Metabolic Pathway Databases and Tools

Speaker and Schedule Update

PMN (Peifen Zhang) KEGG (auto-slide show) MetaCrop (cancelled)



Peifen Zhang Carnegie Institution For Science Department of Plant Biology

Where We Are



Who We Are





PMN:

- Sue Rhee (*PI*)
- Kate Dreher (*curator*)
- A. S. Karthikeyan (*curator, alumni*)
- Anjo Chi (*tech team*)
- Cynthia Lee (*tech team*)
- Larry Ploetz (*tech team*)
- Shanker Singh (tech team)
- Bob Muller (*tech team*)
- Vanessa Kirkup (*tech team, alumni*)
- Tom Meyer (*tech team, alumni*)

<u>Key Collaborators:</u>

- Peter Karp (MetaCyc, SRI)
- Ron Caspi (MetaCYc, SRI)
- Lukas Mueller (SGN)
- Anuradha Pujar (SGN)

Outline

- General introduction
- Browse/Search/Analyze
- Pathway database creation and curation

Introducing the PMN

- Scope
 - PMN is a collection of plant metabolic pathway databases
 - PMN is a community for data curation
 - Curators, editorial board, ally databases, researchers
- Major goals
 - Create metabolic pathway databases for plants
 - For individual plant species
 - e.g. AraCyc (Arabidopsis thaliana)
 - e.g. PoplarCyc (Populus trichocarpa)
 - Combining data for all plant species PlantCyc
 - Create a computational prediction "pipeline":
 - Start with protein sequences for a specific plant species
 - End with a comprehensive set of predicted enzyme functions and associated metabolic pathways

PMN Databases

- AraCyc, PoplarCyc, and more to come
 - Single-species
 - Comprehensive collection of pathways in a particular species
 - Complete collection of enzymes, known or predicted, in that species
- PlantCyc
 - Multiple-species
 - Comprehensive collection of pathways for all plants
 - Representative collection of known enzymes in plants

PMN Database Content Statistics

	PlantC	<mark>yc 4</mark> .0	AraCyc 7.0	PoplarCyc 2.0	
Pathways	68	5	369	288	
Enzymes	110	58	5506	3420	
Reactions	292	29	2418	1707	
Compounds	296	6	2719	1397	
Organisms	343		1	1*	

Valuable plant natural products, many are specialized metabolites that are limited to a few species or genus.

- medicinal: e.g. artemisinin and quinine (treatment of malaria), codeine and morphine (pain-killer), ginsenosides (cardio-protectant), lupenol (antiinflammatory), taxol and vinblastine (anti-cancer)
- industrial materials: e.g. resin and rubber
- food flavor and scents: e.g. capsaicin and piperine (chili and pepper flavor), geranyl acetate (aroma of rose) and menthol (mint).

Other Plant Databases Accessible From PMN

Database	Species	Source	Curation status
RiceCyc ***	Rice	Gramene	some curation
SorghumCyc	Sorghum	Gramene	no curation
MedicCyc ***	Medicago	Noble Foundation	some curation
LycoCyc ***	Tomato	Sol Genomics Network	some curation
PotatoCyc	Potato	Sol Genomics Network	no curation
СарСус	Pepper	Sol Genomics Network	no curation
NicotianaCyc	Tobacco	Sol Genomics Network	no curation
PetuniaCyc	Petunia	Sol Genomics Network	no curation
CoffeaCyc	Coffee	Sol Genomics Network	no curation

