

PMN, A Unified Resource For Plant Metabolism

Peifen Zhang

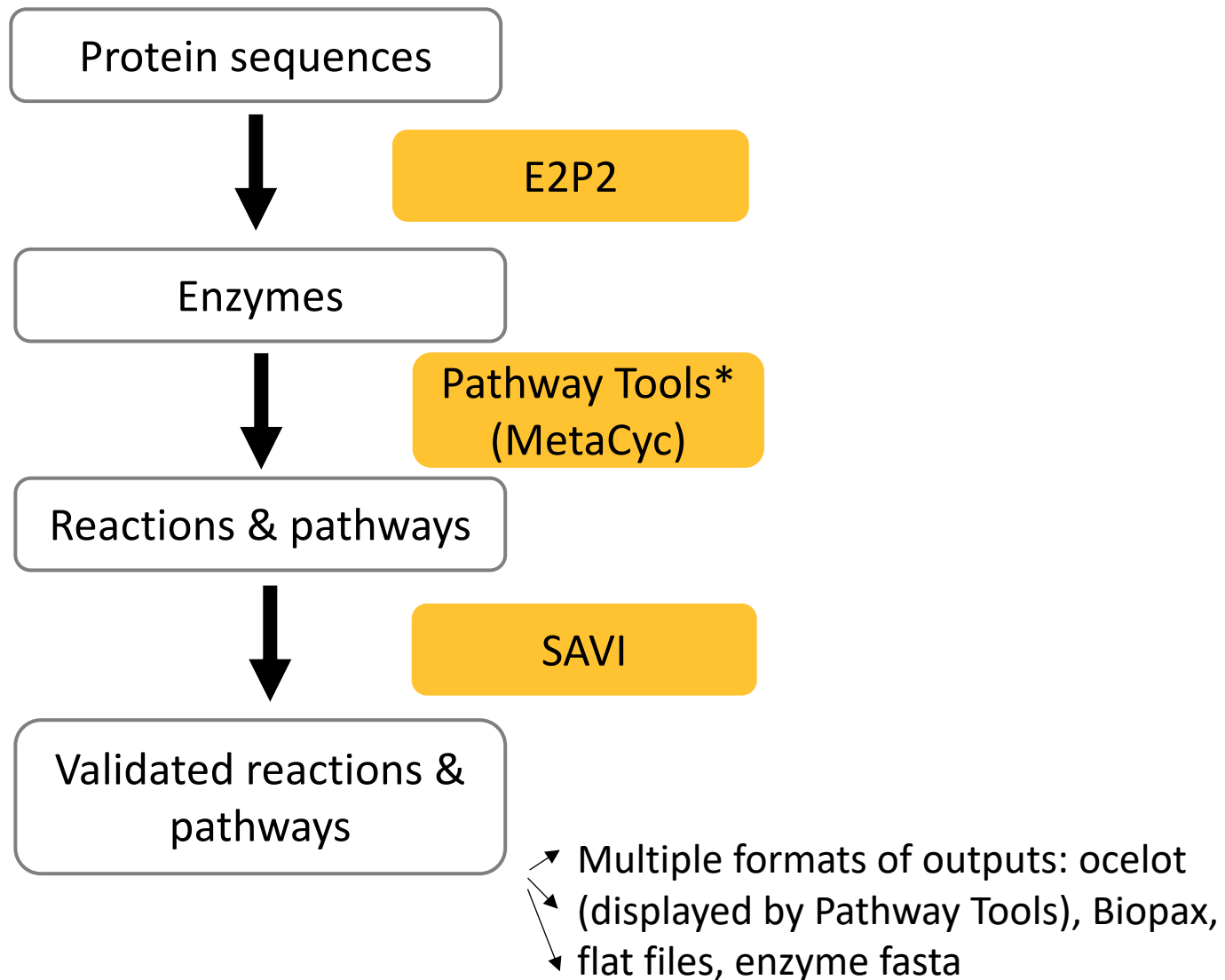
Carnegie Institution for Science

1/17/2017

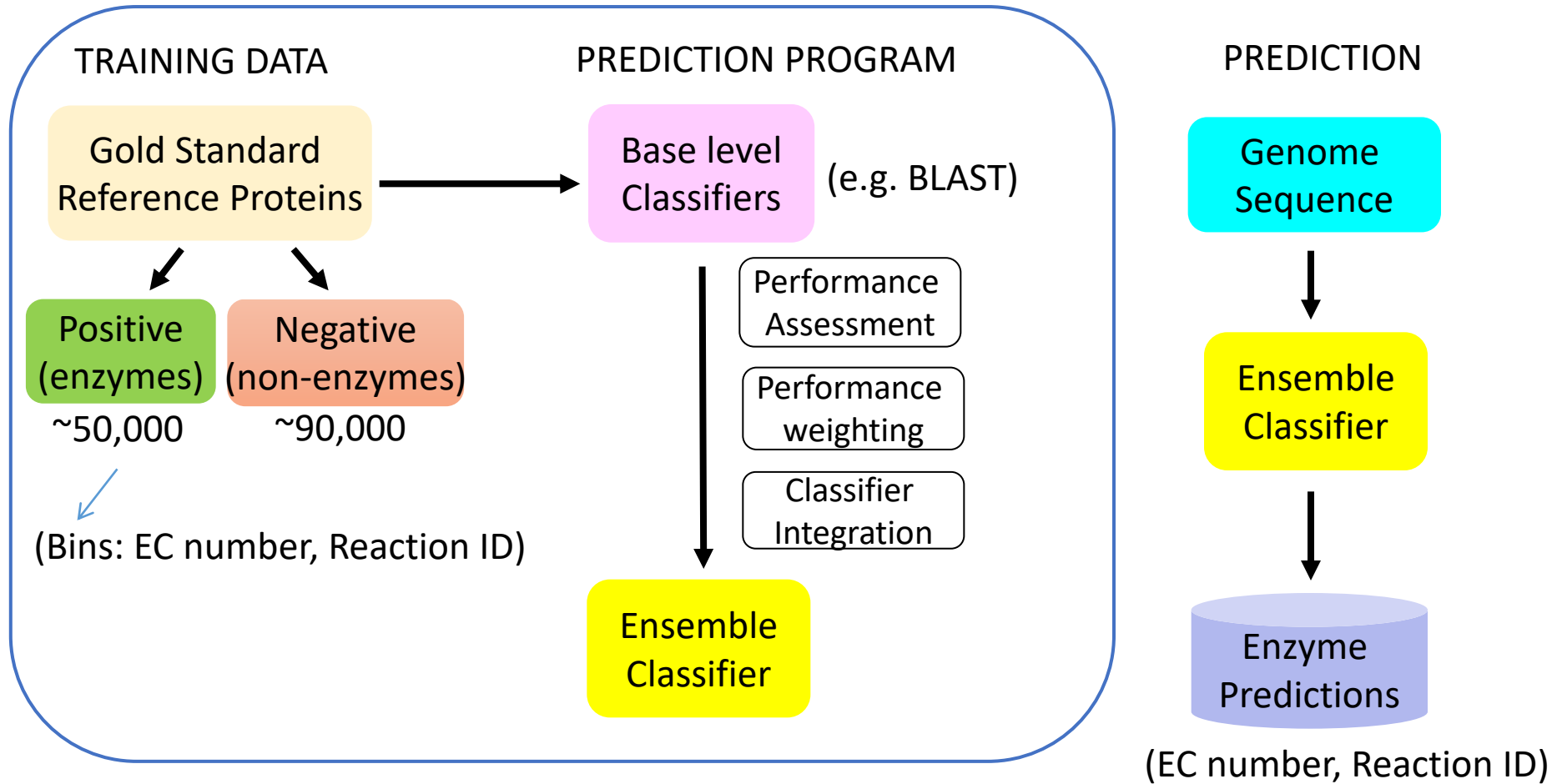
The PMN Resources (plantcyc.org)

- A pipeline for high quality enzyme function prediction and metabolic pathway inference from sequenced plant genomes
- Metabolic pathway databases for 22 plant genomes
- A pan-plant metabolic pathway database PlantCyc
- Experimental information of enzymes and pathways for over 400 plant species

