

AraCyc: Birth and Growth

Peifen Zhang

The Arabidopsis Information Resource

Outline

- The PathoLogic build of AraCyc
- The manual curation and results
- Systematically updating AraCyc

What is AraCyc

- The Biochemical pathway and enzyme database for Arabidopsis
- Open Database
- Created and supported by Pathway Tools
- Manually curated

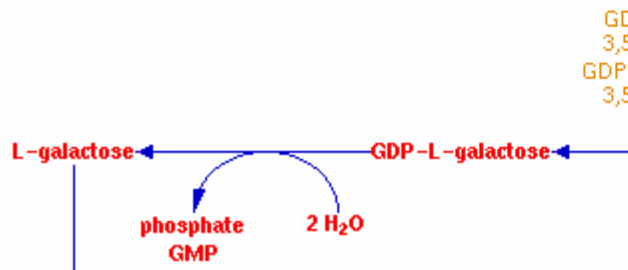
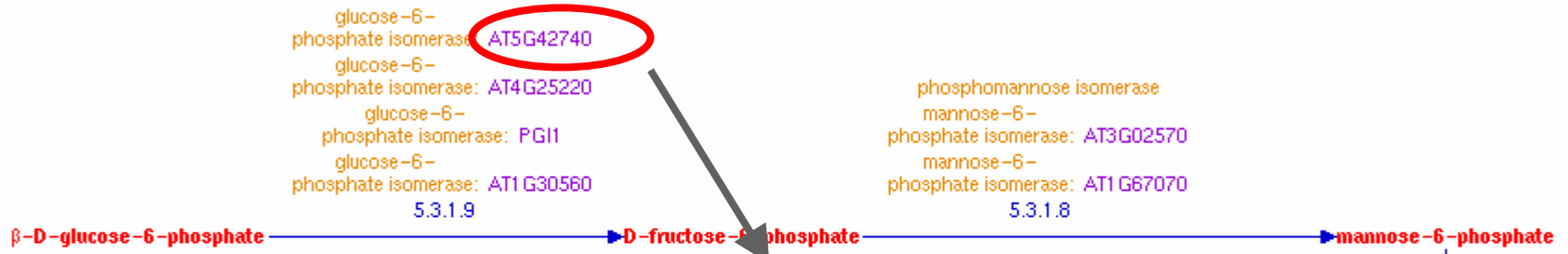
AraCyc Pathway: ascorbate biosynthesis



More Detail

Less Detail

Cross-Species Comparison



Home | About Us | Sitemap | Contact | Help | Order | Login

Search | Tools | Arabidopsis Info | News | Portals | FTP | Stocks

Gene Search

Locus: AT5G42740

Date last modified 2003-05-02

TAIR Accession Locus:2165462

Representative Gene Model [AT5G42740.1](#)

Other names: MJB21.12, MJB21_12

Annotations ?	Category	Relationship Type ?	Keyword ?
GO Biological Process	involved in	none	gluconeogenesis , glycolysis
		none	gluconeogenesis , glycolysis
GO Cellular Component	located in		cellular component unknown
GO Molecular Function	has		glucose-6-phosphate isomerase activity
		none	glucose-6-phosphate isomerase activity

[Annotation Detail](#)

RNA Data

Two-channel Arrays	array element name ?	avg. log ratio (std. error)	avg. intensity (std. error)	expression viewer	spot history (SMD)

from that found in mammals. It proceeds via GDP-D-mannose, L-galactose, and L-galactono-1,4-lactone. The intermediate GDP-mannose is also used for cell wall carbohydrate biosynthesis and protein glycosylation, in addition to ascorbic acid biosynthesis.