*** Significant numbers of genes from these databases have been integrated into PlantCyc

Browse/Search/Analyze

Browsing the PMN Data



Browsing Pathways

Pathways			
Eliosynthesis	(613 instances)		
Amines an	d Polyamines Blosynthesis (14 instances)		
+ Amino act	ds Blosynthesis (52 instances)		
Aminoac			
+ Aromatic	<i>PlantCyc</i> Pathways Clas	s: Alkal	oids Biosynthesis
+ Carbohyd	Summary:		
Cell struc	This class contains biosynthetic pathways of alkale function as defense compounds. Many alkaloids, is receptors of neurotransmitters and have pharmac	oids. Most a ncluding mo ological acti	alkaloids contain cyclic nitrogen. 1 orphine and cocaine, have a high ; ivities.
+ Hormone	Parent Classes: Nitrogen-Containing Secondary Compounds Biosyr	nthesis	
 Nucleosid Other Bio Secondar Siderophi 	Child Classes: Betalaine alkaloids (8), Indole alkaloids (5), Isoquinoline and Benzylisoquinoline alkaloids (9), Peptide alkaloids (0), Bolykotido alkaloids (0)		
Degradation Detoxification Generation Superpathw Transport Pa	Polyketide alkaloids (0), Purine alkaloids (4), Pyrrolidine, Piperidine and Pyridine alkaloids (3), Pyrrolizidine alkaloids (0), Quinoline alkaloids (2), Quinolizidine alkaloids (1), Terpenoid Alkaloids Biosynthesis (3),	Instances (S)-re berbe bisbe dehy lauda	: eticuline biosynthesis I , erine biosynthesis , enzylisoquinoline alkaloid biosynthe droscoulerine biosynthesis , mine biosynthesis ,

morphine biosynthesis , palmatine biosynthesis , sanguinarine and macarpine biosynthesis

Quick Search

• Quick search bar

PMN Plant Metabolic Network			and and					
				Choline		PlantOyc 🖌	search]=)
About PMN	Search	Tools	Downloads	Useful Sites	Submit D	AreOyc PosterCut	Fee	dback
						Google Parki		

Searching in PMN databases

The query choline matched the following objects:

Pathways	Pathway pages cor pathway, of chromo genes, and of regul	Proteins	Protein pages citations, sub- activators, and depiction of ro	contain: Detailed unit structure; col d inhibitors (for en	d comment factors, izymes),	is and	_
 choline choline choline shasphi 	biosynthesis I biosynthesis II biosynthesis II tidylcheline biosyn	• chelia • chelia • chelia	e kinase (pol) e kinase (pol) e kinase (pol)	Compounds	Compour structura reactions compour	nd pages conta l information, a and pathways id participates	ain: compound and links to all s in which the
 ahasaha ahasaha ahasaha superpa superpa 	itidyicheline biosyn itidyicheline biosyn itidyicheline biosyn thway of phosphatid thway of choline bi	 chelin chelin chelin chelin chelin chelin chelin 	ve kinase (pol) ve kinase (pol) ve kinase (pol) ve kinase (pol) ve kinase (pol)	 16.0-18:1- 16.0-18:2- 16.0-18:3- 18.1-18:1- 	PC /1-16 PC /1-16 PC /1-16 PC /1-18	Reactions	Reaction pages contain: reaction equation wit chemical structures, links to all enzymes that catalyze the reaction, and all pathways in which the reaction participates.
Genes de pro co	ne pages contain: c piction of its operon, duct information is i responding protein (chelia chelia chelia chelia chelia chelia chelia chelia chelia 	ve kinase (pol) ve kinase (pol) ve kinase (pol) ve kinase (pol) INE MONOOX ve monooxypo	 18.1-18.2- 18.1-18.3- 18.2-18.1- 18.2-18.2- 3-chosolog 18.2-18.3- 	PC (1-18 PC (1-18 PC (1-18 PC (1-18 xholine PC (1-18	 <u>1-O-sina</u> <u>+ B-D-gl</u> <u>cheline</u> <u>a 1.2-du</u> phospha 	spoyl-B-D-glucose + chelline = O-sinappy/choin ucose + ATP = phosphoryl-choline + ADP scylglycerol + CDP-choline = a stidy/choline + CMP
• choline	manaaxygenase	 chelin expres chelin AT2G 	ve moncoxyge used ve phosphate i 12260	 18.3-18.1- 18.3-18.2- 18.3-18.3- choline 	PC (1-18 PC (1-18 PC (1-18	 a phosp glycerop a phosp shoine phospho 	hatidyicheline + H ₂ O = a 1-acyi-2-tyso- hosphocholine + a carboxylate hatidyicheline + H ₂ O = an L-phosphatidate + aryl-choline + CTP = CDP-choline + diphospha
		• cheli AI4G • cheli LOC_ • cheli	ve-phosphate i 15130 ve-phosphate i Os02g07720.1 vephosphate c	 So-3-gives a 1-acyl-2 a 1-akyl-2 a 1-akyl-2 a 1-akyl-2 a 1-lyso-2 a 1-organy 	erophospho -lyso-glyce -acetyl-an- o-glycero- -acyl-an-gly (1-2-acyl-an-	ocholine rophosphocho glycero-3.pho 3.phosphochol ycero-3.phospl o-glycero-3.pho	dine sphocheline hocheline sphocheline