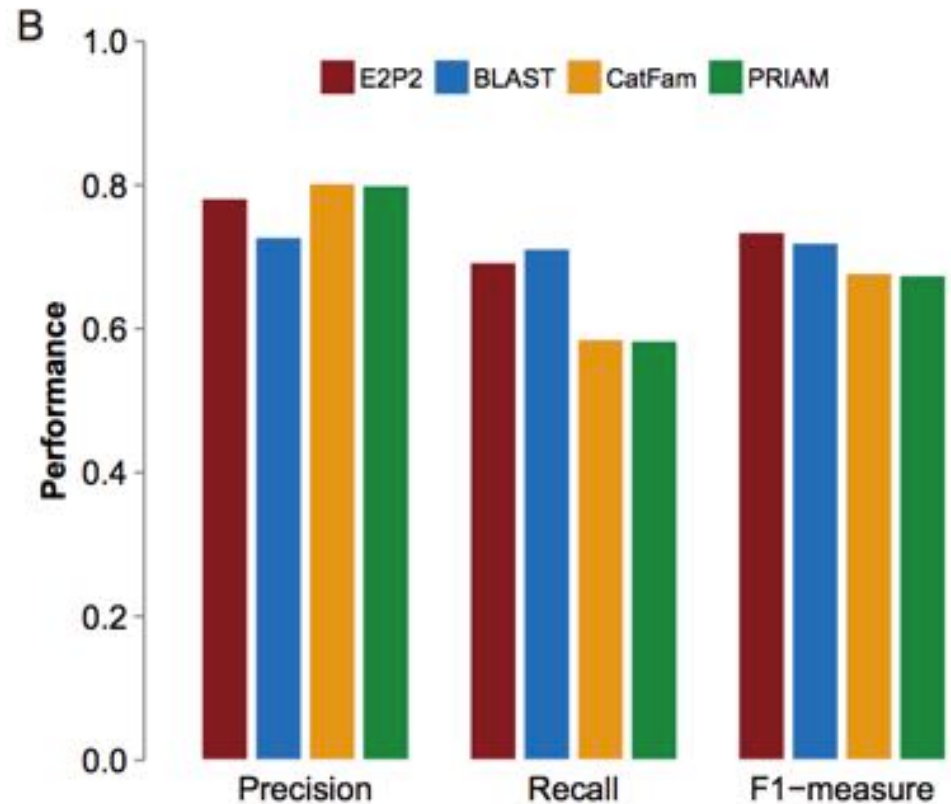
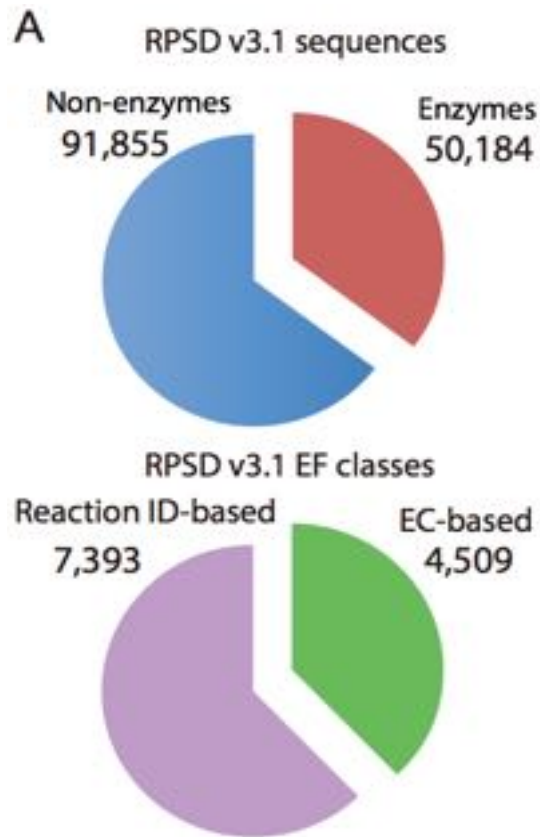
The PMN Pipeline



Ensemble Enzyme Prediction Program (E2P2)

