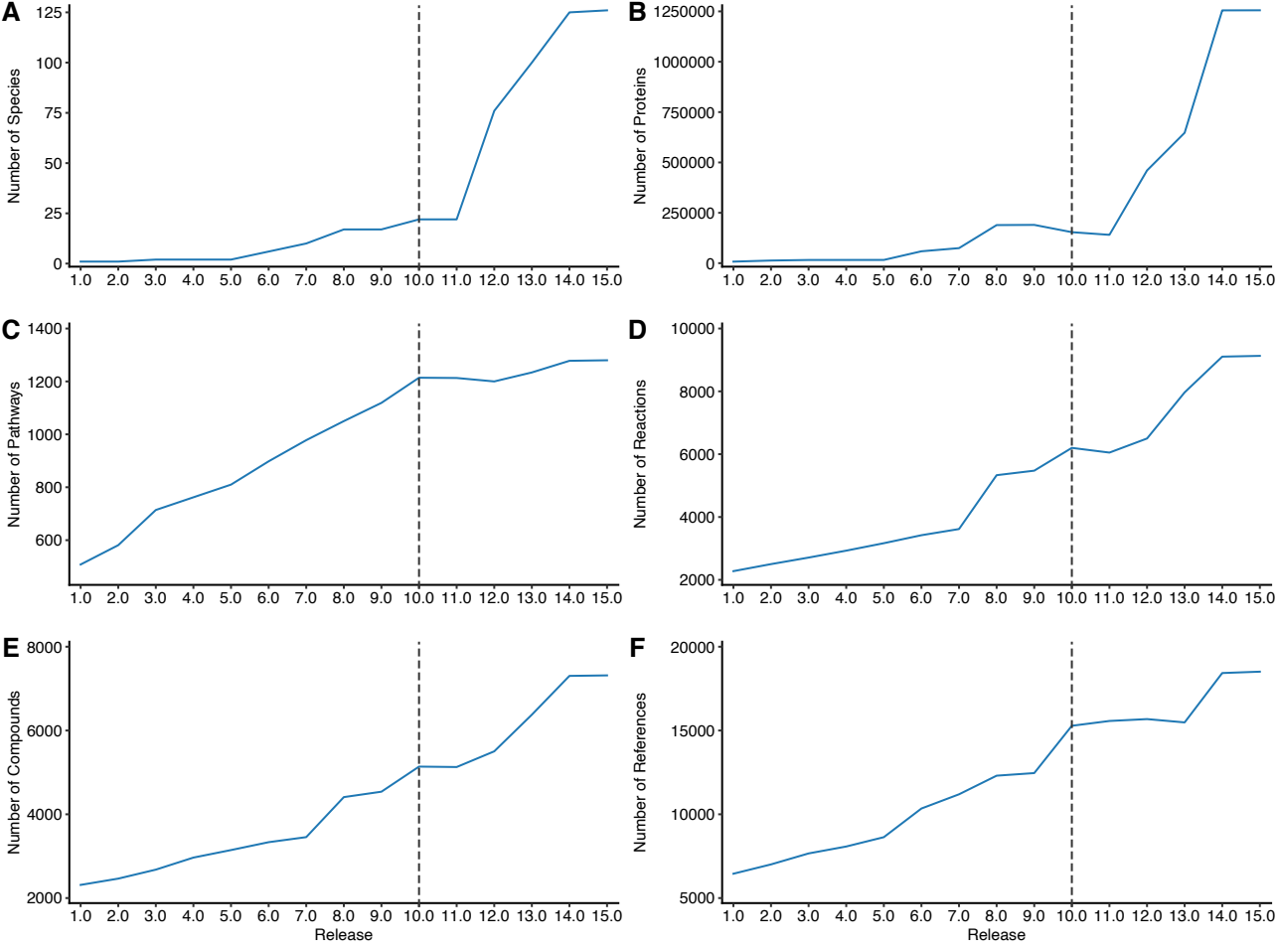

A
Number of Species
125
100
75
50
25
0
B
Number of Proteins
1250000
1000000
750000
500000
250000
0
C
Number of Pathways
1400
1200
1000
800
600
D
Number of Reactions
10000
8000
6000
4000
2000
E
Number of Compounds
8000
6000
4000
2000
F
Number of References
20000
15000
10000
5000
1.0 2.0 3.0 4.0 5.0 6.0 7.0 8.0 9.0 10.0 11.0 12.0 13.0 14.0 15.0
Release

Figure S1
Growth of PMN content over time. PMN version is shown in the X-axis of each plot, and number of database objects in the Y-axis. The vertical dotted lines indicate PMN release 10, the subject of the previous PMN paper (Schläpfer et al., 2017)