- 0

1±

a 2-lysophosphatidylcheline

Specific Data Type Search, Pathway Search

For example, find pathways that includes a specific intermediate ornithine

· Search for pathway by ma	une	ane:	
Examples: "glycolysis", "arginin	e biosynthesis'		
Search/Filter by ontology			
Search/Filter by number	Arabidopsis thaliana c	ol Ouerv	Results
· Search/Filter by substrat			
Enter up to four substant	You searched for all pathways that involve the	substrate L-ornit	hine.
Localitie	Your query returned 11 results.		
r-ompane	Pathway Name 🔺 🗸	#Reactions $ riangle abla abla$	Evidence 🛆 🔽
(A) (A)	arginine biosynthesis I	9	Traceable author statement
	arginine biosynthesis II (acetyl cycle)	9	Traceable author statement
	arginine degradation VI (arginase 2 pathway)	4	Inferred from experiment
	citruline biosynthesis	8	Inferred by curator
Charles and and the second	citruitine degradation	2	Inferred from direct assay
Enter up to four compound n	omithine biosynthesis	5	inferred by a human based on computational e
name, as no substring matche	proline biosynthesis II (from arginine)	6	Inferred by a human based on computational e
Examples: "L-tryptophan", "py	proline biosynthesis III	4	Inferred from experiment
Search/Filter by evidence	putrescine biosynthesis IV	4	Inferred from direct assay
Search (Filter by publicat	superpathway of citrulline metabolism	16	Inferred by curator
- association of parameter	urea cyde	5	Inferred from mutant phenotype Inferred from direct assay Inferred from genetic interaction

A Typical Pathway Detail Page



A Typical Pathway Detail Page



Conventions Used in Curation and Data Presentation

- A pathway, as drawn in the text books, is a functional unit, regulated as a unit
- Pathway displayed is expected to operate as such in the individual species listed



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

Superclasses: Biosynthesis -> Secondary Metabolites Biosynthesis -> Nitrogen-Containing Secondary Compounds Biosynthesis -> Alkaloids Biosynthesis -> Purine alkaloids -> Caffeine Biosynthesis

Species Data Available for: Camellia sinensis , Camellia sinensis assamica , Camellia taliensis , Coffea arabica , Coffea canephora

Conventions Used in Curation and Data Presentation

- Pathway, as drawn in the text books, is a functional unit, regulated as a unit
- Pathway displayed is expected to operate as such in the individual species shown
- Alternative routes that have been observed in different organisms are curated separately as pathway variants

PlantCyc Pathway: caffeine biosynthesis I





Conventions Used in Curation and Data Presentation

- Pathway, as drawn in the text books, is a functional unit, regulated as a unit
- Pathway displayed is expected to operate as such in the individual species shown
- Alternative routes that have been observed in different organisms are curated separately as pathway variants
- Mosaics combined of alternative routes from several different species are curated as Superpathways
- Connected pathways, extended networks, are curated as Superpathways



PlantCyc Pathway: superpathway of sucrose and starch metabolism I

Linking to Other Data Detail Pages



Compound Detail Pages

Synonymu: citicoline , citicholine , cidilos , cyticholine , cytidine 5' dehosphocholine , cytidine dehosphate choline

Superclasses: a nucleic acid component -> a base derivative a nucleic acid component -> a pyrimidine related compound

