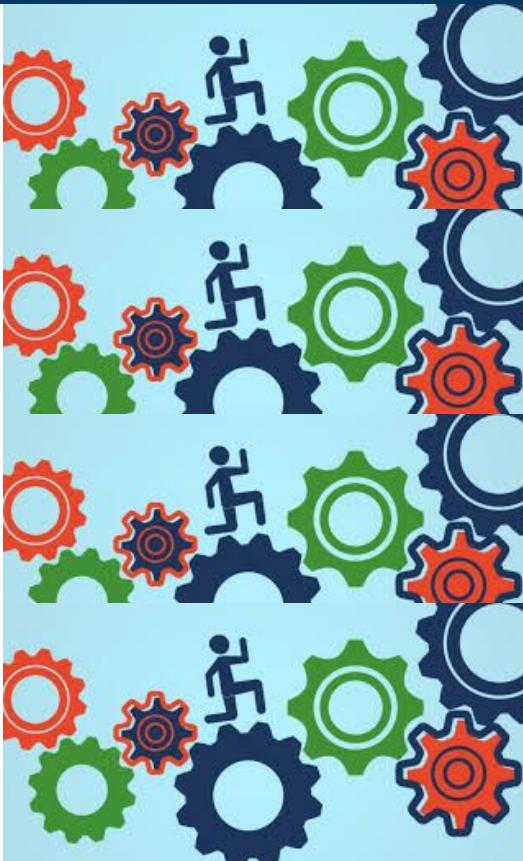




Extending the Gene Ontology for Biological Network Modeling

Pascale Gaudet
SIB Swiss Institute of Bioinformatics
GO Central

Gene Ontology Knowledgebase: Contents and optimization



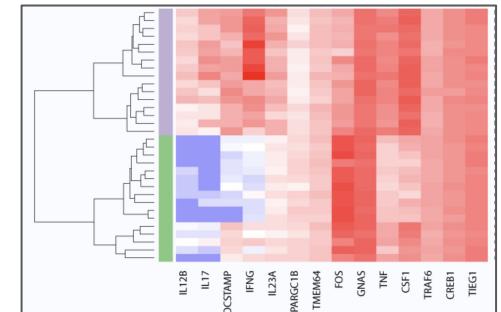
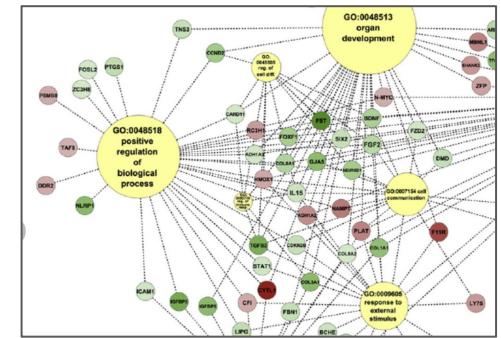
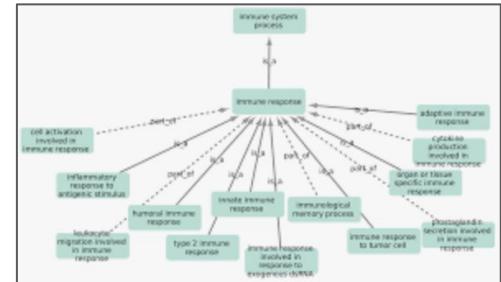
The Gene Ontology (GO)

Provides a computational model of **gene function** at every level:

- molecular activities
- signaling and metabolic pathways
- cellular organism-level systems

Consists of:

- 1) **ontology**: a framework to describe the biological concepts
- 2) **annotations**: associations between genes and GO terms



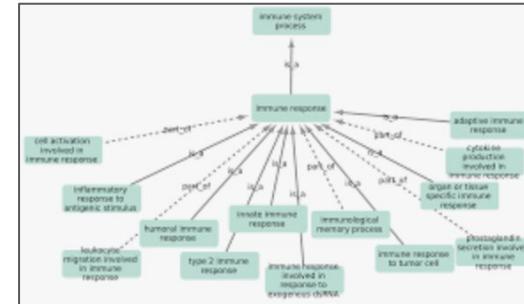
Current GO knowledgebase content

January 2026 release

Ontology

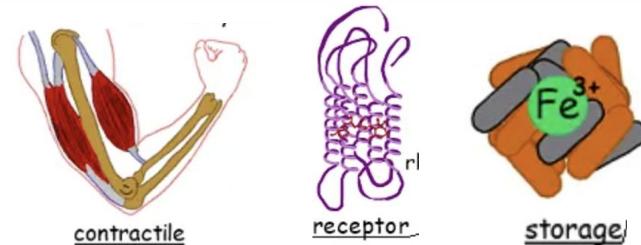
38,739 terms

- 24,547 BP
- 10,123 MF
- 4,069 CC



Annotations

- > 1 M annotations based on experimental data
- Millions of annotations by phylogenetic and electronic methods (800M in QuickGO/UniProt)



The 3 aspects of GO

Molecular Function (MF)

- ***Biochemical activity*** that a gene product performs
- Examples: enzymatic activity, adaptor activity, transcription factor activity

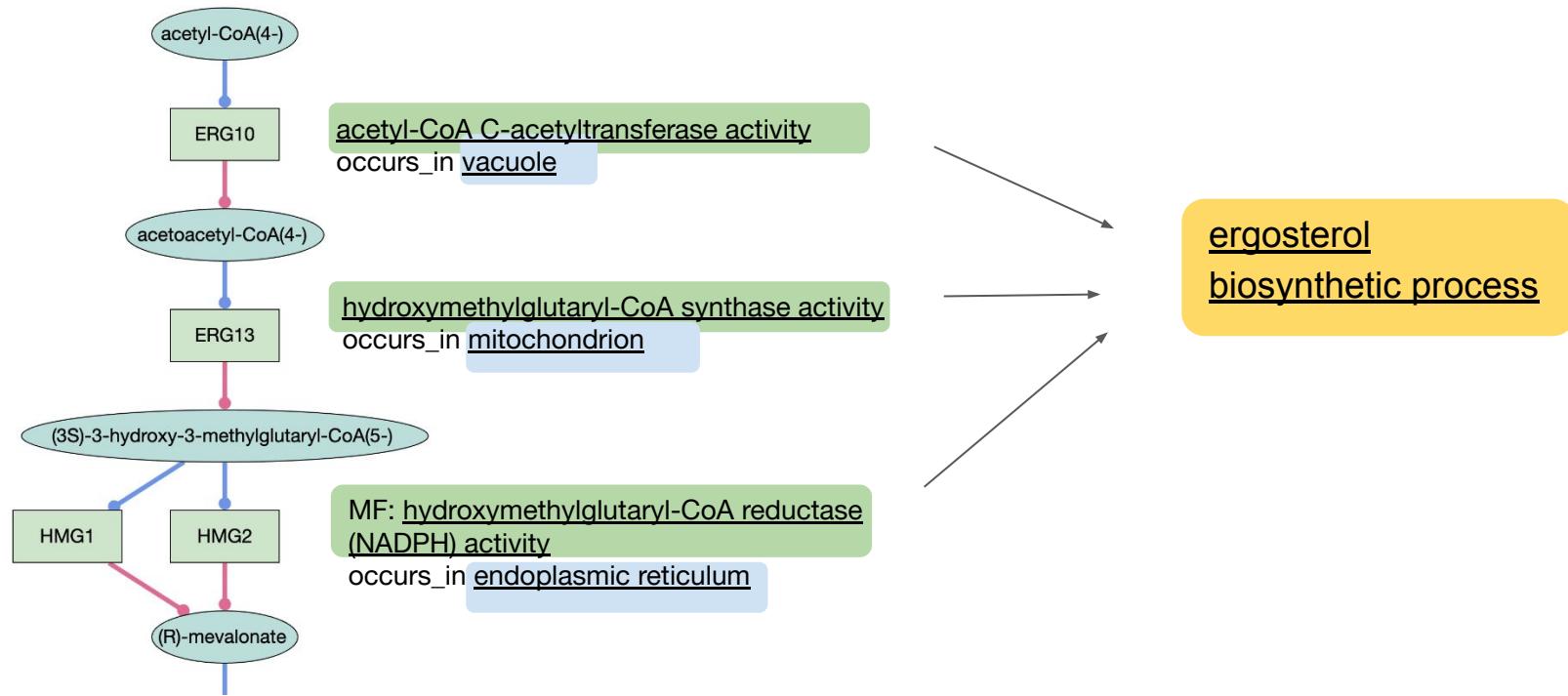
Cellular Component (CC)

- Cellular location where a gene product performs its activity (MF)
- Example: nucleus, cytoskeleton, ribosome

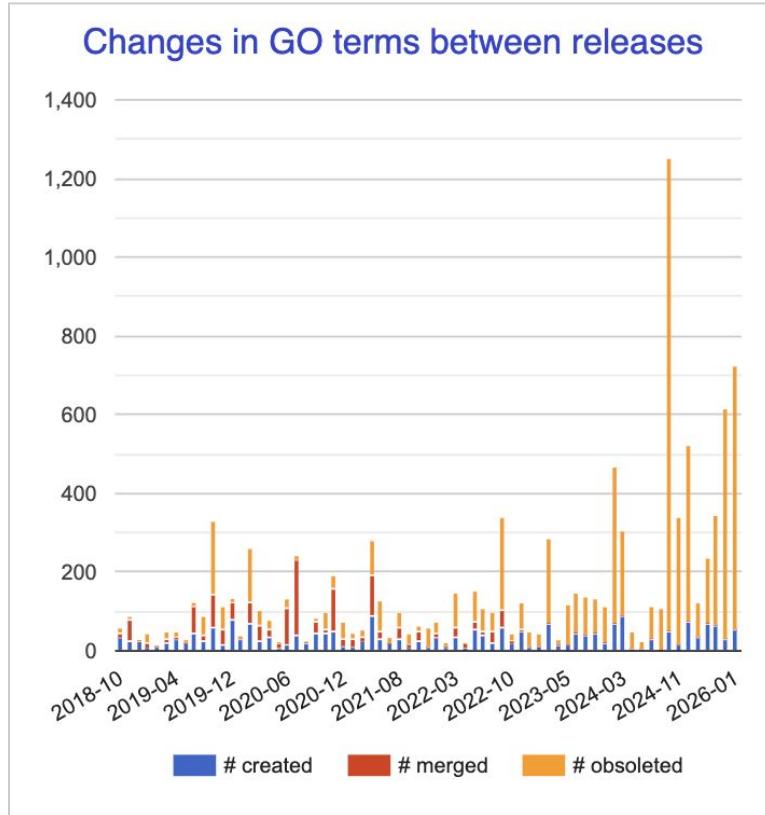
Biological Process (BP)