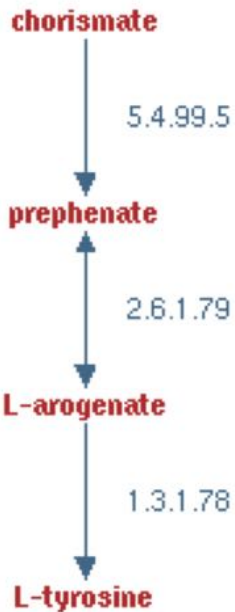


Training Data and Performance of E2P2



Pathway Inference by Pathway Tools

Known pathway
in reference
MetaCyc

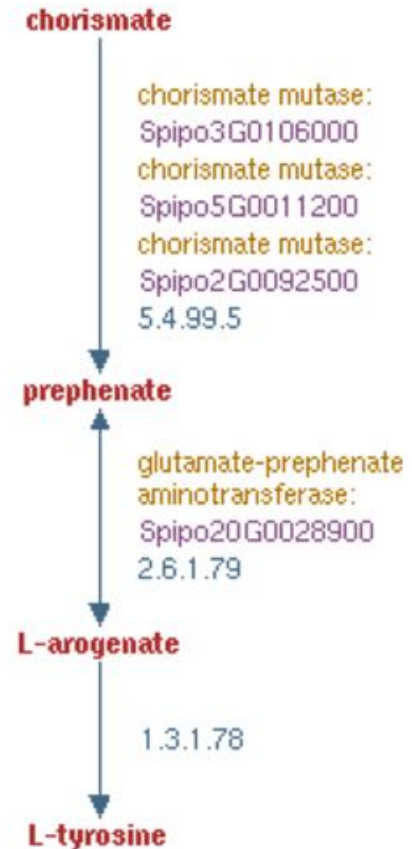


E2P2 enzyme predictions
of the Spirodela genome

Gene - reaction ID
Gene - reaction ID
Gene - reaction ID
.....
.....



Inferred pathway
in SpirodelaCyc



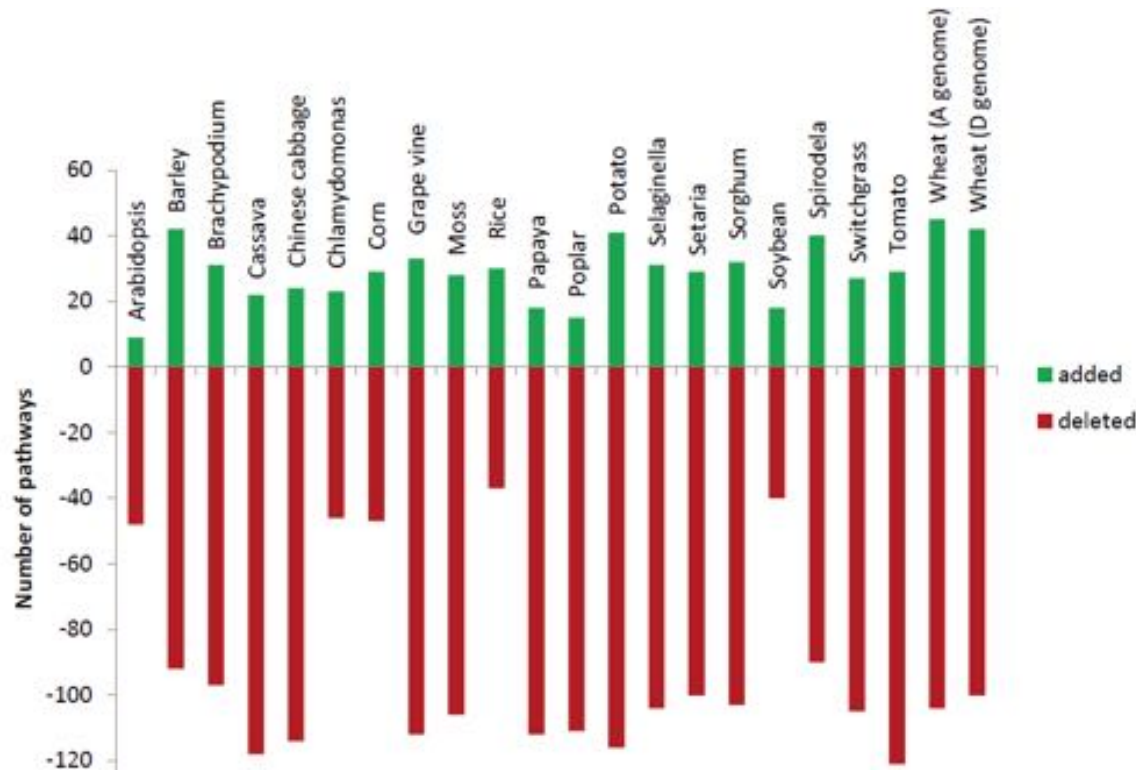
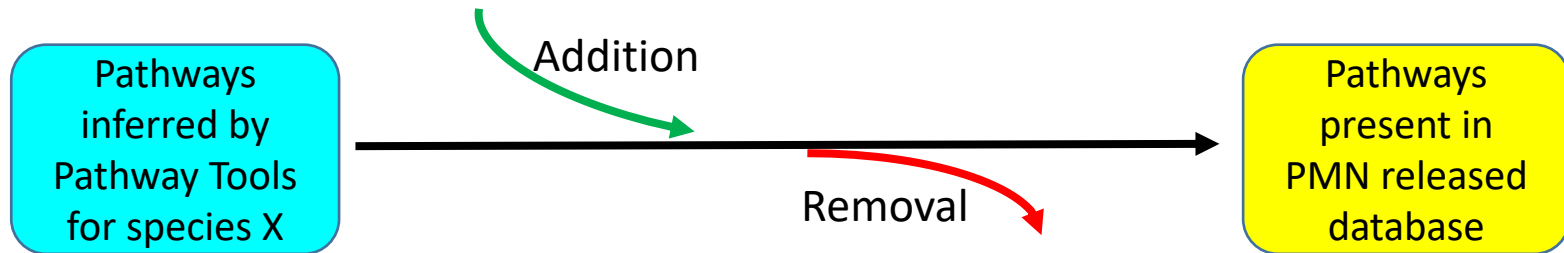
Semi-automated Validation of Pathway-Inference (SAVI)