Empirical Formula: C14HerN4D1-P2

Molecular Weight: 499 34 datures



Molecular Weight / Formula

Smiles: C(OP(C)(=0)OP(O)(=0)OCC(N+)(C)(C)C1(OC(C)(0)C10)(2)(i)*O(nc(N)cc2))

Unification Links: CAS:987-78-0

Gibbs Energy of Formation (Iscal/mol, estimated): +116.7

In Pathway Reactions as a Reactant:

phospholipid biosynthesis

a 1,2-diacalgluceral + CDP-choine + a phosphatich/choine + CMP pholine boxynthesis II.

a 1.2-disculatoreal CDP-choine a shouphatidy/choine + CMP

In Pathway Reactions as a Product

shoughoigid biorecthesis. phosphoryl-choine + CTP = CDP-choine + dohosphate

photophonyl choine + CTP = CDP choine + dehotophate

Smiles / InChI

Appears as Reactant

Appears as Product

Enzyme detail pages

Arabidopsis Enzyme: phosphatidyltransferase



phosphatidylalcohols although with slight differences regarding the substrate preference. AsAAPT2 showed a higher preference for CDP-choline over CDP-ethanolamine in comparison to AsAAPT1 and was also inhibited to a lesser degree by Ca²⁴ and Cytidine monophosohate (CMP) than AtAAPT1 | Goode(9) | Both enzymes (AtAAPT1, AtAAPT2) were able to catalyze the reverse reaction supporting the proposal that diacylightersil, involved as substrate in both PC and triacylightersil biosynthesis (triacylightersil biosynthesis), is in equilibrium with PC and maintains this concluded from the Southern blotting patterns. Both polypetides contain seven membrane spanning regions as shown by their hydropathy files. This finding is in agreement with other AAPTs isolated so far and confirms their localization in membranes | Goode99 1

Inhibitors, Kinetic Parameters, etc.

Inhibitors (Allosteric): CMP [Goode99]

inhibitors (Unkmech) Call [Gooder99]

Primary Physiological Regulators of Enzyme Activity. CMP

References

Geodel99. Goode JH, Devery RE, (1999) "Characterization of animalization photosphotransferances from Anabidopain chalane and psychem." Plant Physiol. Biochem. (1999), 3781, 445-457.

Stack#5: Stack CR, Roughan PG, Browse JA, Gardiner SE, (1995) "Some properties of challmeshosphotransferace from developing sufficient catyledons," Biochim, Biophys. Acta (1995), 833, 438-448.

The Global Overview Map

About PMN	Search	Texalm	Downloads	Useful Sites	Submit Data	Help	Feedback	
		Tools	Overview			01110		
10000	1000		ANT					
Introduc	tion	Metabolic (Ce	ellular) Overview		News			
The Plant M	stabolic Network	Old Met	abolic Map	or g databases	New PIIN re	leases!		
and blochem	lists with a comm	A110-15-00-0		ant metabolic				
pathway data plant bioche	wbases. A central mical pathway d	Compara	tive Analysis	ation from the	Please come search for your favorite pathway using the newest version of the			
literature and	Iterature and computational a		ports I	, compounds,	Pathway Tools software in the newest			
reactions, an	d patriways involu	Pathway Too	is Biocyc Bundle	lism.	versions of a	iur PLIN databa	ISES	
PINN is funde by an Editor	ed by the National rial Board compe	Tu	torials I	769), governed icientists, and	Plan Arac	Cyc 4.0 yc 7.0		

The Global Overview Map



Visualizing Omics Data



Visualize and interpret large scale omics data in a metabolism context:

- Gene expression data
- Proteomic data
- Metabolic profiling data
- Reaction flux data

Input File Format For The Omics Viewer

Tab-delimited text file

At1g77760	1.15	2.3	3.2	2.15	1.53	1.75
At2g13360	0.7	-0.53	0	-0.73	0.03	-0.72
At3g10230	-1.1	-0.05	1.05	1.15	1.25	0.05
At3g10230	-0.65	-0.58	1.13	1.23	0.67	-0.12
At3g01120	-1.08	-0.15	-1.2	-1.15	-1.15	-0.58
At3g01500	0.07	-0.72	-0.68	-1.4	-1.93	-9.23
At3g02470	0.03	-0.53	0.58	1.28	0.55	1.4
At3g02470	0.55	-0.12	0.62	0.65	-0.05	1.22
At3g02580	0.6	-0.55	0.08	0.55	-2.2	-1.65