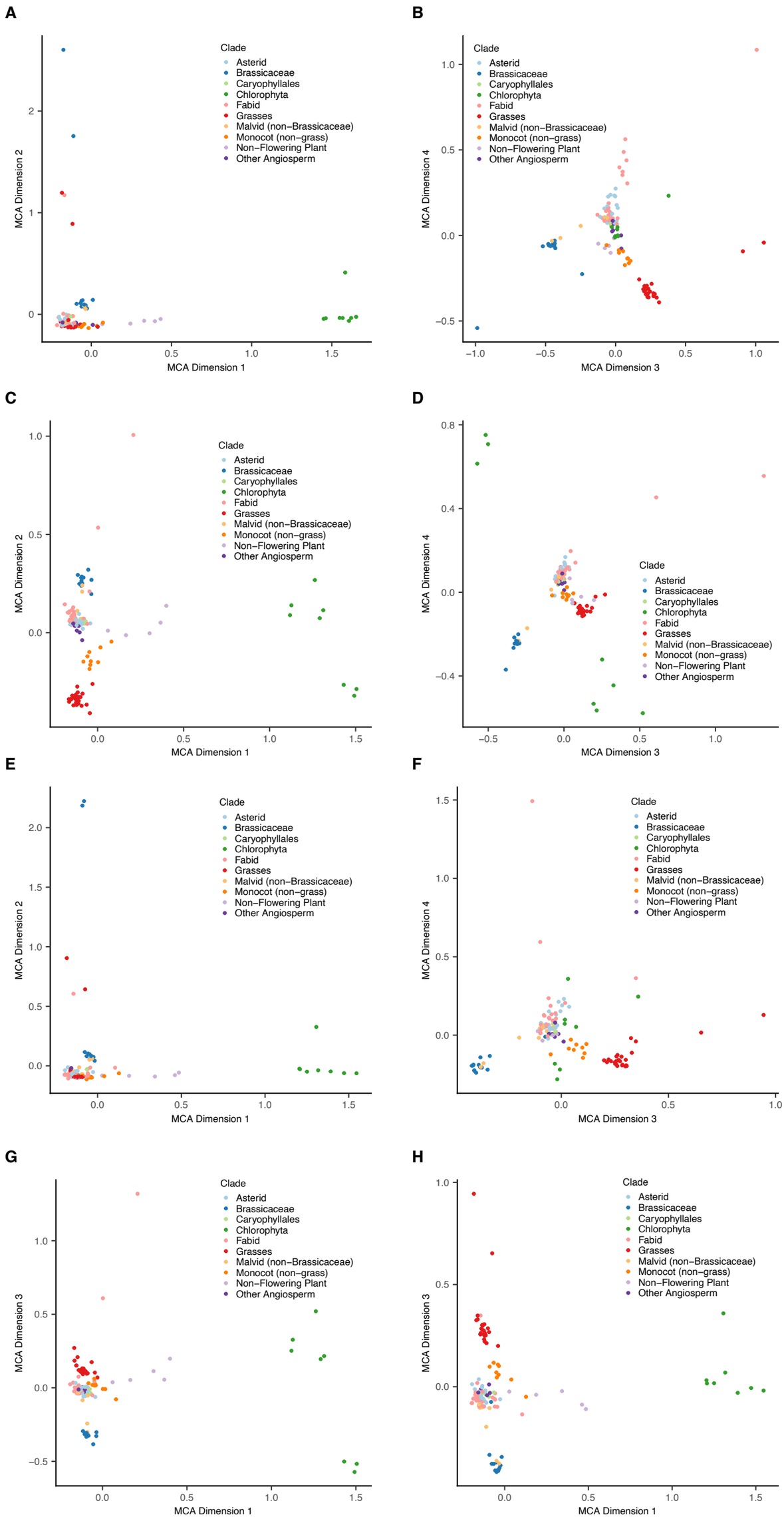

A
B
C
D
E
F
G
H
Clade
Asterid
Brassicaceae
Caryophyllales
Chlorophyta
Fabid
Grasses
Malvid (non−Brassicaceae)
Monocot (non−grass)
Non−Flowering Plant
Other Angiosperm
MCA Dimension 1
MCA Dimension 2
MCA Dimension 3
MCA Dimension 4

Figure S2, Related to Figure 3
Additional MCA plots on pathways (A, B), reactions (C, D, G), and compounds (E, F, H). A, C, and E show MCA dimensions 1 and 2; B, D, and F show MCA dimensions 3 and 4, and G and H show dimensions 1 and 3 of the reaction and compound MCAs.

