Introduction to the Plant Metabolic Network: Data and Tools for Analysis and Discovery

kate dreher
biocurator / plant molecular biologist

The Carnegie Institution for Science
Stanford, CA
Plant metabolism

- Plants provide crucial benefits to humanity and the ecosystem

- A better understanding of plant metabolism may contribute to:
  - More nutritious foods
  - More pest-resistant plants
  - Higher photosynthetic capacity and yield in agricultural and biofuel crops
  - New medicines
  - . . . many more applications

- These efforts benefit from access to high quality plant metabolism data
Introduction

The Plant Metabolic Network (PMN) provides a broad network of plant metabolic pathway databases that contain curated information from the literature and computational analyses about the genes, enzymes, compounds, reactions, and pathways involved in primary and secondary metabolism in plants.

The PMN is funded by the National Science Foundation (Grant #: 1026003 and 0540769), governed by an Editorial Board composed of internationally renowned scientists, and executed at the Carnegie Institution for Science, Department of Plant Biology.

News

Come see the PMN at PAG in January

The PMN invites you to come to our workshop at the Plant and Animals Genomics Conference (PAG '09) on Sunday, January 13 in San Diego.

The PMN, the Soil Genomics Network (SoiCyc), and PMN users will present an overview of our databases and examples of their applications to ongoing research.

Early registration ends on November 1.

We hope to see you there!

Additional exciting news . . .
Outline

- PMN overview
- New metabolic database creation at the PMN
- Community data submissions
- PMN data analysis tools
- Downloadable software
Plant Metabolic Network goals

- Transform published results into data-rich metabolic pathways
- Create and deploy improved methods for predicting enzyme function and metabolic capacity using plant genome sequences
- Facilitate data analysis
- Support research, breeding, and education
- Provide public resources:
  - PlantCyc
  - AraCyc
  - 9 additional species-specific databases

www.plantcyc.org
Plant Metabolic Network collaborators

- SRI International – BioCyc project
  - Provide Pathway Tools Software
  - Maintain and update MetaCyc

- Other collaborators / contributors include:
  - MaizeGDB
  - GoFORSYS
  - SoyBase
  - Sol Genomics Network (SGN) / Boyce Thompson Institute
  - Gramene
  - MedicCyc / Nobel Foundation
  - PlantMetabolomics group
  - . . . and more
PMN data content statistics

- Latest PMN release – August 2012

<table>
<thead>
<tr>
<th></th>
<th>Pathways</th>
<th>Enzymes *</th>
<th>Reactions</th>
<th>Compounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>PlantCyc 7.0</td>
<td>978</td>
<td>34663</td>
<td>3617</td>
<td>3455</td>
</tr>
</tbody>
</table>

Over 400 species of plants appear in PlantCyc
PMN data content statistics

- New and updated species-specific databases

<table>
<thead>
<tr>
<th>Database</th>
<th>Pathways</th>
<th>Enzymes *</th>
<th>Reactions</th>
<th>Compounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>AraCyc 10.0</td>
<td>540</td>
<td>7127</td>
<td>3418</td>
<td>3323</td>
</tr>
<tr>
<td>CassavaCyc 2.0</td>
<td>351</td>
<td>7178</td>
<td>2137</td>
<td>1538</td>
</tr>
<tr>
<td>ChlamyCyc 2.0</td>
<td>283</td>
<td>2157</td>
<td>1688</td>
<td>1130</td>
</tr>
<tr>
<td>CornCyc 2.0</td>
<td>373</td>
<td>10527</td>
<td>2198</td>
<td>1752</td>
</tr>
<tr>
<td>GrapeCyc 2.0</td>
<td>359</td>
<td>5341</td>
<td>2143</td>
<td>1571</td>
</tr>
<tr>
<td>MossCyc 1.0</td>
<td>335</td>
<td>5664</td>
<td>2018</td>
<td>1409</td>
</tr>
<tr>
<td>PapayaCyc 1.0</td>
<td>370</td>
<td>3898</td>
<td>2165</td>
<td>1557</td>
</tr>
<tr>
<td>PoplarCyc 5.0</td>
<td>381</td>
<td>8691</td>
<td>2256</td>
<td>1613</td>
</tr>
<tr>
<td>SelaginellaCyc 1.0</td>
<td>329</td>
<td>3958</td>
<td>1986</td>
<td>1403</td>
</tr>
<tr>
<td>SoyCyc 3.0</td>
<td>441</td>
<td>13055</td>
<td>2566</td>
<td>2007</td>
</tr>
</tbody>
</table>