- Manual curation of rules for each reference pathway
 - **Ubiquitous Plant Pathway**
 - **Accept-if-Predicted Pathway**
 - **Conditionally Accepted Pathway**
 - Key reaction(s)
 - Taxonomic range
 - **Non-PMN Pathway**
 - **Manually Check Pathway**
- Rule-based automatic screening of inferred pathways

Examples of missing pathways and removed pathways (e.g. TomatoCyc)

- Missing then added
 - photorespiration
 - sulfate assimilation pathway
- Removed
 - protein ubiquitylation
 - glycogen degradation
 - serotonin and melatonin biosynthesis
 - C4 photosynthetic carbon assimilation cycle
 - capsaicin biosynthesis
 - dehydroscoulerine biosynthesis

Semi-automated Validation of Pathway-Inference (SAVI)



22 species-specific databases

PlantCyc

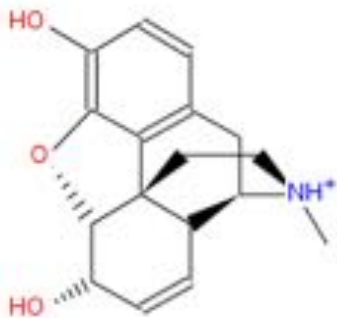
Green

Each database:

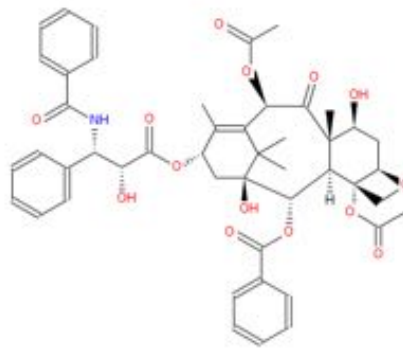
- 2,768 ~ 14,006 Enzymes
- 373 ~ 619 Pathways

Lower level

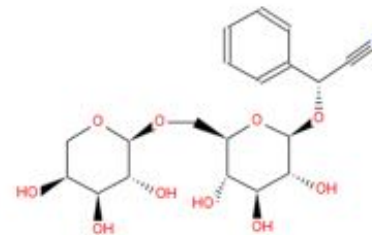
- 400+ species
- 140,581 Enzymes
3,376 with EV-EXP
- 1,213 Pathways
886 with EV-EXP



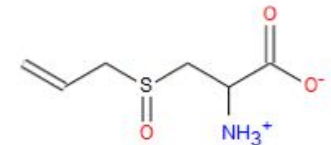
Morphine
(*Papaver somniferum*)



Taxol
(*Taxus brevifolia*)



Vicianin
(*Vicia sativa*)



Alliin
(*Allium sativum*)

http://plantcyc.org

Home

About PMN -

PMN Data -

Help -



PlantCyc

Enter terms then hit S



PLANT METABOLIC PATHWAY DATABASES

The PMN currently houses one multi-species reference database called PlantCyc and 22 species/taxon-specific databases.