Citations: [[Wheeler98](#), [Smirnoff01](#)]

Unification Links: [MetaCyc:PWY-882](#)

References

[Smirnoff01](#): Smirnoff N, Conklin PL, Loewus FA (2001). "BIOSYNTHESIS OF ASCORBIC ACID IN PLANTS: A Renaissance." *Annu Rev Plant Physiol Plant Mol Biol* 52:437-467. PMID: 11337405

[Wheeler98](#): Wheeler GL, Jones MA, Smirnoff N (1998). "The biosynthetic pathway of vitamin C in higher plants." *Nature* 393(6683):365-9. PMID: 9620799

[Download Genes](#)

Query Pathway: ascorbate biosynthesis

Gene name	Reaction id	Reaction EC	Enzymatic activity	Evidence
AT5G28840	RXN-1882	5.1.3.18	GDP-mannose 3,5-epimerase	NIL
AT5G28840	RXN-1882	5.1.3.18	GDP-mannose 3,5-epimerase	EV-EXP
AT5G28840	RXN-1882	5.1.3.18	GDP-D-mannose-3,5-epimerase	EV-EXP
AT2G39770	RXN-1881	2.7.7.13	GDP-D-mannose pyrophosphorylase	EV-EXP
AT3G47930	GALACTONOLACTONE-DEHYDROGENASE-RXN		1.3.2.3 galactonolactone dehydr	
AT2G45790	PHOSMANMUT-RXN	5.4.2.8	phosphomannomutase	EV-COMP
AT3G02570	MANNPISOM-RXN	5.3.1.8	mannose-6-phosphate isomerase	EV-COMP
AT1G67070	MANNPISOM-RXN	5.3.1.8	mannose-6-phosphate isomerase	EV-COMP
AT5G42740	PGLUCISOM-RXN	5.3.1.9	glucose-6-phosphate isomerase	EV-COMP
AT4G25220	PGLUCISOM-RXN	5.3.1.9	glucose-6-phosphate isomerase	EV-COMP
AT4G24620	PGLUCISOM-RXN	5.3.1.9	glucose-6-phosphate isomerase	EV-COMP
AT1G30560	PGLUCISOM-RXN	5.3.1.9	glucose-6-phosphate isomerase	EV-COMP

The Computational Build

- As of 2004, the Arabidopsis genome contained 7900 genes annotated to the GO term 'catalytic activity'
- 4900 enzymes in small molecule metabolism (19% of the total genome)
- PathoLogic inferred 219 pathways and mapped 940 (19% enzyme-coding) genes to the pathways

Pathway Hole-filler Results

- Before: 528 holes (62% of total reactions)
- Additional holes filled: 58 (11% of total holes)

Category of holes	Holes	Portion to all holes	Holes with sequence retrieved	Holes-filled	Hole-filling rate	Maximum percentile of hole-filling
EC1 (oxidoreductase)	114	21.6%	74	12	16.2%	3.5%
EC2 (transferase)	128	24.2%	98	15	15.3%	3.7%
EC3 (hydrolase)	41	7.8%	32	2	6.3%	0.5%
EC4 (lyase)	40	7.6%	29	2	6.9%	0.5%
EC5 (isomerase)	22	4.2%	15	2	13.3%	0.6%
EC6 (ligase)	17	3.2%	14	9	64.3%	2.1%
No EC assignment	166	31.4%	50	16	32.0%	10.1%
Sum	528	100.0%	312	58	-	20.9%

Manual Curation Procedure

- Validating predicted pathways from the literature
 - A pathway that is described in the Arabidopsis literature (EV-EXP)
 - A pathway whose critical metabolites are described in the Arabidopsis literature (EV-COMP)
 - A pathway that has reactions unique to it and having genes assigned to the unique reactions (EV-COMP)
- Curating plant pathway variants
- Curating missing pathways

Manual Curation Results

- Deleted 118 non-plant pathways:
 - Pathways not operate in plants at all
 - Fermentation (*purine fermentation*)
 - Non-plant metabolites (*glycogen biosynthesis*)
 - Other (*glucose and glucose-1P degradation*)
 - Pathway variants of bacteria-origin
 - Amino acid metabolism (e.g. *cysteine biosynthesis II*)
 - Sugar metabolism (*trehalose biosynthesis II, III*)
 - Fatty acid/lipid metabolism (*glycerol degradation II*)
 - Nutrient assimilation (*sulfate assimilation I, II*)
 - Energy metabolism (TCA cycle variants)