* The term "enzyme" refers to both monomers and complexes found in the databases.
PMN data sources

- In PlantCyc and AraCyc, many enzymes have experimental evidence
PMN data sources

- New databases rely primarily on computational predictions

- E2P2 = Ensemble Enzyme Prediction Pipeline
  - Used to predict enzyme function from protein sequence
  - Developed and maintained at the PMN
E2P2 methodology

1) Relies on an in-house, manually curated gold-standard set of proteins called the Reference Protein Sequence Data Set (RPSD)

- Proteins with known enzymatic function
- Proteins with NO enzymatic activity

![Graph showing comparison between RPSD 1.0 and RPSD 2.0](image)
E2P2 methodology

2) Train the pipeline

- Use protein functional prediction programs on sampled data from RPSD
- Current programs:
  - BLAST
  - PRIAM
  - CatFam
- Predict the function of known proteins in the RPSD
- Learning consists of quantifying how well each program can predict each EC class

\[
\text{maltose} + \text{phosphate} = \text{D-glucose} + \text{beta-D-glucose 1-phosphate}
\]

\[EC_{2.4.1.8} = \text{Maltose Phosphorylase}\]
E2P2 methodology

3) Performances of the individual programs on each EC are stored as a scoring matrix and used as weights when calculating a final prediction.

- Record how well BLAST, CatFam, and PRIAM can correctly predict each type of EC for known enzymes to avoid false positive EC assignments to non-enzymes.

Final predictions are based on integrating preliminary predictions from the individual programs.

Multiple integration schemes were tested:
- E2P2 version 1 uses the average weighted integration scheme.

<table>
<thead>
<tr>
<th>Integration scheme</th>
<th>Final EC prediction is based on:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plurality</td>
<td>The EC class that appears most in the set.</td>
</tr>
<tr>
<td>Majority</td>
<td>The EC class with the absolute majority of predictions in the set.</td>
</tr>
<tr>
<td>Maximum weight</td>
<td>The EC class predicted by the classifier with the highest performance on that class.</td>
</tr>
<tr>
<td>Average weight</td>
<td>The EC class with the highest average performance among classifiers predicting that class.</td>
</tr>
<tr>
<td>Bayesian</td>
<td>The EC class with the highest posterior probability of being the true class given the predicted set of classes.</td>
</tr>
</tbody>
</table>
E2P2 methodology

- Integrating predictions: simplified hypothetical example:
  - Unknown Protein A
    - BLAST result: 2.4.1.115
      - From the E2P2 scoring matrix: For 2.4.1.115 - BLAST is 40% accurate
    - CatFam result = 2.4.1.128
      - For 2.4.1.128 - CatFam is 20% accurate
    - PRIAM result = 2.4.1.116
      - For 2.4.1.116 - PRIAM is 84% accurate
  
E2P2 decision: Protein A = 2.4.1.116
E2P2 implementation and PGDB creation

- Retrieve proteome sequence from an annotated genome
  - Phytozome is the primary source

- Run E2P2 pipeline on proteome data to identify enzyme sequences and classify them with EC numbers.

- Load the proteins and EC assignments into Pathologic to predict metabolic pathways

- Generate an "unvalidated" species-specific database
Pathway validation at the PMN

- Semi-automated validation / incorporation (SAVI) pathway review

- Step 1: Check for UPPs ("Universal" Plant Pathways)
  - Pathways believed to be found in "all" land plants
    - e.g. Photosynthesis, glycolysis, etc.
  - 215 UPP pathways during PMN 7 release (August 2012)
  - UPP decision rules
    - If the pathway was predicted -> KEEP it
    - If the pathway was not predicted -> ADD it
Step 2: Check for NPPs (non-plant/non-PMN pathways)

- Pathways believed to be present only outside of the plant kingdom
  - Glycogen biosynthesis

- Pathways describing processes not covered by the PMN
  - protein ubiquitination, etc.