- [PMN Content Statistics](#) - stats about pathways, enzymes, reactions, compounds, and more
- [PMN Database Overview](#) - description of data, algorithms, and software used to generate the databases
- [PMN Release Notes](#) - highlights of the most recent PMN release

**Please make sure that your favorite database is selected in the Quick Search bar as you explore the PMN resources

[PlantCyc](#) [[More Information](#)] [[Content Statistics](#)]

PlantCyc provides access to manually curated or reviewed information about shared and unique metabolic pathways present in over 350 plant species.

- Web

Google PMN

- Databases

✓ PlantCyc

AraCyc

BarleyCyc

BrachypodiumCyc

CassavaCyc

ChineseCabbageCyc

ChlamyCyc

CornCyc

GrapeCyc

MossCyc

OryzaCyc

PapayaCyc

PoplarCyc

PotatoCyc

SelaginellaCyc

SetariaCyc

SorghumBicolorCyc

SoyCyc

SpirodelaCyc

SwitchgrassCyc

TomatoCyc

WheatACyc

WheatDCyc

11.0
BASED!

used to announce the release. Highlights include the curated information of 100 steps in over 50 pathways. To see an example, follow

plantcyc.org/PLANT/NEW-

PLANT/NEW-PLANT/NEW-PLANT/NEW-

PMN use case-1: my gene

- Search for the gene
 - find reaction
 - find homologs
 - find pathway
 - find upstream and downstream genes of the pathway
 - compare the pathway to other species

GRMZM2G048121

Quick Search

Gene Search

Searching *Zea mays mays* [change organism database](#)

Search Results for GRMZM2G048121

using database *Zea mays mays* [what is this?](#)[Genes \(1\)](#) | [Proteins \(1\)](#)

Genes Gene/Gene Product pages contain: chromosomal location of gene; depiction of its operon; link to genome browser; detailed summaries and citations; subunit structure (for protein complexes); cofactors, activators, and inhibitors (for enzymes); depiction of regulon (for transcriptional regulators); protein features.

- [GRMZM2G048121_P01](#)

[Turn into a SmartTable](#)

Proteins Gene/Gene Product pages contain: chromosomal location of gene; depiction of its operon; link to genome browser; detailed summaries and citations; subunit structure (for protein complexes); cofactors, activators, and inhibitors (for enzymes); depiction of regulon (for transcriptional regulators); protein features.

- [GRMZM2G048121_P01](#)

[Turn into a SmartTable](#)

Alternative searches:

- [Full text search for GRMZM2G048121 on all pages in this database using Google](#)
- [Full text search for GRMZM2G048121 on all pages of this website using Google](#)



Add to SmartTable

Zea mays mays Reaction: 1.1.1.285

Superclasses: Reactions Classified By Conversion Type → Simple Reactions → Chemical Reactions
Reactions Classified By Substrate → Small-Molecule Reactions

EC Number: 1.1.1.285

Enzymes and Genes:

GRMZM2G048121_P01: GRMZM2G048121_P01
GRMZM2G132875_P02: GRMZM2G132875_P02
GRMZM2G060952_P01: GRMZM2G060952_P01
GRMZM2G087507_P01: GRMZM2G087507_P01
GRMZM2G415579_P01: GRMZM2G415579_P01
GRMZM2G126062_P01: GRMZM2G126062_P01
GRMZM2G000268_P03: GRMZM2G000268_P03
deoxymugineic acid synthase: ZmDMAS1

In Pathway: 2'-deoxymugineic acid phytosiderophore biosynthesis

Zea mays mays

Reaction:
1.1.1.285

OPERATIONS

- Show on Cellular Overview
 - Download atom mapping(s) for this reaction
 - Download Genes
- ### Comparison Operations
- Show this reaction in another database
 - Change organisms/databases for comparison operations
 - Search for this reaction in other databases
 - Species Comparison