Manual Curation Results (*cont.*)

- Added 127 new pathways:
 - Pathway variants of plant-origin
 - Plant pathways that were not in MetaCyc
 - Plant metabolites (hormone, chlorophyll, cell wall, secondary metabolites)
 - Plant-specific metabolism (photorespiration, xanthophyll cycle)
 - Universal pathways that were not in MetaCyc
 - fatty acid alpha-oxidation, fatty acid omega-oxidation, glycolipid biosynthesis, etc

Current Status of AraCyc

	AraCyc (2.6)
Total pathways	228
Pathways with literature supports	201 (88%)
Total enzymes	4909
Enzymes of pathways	1808
Enzymes with comments	150 (8%)
Pathway holes	38%
Total compounds	1133
Compounds with structures	1114 (98%)
Total citations	1572

Expert Curation Jamborees

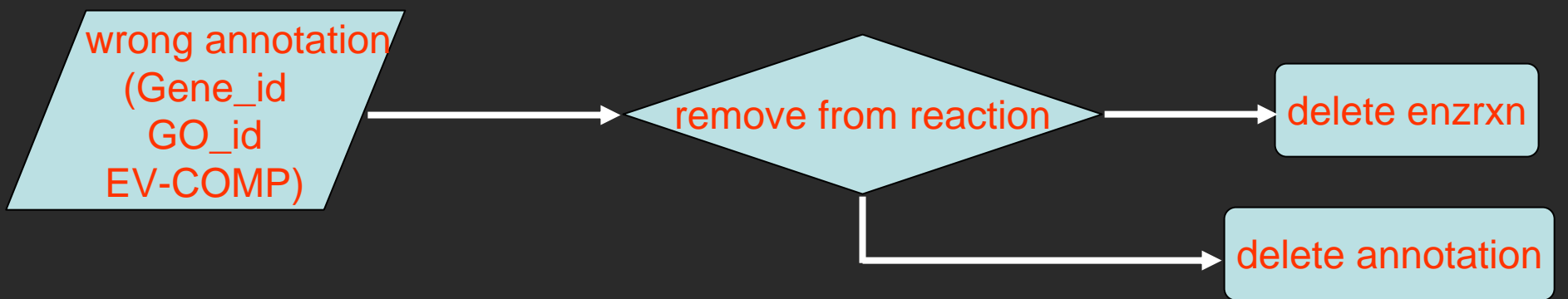
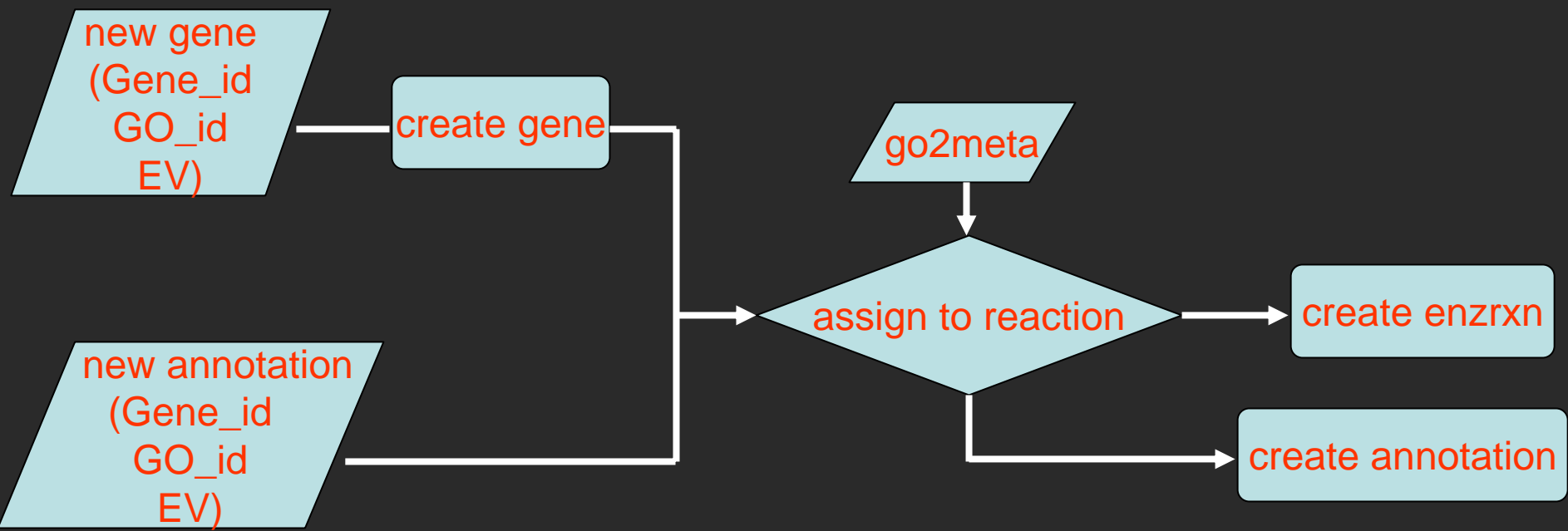
- 2X 1.5 days
- Session 1:
 - Clint Chapple from Purdue (phenylpropanoids)
 - Ed Cahoon from Danforth (lipids)
 - Maor Bar-Peled from CCRC (carbohydrates)
- Session 2:
 - Peter Facchini from U. Calgary (alkaloids)
 - Jon Page from NRC Saskatoon (terpenoids)
 - David Seigler from U. Illinois (flavonoids and polyketides)
- pathways to be updated as a result: 25
- pathways to be added as a result: 40 (previously unknown to curators)

