Aligning MouseCyc with GO

Case Study and Work in Process:Ceramide Biosynthesis

MGI

About Help FAQ

Tour our new Quick Search

Mouse Genome Informatics

Search ▾ Download ▾ More Resources ▾ Submit Data Find Mice (IMSR) Analysis Tools Contact Us

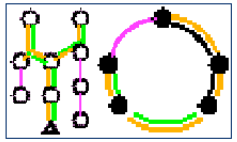
Keywords, Symbols, or IDs Quick Search

Explore MGI All Search Tools

Genes Phenotypes Expression

Biochemical Pathways

MouseCyc is a database of curated biochemical pathways data for the laboratory mouse that can be integrated with functional and phenotypic data from MGI.



Pathway glyphs, colorized with experimental data

Access Data

Search the MouseCyc database

Browse the pathway ontology in the MouseCyc database

Browse all pathways represented in MouseCyc

Browse a list of all genes in MouseCyc

Browse a list of all proteins in MouseCyc according to their activity

Browse a list of all compounds represented in MouseCyc

Overlay experimental data onto the Pathway Tools Omics Viewer
Example OmicsViewer files

Generate a mouse metabolic map

FAQs

How do I ...

.. see a list of all pathways in MouseCyc? FAQ

.. search MouseCyc for a specific compound? FAQ

.. display my mouse gene expression data in the context of mouse pathways? FAQ

.. compare mouse pathways to human pathways? FAQ

More FAQs

Biochemical Pathways Includes

24 Mar 2010

347	Pathways
2,018	Enzymatic reactions
9	Transport reactions
39,629	Polypeptides
6	Protein complexes
3,000	Enzymes
61	Transporters
1,337	Compounds

More...

About MouseCyc | Help Documents | Collaborators | Other Links

The availability of the complete genome sequence for the laboratory mouse provides a powerful platform for predicting genes and other genome features. However, building a catalog of genome annotations is just the beginning for biology in a "post-genome" era. To derive new insights into fundamental biological processes using complete genome sequences will require understanding how genome features interact in pathways and networks in the cell and how perturbations of these interactions contribute to disease processes. Toward this end, we have implemented a new database of curated biochemical pathways for the laboratory mouse called MouseCyc.

The MouseCyc database represents a significant advance for biomedical researchers wanting to access mouse genetic and genomic data in the context of physiological and cellular processes. The initial focus for the development of MouseCyc is on metabolism and includes such cell level processes as biosynthesis, degradation, energy production, and detoxification. MouseCyc differs from existing pathway databases and software tools because of the extent to which the pathway information in MouseCyc is integrated with the wealth of biological knowledge for the laboratory mouse that is available from the Mouse Genome Informatics (MGI) database.

MouseCyc is supported by NIH HG003622.

Personnel:

Carol J. Bult, Ph.D., Principal Investigator
Alexei Evsikov, Ph.D., Research Scientist/Scientific Curator
Mary Dolan, Ph.D., Bioinformatics Analyst

Past personnel:
Emily Patek, Graduate Student

Mus musculus Pathways

How to Navigate: A class hierarchy (ontology) allows you t
Clicking on a class will display a page containing its instan

Summary: This class is the root of a classification

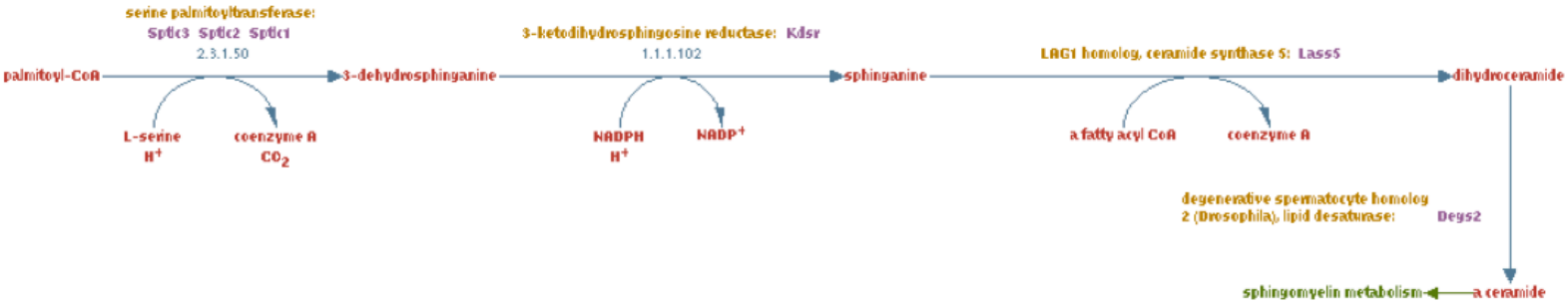
- Pathways
  - Biosynthesis (164)
  - Degradation/Utilization/Assimilati
  - Detoxification (3)
  - Generation of precursor metabolit
  - Superpathways (22)

- Biosynthesis (164)
  - Amines and Polyamines Biosynthesis (11)
  - Amino acids Biosynthesis (29)
  - Aminoacyl-tRNA Charging (2)
  - Carbohydrates Biosynthesis (17)
  - Cell structures Biosynthesis (4)
  - Cofactors, Prosthetic Groups, Electron Carriers Biosynthesis (40)
  - Fatty Acids and Lipids Biosynthesis (36)
  - Hormones Biosynthesis (8)
  - Metabolic Regulators Biosynthesis (2)
  - Nucleosides and Nucleotides Biosynthesis (14)
  - Other Biosynthesis (3)
  - Secondary Metabolites Biosynthesis (6)

- Degradation/Utilization/Assimilation (124)
  - Alcohols Degradation (2)
  - Aldehyde Degradation (3)
  - Amines and Polyamines Degradation (9)
  - Amino Acids Degradation (43)
  - Aromatic Compounds Degradation (5)
  - C1 Compounds Utilization and Assimilation (6)
  - Carbohydrates Degradation (19)
  - Carboxylates Degradation (8)
  - Cofactors, Prosthetic Groups, Electron Carriers Degradation (1)
  - Degradation/Utilization/Assimilation - Other (6)
  - Fatty Acid and Lipids Degradation (13)
  - Inorganic Nutrients Metabolism (5)
  - Nucleosides and Nucleotides Degradation and Recycling (5)
  - Secondary Metabolites Degradation (2)

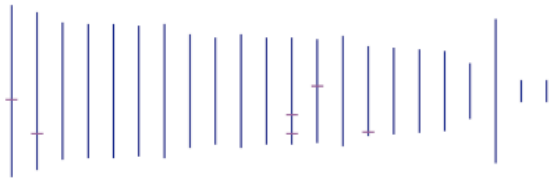
Mus musculus Pathway: ceramide biosynthesis

More Detail | Less Detail | Cross-Species Comparison



If an enzyme name is shown in bold, there is experimental evidence for this enzymatic activity.