- 164 NPP pathways used during PMN 7 release (August 2012)

- NPP decision rule:
  - If the pathway was predicted -> REMOVE it
Pathway validation at the PMN

- **Step 3: Manually validate remaining pathways**
  - Keep pathways with experimental evidence for enzymes or key compounds in the given species
  - Keep pathways with enzymes catalyzing "unique" reactions
    - Sometimes results in the inclusion of biologically unlikely pathways
    - Clearly noted with a computational evidence code
    - ** May be removed as a criteria for pathway retention in future PMN releases
Pathway validation at the PMN

- Future SAVI pipeline improvements
  - Add more automated decision rules:
    - Reactions that distinguish between two pathway variants (Primary metabolism)
    - Taxonomic range (Secondary metabolism)
Expert input welcome!!

- To submit data, report an error, volunteer to validate or ask a question
  - Send an e-mail: curator@plantcyc.org
  - Use our feedback form:
  - Visit the PMN poster – P0929 - Monday - 3:00 - 4:30 PM
  - Meet with me at the end of the workshop
  - Schedule an individual meeting with me at PAG (or ASPB)
We thank you publicly!

Together, we can make valuable, high quality databases
PMN databases - part of the BioCyc / MetaCyc family

- Databases at the PMN have:
  - Standard BioCyc display features
  - Standard BioCyc analyses tools
  - PMN-specific features
**PlantCyc Enzyme: phosphatidyltransferase**

**Synonyms**: aminoalcoholphosphotransferase

**Species**: Arabidopsis thaliana

**Gene**: AT5G35660

**Sequence Length**: 1145 AAs

**Unification Links**: Entrez: AAC61769, Phytozone Plant Orthologs: AT5G35660.1

**Gene-Reaction Schematic**:

- **Enzyme**: At-ATG135660
- **Gene**: At-ATG25565
- **Compound**: Mt-AW774344
- **Gene**: Mtr-AAP31
- **Glycine max**: Mt-AW986770
- **Medicago truncatula**: Mt-BG146449

**MultiFun Terms**: UNCLASSIFIED

**Enzymatic reaction of**: cholinephosphotransferase (phosphatidyltransferase)

\[1,2-diacylglycerol + CDP-choline \rightleftharpoons \text{a phosphatidylcholine} + \text{CMP}\]

The reaction direction shown, that is, \( A + B \rightleftharpoons C + D \) versus \( C + D \rightleftharpoons A + B \), is in accordance with the Enzyme Commission system.

Reversibility of this reaction is unspecified.

**In Pathways**: choline biosynthesis, phosphatidylcholine biosynthesis

**Summary**: The enzyme is bifunctional that utilizes CDP-choline and CDP-ethanolamine in synthesizing phosphatidylcholine and phosphatidylethanolamine, respectively. Its cholinephosphotransferase activity is greater than its ethanolaminephosphotransferase activity.

**Inhibitors (Unknown)**: CMP [Goode99], Ca^{2+} [Goode99]
General searching at the PMN

- Pathway Tools quick search bar
Specialized search options
Comparative analyses
Comparative analyses

Select which set(s) of comparative-analysis tables you wish to generate:

- Organisers: breakdown by principal data types.
- Reactions: breakdown by type of substrate, by EC Number, by number of isoforms, etc.
- Pathways: breakdown by pathway class, information on pathway roles.
- Compounds: small molecules that act as substrates, enzyme activators/inhibitors/cofactors.
- Genes/Proteins: breakdown of genes by product type, ontology, annotations; breakdown of proteins with activating/inhibiting/cofactor, multifunctional enzymes.

Specify List of Organism Databases

Select One or More Databases:

7 databases currently selected.
- Arabidopsis thaliana
col
- Glycine max
- Manihot esculenta esculenta
- PlantCyc
- Populus trichocarpa
- Vitis vinifera
- Zea mays mays
Comparative analyses