Omics Viewer: Color Coding Gene Expression Levels



Red: Enhanced expression over my threshold (i.e. 2-fold change)Yellow: Repressed expression over my thresholdBlue: Not significantly changed over my threshold

Omics Viewer: generating a table of individual pathways exceeding certain threshold

Pathway	Pathway Diagram	Enzymes, Genes, and Enzyme Cellular	Locations
methionine salvage pathway	Sadarooys L -methodage - Color - Di	1-aminocyclopropane-1-carboxylate synthase	ACS9
	ACTION ACTION	1-aninocyclopropane-1-carboxylate synthase	ACS11
	Lowfrong TIG17300 AI2G36860 S-outputtionaire	1-aninocyclopropane-1-carboxylate synthese	AC56
		1-antinocyclopropane-1-carboxytate synthese	AC\$7
	AT4G34840 A74G38800	1-antinocyclopropane-1-carboxylate	AC58
	2-cio-4 mity/factorian	1-aninocyclopropane-1-carboxylate	ACS5
	ATT G49820	1-aminocyclopropane-1-carboxylate	AC54
	1,2-dtydrosy-3-keto 5-methylthioethyle -1-shoseheth	1-aminocyclopropane-1-carboxylate	ACS2
		catalytic//methythioadenosine successidate	AT46388
		methylthioadenosine sudeosidate	AT46348
	settyttio-1-shoshopentere 5-mettyttionkakoe-1-shoshole	5-methyl-5-thioribose kinase	AT16498
		methionine adepointmentferase	AT26368
		methionine adenosyltransferase	AT36173
		methionine adenosyltransferase	AT4GD18
	5-(net sitted) 2,3-du-operity prosphete	methionine adenosyltransferase	SAM-F
purine nucleotide	nome-5-phosphate	WP dehydrogenase AT1679470	
metabolism	ATTACASH INTER	WP dehydrogenate AT1G15350	
(phosphotransfer and nucleotide modification)	ATI G56290 ATI G10440 ATI G56290 ATI G10440 ATI G56200 ATI G10440 ATI G56200	GMP synthetase / GMP ATIGA3660 synthese (glutamine- hydrolyding)	2 -
1	ACTICIDATE ATSOCOME	guanylate kinase AT2G41880	8 👔
1	ATIGETERO ATIGETERO ATIGETERO	guanylate kinase AT3G57550	2.5
	AISGIST/D AISGR/DHD	adesylosuccinate synthese AT3G57618	2 .
	ATSGROMO ATSGROUPD ATSGROU	adenytosuccinate type AT4G18448	1
		adenylosuccinate lyase AT1G36280	-
	A73G23580 A43g27080 A73G83310 A71G17410	adenylate kinase//nudeotide_AT2G37250_ Kinase	chiords

Comparing Across Species

Overview of the AraCyc Metabolic Map

This diagram provides a schematic of all pathways of AraCyc metabolism in the aracyc database. Nodes represent metabolites, with shape indicating class of metabolite (see key to right). Lines represent reactions. Move the mouse over a metabolite plife icon to navigate to the metabolite page or a related pathway page. Select one or more organisms: Instructions 🗖 PlantCyc 🗹 Populus trichocarpa Pathway Tools query page Omics Viewer: Paint omics data onto this diagram Species Comparison: Highlight reactions shared with other organism Clear All Submit *********

Comparing Across Species



Comparing Across Species

About PMN	Search	Tools	Download	8	Useful Sites	Submit
		Tools	Overview			1
Internation		8	LAST			
Introduction		Metabolic (C	Metabolic (Cellular) Overview			N
The Plant M	etabolic Network	Old Me	tabolic Map	. •	ng databases	Net
and biochemists with a comm		ОМІС	OMICs Viewer +			
pathway data plant bloche	mical pathway d	Contains	Bre-Andysia.		ation from the	pat
interature and computational		Re	ports		, compounds,	Pat
reactions, an	id pathways invol	Pathway Too	is Biocyc Bund	ke	lism.	ven
PMN is fund	ed by the Nationa	Tu	torials		769), governed	