Locations of Mapped Genes:



Superclasses: Biosynthesis -> Fatty Acids and Lipids Biosynthesis
Degradation/Utilization/Assimilation -> Fatty Acid and Lipids Degradation

Summary:

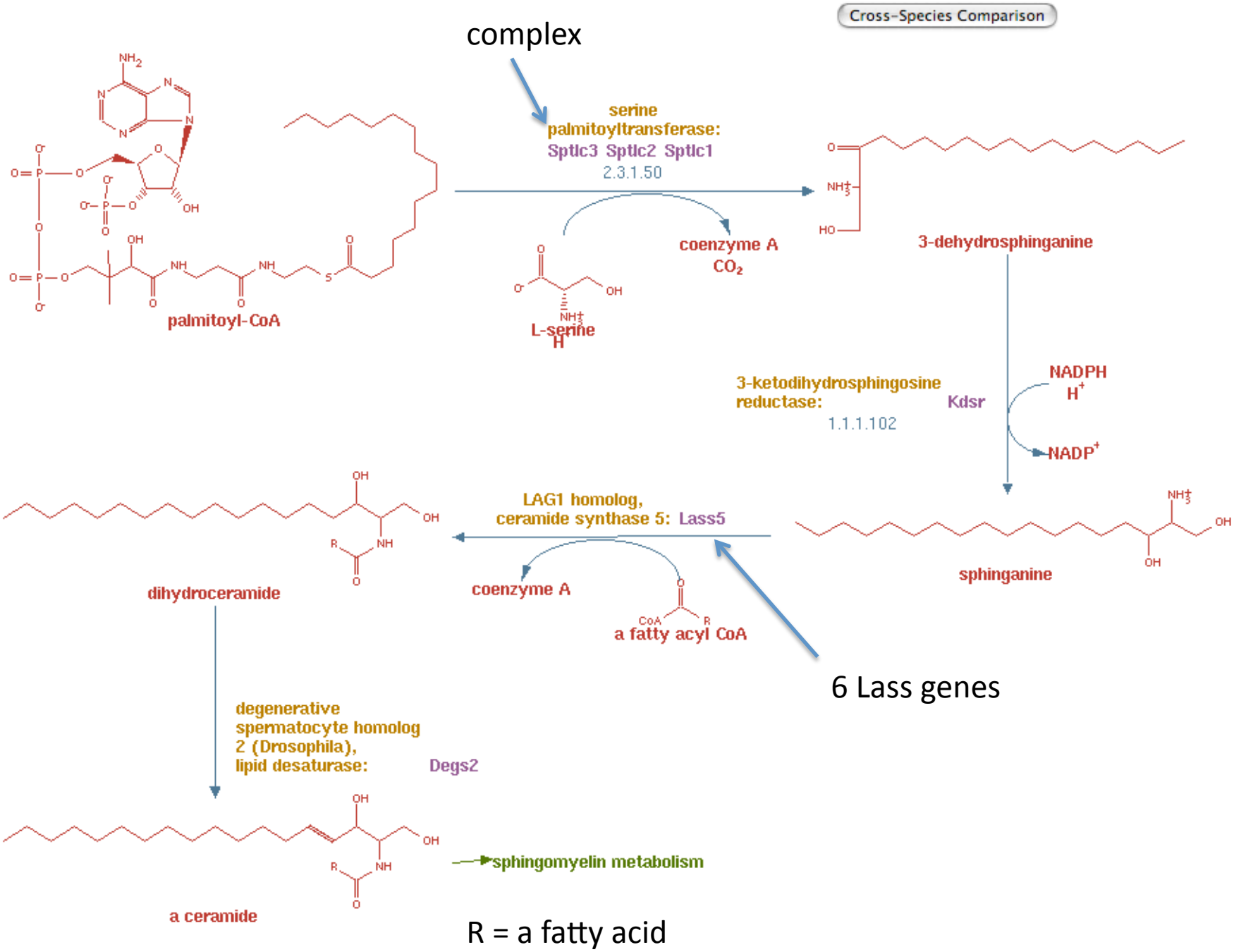
Ceramide receives a lot of attention as a potential endogenous mediator of apoptosis in response to cytokines, antigens, anticancer drugs, or environmental stress inducers [Hannun00]. stimuli can induce sphingomyelin hydrolysis accompanied by the accumulation of ceramide, which occurs before morphological changes. In addition, administration of exogenous ceramid mimic apoptotic cell death. These observations led to the proposal that ceramide has a key role in apoptosis signaling [Cuvillier02].