<table>
<thead>
<tr>
<th>Number of Enzymes per Reaction</th>
<th>ArcCyc col</th>
<th>PlantCyc</th>
<th>V.很满意</th>
<th>G. max</th>
<th>P. trichoceros</th>
<th>Z. macrospora</th>
<th>Z. mays sp.</th>
<th>A. thaliana excelentes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>875 (75%)</td>
<td>840 (75%)</td>
<td>795 (75%)</td>
<td>151 (75%)</td>
<td>215 (55%)</td>
<td>147 (115%)</td>
<td>271 (155%)</td>
<td>147 (55%)</td>
</tr>
<tr>
<td>2</td>
<td>497 (135%)</td>
<td>262 (135%)</td>
<td>247 (135%)</td>
<td>128 (135%)</td>
<td>238 (135%)</td>
<td>156 (135%)</td>
<td>180 (115%)</td>
<td>156 (135%)</td>
</tr>
<tr>
<td>3</td>
<td>149 (21%)</td>
<td>115 (21%)</td>
<td>114 (21%)</td>
<td>115 (21%)</td>
<td>147 (21%)</td>
<td>115 (21%)</td>
<td>182 (21%)</td>
<td>115 (21%)</td>
</tr>
<tr>
<td>4</td>
<td>184 (65%)</td>
<td>88 (65%)</td>
<td>197 (65%)</td>
<td>114 (65%)</td>
<td>145 (65%)</td>
<td>90 (65%)</td>
<td>139 (65%)</td>
<td>90 (65%)</td>
</tr>
<tr>
<td>5</td>
<td>119 (45%)</td>
<td>70 (45%)</td>
<td>59 (45%)</td>
<td>104 (45%)</td>
<td>113 (45%)</td>
<td>46 (45%)</td>
<td>87 (55%)</td>
<td>46 (45%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EC Category</th>
<th>ArcCyc col</th>
<th>PlantCyc</th>
<th>V.很满意</th>
<th>G. max</th>
<th>P. trichoceros</th>
<th>Z. macrospora</th>
<th>Z. mays sp.</th>
<th>A. thaliana excelentes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 -- Oxidoreductases</td>
<td>817 (35%)</td>
<td>714 (35%)</td>
<td>459 (35%)</td>
<td>394 (35%)</td>
<td>474 (35%)</td>
<td>470 (35%)</td>
<td>476 (35%)</td>
<td>476 (35%)</td>
</tr>
<tr>
<td>2 -- Transferases</td>
<td>781 (32%)</td>
<td>776 (32%)</td>
<td>481 (32%)</td>
<td>407 (32%)</td>
<td>497 (32%)</td>
<td>492 (32%)</td>
<td>489 (32%)</td>
<td>489 (32%)</td>
</tr>
<tr>
<td>3 -- Hydrolases</td>
<td>486 (17%)</td>
<td>381 (17%)</td>
<td>286 (17%)</td>
<td>256 (17%)</td>
<td>276 (17%)</td>
<td>211 (17%)</td>
<td>211 (17%)</td>
<td>211 (17%)</td>
</tr>
<tr>
<td>4 -- Lyases</td>
<td>189 (8%)</td>
<td>112 (8%)</td>
<td>109 (8%)</td>
<td>134 (8%)</td>
<td>114 (8%)</td>
<td>124 (8%)</td>
<td>122 (8%)</td>
<td>122 (8%)</td>
</tr>
<tr>
<td>5 -- Dehydrases</td>
<td>94 (45%)</td>
<td>77 (45%)</td>
<td>57 (45%)</td>
<td>72 (45%)</td>
<td>59 (45%)</td>
<td>59 (45%)</td>
<td>56 (45%)</td>
<td>56 (45%)</td>
</tr>
<tr>
<td>6 -- Isomerases</td>
<td>131 (75%)</td>
<td>93 (75%)</td>
<td>85 (75%)</td>
<td>85 (75%)</td>
<td>75 (75%)</td>
<td>74 (75%)</td>
<td>79 (75%)</td>
<td>79 (75%)</td>
</tr>
<tr>
<td>Total reactions with full or partial EC Numbers</td>
<td>2,350</td>
<td>2,071</td>
<td>1,478</td>
<td>1,868</td>
<td>1,531</td>
<td>1,531</td>
<td>1,531</td>
<td></td>
</tr>
</tbody>
</table>
Comparative analyses