- Series of events making up a ***biological module or program***, accomplished by a specific set of molecular functions, in a specific order
- Examples: glycolysis, transcription, photosynthesis

Example: function, component and process



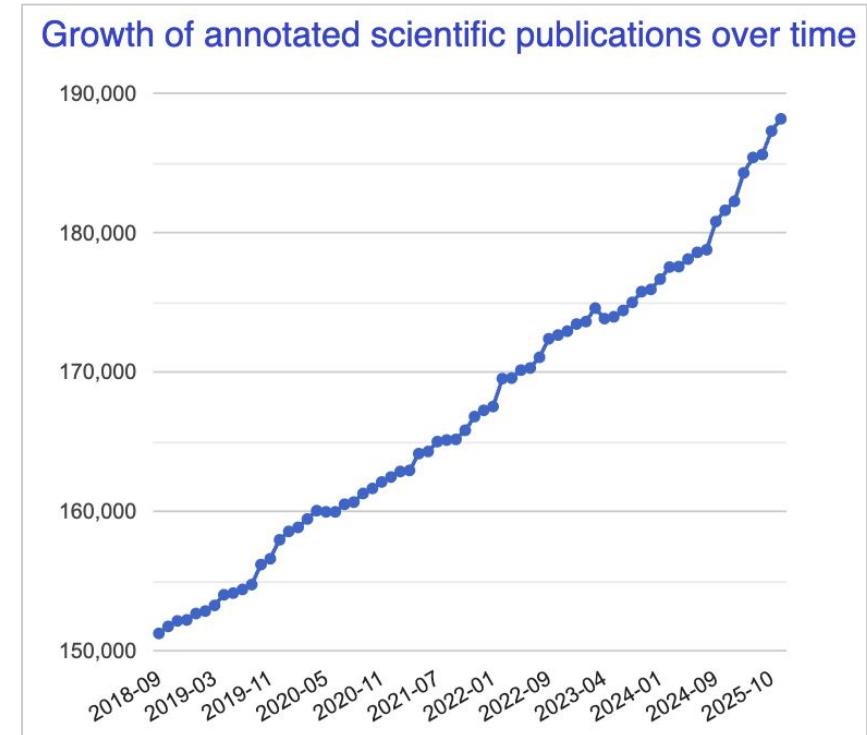
GO is dynamic



- GO aims to represent the current state of knowledge in biology
- Constantly revised as knowledge evolves
- Most change requests come from biocurators when annotating papers
- Large refactorings are made collaboratively with domain experts in particular areas of biology

Annotations are added and removed to represent current state of knowledge

- A GO annotation is a statement that links a gene product and a GO term
- These are linked via relations from the Relations Ontology
- GO aims to represent the **current state of knowledge** in biology, hence it is constantly revised and expanded as biological knowledge accumulates



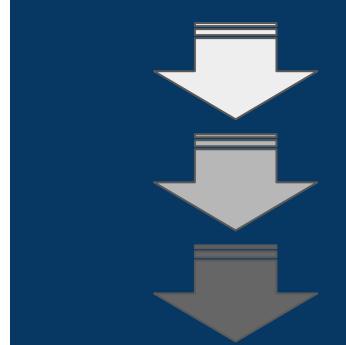
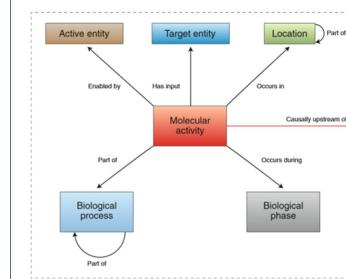
GO annotations

- Each gene can have any number of annotations to represent:
 - its function(s) and substrates
 - the cellular locations where it is active
 - the roles it plays in any process
 - each paper and evidence
- Both a feature and a limitation



Gene Product Symbol Qualifier GO Term Evidence Reference

Gene Product	Symbol	Qualifier	GO Term	Evidence	Reference
UnP00KB42697	DRP1A	acts_upstream_of, or_within	GO:0009111 (cytokinesis by cell plate formation)	ECO:0000314 (I)	PMID:12071066
UnP00KB42697	DRP1A	acts_upstream_of, or_within	GO:0009783 (embryo development ending in seed dormancy)	ECO:0000314 (I)	PMID:12071066
UnP00KB42697	DRP1A	acts_upstream_of, or_within	GO:0009020 (cell plate formation involved in plant-type cell wall biogenesis)	ECO:0000314 (I)	PMID:12071066
UnP00KB42697	DRP1A	acts_upstream_of, or_within	GO:0010051 (xylem and phloem pattern formation)	ECO:0000314 (I)	PMID:15603323
UnP00KB42697	DRP1A	acts_upstream_of, or_within	GO:0010091 (xichrome branching)	ECO:0000314 (I)	PMID:12071066
UnP00KB42697	DRP1A	acts_upstream_of, or_within	GO:0047686 (root hair initiation)	ECO:0000314 (I)	PMID:27251533

From function to GO annotation

 **P46934 · NEDD4_HUMAN**

Proteinⁱ | E3 ubiquitin-protein ligase NEDD4 ➤ **Molecular activity**

Geneⁱ | NEDD4

Functionⁱ

E3 ubiquitin-protein ligase which accepts ubiquitin from an E2 ubiquitin-conjugating enzyme in the form of a thioester and then directly transfers the ubiquitin to targeted substrates. Specifically ubiquitinates 'Lys-63' in target proteins (PubMed:[19920177](#), PubMed:[21399620](#), PubMed:[23644597](#)).

Involved in the pathway leading to the degradation of VEGFR-2/KDFR, independently of its ubiquitin-ligase activity. Monoubiquitinates IGF1R at multiple sites, thus leading to receptor internalization and degradation in lysosomes (By similarity).

Many substrates with roles in various cellular processes

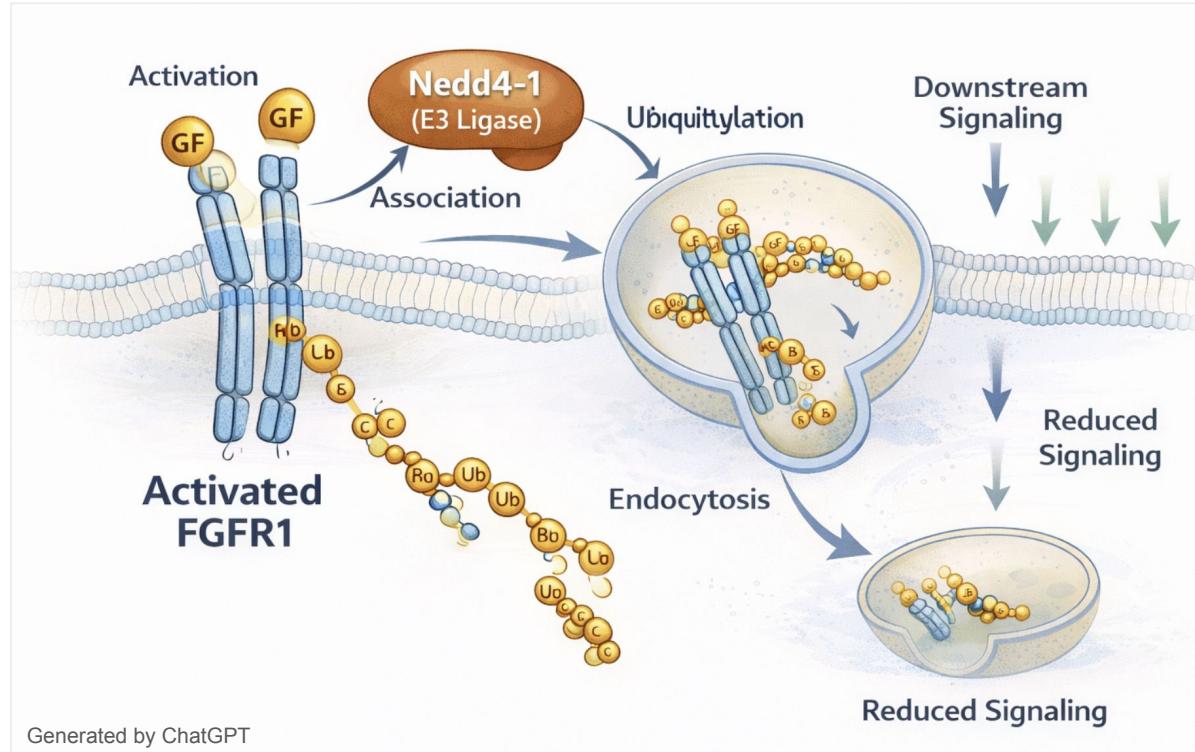
Ubiquitinates TNK2 and regulates EGF-induced degradation of EGFR and TNF2 (PubMed:[20086093](#)).

Ubiquitinates BRAT1 and this ubiquitination is enhanced in the presence of NDFIP1 (PubMed:[25631046](#)).