Credits:

Created 31-Jul-2008 by Evsikov A , JAX

References

Cuvillier02: Cuvillier O (2002). "Sphingosine in apoptosis signaling." Biochim Biophys Acta 1585(2-3);153-62. PMID: 12531549

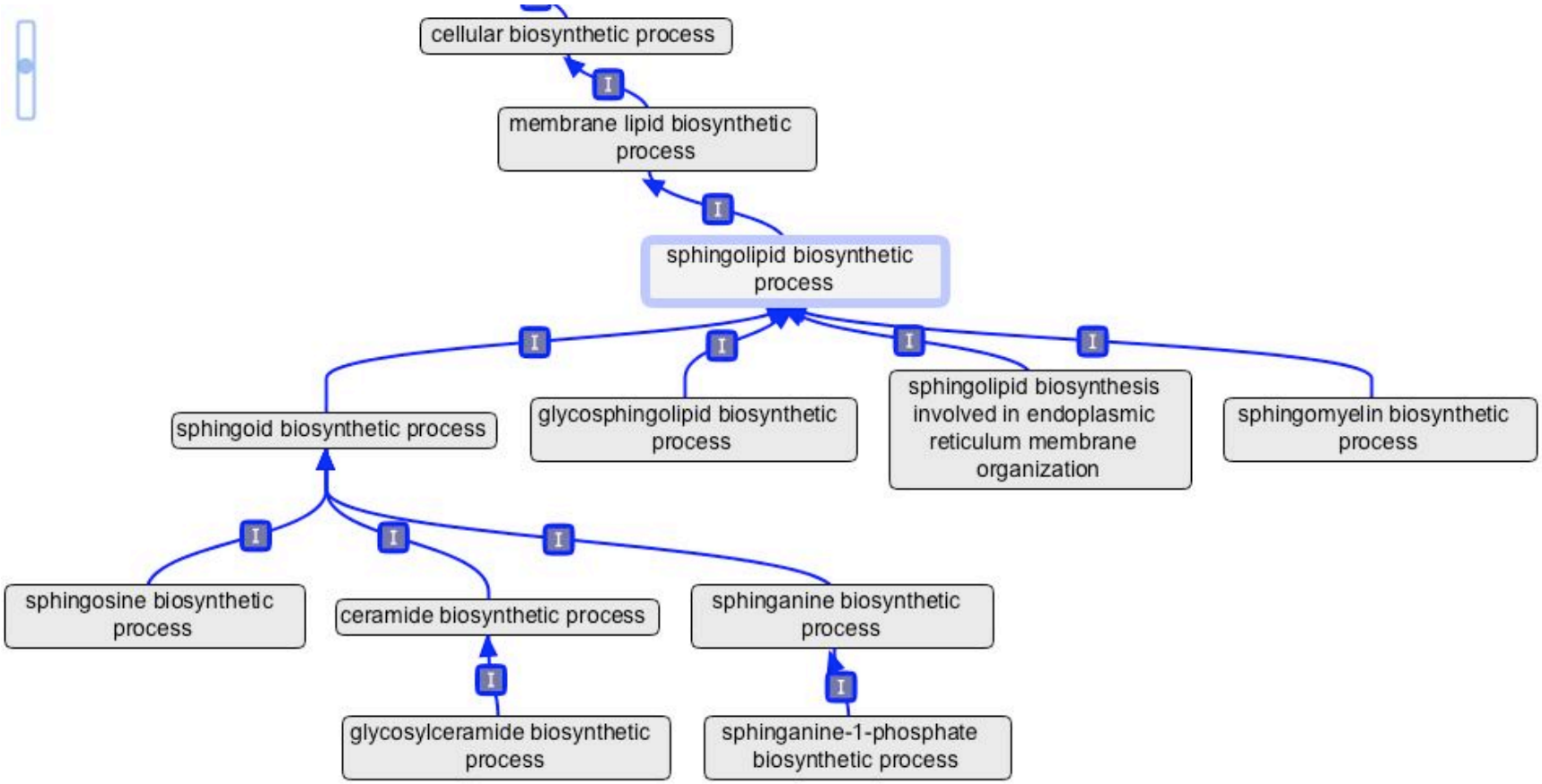

Cross-Species Comparison
complex
serine
palmitoyltransferase:
Sptlc3 Sptlc2 Sptlc1
2.3.1.50
coenzyme A
CO_2
L-serine
H
palmitoyl-CoA
3-dehydrosphinganine
3-ketodihydrosphingosine
reductase:
Kdsr
1.1.1.102
NADPH
H^+
$NADP^+$
sphinganine
LAG1 homolog,
ceramide synthase 5: Lass5
coenzyme A
a fatty acyl CoA
dihydroceramide
6 Lass genes
degenerative
spermatocyte homolog
2 (Drosophila),
lipid desaturase:
Degs2
sphingomyelin metabolism
a ceramide
R = a fatty acid

Searching on ceramide biosynthetic process did not return the expected suspects

- Lack of appropriate annotations
  - Add where possible
  - Lack of mouse literature: can sometimes use ISO to pull in
- Structure of GO graph may not allow annotations to “travel”

What gets measured..