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

- Reactions: breakdowns by type of substrate, by EC Number, by number of isozymes, etc.
- Pathways: breakdown by pathway class, information on pathway holes.
- Compounds: small molecules that act as substrates, enzyme activators/inhibitors/cofactors.
- Proteins: breakdown of protein complexes by type and number of subunits, number of enzymes.
- Crthologs: proteins shared among organisms or unique to an organism.
- Transporters: proteins that facilitate the movement of compounds across cell membranes.
- Transcription Units: number of genes per transcription unit, number of operons per pathway.

Select one or more organism databases:

Arabidopsis thaliana col PlantCyc Populus trichocarpa

Submit	Clear All
--------	-----------

Pathways

Table 1: Breakdown of Pathways by Pathway Class

This table present:	Pathway Class: Biosynthesi	is - Hormones Biosyn	thesis AraCy	c col P. trichocarpa
down further to sh	cis-zeatin biosynth Organism	Evidence Glyph	Enzymes and G	Senes for brassinosteroid biosynthesis II
will see only those	ent-kaurene biosyr AraCyc co	0	EC# 5.3.3.1	Δ ⁵⁻³ -ketosteroid isomerase:
	<i>trans-</i> zeatin biosyr	00	EC# 1.1.1.145	38-hydroxylsteroid dehydrogenase:
	abscisic acid biosyr	09	EC# 1.3.99	sterol 5-alpha reductase: DET2
Biosynthesis	abscisic acid glucos	99	EC# 1.14.13	sterold 22g-hydroxylase: DWF4
- Amines and Poly	aldehyde oxidation	00	9XN-712	None
- Amino acids Bios	brassinosteroid bio	0	FC# 1 14 13	stand 22s budgedata: DWEd
- Aminoacyl-tRNA	brassinosteroid bio		SC# 1.14.13.	stand 22e hydroxiana DWE4
- Aromatic Compo	brassinosteroid bio		CC# 1.14.13	steroid 220-hydroxyase, DWF4
- Carbohydrates B	cvtokinins 7-N-gluc		EC# 1.14.13	steroid 220-hydroxytase: DWF4
- Cell structures B	cytokinins 9-N-gluc		001-4220	SAX1: SAX1
- Cofactors, Prost	cytokinins-Q-ducor		EC# 1.3.99	sterol 5-alpha reductase: DET2
- Fatty Acids and	athylana biosynthe		R00N-4228	None
- Hormones Biosyr			EC# 1,14,-,-	23alpha hydroxylase / cathasterone 23o-hydroxylase: CBB3
- Metabolic Regula	P. trichoc	arpa 🖸	EC# 5.3.3.1	None
- Nucleosides and	gibberellin biosynti		EC# 1.1.1.145	None
- Other Biosynthe	gibberellin biosyntl	14	FC# 1 3 99 .	None
- Polysaccharides	gibberellin biosyntl	IA	50811413	man commencer 1/51,820/24
Secondary Motal	hydroxyjasmonate	10	Con terretor	monooxygenase: JGI-796803
- Secondary Metal	IAA biosynthesis I	2	RXN-712	None
- Secondary Metal	IAA biosynthesis II	0	FC# 1.14.13.	monocoverenese: IGL820024
- Siderophore Bio:	IAA biosynthesis V			monooxygenase: JGI-796803
	indole-3-acetyl-ami		EC# 1.14.13	monooxygenase: JGI-820024
	jasmonic acid biosy		and the state of t	monooxygenase: JGI-796803
	jasmonoyl-amino a		EC# 1.14.13	monooxygenase: JGI-820024
				monooxygenase: JGI-796803

PMN E

PMN	° -	
About PMN	Search	Tools
	Databases O	werview
	Compounds	8
	Genes/Prote	nins
	Reactions	
	Pathways	
	Advanced Q	sery
	Browse/Oel	ologies +
	Google the P	IMN
	BLAST	A Constant
	Species in th	e PMN
	Bearth Dat AlantCyc chaoge	abase:

Blad BLADT ^{**} progr Defectes: Input svery sequence	BLASTP AA guery AA da PaertCyc Engrees Extension Engrees experimental or	1
Dufasets: Input guery sequence	BLASTP AA query AA de PlentCyc Encyrnes Poderesce Encyrnes experimental or	1
Datasats: Input Suery sequence	experimental or	1
grev redneuro	• experimental or	
	• experimental or	
Raw, FAGTA, O	plants only	
Advanced BLASS 1	a summer of party	
nonanced BLASS ¹⁴⁴ repard Results Return Opt		
Automiced BLASS [®] organid Results Retem Opti Output Title	ptonal, will be added to output for your use	
Automiced BLASS ** organd Results Retem Opt Output Title:	nes. prional, will be added to output for your use	
Results Return Opt Output Title 0 S-mail address.	ees ptional, will be added to output for your use	

Online Tutorials

About PMN	Search	Tools	Downloads	Useful Sites	Submit Data	Help	feedback		
100 C C C C C C C C C C C C C C C C C C		Tools	Overview		and o etch	1	1.		
distant in a			LAST		history				
introduc	tion	Metabolic (C	ellular) Overview		News				
The Plant M	etabolic Network	Old Me	tabolic Map 1	ang databases	New Pfill	New PIIN releases!			
and biochen	nists with a comm	OMIC	s Viewer I	plant metabolic			Contractory and		
plant bioch	emical pathway d	Compara	five Analysis	ation from the	pathway	using the newest	version of the		
Iterature an	d computational :	Ĥ	ports I	C compounds.	Pathway	Tools software	in the newest		
reactions, at	Plin is funded by the National by an Editorial Board comp		is Biocyc Bundle	asin. Versions		or our many databases.			
PENIS fund by an Edite			Tanatata 🕨 🕨		Tutorials Overview				
executed at	the Carnegie Institu	tion for Science	Department of PI	e Searchin	g the PMN	HarCyc 2.0			
CARNEGH	CARNEGH				Understanding PMN Detail Pages		tes to the undefed debetween		
SCIENCE				Using the Omics Viewer		ented at the second Banfl			
				Using the Metabolic (Cellular) Over					
				Using the Old	Using the Old Metabolic Map		exciting news		
Plant M	etabolic Path	way Datab	oases	Using the Advan	iced Query Page	0200000000			
				Assigning Evider	nce Codes in PMI				
PlantCyc		earch U	lore into	Bulk Dov	miceding	the Month			
PlantCyc pr	ovides access to	manually curat	ed or reviewed in	Submitting	Data to PMN				

Data Downloads

PMN Data Downloads To download all of the plant metabolic pathway databases hosted by the PMN, including AraCyc. please fill out and submit the license agreement form install and The complete download, together with the Pathway Tools softwar Complete iddition, all have a local copy of AraCyc/PlantCyc, just like the one visible from databases of the BioCyc formatifiat files and a complete biopax.owl file are a download. You can get more details about the format of all the file types from Additional custom flat files containing data extracted from the latest PMN release can be freely downloaded without a license PlantCyc pathways (plantcyc_dump) PlantCyc compounds diote This file pounds used in PlantCyc pathways) Custom AraCyc pathways (aracyc_dump) flat files AraCyc compounds (Nete: This file One) compounds used in AraCyc pathways) PoplarCyc pathways (applarcyc_dump) Popla/Cyc compounds Note: This file ONLY contains compounds used in Popla/Cyc pathways) Custom The two BLAST sets provided by the aded without registration. BLAST PlantCyc Enzymes dataset Reference Enzymes

Download and Install a Local Copy of the PMN Databases

- Run robust live database query by scripts, via Perl, Java, LISP interfaces
- Edit with private data
- Access to additional features not available on web mode
- Free, open database license
- Pathway Tools Software (SRI)