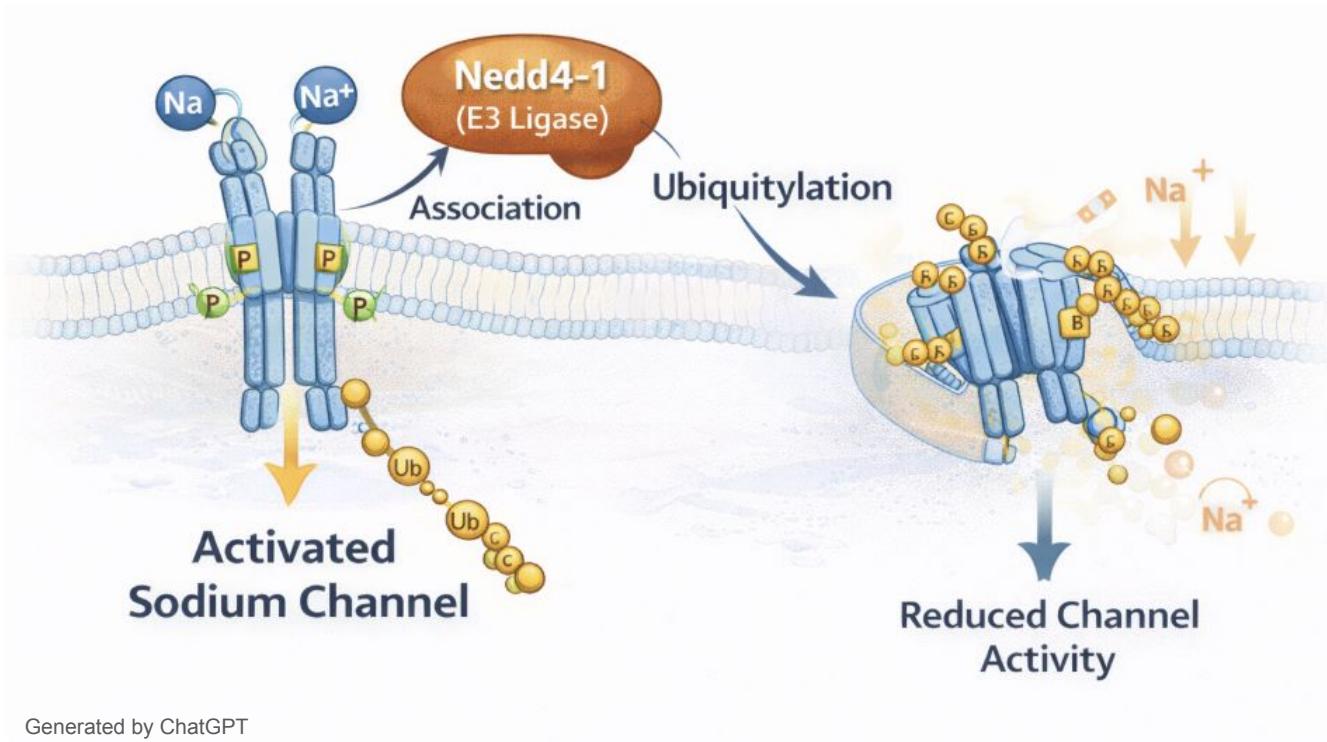
Ubiquitinates DAZAP2, leading to its proteasomal degradation (PubMed:[11342538](#)).

Ubiquitinates POLR2A (PubMed:[19920177](#)).

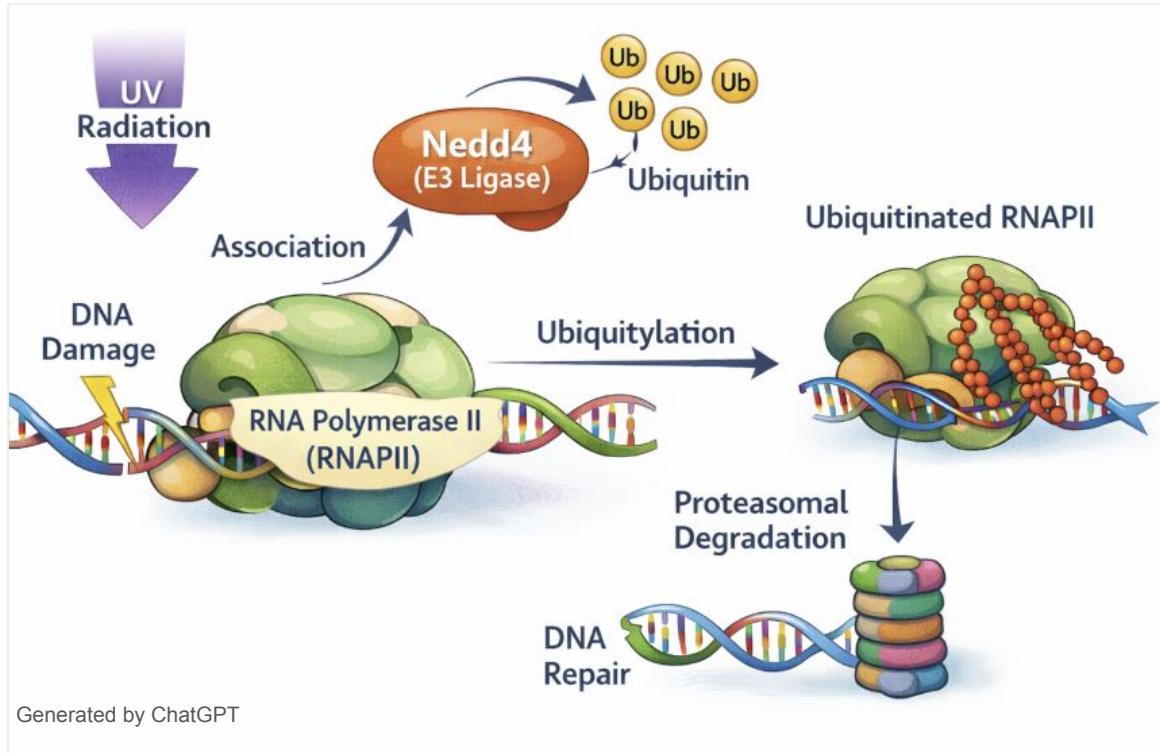
NEDD4 downregulates FGFR1 signaling



NEDD4 downregulates Na^{2+} channels



NEDD4 inhibits gene transcription in the presence of DNA damage



Standard annotations are unconnected and lack context

Molecular Function

- ubiquitin-protein ligase activity
- sodium channel inhibitor activity
- RNA polymerase binding
- fibroblast growth factor binding

Cellular Component

- cytosol
- plasma membrane
- nucleus

Biological Process

- ubiquitin-dependent protein catabolic process
- positive regulation of endocytosis
- negative regulation of sodium ion transport
- cellular response to UV
- negative regulation of transcription from RNA polymerase II promoter
- negative regulation of fibroblast growth factor receptor signaling pathway

Dissecting the biology

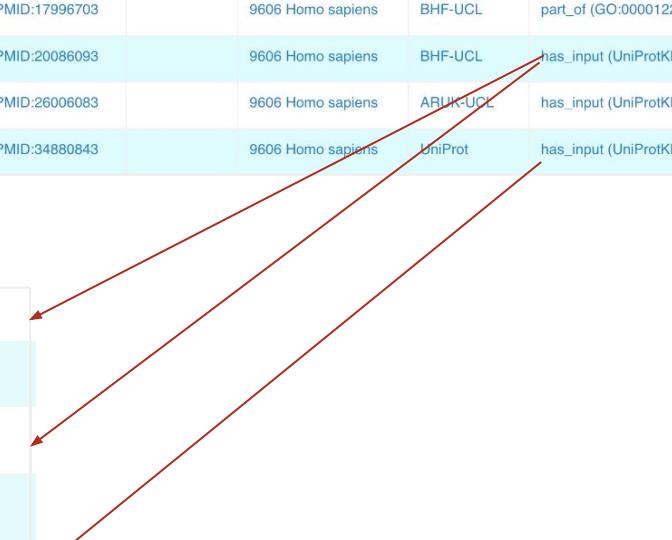
Molecular function (33 annotations + 75 'protein binding')

substrates

UniProtKB:P46934	NEDD4	enables	GO:0061630    ubiquitin protein ligase activity	ECO:0000314  IDA	PMID:17116753		9606 Homo sapiens	ARUK-UCL	has_input (UniProtKB:Q9NZ52)
UniProtKB:P46934	NEDD4	enables	GO:0061630    ubiquitin protein ligase activity	ECO:0000314  IDA	PMID:17996703		9606 Homo sapiens	BHF-UCL	part_of (GO:0000122) more...
UniProtKB:P46934	NEDD4	enables	GO:0061630    ubiquitin protein ligase activity	ECO:0000314  IDA	PMID:20086093		9606 Homo sapiens	BHF-UCL	has_input (UniProtKB:Q07912)
UniProtKB:P46934	NEDD4	enables	GO:0061630    ubiquitin protein ligase activity	ECO:0000314  IDA	PMID:26006083		9606 Homo sapiens	ARUK-UCL	has_input (UniProtKB:P08183) more...
UniProtKB:P46934	NEDD4	enables	GO:0061630    ubiquitin protein ligase activity	ECO:0000314  IDA	PMID:34880843		9606 Homo sapiens	UniProt	has_input (UniProtKB:Q7Z434) more...