<table>
<thead>
<tr>
<th>EC Category: 4 — Lysates</th>
<th>AnaCyc cell</th>
<th>FerretCyc</th>
<th>V. vinifera</th>
<th>G. max</th>
<th>R. trichocarpa</th>
<th>Z. mays</th>
<th>M. esculenta</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1.1.2: 3-ketoacid + H⁺ → 2-methylamine + CO₂</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.1.1.3: D-ribose-1,5-bisphosphate + O₂ → 2-phosphoglycerate + 3-phosphoglycerate + 2-H⁺</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>4.1.1.4: L-serine + H⁺ → ethanolamine + CO₂</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.1.1.5: UDP-D-glucuronic acid + H⁺ → UDP-D-glucose + CO₂</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.1.1.15: L-glutamate + H⁺ → CO₂ + 4-aminobutirate</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.1.1.17: L-ornithine + H⁺ → CO₂ + putrescine</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.1.1.18: L-lysine + H⁺ → CO₂ + cadaverine</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Sequence-based search options

- experimental or computational support
- plants only

- experimental support
- all kingdoms
PlantCyc Pathway: ethylene biosynthesis from methionine

methionine → L-methionine

methionine adenosyltransferase (AT): At-SAM-1

H₂O ATP diposphosphate phosphate

S-adenosyl-L-methionine

1-aminoacyclopropane-1-carboxylate synthase (ACS): At-ACS1
1-aminoacyclopropane-1-carboxylate synthase (ACS): At-ACS8
1-aminoacyclopropane-1-carboxylate synthase (ACS): At-ACS11
1-aminoacyclopropane-1-carboxylate synthase (ACS): At-ACS10
1-aminoacyclopropane-1-carboxylate synthase (ACS): At-ACS5
1-aminoacyclopropane-1-carboxylate synthase (ACS): At-ACS4
1-aminoacyclopropane-1-carboxylate synthase (ACS): At-ACS2

S-methyl-S-thioselenocysteine

1-aminoacyclopropane-1-carboxylate oxidase (ACO): At-ACO
1-aminoacyclopropane-1-carboxylate oxidase (ACO): At-ACO

ethylene

2 H₂O L-dehydro-ascorbate hydrogen cyanide CO₂

1-aminocyclopropane-1-carboxylate
Working with Groups

- Opportunities
  - Create custom data sets
  - Examine experimental results
  - Perform enrichment analyses
  - Share data

- Requires free registration
Register at the PMN

- Registration and login available from every data and search page
  - **Not available on the home page yet.**

![Image of login page]

![Image of create account page]

Sign up now!
Register at the PMN

- Set your preferences

- Login . . .and start working with groups!!
Work with groups
Work with groups

Welcome to Webgroups

Web Groups is a new feature on the BioCyc web sites. A Web Group is a collection of BioCyc objects, such as genes or pathways, together with associated data, that can be created, edited, manipulated, and shared on the web. For more information, see the documentation.

Special groups

- All compounds of PlantCyc
- All enzymes of PlantCyc
- All genes of PlantCyc
- All pathways of PlantCyc
- All proteins of PlantCyc
- All reactions of PlantCyc
- All organisms
### Work with groups

You have read-only access to this group, but you can make your own writeable copy of it.

**Group: All compounds of PlantCyc**

3456 rows of compounds

<table>
<thead>
<tr>
<th>Compounds</th>
<th>Structure of Compound</th>
</tr>
</thead>
<tbody>
<tr>
<td>(+)-(4R)-limonene</td>
<td><img src="image" alt="Limonene Structure" /></td>
</tr>
<tr>
<td>(+)-(S)-carvone</td>
<td><img src="image" alt="Carvone Structure" /></td>
</tr>
</tbody>
</table>

**Owner:** Kate Greiner, **Created:** 09-Jan-2013 13:59:55
Create groups from searches

Pathways

- Activation/Inactivation/Interconversion (3)
- Biosynthesis (258)
  - Amines and Polyamines Biosynthesis (5)
  - Amino Acids Biosynthesis (49)
  - Aminoacyl-tRNA Charging (1)
  - Aromatic Compounds Biosynthesis (7)
  - Carbohydrates Biosynthesis (23)
  - Cell Structures Biosynthesis (6)
  - Cofactors, Prosthetic Groups, Electron Carriers Biosynthesis (51)
  - Fatty Acids and Lipids Biosynthesis (44)
  - Hormones Biosynthesis (12)
  - Metabolic Regulators Biosynthesis (2)
  - Nucleosides and Nucleotides Biosynthesis (16)
  - Other Biosynthesis (3)
  - Secondary Metabolites Biosynthesis (61)
    - Fatty Acid Derivatives Biosynthesis (1)
    - Nitrogen-Containing Secondary Compounds Biosynthesis (6)
    - Phenylpropanoid Derivatives Biosynthesis (22)
    - Phytoalexins Biosynthesis (2)
    - Sugar Derivatives Biosynthesis (6)
    - Terpenoids Biosynthesis (24)
    - Acetaldehyde biosynthesis II
    - Oxalate biosynthesis
- Degradation/Utilization/Assimilation (82)
- Denitrogenification (6)
Create groups from searches
Create groups from searches

