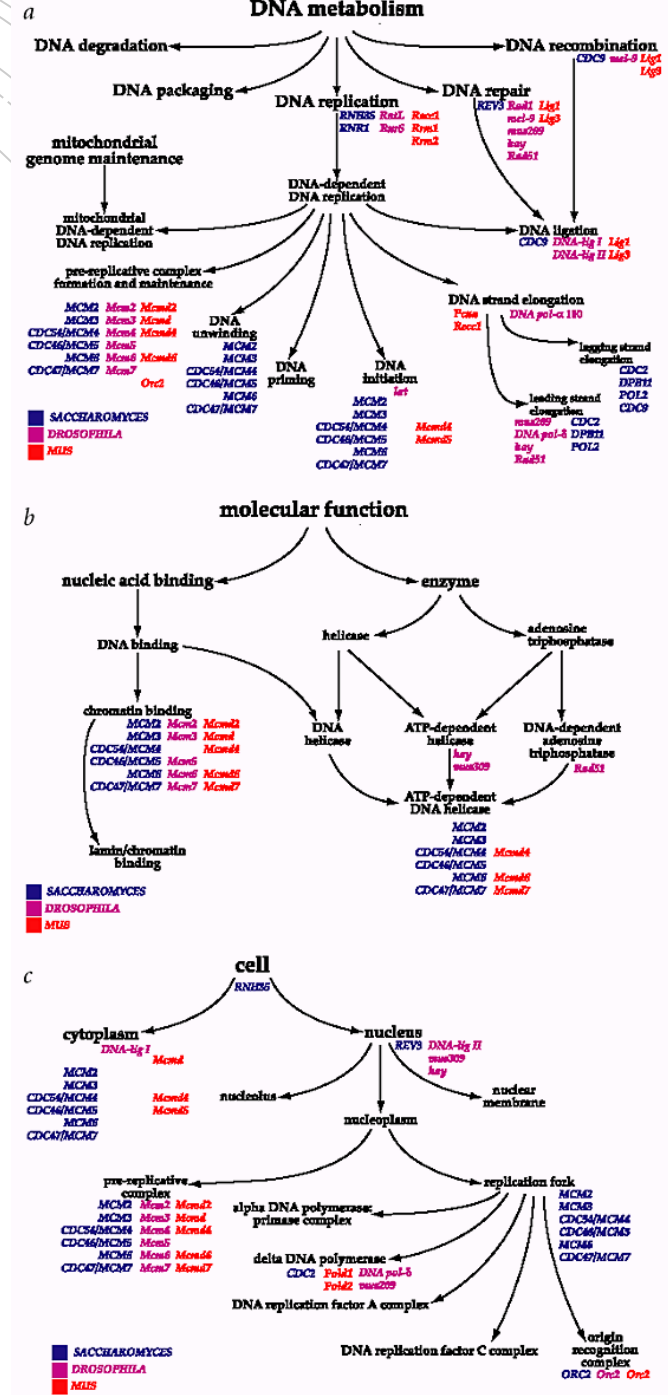
67k Accesses | 39k Citations | 118 Altmetric | Metrics

[https://www.nature.com/articles/ng0500\\_25](https://www.nature.com/articles/ng0500_25)