The MetSearch Tool

- + [P Biosynthesis](#) , [5400 Hits.]
 - + [P Cofactor-Biosynthesis](#) , [No Hits.]
 - + [P Vitamin-Biosynthesis](#) , [No Hits.]
 - [P vitamin E biosynthesis](#) , [3 Hits.]
 - [P p-hydroxyphenylpyruvate dioxygenase](#) , [11 Hits.]
 - [P gamma-tocopherol methyltransferase](#) , [8 Hits.]
 - [P tocopherol cyclase](#) , [9 Hits.]
 - [P MPBQ/MSBQ methyltransferase](#) , [1 Hits.]
 - [P homogentisate phytyltransferase](#) , [6 Hits.]
 - [P lactoylglutathione lyase](#) , [No Hits.]
 - [P lactoylglutathione lyase](#) , [No Hits.]
 - [P delta-tocopherol](#) , [22 Hits.]
 - [P gamma-tocopherol](#) , [78 Hits.]
 - [P 2,3-dimethyl-6-phytyl-1,4-benzoquinone](#) , [1 Hits.]
 - [P 2-methyl-6-phytyl-1,4-benzoquinone](#) , [1 Hits.]
 - [P homogentisate](#) , [28 Hits.]
 - [P O2](#) , [704 Hits.]
 - [P p-hydroxyphenylpyruvate](#) , [13 Hits.]
 - [P phytyl pyrophosphate](#) , [3 Hits.]
 - [P S-adenosyl-L-methionine](#) , [174 Hits.]

AraCyc Maintenance and Updates

- Updating genes and annotations from the TAIR database
 - New enzyme-coding genes
 - Updated gene function annotations
 - Updated evidence of annotation



AraCyc Maintenance and Updates (*cont.*)

- Propagating Updates from MetaCyc
 - Compounds (updates name, structure)
 - Reactions (updates EC number)
 - Pathways (identify differences)

Future Plans

- Enhance pathway comments (ca. 60 pathways)
- Enhance superpathway curation
- Extensive curation of enzyme functional properties
- Curate transporters and transport reactions (ca. 210 transporters described in the literature)

Acknowledgments

Sue Rhee (PI)

Hartmut Foerster (curator)

Christophe Tissier (curator)

Aleksey Kleytman (curator assistant)

Thomas Yan (programmer)

Lukas Muller (Cornell, previous contributor)

The SRI Pathway Tools and MetaCyc groups

NIH, NSF, Pioneer Hi-Bred