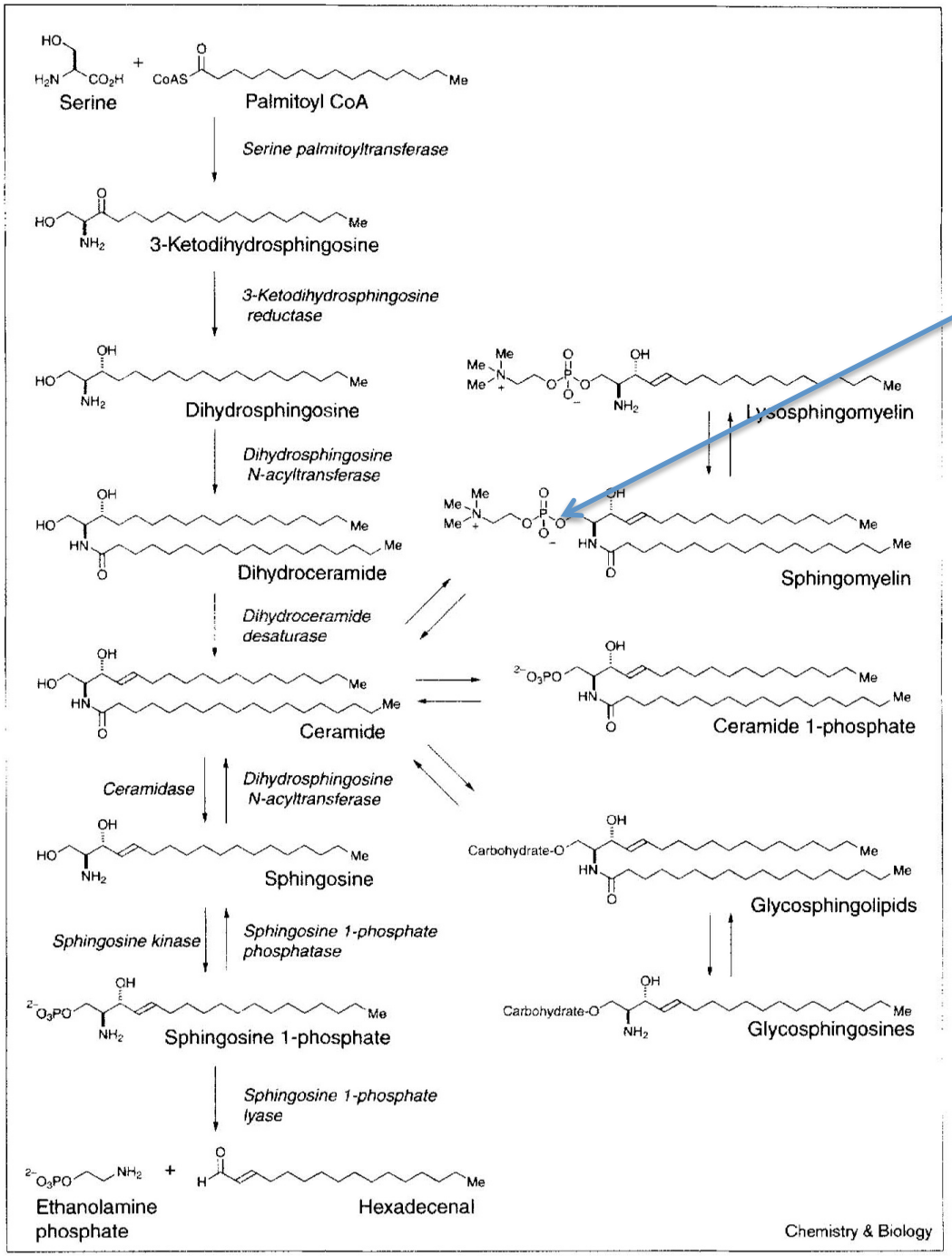
- Knock out of Sptlc1 followed by analysis of sphingomyelin levels; annotate to involvement in sphingomyelin biosynthesis; no link to ceramide biosynthesis (old GO version)
- BUT: since need ceramide to get sphingomyelin….
- Same old problems..
  - Beginning and End definitions
  - One step “transitions”



Typical experiments use knockouts and study levels of sphingosines and sphinanines

Sometimes it helps to define the species

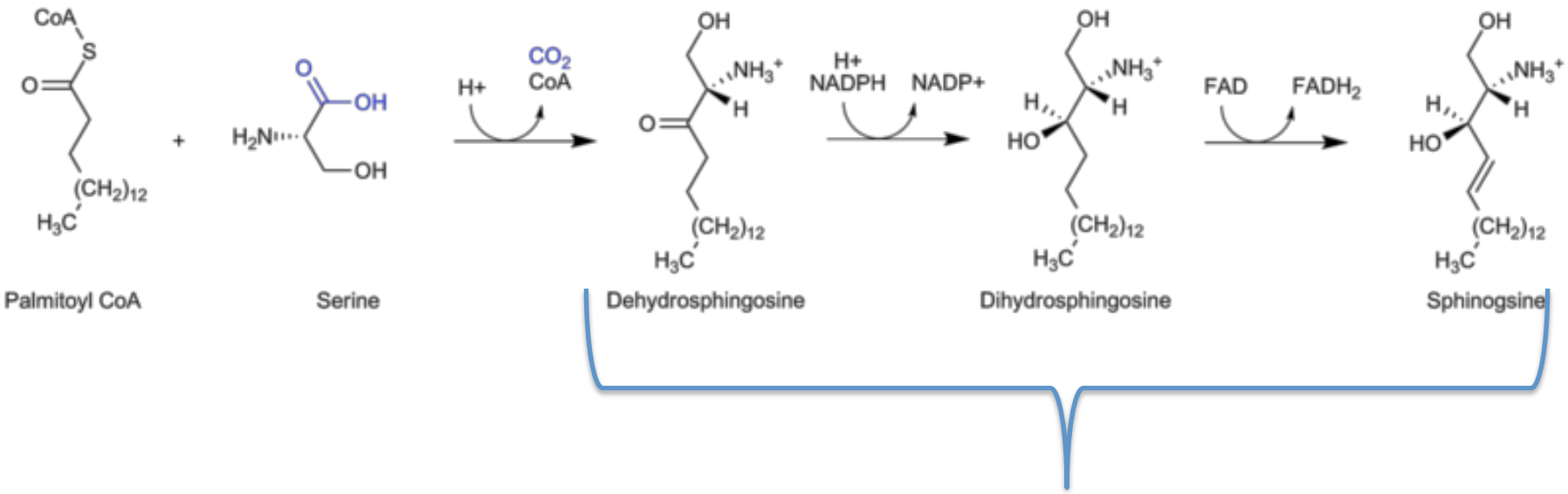
- Sphingosine: a palmitolylated serine derivative
- Ceramide: a sphingosine + fatty acid
- **Sphinogomyelin: a phosphocholinyl ceramide**
- **Glycosphingosine: a glycosylated ceramide**



Sphingomylelin is a "phosphocholinylated ceramide

IUPAC confusion

- Dihydrosphingosine == sphinganine (IUPAC)
- Sphingoids differing from sphinganine in configuration around C2 or C3 should be named not as derivatives of sphinganine…
- Sphinganine specifies the D-erythro configuration.



Sphingoids?

Protein Ontology

Introduction

Why we need

PRO encompasses

- a sub-ontology of proteins based on evolutionary relatedness (ProEvo) and
- a sub-ontology of the multiple protein forms produced from a given gene (ProForm)
  - Splice variants leading to different protein sequences
  - Alternative start codon selection
  - Amino acid modifications
  - Cleaved forms
  - Any combination of the above

GO

- Introduction of column 17 allows for identification of protein variant being annotated up to a point (isoform ids)
  - Q8K428-1, -2, etc
- Column 16 could hold Psi-Mod id to indicate a modified protein form; no position info.