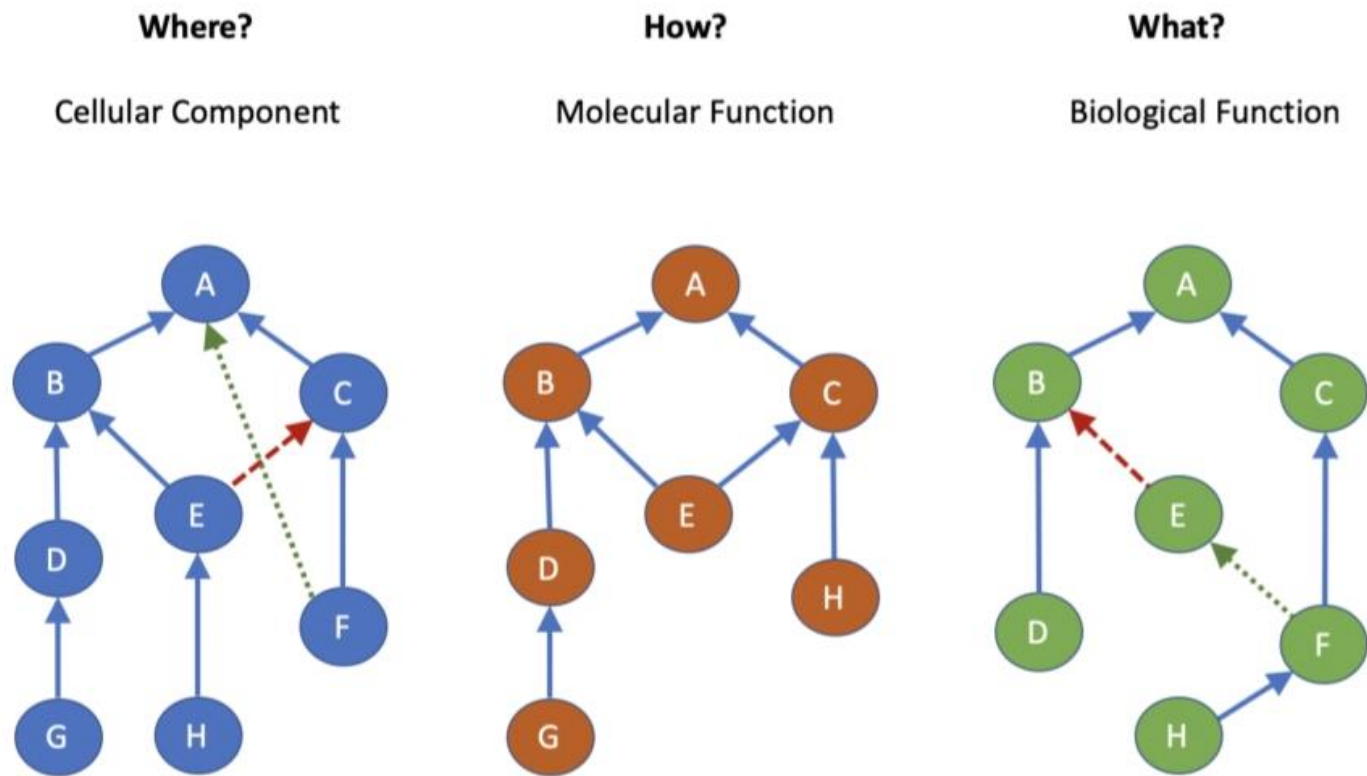




|      |        |
|------|--------|
| 2021 | 277681 |
| 2020 | 42252  |
| 2019 | 17857  |
| 2018 | 20934  |
| 2017 | 34328  |
| 2016 | 21792  |
| 2015 | 17677  |
| 2014 | 22617  |
| 2013 | 57967  |
| 2012 | 13701  |
| 2011 | 20021  |
| 2010 | 12234  |
| 2009 | 11904  |
| 2008 | 8298   |
| 2007 | 4470   |
| 2006 | 7192   |
| 2005 | 4060   |
| 2004 | 3017   |
| 2003 | 8467   |
| 2002 | 977    |
| 2001 | 1748   |
| 2000 | 5      |



Каждый из приведенных выше подграфов независим от других. При выполнении функциональной аннотации нам необходимо обрабатывать каждый подграф отдельно. В некотором смысле мы выполняем функциональную аннотацию трижды, по одному разу для каждой категории. Один и тот же генный продукт может иметь меньше или больше аннотаций в разных категориях.



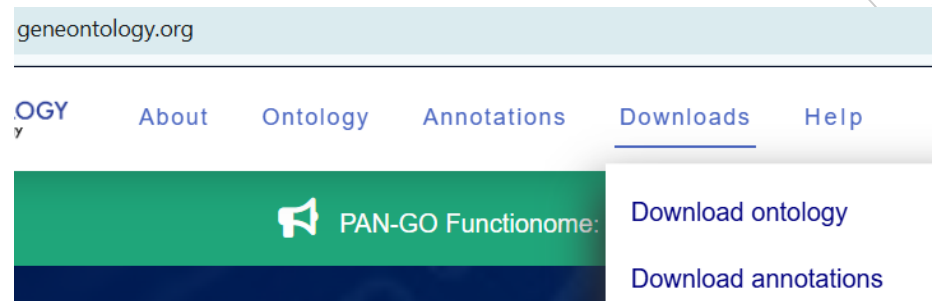
*Общее правило, помогающее различать биологический процесс и молекулярную функцию, заключается в том, что процесс должен иметь более одного определенного этапа.*

# Общие правила:

- ❖ Каждый генный продукт может иметь записи в каждой из различных категорий
- ❖ Каждый генный продукт может иметь несколько записей в каждой из категорий
- ❖ Листовые узлы (узлы без дочерних элементов) являются наиболее точным представлением знаний

Данные GO доступны для каждого организма на странице загрузки GO:

<https://geneontology.org/>



Базовые данные для онтологии генов состоят из двух файлов:

1. Файл определения онтологии генов.
2. Файл ассоциаций генов.

Эти два фрагмента информации связывают термины онтологии с названиями продуктов генов. Файлы GO удивительно малы.

## 1. Файл определения онтологии генов.

Эту информацию можно рассматривать как биологические знания, сведенные к простым и кратким определениям, закодированные в доступный для вычислительных целей формат, например:

А) «*id: GO:0000002 name: mitochondrial genome maintenance namespace: biological\_process def: "The maintenance of the structure and integrity of the mitochondrial genome; includes replication and segregation of the mitochondrial chromosome."*».