<table>
<thead>
<tr>
<th>Pathway Name</th>
<th>#Reactions</th>
<th>Evidence</th>
<th>Enzymes of pathway</th>
</tr>
</thead>
<tbody>
<tr>
<td>flavonoid biosynthesis</td>
<td>5</td>
<td>Inferred by a human based on computational evidence</td>
<td>cassava4.1_023640m</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cassava4.1_004825m</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cassava4.1_009910m</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cassava4.1_004643m</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cassava4.1_004744m</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cassava4.1_005052m</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cassava4.1_026016m</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cassava4.1_004055m</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cassava4.1_032243m</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cassava4.1_033129m</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cassava4.1_004723m</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cassava4.1_027178m</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cassava4.1_005114m</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cassava4.1_030552m</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>cassava4.1_004658m</td>
</tr>
</tbody>
</table>
Create groups from searches

<table>
<thead>
<tr>
<th></th>
<th>proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>cassava4.1_000277m</td>
</tr>
<tr>
<td>2</td>
<td>cassava4.1_000330m</td>
</tr>
<tr>
<td>3</td>
<td>cassava4.1_001850m</td>
</tr>
<tr>
<td>4</td>
<td>cassava4.1_001945m</td>
</tr>
<tr>
<td>5</td>
<td>cassava4.1_001987m</td>
</tr>
<tr>
<td>6</td>
<td>cassava4.1_002130m</td>
</tr>
<tr>
<td>7</td>
<td>cassava4.1_002175m</td>
</tr>
<tr>
<td>8</td>
<td>cassava4.1_002176m</td>
</tr>
<tr>
<td>9</td>
<td>cassava4.1_002229m</td>
</tr>
<tr>
<td>10</td>
<td>cassava4.1_002358m</td>
</tr>
<tr>
<td>11</td>
<td>cassava4.1_002363m</td>
</tr>
<tr>
<td>12</td>
<td>cassava4.1_002424m</td>
</tr>
</tbody>
</table>

Choose a property: Catalysts, Citations, Comment, Common Name, CREATION-DATE, CREATOR, Credits, Dblinks, Gene, Molecular Weight, From-Sequence, Synonyms.
Create groups from searches

<table>
<thead>
<tr>
<th>Proteins</th>
<th>Catalyzes</th>
<th>Pathways of enzyme</th>
<th>Reactions catalyzed by enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>cassava4.1_033512m</td>
<td>caffeoyl-CoA O-methyltransferase</td>
<td>phenylpropanoid biosynthesis suberin biosynthesis</td>
<td>caffeoyl-CoA + S-adenosyl-L-methionine → feruloyl-CoA + S-adenosyl-L-homocysteine + H⁺</td>
</tr>
</tbody>
</table>

Search term: cassava4.1_033512m
Search target: families
Member filtering: none
Search type: Symbols/Identifiers/Deflines

NODE DESCRIPTION

<table>
<thead>
<tr>
<th>NODE</th>
<th>DESCRIPTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>41</td>
<td>Fabid</td>
</tr>
<tr>
<td>1</td>
<td>Manihot esculenta cassava4.1_033512m</td>
</tr>
</tbody>
</table>
Upload groups
Upload groups
Perform enrichment analyses