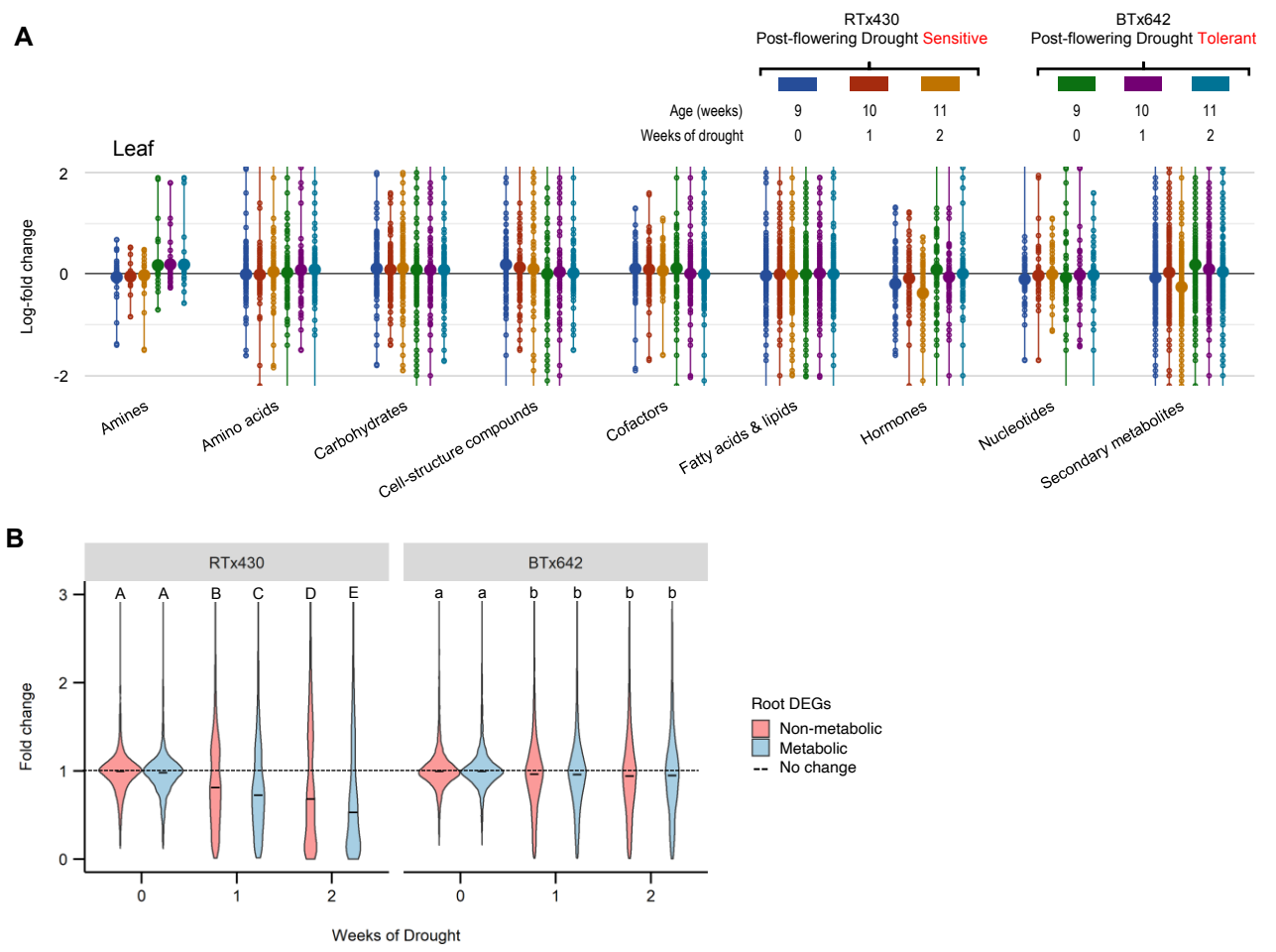
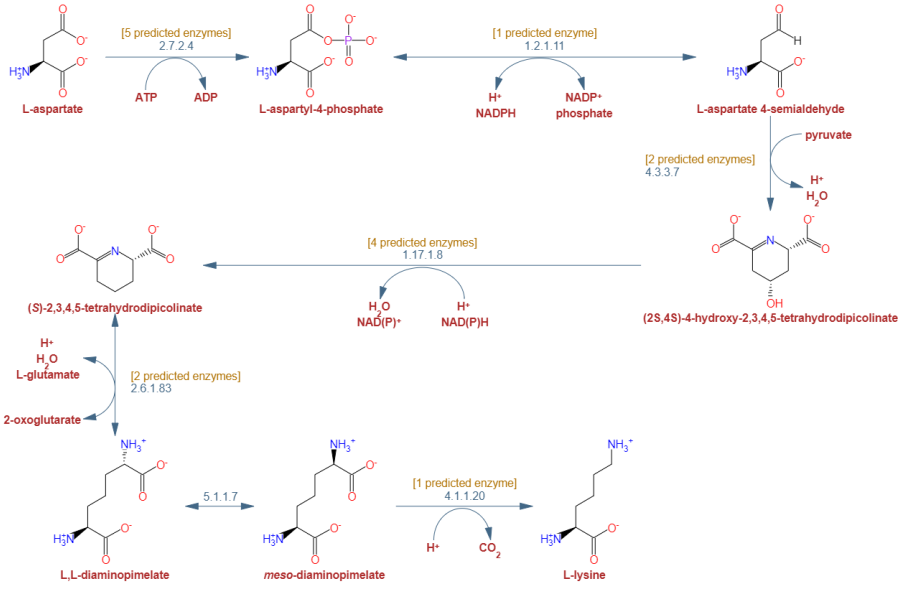
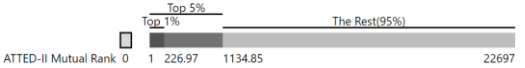