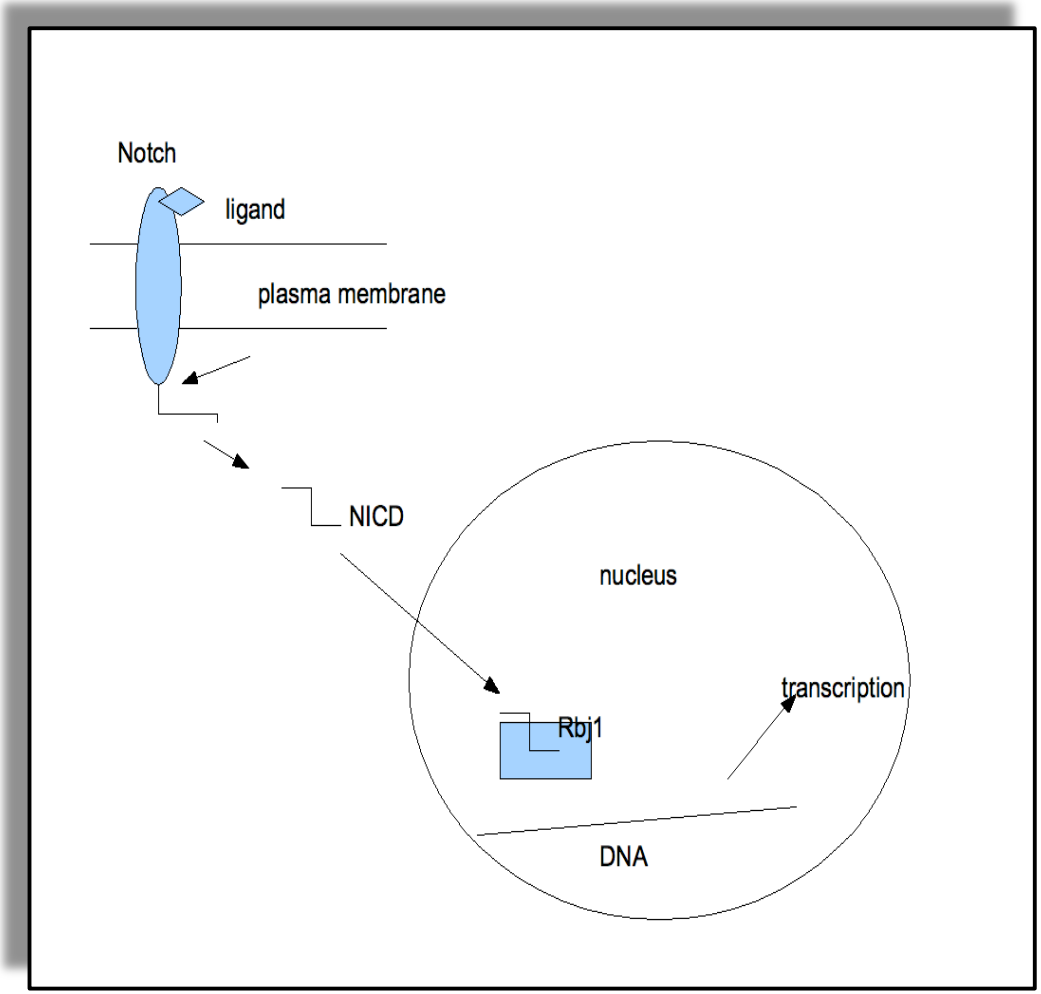
Notch1 whole vs piece

Full protein (Q8K428)is transmembrane
(integral to plasma membrane)
Binds extracellular ligand
(protein binding)

Intracellular domain:
Cleaved from full to give NCID
(Notch intracellular domain)

NCID:
Nuclear
Binds to RBJ-1 Transcription co-activator

NCID has no id!



- Each specific protein form would have it's own PRO id.
- Example:

⊟	PRO:000001719 *transient receptor potential cation channel TRPM8*
⊟	PRO:000001720 *transient receptor potential cation channel TRPM8 isoform 1*
⊞	PRO:000001723 *transient receptor potential cation channel TRPM8 isoform 1 glycosylated form*
	PRO:000002735 *transient receptor potential cation channel TRPM8 isoform 1 unmodified form*
⊞	PRO:000001721 *transient receptor potential cation channel TRPM8 isoform 2*

Species specific ids to be added:
PRO:00000xxxx, transient receptor potential cation channel TRPM8 isoform 1 glycosylated form, murine

Each species specific id associated with a unique UniProtKB:**Q8R4D5-1 (murine isoform 1),** and defined by cross products to Psi-Mod, etc.
So,cleaved fragments would have IDs and be related to the uncleaved form.

⊟	PRO:000011331 *neurogenic locus notch protein 1*
	PRO:000020763 *neurogenic locus notch protein 1 proteolytic cleavage product*

These could become "objects" for annotation.

- A PRO Annotation File (PAF) similar to a GAF is maintained and holds isoform and variant specific GO annotations for the PRO object

- Species specific ids to be generated for modified forms, etc.

- Concern: Overlap/redundancy of annotations
  - E.G. annotations to

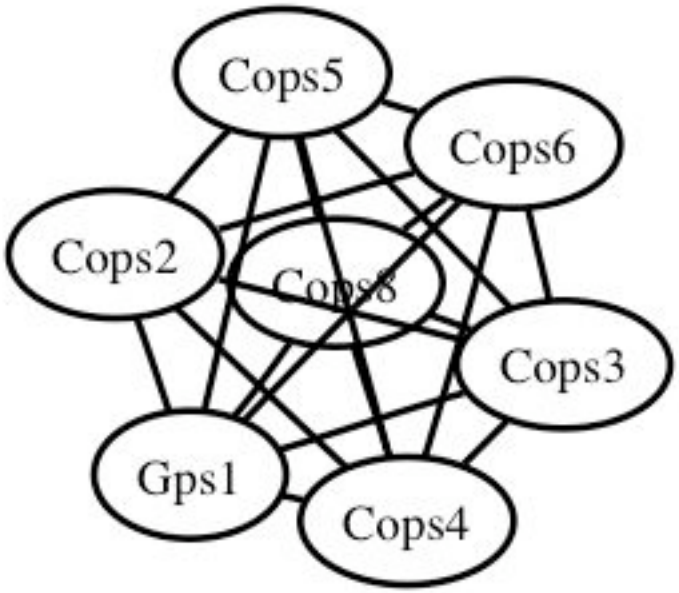
  " transient receptor potential cation channel TRPC2 isoform 2 unmodified form" (PRO:000002739) already exist as annotations to