Biological process (52 annotations)

UniProtKB:P46934	NEDD4	involved_in	GO:0000122    negative regulation of transcription by RNA polymerase II
UniProtKB:P46934	NEDD4	involved_in	GO:0006511    ubiquitin-dependent protein catabolic process
UniProtKB:P46934	NEDD4	involved_in	GO:0006974    DNA damage response
UniProtKB:P46934	NEDD4	involved_in	GO:0007041    lysosomal transport
UniProtKB:P46934	NEDD4	involved_in	GO:0010766    negative regulation of sodium ion transport
UniProtKB:P46934	NEDD4	involved_in	GO:0016567    protein ubiquitination



Early 2000's solution: Precomposition

Precomposed terms aimed to provide more expressive terms

cellular response to UV + negative regulation of transcription
from RNA polymerase II promoter

=

negative regulation of transcription from RNA polymerase II
promoter in response to UV-induced DNA damage

Examples obsolete pre-composed terms

- regulation of planar cell polarity pathway involved in axis elongation
- regulation of transcription involved in G1/S transition of mitotic cell cycle
- cell morphogenesis involved in differentiation
- regulation of the force of heart contraction involved in baroreceptor response to decreased systemic arterial blood pressure
- regulation of cytosolic calcium ion concentration involved in phospholipase C-activating G protein-coupled signaling pathway

Precomposition

cellular response to UV + negative regulation of transcription
from RNA polymerase II promoter

- Combinatorially unmanageable
- Complexity in the ontology leads to errors and inconsistencies
- Most have already been obsoleted

... response to UV-induced DNA damage

Annotation extensions

- AE provide additional context for annotations by linking MF, BP, CC annotations, as well as inputs/substrates, cell types and anatomical structures

<input type="checkbox"/> Gene/product	Gene/product name	Annotation qualifier	GO class (direct)	Annotation extension
<input type="checkbox"/> NEDD4	E3 ubiquitin-protein ligase NEDD4		ubiquitin protein ligase activity	part of negative regulation of transcription by RNA polymerase II part of DNA damage response

Annotation extensions

- AE provide additional context for annotation
- Annotation extensions are not easily exploitable by GO tools
- Limited scope because extensions are not easily exploitable by GO tools
- Not amenable to pathway-like representation

Annotation extensions ≠ pathways

Gene/product	Gene/product name	Annotation qualifier	GO class (direct)	Annotation extension
<input type="checkbox"/>	NEDD4	E3 ubiquitin-protein ligase NEDD4	ubiquitin protein ligase activity	part of negative regulation of transcription by RNA polymerase II part of DNA damage response
allowed extension			negatively regulates RNA polymerase activity	
disallowed extension (logic too complex for format)			negatively regulates RNA polymerase activity of POL2A	

GO causal activity models (GO-CAMs)

Provides a system to extend GO annotations

 **provides context**

- CC: in which cellular component, cell, tissue the function/process take place
- MF: capture substrates
- BP: which BP is a MF part of
- BPs can also be part of broader BPs

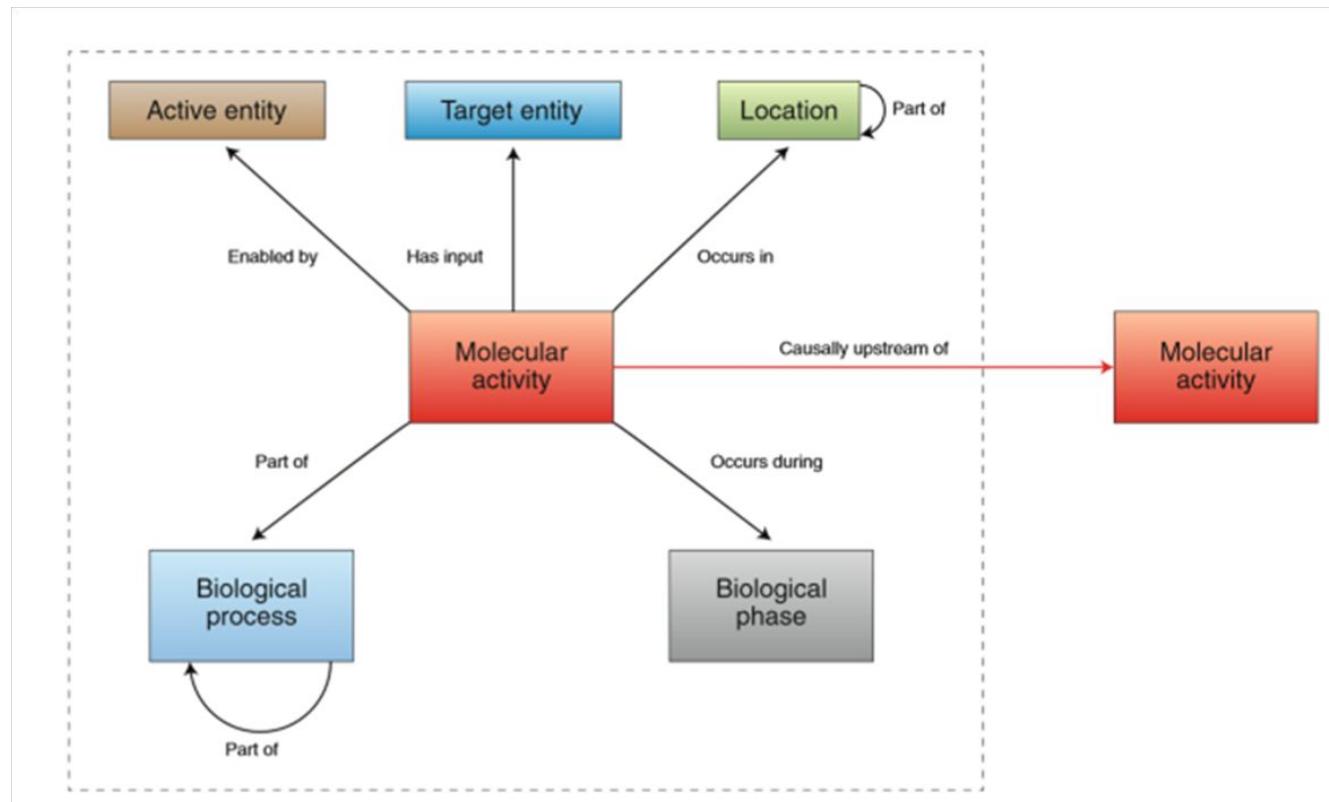
 **defines causal connections between activities (MFs)**

- allows network representation
- enables pathway visualization & analysis

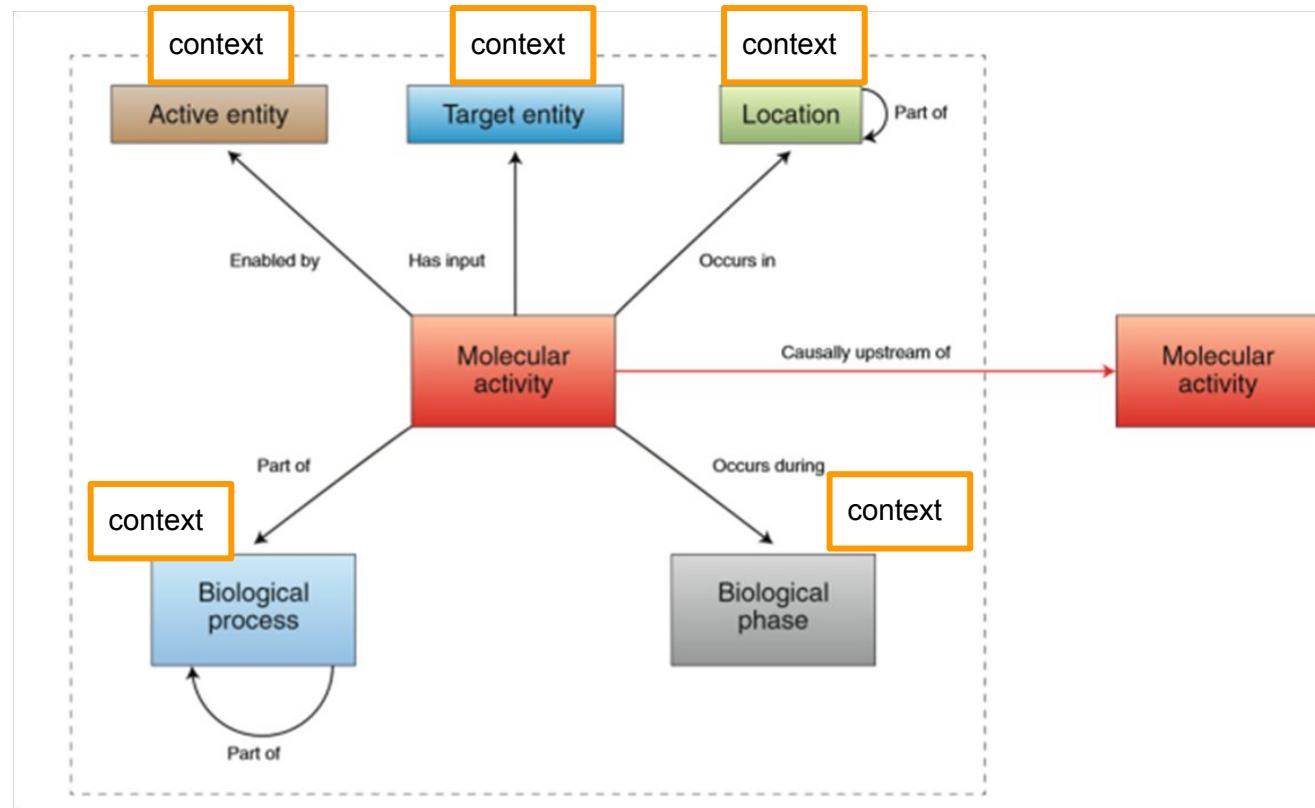
Published in Thomas & al, [Nat Genet. 2019](#)

See also [GO documentation](#)