Developing The PMN

Creating Single-Species Databases

- New sets of DNA sequences become available
 - Genomes are sequenced
 - Large EST data sets are created
 - Unigene builds are generated
- PMN pipeline predicts enzyme functions
 - Based on sequence similarity to known enzymes, enzymes with experimental or literature support
- Set of predicted enzymes is used to predict metabolic pathways
 - The pathway prediction software (Pathway Tools) uses:
 - Enzyme functional annotations
 - A réference set of pathways (e.g. PlantCyc)
- Curators validate predicted pathways in the new database
 - Remove incorrect information and add additional data



Identifiers Used in Automated Enzyme Annotation and Enzyme to Pathway Mapping

- Complete EC number
 e.g. 2.1.1.128
- Unique PlantCyc reaction id, when complete EC is not assigned
 - e.g. RXN-0981
- GO term id
 - e.g. GO:560010



Manual Curation

- Who
 - Curators identify, read and enter information from published journal articles
- What
 - Add missing pathways
 - Update existing pathways
 - Create new reactions
 - Add compound structures
 - Add missing enzymes
 - Curate enzyme properties, kinetic data
 - Remove false-positive pathway predictions
 - Remove false-positive enzyme annotations

Submitting Data To Us



Contact Us

- Contact info
- FeedbackForm
- Submit Data

Please help us expand the content of the PMN databases, including AraCyc and PlantCyc!

- · We welcome any and all new data submissions related to plant biochemical pathways.
- · You can also carrect an existing pathway
 - Not sure what types of information to include? Try using one of our data submission / correction forms below.
- Or, just use our Feedback Form or send an e-mail to curator@plantcyt.orp and append your data as an attachment

Data submission / correction forms

""" You do not need to fill in all the columns. (See our lutorial for help)

We will be happy to accept whatever information you can easily provide

Compound form

The second s

Community Gratitude

About Plan	Search	Tools	Downloads	Useful Sites	Submit Data	The last	Feedback	
Project Overstew		Non Constant	Succession	-				
Decumentation	- 1 C	PMN Contributors						
Presentations	8	1111111		011000000000	2002 A 1997	1922/2011/1	0,00000	
Statistics	-	PMN contribution	dors from around 8	he world have adde	d to or helped to imp build any part of the	move the content ends	R of AraCyr.	
Rainane Boten	D45	Franklin, an	o ere or er r recitor	-derved descape	o that are part of the	-		
Contact Info Brown		In addition to individuaris h Ptan8Cyc-der	the active contribut ave contributed sig fired databases the	Sons from the PMN nific entry in improvi (are platt of the PM	editorial board and ing the content of Pla N.	MAN COLLECTER INCYC, AraCyc a	nd the other	
Collaboratore Contributore	K0.	Some Some Some	e contributors have e have penerousty r e have attended cur	contacted us with a esponded to appe- atorial samborees.	uppestions and new als for help from the	siona. curators		
Editorial Board	\$.							
		Contributous	for FSR release.	Auso 2019				
		 Lodg Devic Nistal Andro Zutro Jung Peter Radir Berris Mitris Brain Churi Elevic Zhiyo 	er Deenhuns - Tech (Dison - Durham U Is Dudaenis - Pard s) Guranowski - Th a He - Chinese Acs Im Hur - University- Landquist - Come n Sadre - RWTM As d Schneider - Max F mori Seo - RIKEN, , und Tenhaken - Univer nore Warbei - The C ng Xe - Shanghari	mache Universitat i niversity, UK e University, USA e University of Life demy of Sciences, of California, Los A f University, USA chen University, Gr Nanck (rotitule for O Appan iversitat Satzburg, ity of Decently, of Na Anversity of Traditi	Braunschweig, Gent Sciences, Poland China ngolini, USA immical Ecology Bi Austha i w York, USA onal Chinese Medici	uany natenberg Carl ne, China	puş Germeny	
		Contribution	for prior Phillippin	atas Sala (Johanady 1/5)	4			
		Jose Richi	Alonso - North Car ard Amasino - Unive	olina State Universi erady of Misconsin	ac USA USA			

www.plantcyc.org



curator@plantcyc.org

meet Kate, Booth# 219