  Q9R244-2 in column 17 for murine Trpc2

Complexes

GO and PRO

historical



Murine COPS complex as defined by
GO:0005515 annotations and WITH field

GO:0008180 signalosome

Examples of Protein Complex Representation in PRO

- methylcrotonoyl-CoA carboxylase complex, mitochondrial (mouse)
- Isocitrate dehydrogenase complex, mitochondrial (mouse)
- Serine palmitoyl transferase complex (SPT) (mouse)
- *Mitochondrial cytochrome oxidase: Respiratory chain complex IV (in progress)*

methylcrotonoyl-CoA carboxylase complex, mitochondrial (mouse)

- Complex composed of two subunits
  - None of the subunits have isoforms
- **Methylcrotonyl CoA carboxylase** (MCC)
  - biotin-requiring enzyme
  - located in the mitochondria
  - uses bicarbonate as a carboxyl group source to catalyze the carboxylation of a carbon adjacent to a carbonyl group performing the 4th step in processing leucine, an essential amino acid.
- Point mutations and deletion events in the genes lead to MCC deficiency, an inborn error of metabolism
  - vomiting, metabolic acidosis, very low plasma glucose concentration, and very low levels of carnitine in plasma.

methylcrotonoyl-CoA carboxylase complex, mitochondrial : **subunit 1**

```
[Term]
id: PRO:000025354
name: methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial (mouse)
def: "A methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial that is encoded in the
genome of mouse." [PRO:CB]
comment: Category=organism-gene.
xref: UniProtKB:Q99MR8
intersection_of: PRO:000010233 ! methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial
intersection_of: only_in_taxon taxon:10090 ! Mus musculus
```

methylcrotonoyl-CoA carboxylase complex, mitochondrial : **subunit 2**

```
[Term]
id: PRO:000025357
name: methylcrotonoyl-CoA carboxylase beta chain, mitochondrial (mouse)
def: "A methylcrotonoyl-CoA carboxylase beta chain, mitochondrial that is encoded in the genome of
mouse." [PRO:HJD]
comment: Category=organism-gene.
xref: UniProtKB:Q3ULD5
intersection_of: PRO:000010234 ! methylcrotonoyl-CoA carboxylase beta chain, mitochondrial
intersection_of: only_in_taxon taxon:10090 ! Mus musculus
```

methylcrotonoyl-CoA carboxylase **complex,** mitochondrial

[Term]
id: PRO:XXXXXXXX
name: methylcrotonoyl-CoA carboxylase complex, mitochondrial (mouse)
def: "A heterodimeric complex having 3-methylcrotonyl-CoA carboxylase activity. The alpha subunit has a covalently bound biotin essential for the ATP-dependent carboxylation. The beta subunit possess carboxyltransferase activity which presumably is essential for binding to 3-methylcrotonyl-CoA.[ref]"
xref: EC 6.4.1.4
relationship: is_a: GO:0002169! 3-methylcrotonlyl-CoA carboxylase complex, mitochondrial
relationship: has_part: ! PRO:000025354 methylcrotonoyl-CoA carboxylase subunit alpha, mitochondrial (mouse)
relationship: has_part: ! PRO:000025357 methylcrotonoyl-CoA carboxylase beta chain, mitochondrial (mouse)

Isocitrate dehydrogenase complex, mitochondrial (NAD+) (mouse)

- Part of the TCA cycle
- Complex composed of multiple subunits
  - alpha (two isoforms)
  - beta
  - gamma

Two other types:

1. Cytosolic complex: homodimer
2. Mitochondrial (NADP+): homdimer

Isocitrate dehydrogenase complex, mitochondrial (mouse): alpha subunit

```
[Term]
id: PRO:000025358
name: isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial (mouse)
def: "An isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial that is encoded in the genome
of mouse." [PRO:HJD]
comment: Category=organism-gene.
xref: UniProtKB:Q9D6R2
intersection_of: PRO:000008889 ! isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial
intersection_of: only_in_taxon taxon:10090 ! Mus musculus
```

Isocitrate dehydrogenase complex, mitochondrial (mouse): beta subunit and
Isocitrate dehydrogenase complex, mitochondrial (mouse):gamma subunit

have similar stanzas

Isocitrate dehydrogenase complex, mitochondrial (mouse): alpha subunit, isoform 1

[Term]
Id:PRO:000025355
name: isocitrate dehydrogenase 3 (NAD+) alpha, isoform 1
xref: Q9D6R2-1
def: " "
intersection_of: PRO:000025358! isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial (mouse)
intersection_of: only_in_taxon taxon:10090 ! Mus musculus

Isocitrate dehydrogenase complex, mitochondrial (mouse): alpha subunit, isoform 2

[Term]
Id:PRO:000025356
name: isocitrate dehydrogenase 3 (NAD+) alpha, isoform 2
xref: Q9D6R2-2
def: " "
intersection_of: PRO:000025358! isocitrate dehydrogenase [NAD] subunit alpha, mitochondrial (mouse)
intersection_of: only_in_taxon taxon:10090 ! Mus musculus

Isocitrate dehydrogenase complex, mitochondrial (mouse)
the canonical complex

[Term]
id: PRO:*111111*
name: isocitrate dehydrogenase [NAD] complex, mitochondrial (mouse)
dbxref:
def: “ “
is_a: GO:0005962! mitochondrial isocitrate dehydrogenase complex (NAD+)
has_part: PRO:000008889! isocitrate dehydrogenase [NAD] subunit alpha,
mitochondrial **
has_part: PRO:000008890! isocitrate dehydrogenase [NAD] subunit beta,
mitochondrial
has_part: PRO:000008891! isocitrate dehydrogenase [NAD] subunit gamma,
mitochondrial