1.1.1.285-RXN



EC-1.1.1.285

Gene ID	Gene	Accession	Gene name	Enzymatic activity	Evidence	
GDQC-107929		GRMZM2G048121_P01		GRMZM2G048121_P01	GRMZM2G048121_P01	EV-COMP
GDQC-107573		GRMZM2G132875_P02		GRMZM2G132875_P02	GRMZM2G132875_P02	EV-COMP
GDQC-114339		GRMZM2G060952_P01		GRMZM2G060952_P01	GRMZM2G060952_P01	EV-COMP
GDQC-111722		GRMZM2G087507_P01		GRMZM2G087507_P01	GRMZM2G087507_P01	EV-COMP
GDQC-104580		GRMZM2G415579_P01		GRMZM2G415579_P01	GRMZM2G415579_P01	EV-COMP
GDQC-104554		GRMZM2G126062_P01		GRMZM2G126062_P01	GRMZM2G126062_P01	EV-COMP
GDQC-114475		GRMZM2G000268_P03		GRMZM2G000268_P03	GRMZM2G000268_P03	EV-COMP
G-10667	ZmDMAS1		deoxymugineic acid synthase			EV-EXP



3''-deamino-3''-oxonicotianamine

Add to SmartTable

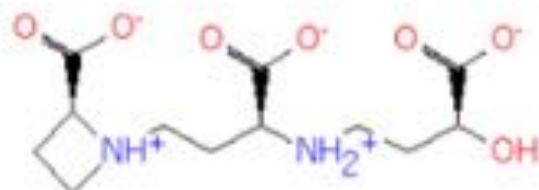
Zea mays Compound: 2'-deoxymugineate

Synonyms: 1-azetidinebutanoic acid, 2'-deoxymugineic acid

Chemical Formula: $C_{12}H_{19}N_2O_7$

Molecular Weight: 303.29 Daltons

Monoisotopic Molecular Weight: 306.1427010713 Daltons



SMILES: C1(CC(C([O-])=O)[N+])1CCC([N+])CCC(O)C(=O)[O-]C(=O)[O-]

InChI: InChI=1S/C12H20N2O7/c15-9(12(20)21)1-4-13-7(10(16)17)2-5-14-6-3-8(14)11(18)19/h7-9,13,15H,1-6H2,(H,16,17)(H,18,19)(H,20,21)/p-1/t7-,8-,9-/m0/s1

InChIKey: InChIKey=CUZKLRTTYZOCSD-CIUDSAMLSA-M

Unification Links: ChEBI:58487, HMDB:HMDB33909, KEGG:C15485, PubChem:25244997

Reactions known to consume the compound:

hydroxylated mugineic acid phytosiderophore biosynthesis :



Reactions known to produce the compound:

2'-deoxymugineic acid phytosiderophore biosynthesis :



Organism Evidence Glyph Enzymes and Genes for hydroxylated mugineic acid phytosiderophore biosynthesis

AraCyc col		This pathway is not marked as present in this organism. No enzymes or genes have been identified for this pathway
------------	--	--

H. vulgare		EC 1.14.11.24 mugineic acid synthase: ids3 MLOC_33486.1: MLOC_33486.1 MLOC_80402.1: MLOC_80402.1 MLOC_17126.1: MLOC_17126.1 MLOC_76208.1: MLOC_76208.1 MLOC_48351.2: MLOC_48351.2 MLOC_1079.1: MLOC_1079.1 MLOC_19745.1: MLOC_19745.1
		EC 1.14.11.25 mugineic-acid 3-dioxygenase: ids2
		EC 1.14.11.25 mugineic-acid 3-dioxygenase: ids-2 mugineic-acid 3-dioxygenase: ids2

O. sativa Japonica Group		EC 1.14.11.24 LOC_Os07g07410.1: LOC_Os07g07410.1
		EC 1.14.11.25 None
		EC 1.14.11.25 None

P. patens		This pathway is not marked as present in this organism. No enzymes or genes have been identified for this pathway
-----------	--	--

S. bicolor		EC 1.14.11.24 Sobic.008G109700.1.p: Sobic.008G109700.1.p Sobic.008G109400.1.p: Sobic.008G109400.1.p Sobic.002G046400.1.p: Sobic.002G046400.1.p Sobic.008G109300.1.p: Sobic.008G109300.1.p Sobic.004G348400.1.p: Sobic.004G348400.1.p Sobic.008G109600.1.p: Sobic.008G109600.1.p
		EC 1.14.11.25 None
		EC 1.14.11.25 None