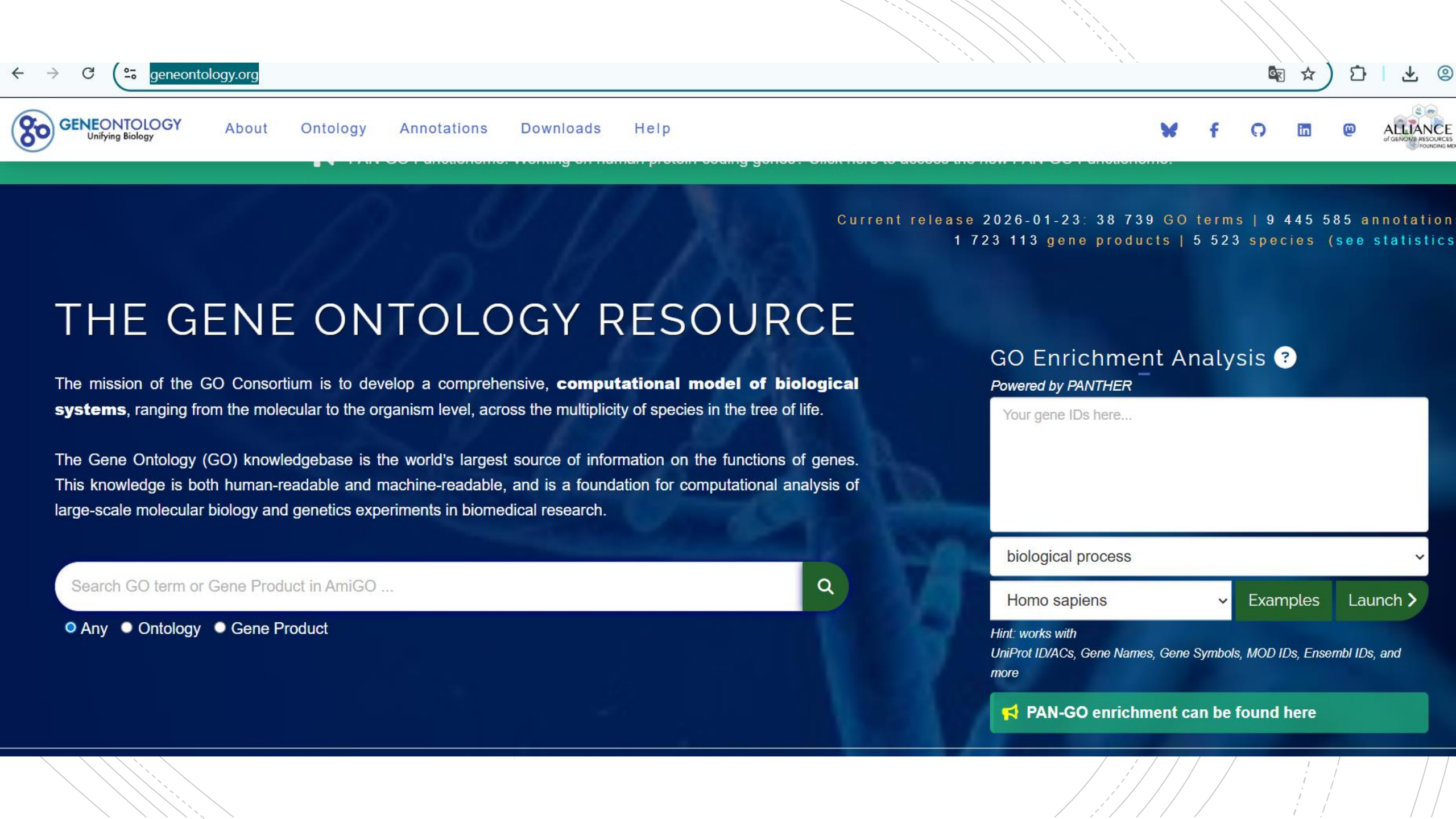
Б) «*id: GO:0005886 name: plasma membrane namespace: cellular\_component alt\_id: GO:0005904 def: "The membrane surrounding a cell that separates the ..."*»

## 2. Файл ассоциаций генов.

Эти два фрагмента информации связывают термины онтологии с названиями продуктов генов:

«*IGLV3-9 ... GO:0005886*»

Файлы GO удивительно малы.



📄 PAN-GE 1 annotation: Working on human protein-coding genes. Click here to access the new PAN-GE 1 annotation.

Current release 2026-01-23: 38 739 GO terms | 9 445 585 annotation  
1 723 113 gene products | 5 523 species (see statistics)

# THE GENE ONTOLOGY RESOURCE

The mission of the GO Consortium is to develop a comprehensive, **computational model of biological systems**, ranging from the molecular to the organism level, across the multiplicity of species in the tree of life.

The Gene Ontology (GO) knowledgebase is the world's largest source of information on the functions of genes. This knowledge is both human-readable and machine-readable, and is a foundation for computational analysis of large-scale molecular biology and genetics experiments in biomedical research.



☒ Any ☐ Ontology ☐ Gene Product

## GO Enrichment Analysis ?

Powered by PANTHER

Hint: works with  
UniProt ID/ACs, Gene Names, Gene Symbols, MOD IDs, Ensembl IDs, and more

📌 PAN-GO enrichment can be found here