<table>
<thead>
<tr>
<th>Pathways</th>
<th>p-values</th>
<th>Matches</th>
</tr>
</thead>
<tbody>
<tr>
<td>detoxification of reactive carbonyls in chloroplasts</td>
<td>1.86815e-5</td>
<td>AKR4C9, AOR, ChlADR</td>
</tr>
<tr>
<td>IAA biosynthesis VII</td>
<td>3.9579987e-5</td>
<td>IBR3, IBR1</td>
</tr>
<tr>
<td>flavonol biosynthesis</td>
<td>4.4805877e-4</td>
<td>AT4G22880, GA20ox4, CYP82G1, CYP98A9</td>
</tr>
<tr>
<td>gibberellin biosynthesis II (early C-3 hydroxylation)</td>
<td>8.153568e-4</td>
<td>GA20ox3, GA20ox4</td>
</tr>
<tr>
<td>starch degradation II</td>
<td>0.001082974</td>
<td>BAM3, SEX4</td>
</tr>
<tr>
<td>fatty acid α-oxidation</td>
<td>0.001082974</td>
<td>AT3G48000, DOX2</td>
</tr>
<tr>
<td>gibberellin biosynthesis I (non C-3, non C-13 hydroxylation)</td>
<td>0.001082974</td>
<td>GA20ox4, GA20ox3</td>
</tr>
</tbody>
</table>
Download and share groups
Interpreting quantitative Omics results
Cellular Overview of Arabidopsis thaliana co1

Pan to rightto pulldown the entire diagram by holding the left mouse button, click on an object for more info, right click for settings, and press any key to hide.

Display Cellular Overview

Overflow Experimental Data (Omics Viewer)

- Highlight Pathway(s)
- Highlight Reaction(s)
- Highlight Gene(s)
- Highlight Enzyme(s)
- Highlight Compound(s)
- Highlight/Omics View Based on Names and Frame IDs
- Clear All Highlighting
- Show Legend
- Generate Bookmark for Current Cellular Overview

Help
Data analysis with the Metabolic Map / Omics Viewer

- Display experimental data on metabolic map

  - Data types:
    - Genes - transcriptomics
    - Enzymes - proteomics
    - Reactions - fluxomics
    - Compounds - metabolomics

- Data inputs:
  - Single or multiple values for each object
  - Absolute or relative values
Data and software downloads

- Install a local copy of the Pathway Tools software
Data and software downloads

- Download the complete database files with a FREE license.
Pathway Tools – desktop version

- Work offline
- Create new pathways
- Modify existing pathways
- See Omics Viewer results on pathway pages
On the horizon . . .

- Develop a better enzyme annotation pipeline
- Predict plant metabolic pathways for additional species with predicted proteomes
  - On the way . . . :
    - rice
    - barley
    - sorghum
    - Brachypodium
    - Setaria
- Solicit help from experts for pathway validation
- Increase coverage of secondary metabolism from experimental evidence
Plant metabolic NETWORKING

- Please use our data
- Please use our tools
- Please help us to improve our databases!
- Please contact us if we can be of any help!

curator@plantcyc.org

www.plantcyc.org
# PMN Acknowledgements

**Curator:**
- kate dreher

**Post-docs:**
- Lee Chae
- Ricardo Nilo Poyanco
- Chuan Wang

**Interns**
- Caryn Johansen

**Tech Team Members:**
- Bob Muller (*Manager*)
- Larry Ploetz (*Sys. Admin*)
- Shanker Singh
- Bill Nelson

**Rhee Lab Members:**
- Flavia Bossi
- Hye-in Nam
- Taehyong Kim
- Meng Xu

---

**Peifen Zhang (*Director and curator*)**

**PMN Alumni:**
- A. S. Karthikeyan (*curator*)
- Christophe Tissier (*curator*)
- Hartmut Foerster (*curator*)
- Eva Huala (*co-PI*)
- Tam Tran (*intern*)
- Varun Dwaraka (*intern*)
- Damian Priamurskiy (*intern*)
- Ricardo Leitao (*intern*)
- Michael Ahn (*intern*)
- Purva Karia (*intern*)
- Anuradha Pujar (*SGN curator*)

**Tech Team Alumni**
- Anjo Chi
- Cynthia Lee
- Tom Meyer
- Vanessa Kirkup
- Chris Wilks
- Raymond Chetty

---

**Collaborators:**

**SRI**
- Peter Karp
- Ron Caspi
- Hartmut Foerster
- Suzanne Paley
- SRI Tech Team

**MaizeGDB**
- Mary Schaeffer
- Lisa Harper
- Jack Gardiner
- Taner Sen

**ChlamyCyc**
- Patrick May
- Dirk Walther
- Lukas Mueller (*SGN*)
- Gramene and MedicCyc
We're here to help . . .

curator@plantcyc.org

www.plantcyc.org

Poster P0929:
Monday: 3:00 – 4:30 PM