Isocitrate dehydrogenase complex, mitochondrial (mouse)
the complex with isoform 1 of the alpha subunit

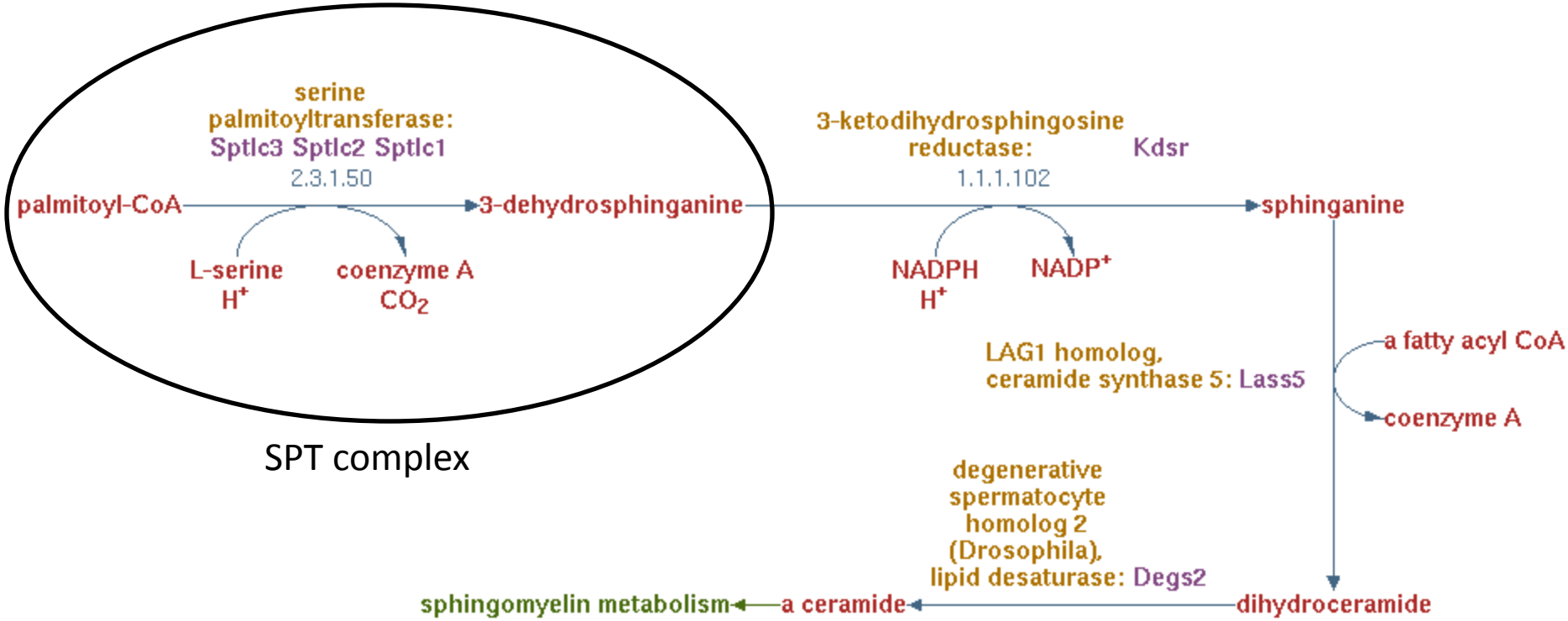
[Term]
id: PRO:222222
name: isocitrate dehydrogenase [NAD] complex, mitochondrial (mouse)
def: “ “
is_a: GO:0005962! mitochondrial isocitrate dehydrogenase complex (NAD+)
Is_a: PRO:*111111*! isocitrate dehydrogenase [NAD] complex, mitochondrial (mouse)
has_part: PRO:000025359! isocitrate dehydrogenase [NAD] subunit beta,
mitochondrial (mouse)
has_part: PRO:000025360! !isocitrate dehydrogenase [NAD] subunit gamma, mitochondrial (mouse
has_part: PRO:000025355 ! isocitrate dehydrogenase 3 (NAD+) alpha, isoform 1

Complex with isoform 2 of alpha has similar stanza

Serine palmitoyltransferase complex (SPT) (mouse)

- Involved in ceramide biosynthesis
- Complex composed of two or three subunits
  - Subunits have no known isoforms
- Ceramide receives a lot of attention as a potential endogenous mediator of apoptosis in response to cytokines, antigens, anticancer drugs, or environmental stress inducers [Hannun00]
- *Only showing the representation of the SPT complex in PRO for this example, not the individual components*

Pathway in MouseCyc



SPT complex

Serine palmitoyltransferase complex (mouse)
canonical complex

[Term]
id:PRO:xxxxx
name: serine palmitoyltransferase complex (mouse)
synonyms: 3-oxosphinganine synthetase, acyl-CoA:serine C-2 acyltransferase decarboxylating
xref: EC 2.3.1.50
def: "An enzyme complex that catalyzes the transfer of a palmitoyl on to serine, forming 3-dehydro-D-sphinganine."
has_part PRO:aaaaaa! Serine palmitoyltransferase, long chain base subunit 1 (mouse)
has_part PRO:bbbbb! Serine palmitoyltransferase, long chain base subunit 2 (mouse)
relationship: is_a GO:0017059! serine C-palmitoyltransferase complex

This SPT complex always has subunit 1 and subunit 2. These subunits have are "part of" SPT.

Serine palmitoyltransferase complex A (mouse)
non canonical complex

[Term]
id:PRO:yyyyy
name: serine palmitoyltransferase complex **A** (mouse)
symbol: SPTa
synonyms: 3-oxosphinganine synthetase, acyl-CoA:serine C-2 acyltransferase decarboxylating
dbxref: EC 2.3.1.50
def: "An enzyme complex that catalyzes the transfer of a palmitoyl on to serine, forming 3-dehydro-D-sphinganine."
is_a PRO:xxxxx! serine palmitoyltransferase complex (mouse)
has_part PRO:ccccc! Serine palmitoyltransferase, long chain base subunit 3 (mouse)
is_a GO:0017059! serine C-palmitoyltransferase complex

Because subunit 3 is not always part of SPT, a different complex is represented.