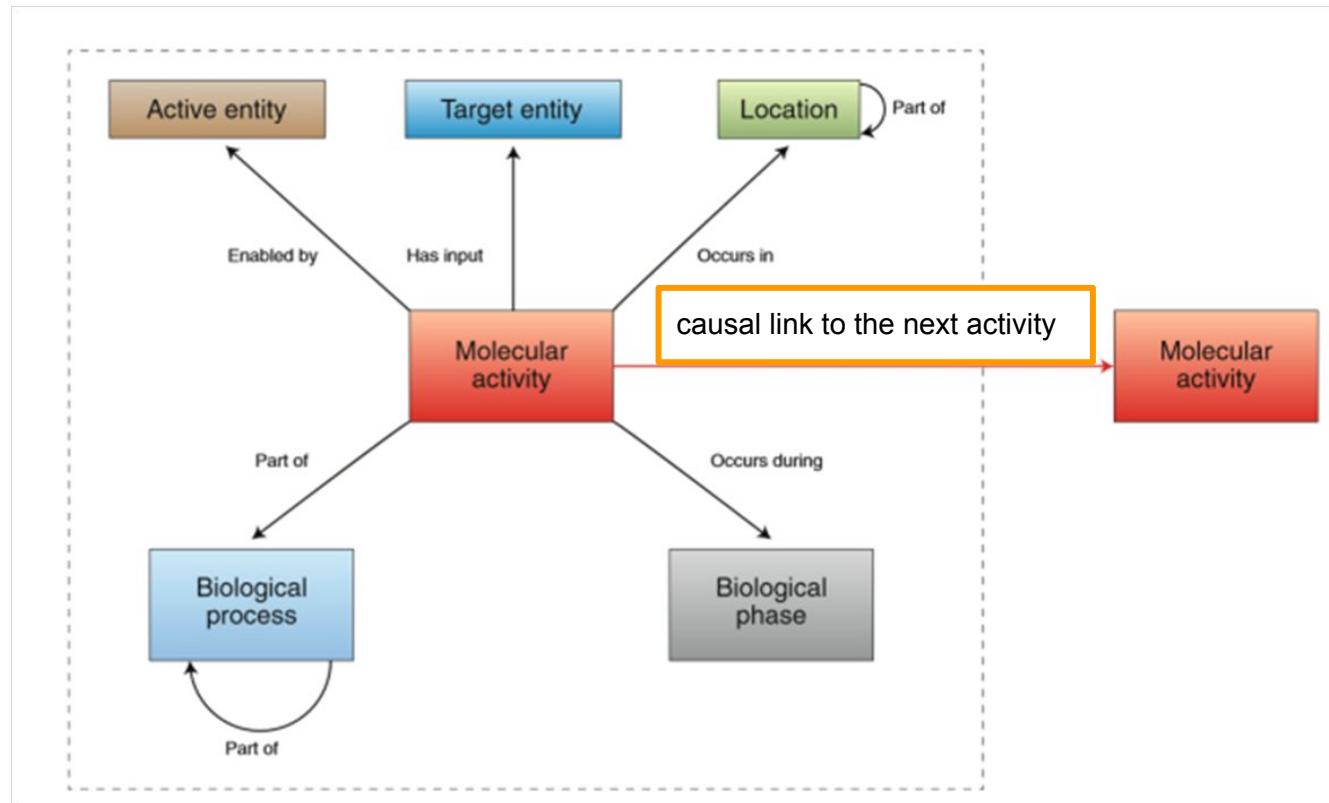
GO-CAM model specifications



GO-CAM model specifications



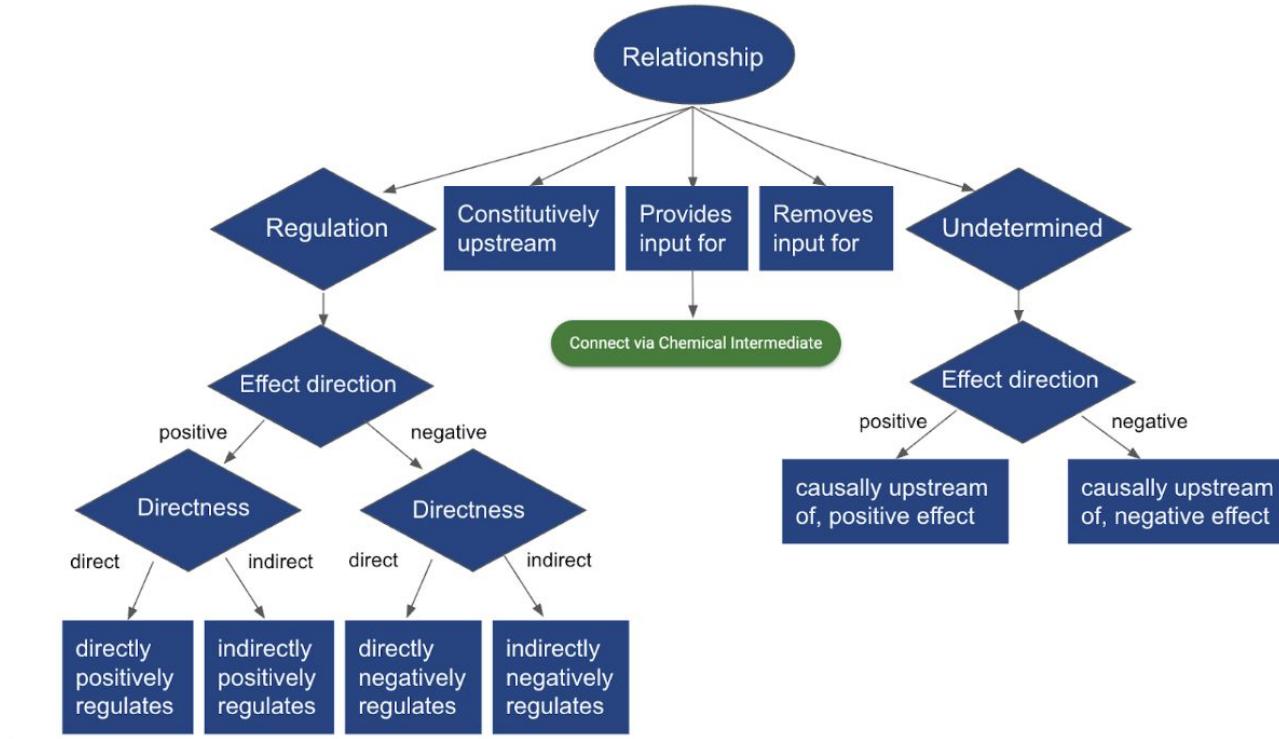
GO-CAM model specifications



Causal relations

- Link a MF to a MF
- Can be positive or negative

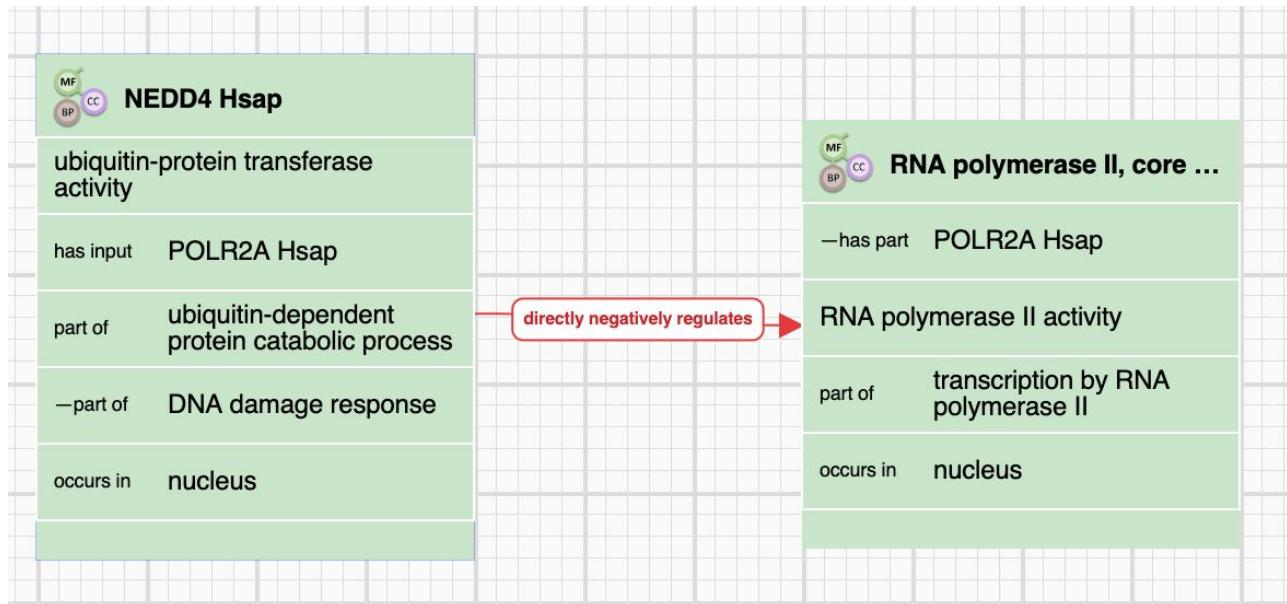
GO-CAM causal relations



See also the [Noctua User Guide](#)

directly negatively regulates	Links two activities when the upstream activity has a negative causal effect (decreasing or inhibiting) on an immediately downstream activity. Immediately means there is no intervening activity. The mechanism by which the upstream activity controls the downstream activity must be known.
directly positively regulates	Links two activities when the upstream activity has a positive causal effect (increasing or activating) on an immediately downstream activity.
indirectly negatively regulates	Links two activities when the upstream activity has a negative regulatory effect (decreasing or inhibiting) on the downstream activity via a larger process (e.g. proteasome-mediated protein degradation) that is reused in many contexts and the curator does not want to reproduce that process in the GO-CAM. The mechanism by which the upstream activity controls the downstream activity must be known.
indirectly positively regulates	Links two activities when the upstream activity has a positive regulatory effect (increasing or activating) on the downstream activity via a larger process (e.g. transcription) that is reused in many contexts and the curator does not want to reproduce that process in the GO-CAM. The mechanism by which the upstream activity controls the downstream activity must be known.
provides input for	Links two successive activities when the product (output) of the upstream activity is the substrate (input) for the downstream activity, and the product is a macromolecule (i.e. DNA, RNA, protein).
removes input for	Links two activities when the upstream activity has a negative causal effect (decreasing or inhibiting) on the downstream activity and the two activities act on or modify the same molecular target at the same site(s).
constitutively upstream of	Links two activities when the upstream activity is required for the downstream activity, but does not regulate the downstream activity.
causally upstream of, negative effect	Links two activities when the upstream activity has a negative causal effect (decreasing or inhibiting) on the downstream activity but the mechanism is not known.
causally upstream of, positive effect	Links two activities when the upstream activity has a positive causal effect (increasing or activating) on the downstream activity but the mechanism is not known.