Z. mays mays		EC 1.14.11.24 GRMZM2G110369_P01: GRMZM2G110369_P01 GRMZM2G146234_P01: GRMZM2G146234_P01 GRMZM2G086231_P01: GRMZM2G086231_P01
		EC 1.14.11.25 None
		EC 1.14.11.25 None



hide

acid
thesis

Data on

ther

ases for

other

oxy
2-exogl

oxy
2-exogl

ss

Use case-2: given a list of genes, find enrichment of pathways

Summary of *Arabidopsis thaliana* col, Strain col, version 14.0

 Authors: Chuan Wang¹, Pascal Schläpfer¹, Taehyong Kim¹, Seung-Hwan Lee¹, Peifen Zhang¹

- My SmartTables
- Public SmartTables
- Special SmartTables

OPERATIONS

ADD TRANSFORM COLUMN

ADD PROPERTY COLUMN

ENRICHMENTS

ADD TRANSFORM COLUMN

ADD PROPERTY COLUMN

Show paged Show all

	Pathways	p-values	Matches
<input type="checkbox"/> 1	flavonoid biosynthesis	3.440792e-6	AT3G51240.1 FL5 AT3G55120.1 AT5G13930.1 4CL3
<input type="checkbox"/> 2	flavonoid biosynthesis (in equisetum)	5.3498883e-5	FL5 AT3G51240.1 AT3G55120.1 AT5G13930.1
<input type="checkbox"/> 3	phenylpropanoid biosynthesis, initial reactions	5.824412e-4	PAL1 PAL2
<input type="checkbox"/> 4	Trehalose Biosynthesis	5.9673557e-4	TPPE TPPG TPPA
<input type="checkbox"/> 5	Organic Solutes Biosynthesis	5.9673557e-4	TPPE TPPG TPPA
<input type="checkbox"/> 6	Trehalose biosynthesis I	5.9673557e-4	TPPE

OPERATIONS

- New ▾
- Export ▾
 - to Spreadsheet File
 - to Fasta File...
 - Export Genes to PortEco Cluster My Genes
 - Export to SDF...
 - Export pathways to Pathway Collage
- Delete ▾
- Column ▾
- Rows ▾
- Paint Data ▾
 - Rename
 - Edit Description
 - Set Operations ...
 - Filter
 - Browse this SmartTable
 - Sharing...
 - Create Frozen Copy...

- Gene
- Gene
- Gene
- Gene
- GO term
- GO term
- GO term
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10

Overlay omics data on pathways

The image shows a screenshot of the PMN (Plant Metabolic Network) web interface. The main content area displays a metabolic pathway diagram for *Arabidopsis thaliana*, Strain col, version 14.0. The pathway is color-coded to represent omics data, with a color scale ranging from -8.180 (purple) to +8.180 (red). The interface includes a search bar at the top right, a navigation menu at the top, and an Omics Viewer Control Panel on the right side. The control panel allows users to adjust the opacity of the omics data overlay and includes buttons for Preferences, Remove Omics Data, and Turn into Group.

PMN

Enter a gene, protein, metabolite or pathway... Quick Search Gene Search
Searching *Arabidopsis thaliana* col change organism database

Home Databases - Search - Metabolism - Analysis - SmartTables - Help -

Omics Viewer / Metabolic Map
Dead-end Metabolites
Checkpoint Reactions

opsis thaliana, Strain col, version 14.0

Cellular Overview of *Arabidopsis thaliana* col

From left to right: zoom in the entire diagram by holding the left mouse button, click on an object for more info, right-click (or click for Mac) for menu

Basic Layer
Cellular Overview

OPERATIONS
Overlay Experimental Data (Omics Viewer) ▾
- Upload Data from File
- Enter/Paste Data from Keyboard
- Import Data from PubEco
- Import Data from GEO

Omics Viewer Control Panel

Control Opacity of Overview

+8.180
+6.360
+4.540
+2.720
+0.900
-0.900
-2.720
-4.540
-6.360
-8.180

Preferences Remove Omics Data Turn into Group ▾

Future Development

- Further enhance accuracy of enzyme function predictions
- Annotate transporters and subcellular location of reactions
- Create 100 pathway databases for sequenced genomes
- Provide online database creation service
- Develop genome-scale metabolic models
- Provide web app for integrated view of metabolism with other information such as gene expression

Team and Funding Sources

- Sue Rhee (PI)
- Pascal Schläpfer (Postdoc)
- Michael Banf (Postdoc)
- Bo Xue (Programmer)
- Garret Huntress (System Admin)
- Peter Karp (Collaborator)