PRO:yyyyy is_a SPT complex so it should inherit the parts of that complex but also includes subunit 3.

Summary of the SPT complex in mouse

SPT Complex (PRO:xxxxx)

Subunit 1 (Sptlc1) is part_of SPT (PRO:000025362)

Subunit 2 (Sptlc2) is part_of SPT (PRO:000025363)

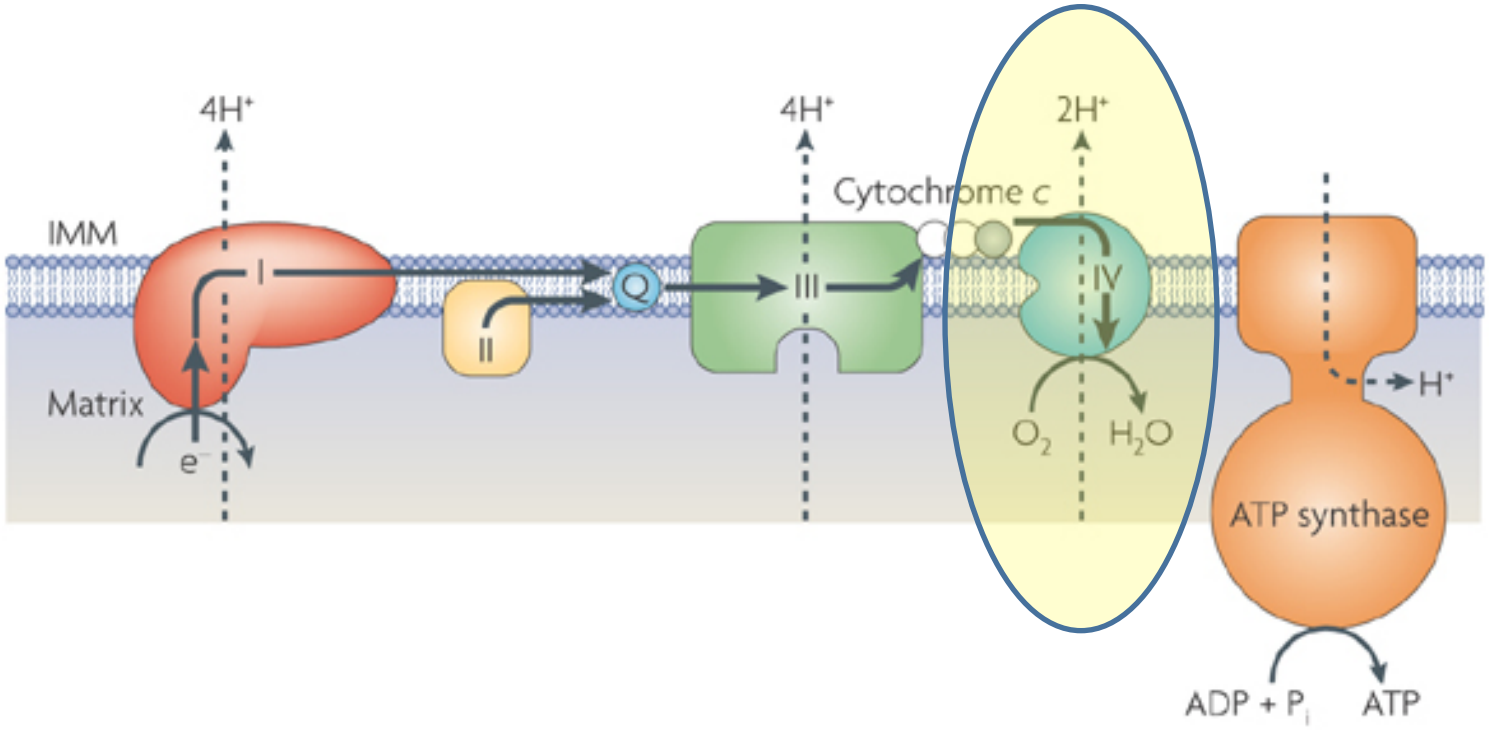
SPT Complex A (PRO:yyyyy)

is_a SPT (PRO:xxxxx)

Subunit 3 (Sptlc3) is part_of SPT A (PRO:000025363)

A more complex complex

Mitochondrial cytochrome oxidase:
<u>Respiratory chain complex IV</u>



Nature Reviews | Molecular Cell Biology

http://www.nature.com/nrm/journal/v9/n7/images/nrm2434-i1.jpg

Respiratory chain complex IV

The respiratory chain complex IV has 13 parts.
Many of those parts (proteins) are encoded by a single gene.
But some of those parts can be encoded by multiple genes depending on the cell/tissue type.

- Mitochondrial encoded parts
  - COX I (mt-Co1) PRO:000025364
  - COX II (mt-Co2) PRO:000025366
  - COX III (mt-Co3) PRO:000025367
- Nuclear encoded mitochondrial parts
  - COX IV (Cox4i1, Cox4i2) (PRO:000025368, PRO:000025369)
  - COX Va (Cox5a) (PRO:000025370)
  - COXVb (Cox5b) (PRO:000025371)
  - COXVIa (Cox6a1, Cox6a2, Cox6a-ps) (PRO:000025372, PRO:000025373, ps is a pseudogene:Is it known to be expressed?
  - COXVIb (Cox6b1, Cox6b2) (PRO:000025374, PRO:000025375)
  - COX Vic (Cox6c) (PRO:000025376)
  - COX VIIa (Cox7a1, Cox7a2, Cox7a2l) (PRO:000025377, PRO:000025378, PRO:000025379)
  - COXVIIb (Cox7c) (PRO:000025383)
  - COXIII (Cox8a, Cox8b, Cox8c) (PRO:000025380, PRO:000025381, PRO:000025382)