GO-CAM model view fo regulation of RNA polymerase by NEDD4



GO-CAM impact on the ontology

Ontology refactoring

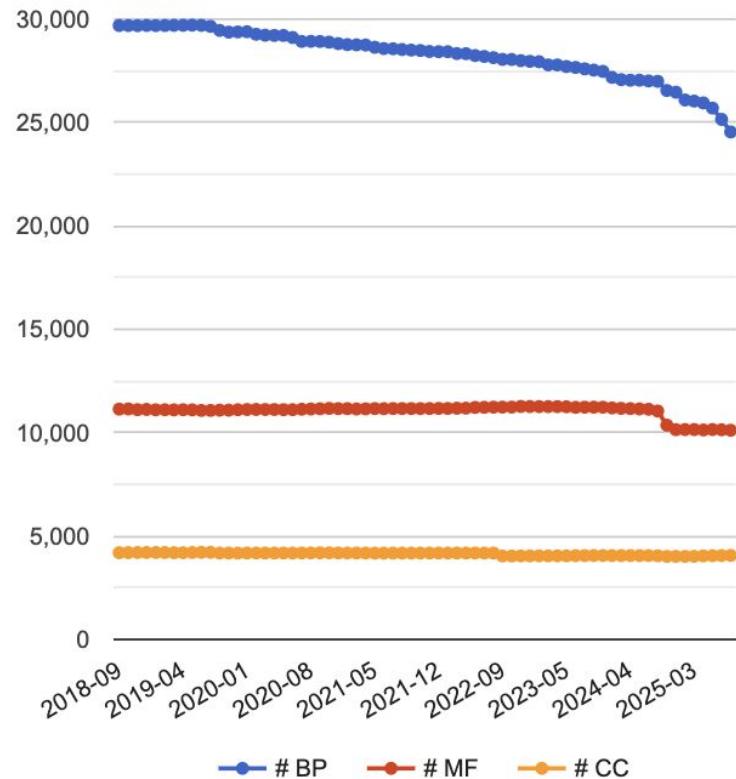
Started a major ontology refactor in 2019 to address issues caused by old ontology practices

- Proactive introduction of terms to anticipate potential curator needs (but never used)
- Excessive intermediate-level grouping terms
- Precomposed terms
- Single step reactions in the BP aspect of the ontology
- Terms that refer to more than one ontological aspect
- Experimental assays and non-physiological substrates

Examples obsoleted terms

Reason for obsoletion	Examples
Phenotype of perturbation experiments	<i>negative regulation of necrotic cell death</i> (GO:0060547)
BP terms that represent a MF	<i>histone H2A acetylation</i> (GO:0043968)
Pre-composed terms that are now represented by GO-CAMs	<i>histone H4 acetylation involved in response to DNA damage stimulus</i> (GO:2000776)
Sub-reaction and reaction mechanisms	<i>formation of peptidyl-cystine persulfide by sulphur transfer from free cysteine</i> (GO:0044526)
Substrates beyond the specificity of known enzymes or not physiologically relevant	<i>4,4'-diapophytoene desaturase (4,4'-diapoly copene-forming)</i> (GO:0140868)

Number of GO terms by aspect



WELCOME TO NOCTUA

Noctua is a web-based, collaborative Gene Ontology (GO) annotation tool developed by the GO Consortium to create standard GO annotations as well as GO-CAMs (Gene Ontology Causal Activity Models).

Noctua allows to view, create and edit GO annotations

The screenshot shows the Noctua web application interface. The top navigation bar includes a user profile (Pascale Gaudet, GO Central), a search icon, a help icon, and the GENEONTOLOGY Unifying Biology logo. The main content area features a dark blue background with a large "WELCOME TO NOCTUA" heading. Below it, a text block describes Noctua as a web-based, collaborative Gene Ontology (GO) annotation tool developed by the GO Consortium to create standard GO annotations as well as GO-CAMs (Gene Ontology Causal Activity Models). Below this, there are two main buttons: "CREATE" (with "STANDARD ANNOTATIONS EDITOR" and "GO-CAM VISUAL PATHWAY EDITOR" sub-options) and "HELP" (with "USER GUIDE" sub-option). The left sidebar, titled "Noctua", contains a "Filter by" section with various dropdowns and checkboxes for filtering annotations by term, obsolete term, gene product, chemical, reference, organism, contributor, and group. It also includes a "Date last modified" section with a "Date Range" checkbox. The main content area shows a table of filtered GO annotations. The table has columns for Title, Saved (with a green checkmark), State (Production or Development), Date Modified (2026-02-04), Contributors (Lionel Breuza or Job Berkhout), and Actions (a blue button). The table includes a search bar at the top and pagination at the bottom.

WELCOME TO NOCTUA

Noctua is a web-based, collaborative Gene Ontology (GO) annotation tool developed by the GO Consortium to create standard GO annotations as well as GO-CAMs (Gene Ontology Causal Activity Models).

CREATE

STANDARD ANNOTATIONS EDITOR

GO-CAM VISUAL PATHWAY EDITOR

HELP

USER GUIDE

Filtered By:

Title	Saved	State	Date Modified	Contributors	Actions
Vesicle fusion in cytotoxic T cell degranulation (Human)	✓	Production	2026-02-04	LB Lionel Breuza	Actions
CD36 transport to the plasma membrane for muscle regeneration (Human)	✓	Production	2026-02-04	LB Lionel Breuza	Actions
NMDAR-triggered AMPAR recruitment to AP2/PICK1 machinery enables clathrin-mediated endocytosis and LTD	✓	Development	2026-02-04	JB Job Berkhout	Actions

Results: 54720

GO CAMs per page: 50

1 - 50 of 54720

Date last modified

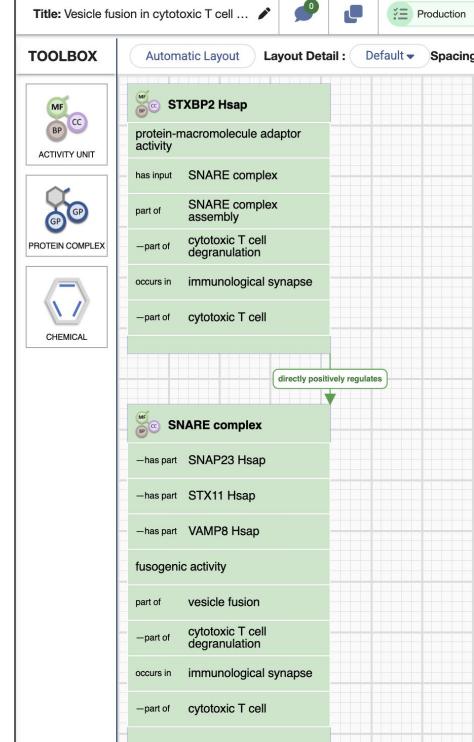
Date Range

Two main workbenches

Standard Annotations Editor (SAE)

GENE PRODUCT	RELATIONSHIP	TERM ^	EVIDENCE	REFERENCE	WITH	EXTENSION	DATE MODIFIED
STXBP2 Hsap UniProtKB:Q15833	enables	protein-macromolecule adaptor activity GO:0030674	direct assay evidence used in manual assertion ECO:0000314	PMID:28265073		<p>has input SNARE complex GO:0031201</p> <p>occurs in immunological synapse GO:0001772</p> <p>part of SNARE complex assembly GO:0035493</p>	Feb 4, 2026
SNARE complex GO:0031201	enables	fusogenic activity GO:0140522	direct assay evidence used in manual assertion ECO:0000314	PMID:28265073		<p>occurs in immunological synapse GO:0001772</p> <p>part of vesicle fusion GO:0006906</p>	Feb 4, 2026

Visual Pathway Editor (VPE)



Standard Annotations Editor - creating annotations

NOT

Gene Product	GP to Term Relation	▼	GO Term	⋮
--------------	---------------------	---	---------	---

Evidence

Evidence	Reference	With/From
----------	-----------	-----------

Extensions

[Add Extension](#)

Comments

[Add Comment](#)

Clear **Save**

Visual Pathway Editor - new annotation

The screenshot illustrates the Visual Pathway Editor interface, specifically the 'new annotation' feature. On the left, a sidebar lists three categories: 'ACTIVITY UNIT' (with a green icon), 'PROTEIN COMPLEX' (with a grey icon), and 'CHEMICAL' (with a blue icon). The main workspace displays two entities: 'STXBP2 Hsap' and 'SNARE complex'.

STXBP2 Hsap (highlighted in yellow):

- MF: protein-macromolecule adaptor activity
- BP: SNARE complex
- CC: part of SNARE complex assembly
- BP: part of cytotoxic T cell degranulation
- BP: occurs in immunological synapse
- BP: part of cytotoxic T cell

SNARE complex (highlighted in yellow):

- MF: directly positively regulates
- BP: SNAP23 Hsap
- BP: STX11 Hsap
- BP: VAMP8 Hsap
- MF: fusogenic activity
- BP: part of vesicle fusion
- BP: part of cytotoxic T cell degranulation
- BP: occurs in immunological synapse
- BP: part of cytotoxic T cell

A red arrow points from the 'Activity Unit Form' for 'STXBP2 Hsap' to the 'Function Description' section of the 'Activity Unit Form' for 'SNARE complex'.

Activity Unit Form (Top):

- Gene Product: STXBP2 Hsap
- Function Description:
 - enabled by (GP)

Activity Unit Form (Bottom):

- Function Description:
 - Molecular Function
 - Evidence
 - Reference
 - With
 - More
- (MF) part of (BP)
- (MF) occurs in (CC)

GO-CAM status

- ~ 2,000 models ('production')
- > 10,000 distinct genes
- 5 major contributing groups: UniProt, PomBase, FlyBase, SGD, MGI

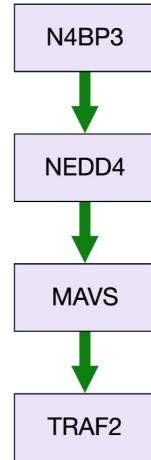
GO-CAM model availability

- GO-CAM model in OWL-type formats (ttl)
- conversion to standard annotations, with extensions if appropriate

<input type="checkbox"/> Gene/product	Gene/product name	Annotation qualifier	GO class (direct)	Annotation extension
<input type="checkbox"/> NEDD4	E3 ubiquitin-protein ligase NEDD4		ubiquitin protein ligase activity	part of negative regulation of transcription by RNA polymerase II part of DNA damage response

GO-CAM Pathway Viewer widget

- Widget can be embedded in any website
- Available in AmiGO, Alliance of Genomes, UniProt, PomBase, FlyBase



GO-CAMs are also available at NDEx

<https://www.ndexbio.org>

“The NDEx Project provides an open-source framework where scientists and organizations can store, share, manipulate, and publish biological network knowledge.”

The NDEx Project provides an open-source framework where scientists and organizations can store, share, manipulate, and publish biological network knowledge. The NDEx website features a search bar, a login/register button, and a featured content section for NDEx IQuery. The main content area includes sections for NDEx Integrated Query and GO-Causal Activity Models (GO-CAMs). The GO-CAMs section displays a diagram illustrating the structure of a GO-CAM, showing how multiple GO annotations (e.g., GO:MF1, GO:MF2, GO:CC1, GO:CC2) are linked through causal relationships (e.g., positively regulates, negatively regulates, part of). The NDEx Integrated Query section provides a detailed description of the tool, its features, and examples of its use.

NDEx Integrated Query

Powered by NDEx and integrated with Cytoscape, IQuery uses selected pathway and a variety of interactome networks to power gene set analysis. The networks come from many different sources and new networks will be continuously added.

By doing many simultaneous queries, we present the biologist with a Google-like experience and immediate actionable results; any IQuery hit can be seamlessly opened in Cytoscape for additional analysis or saved to NDEx for later use. Alternatively, users can explore the master networks that originated those results and perform additional queries or other operations.

Enter your gene set here... 

Examples: [Death Receptors](#) [PTC](#) [Beta Cell](#)

Enter your gene set in the text box above or use one of the

GO-Causal Activity Models

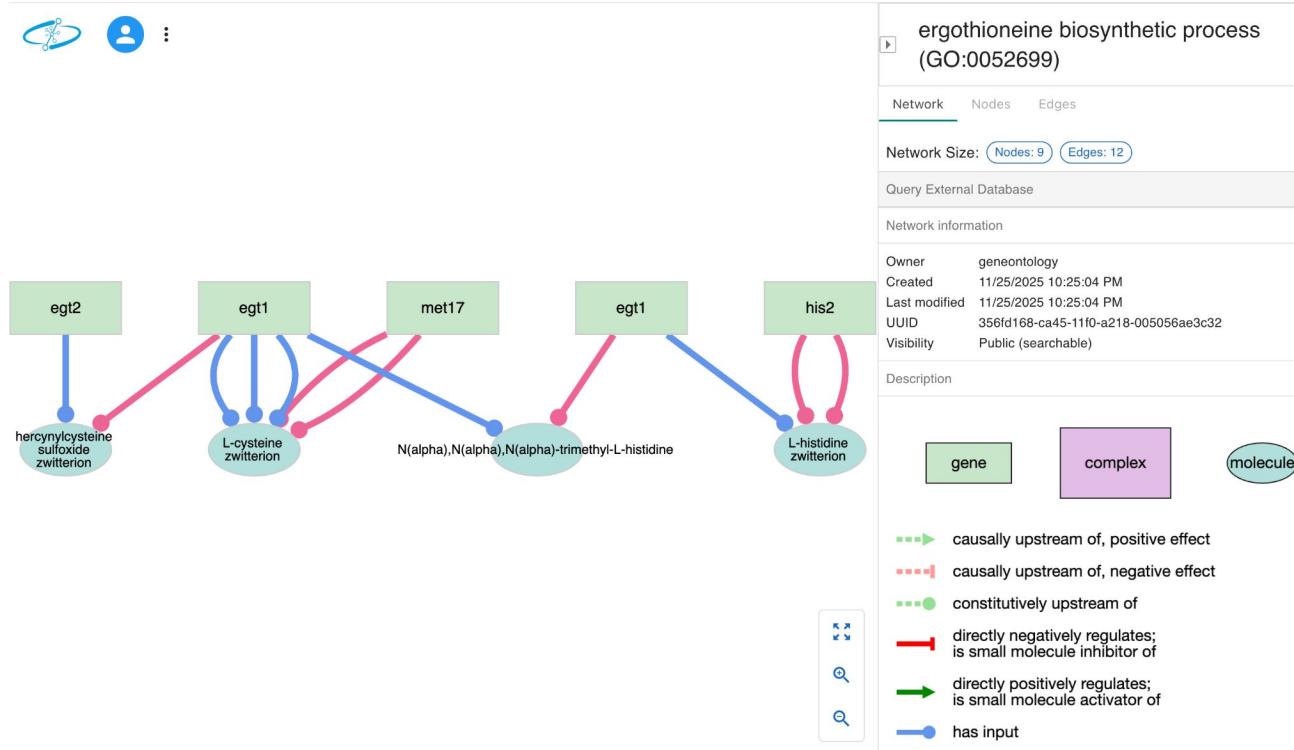
GO-Causal Activity Models (GO-CAMs) use a defined “grammar” for linking multiple GO annotations into larger models of biological function (such as “pathways”) in a semantically structured manner. GO-CAMs are created by expert biocurators from the GO Consortium, using the Noctua Curation Platform. GO-CAMs can be used as “pathway diagrams” that can be browsed and visualized, as full GO-CAM graphs that can be downloaded or as “reduced” pairwise GO annotations, included in the GO annotation files for backward compatibility with the broader GO communities. Human GO-CAMs are available in NDEx.

Featured Content

NDEx IQuery: a multi-method network gene set analysis leveraging the Network Data Exchange

NDEx IQuery is a new web tool for network and pathway-based gene set interpretation.

NDEx GO-CAM pathway viewer



NDEx IQuery

One search finds a variety of pathways and interaction networks relevant to your set of genes. Powered by **NDEx** and integrated with **Cytoscape**.

The logo features a stylized blue molecular structure with a central node and connecting lines, enclosed within a blue oval shape.

NDEx Integrated Query
v1.4



Enter gene list (or click an example below)

Query gene set examples:

[PANCREATIC BETA CELL](#) [DEATH RECEPTORS](#) [COMPLEMENT](#)

Send us [feedback](#).

NDEx Integrated Query uses selected pathway and interactome networks in NDEx to power gene set analysis. The networks come from many different sources and new networks will be continuously added. Do you have a pathway

<https://www.ndexbio.org/iquery>

IQuery provides a form of gene set enrichment analysis by indexing certain NDEx network sets and allowing users to find networks by searching over Human gene sets

NDEx IQuery tests pathway representation for lists of genes

NDEx Integrated Query APAF1 BCL2 BID BIRC2 BIRC3 CASP10 CASP3 CASP6 CASP7 CFLAR CHUK DFFA DFFB FADD GAS2 🔍 Powered by 

CURATED PATHWAYS PATHWAY FIGURES IN德拉 GO CANCER CELL MAP INITIATIVE GO-CAMS INTERACTOMES

Sort by Similarity p-Value Overlap

GO-CAMS: Involvement of CASP7 in apoptosis, its acti... Layout Default  Legend Show

Involvement of CASP7 in apoptosis, its activation by initiator caspases (CASP8, CASP9 and CASP10) and regulation by XIAP and PAK2 (Human)
3 / 6 unique genes Similarity: 0.32 p-Value: **8.95e-4**

CASP10	CASP7	XIAP	APAF1
BCL2	BID	BIRC2	BIRC3
CASP3	CASP6	CFLAR	CHUK
DFFA	DFFB	FADD	GAS2
LMNA	MAP3K14	NFKB1	RELA
RIPK1	SPTAN1	TNFRSF25	TNFSF10
TRADD	TRAF2		

Poxvirus-CRMB inhibition of TNF signaling pathway (Human-Poxvirus)
3 / 5 unique genes Similarity: 0.33 p-Value: **8.95e-4**

FADD TRADD TRAF2

TRAIL/TRAILR1 interaction leading to apoptosis
3 / 6 unique genes Similarity: 0.33 p-Value: **8.95e-4**

FADD TNFSF10 TRADD

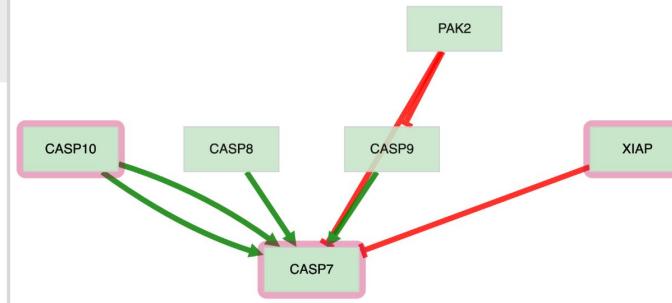
TRAIL/TRAILR2 interaction leading to apoptosis
3 / 5 unique genes Similarity: 0.36 p-Value: **8.95e-4**

FADD TNFSF10 TRADD

TRAF7 modulation of NF-kappa-B via IKBKG and NEMO K-29 ubiquitination (Human)
3 / 7 unique genes Similarity: 0.23 p-Value: **1.23e-3**

CHUK NFKB1 RELA

Cleavage and inactivation of PARP1 by CASP3 and/or CASP7 during the execution phase of apoptosis (Human)



- Query set: NDEx 'Death receptor' test list
- Over-represented GO-CAMs show up on the left
- For each model, hits are highlighted

Exploring AI to enhance GO

GO tasks

- Ontology development
 - create new terms, obsolete terms
 - fix hierarchy to align with external resources (EC, RHEA, etc)
 - apply design patterns
- Annotation
 - add new annotation
 - remove incorrect annotations
 - annotation reviews
 - new GO-CAMs
 - review GO-CAMs
- QC

AI-generated annotation review

Gene Annotation Review Browser	
GENE:	APO1
REVIEW:	Review
LINK:	Link
ID:	Q9XIR4
GENE PRODUCT TYPE:	PROTEIN
REVIEW STATUS:	COMPLETE
GO TERM:	regulation of gene expression
GO ID:	GO:0010468
EVIDENCE:	IMP
PATHWAY:	Pathway
ORIG REF:	PMID:21421812
ORIG TITLE:	APO1 promotes the splicing of chloroplast group II introns and harbors a plant-specific zinc-dependent RNA binding domain.
SUMMARY:	Too general - should use more specific RNA processing term
ACTION:	MODIFY
REASON:	APO1 regulates gene expression through RNA splicing, not transcriptional regulation. The more specific term "chloroplast group II intron splicing" (GO:0000373) better captures the mechanism
REF TEXT:	APO1 promotes the splicing of chloroplast group II introns and harbors a plant-specific zinc-dependent RNA binding domain.
REFIDS:	PMID:21421812
REPLACEMENTS:	GO:0006397:mRNA processing

GO term & evidence

Justification

Suggested GO term

AI-generated new annotation suggestions

Gene Annotation Review Browser

Browse AI reviews

GENE: ELF4

REVIEW: [Review](#)

LINK: [Link](#)

ID: 004211

GENE PRODUCT TYPE: PROTEIN

REVIEW STATUS: DRAFT

GO TERM: transcription regulator complex

GO ID: [GO:0005667](#)

EVIDENCE: IDA

ORIG REF: [PMID:21753751](#)

ORIG TITLE: The ELF4-ELF3-LUX complex links the circadian clock to diurnal control of hypocotyl growth.

SUMMARY: ELF4 is a component of the Evening Complex (EC), a tripartite transcriptional repressor complex. This is documented in ComplexPortal (CPX-1291) and demonstrated experimentally by co-immunoprecipitation and ChIP experiments.

ACTION: NEW

REASON: This annotation captures ELF4's role as part of the Evening Complex, which is central to its function. The EC is a well-characterized transcription regulator complex that represses target genes.

REF TEXT: Here we identify a protein complex (called the evening complex)--composed of the proteins encoded by EARLY FLOWERING 3 (ELF3), ELF4 and the transcription-factor-encoding gene LUX ARRHYTHMO (LUX; also known as PHYTOCLOCK 1)--that directly regulates plant growth | We found that ELF4-HA could co-immunoprecipitate both ELF3 and LUX

REFIDS: [PMID:21753751](#); [PMID:21753751](#)

ISOFORM: false

TAGS:

GO term & evidence

Justification

Conventional workflow for ontology editing

- GO uses Protégé as ontology editing tool, hooked to GitHub actions for QC

The image shows a screenshot of the Protégé ontology editor interface. The interface is divided into several panels:

- Left Panel (1):** Shows the class hierarchy under the 'catalytic activity' class. It includes categories like 'fully formed stage', 'gastrula stage', 'larval stage', 'late embryonic stage', 'life cycle', 'neurula stage', 'organogenesis stage', 'pharyngula stage', 'post-embryonic stage', 'post-juvenile adult stage', and 'process' (with sub-categories like 'biological_process' and 'molecular_function').
- Central Panel (2):** A 'Create a new Class' dialog box. The 'Name' field is filled with 'phycoerythobilin synthase activity'. The 'IRI' field shows the URL 'http://purl.obolibrary.org/obo/GO_0140386'. There are 'Cancel' and 'OK' buttons.
- Bottom Panel (3):** The detailed properties of the newly created class 'phycoerythobilin synthase activity'. The 'Value' field contains the definition: 'Catalysis of the reaction: (32)-phycoerythobilin + oxidized ferredoxin = biliverdin IXa + reduced ferredoxin.' The 'Datatype' is set to 'xsd:string'. The 'OK' button is visible at the bottom right.
- Right Side (Listed 4-10):**
 - 4 add namespace
 - 5 add cross-reference #1
 - 6 add cross-reference #2
 - 7 add synonym
 - etc
 - 8 save
 - 9 commit, push
 - 10 create pull request, merge

Agentic workflow

User request in GitHub

NTR: Phycoerythrobilin synthase [Rhea:25359], EC 1.3.7.6 #31443

 **Closed**  #31452

 **AndreaAuchincloss** opened last week Member ...

Hello,
This reaction is missing from GO, can it be made please?
[Rhea:25359], EC 1.3.7.6
Label:
Phycoerythrobilin synthase

Description:
Catalysis of the reaction: (3Z)-phycoerythrobilin + oxidized ferredoxin = biliverdin IXa + reduced ferredoxin.
PMIDs 18356052, 18662988 and 21050180 show the reaction, the last one in particular shows the product is the 3Z isomer of PEB.
Cyanobacteria do this reaction with 2 proteins (PebA 1.3.7.2 GO:0050617 Q02189) and (PebB 1.3.7.3 GO:0050618 Q02190), but there are a number of viruses (including characterized Q58MU6) that have a protein that does it all in one swoop.
It is a child of GO:0016636.

Editor asks agent to do the task

 **raymond91125** 5 days ago Member ...

@dragon-ai-agent please add the new term phycoerythrobilin synthase activity. Use RHEA:25359 and PMID:18662988, and no others, as the definition references. Add RHEA and EC as exact match general cross references.

  1

Agentic workflow - new term created !

dragon-ai-agent 5 days ago Member ...

I have created the new term phycoerythrobilin synthase activity (GO:7770044).

Term Summary

- ID: GO:7770044
- Name: phycoerythrobilin synthase activity
- Parent: GO:0016636 (oxidoreductase activity, acting on the CH-CH group of donors, iron-sulfur protein as acceptor)
- Definition: "Catalysis of the reaction: biliverdin IXalpha + 2 reduced ferredoxin + 4 H+ = (3Z)-phycoerythrobilin + 2 oxidized ferredoxin."
- Definition references: PMID:18662988, RHEA:25359
- Cross-references: EC:1.3.7.6 (exactMatch), RHEA:25359 (exactMatch)
- Synonyms:
 - "(3Z)-phycoerythrobilin:ferredoxin oxidoreductase activity" (RELATED)
 - "PebS" (RELATED)

Background

Phycoerythrobilin synthase (PebS, EC 1.3.7.6) is a ferredoxin-dependent bilin reductase found in marine cyanophages. The enzyme performs a four-electron reduction of biliverdin IXalpha to (3Z)-phycoerythrobilin in a single step, combining the activities of PebA (EC 1.3.7.2) and PebB (EC 1.3.7.3) that are required in cyanobacteria.

The reaction direction in the definition follows the biosynthetic (physiological) direction, consistent with the enzyme's name as a synthase.

PR: #31452

raymond91125 5 days ago

@dragon-ai-agent please remove the synonyms. They are not required.

1

Editor asks agent to make corrections if needed

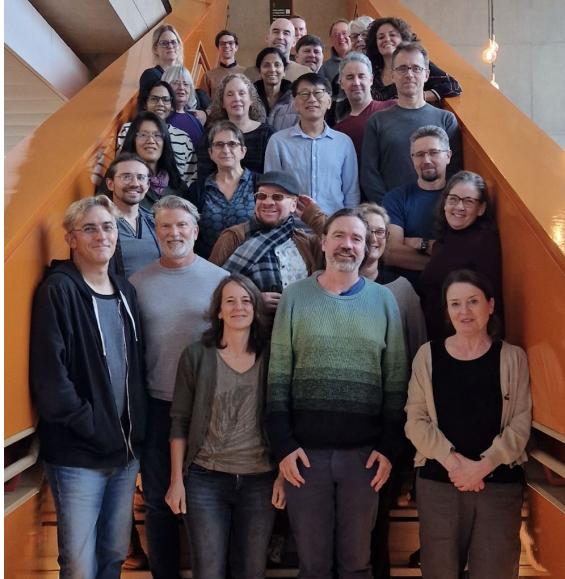
Summary

- GO is a computational model of gene function that aims to be comprehensive, species-agnostic and current
- The ontology is continually updated, with recent major updates to better capture biology rather than experimental results
- Annotations exist in both standard and GO-CAM flavors
- GO-CAMs are now disseminated across various resources
- GO strives to be efficient and accurate

Ressources

- [GO Handbook](#)
- [GO-CAM paper \(2019\)](#)
- [Gene Ontology website](#)
 - [General ontology documentation](#)
 - [General annotation documentation](#)
- [Annotation Relations documentation](#)
- [EBI Gene Ontology tutorial on YouTube \(general\)](#)

Contributors



PIs

Paul Thomas, Chris Mungall, Paul Sternberg

GO Central

Suzi A. Aleksander, James Balhoff, Seth Carbon, Dustin Ebert, Marc Feuermann, Raymond Lee, Huaiyu Mi, Sierra Moxon, Anushya Muruganugan Tremayne, Mushayahama, Kimberly Van Auken

Contributing groups

CGD, CACAO, EcoWiki, ComplexPortal, dictyBase, DisProt, EcoCyc, EcoliWiki, Ensembl, FlyBase, HGNC, HUGO, IntAct, InterPro, JaponicusDB, MGI, PANTHER, PHI-base, PomBase, RGD, Reactome, RHEA, RNACentral, SGD, SignatureScience, TAIR, HPA, Matrisome, SynGO, UniProt, WormBase, ZFIN