A
RTx430
Post-flowering Drought Sensitive
BTx642
Post-flowering Drought Tolerant
Age (weeks)
9
10
11
9
10
11
Weeks of drought
0
1
2
0
1
2
Leaf
Log-fold change
2
0
-2
Amines
Amino acids
Carbohydrates
Cell-structure compounds
Cofactors
Fatty acids & lipids
Hormones
Nucleotides
Secondary metabolites
B
RTx430
BTx642
A A B C D E
a a b b b b
Fold change
3
2
1
0
0
1
2
0
1
2
Weeks of Drought
Root DEGs
Non-metabolic
Metabolic
No change

Figure S3, Related to Figure 4
(A) Omics Dashboard representation of global metabolic biosynthetic gene expression patterns among differentially expressed genes (DEGs) in response to drought in leaf tissues of two sorghum cultivars. (B) Comparison of weekly expression levels between non-metabolic and metabolic genes in response to drought in both cultivars. Letters above the boxes indicate statistically different groups (two-way ANOVA followed by Tukey's HSD test, $\alpha = 0.05$).

A



B



Gene Name	EC Number	Reaction ID	Gene ID	Os01g0927900	Os07g0300900	Os09g0294000	Os08g0342400	Os03g0850400	Os03g0760700	Os04g0254000	Os03g0118600	Os03g0245100	Os02g0436400	Os03g0299900	Os03g0195100	Os02g0440000
Os01g0927900	2.7.2.4	ASPARTATEKIN-RXN	Os01g0927900	0	11560.59	13138.14	3402.7	10223.08	10556.17	5543.26	9563.69	4668.32	6871.98	1965.91	13219.74	1475.71
Os07g0300900	2.7.2.4	ASPARTATEKIN-RXN	Os07g0300900	11560.59	0	1516.98	3805.4	1787.97	3488.99	4891.11	10057.43	2565.21	942.9	890.95	16314.71	1992.58
Os09g0294000	2.7.2.4	ASPARTATEKIN-RXN	Os09g0294000	13138.14	1516.98	0	12974.38	219.15	8285.72	10057.43	12229.46	12478.81	7230.75	4668.32	14023.89	8285.72
Os08g0342400	2.7.2.4	ASPARTATEKIN-RXN	Os08g0342400	3402.7	3805.4	12974.38	0	4948	105.24	2188.74	9892.33	2370.46	588.57	715.69	13705.27	105.24
Os03g0850400	2.7.2.4	ASPARTATEKIN-RXN	Os03g0850400	10223.08	1787.97	219.15	4948	0	761.83	5299.6	8598.82	10140.19	1559.32	2738.2	10639.58	8598.82
Os03g0760700	1.2.1.11	ASPARTATE-SEMIALDEHYDE-DEHYDROGENASE-RXN	Os03g0760700	10556.17	3488.99	8285.72	105.24	761.83	0	1913.53	13382.72	6661.31	605.74	1197.87	12146.14	878.4
Os04g0254000	4.3.3.7	DIHYDRODIPICSYN-RXN	Os04g0254000	5543.26	4891.11	10057.43	2188.74	5299.6	1913.53	0	10306.1	2277.92	1085.88	1116.82	12395.8	306.63
Os03g0118600	1.17.1.8	RXN-14014	Os03g0118600	9563.69	10057.43	12229.46	9892.33	8598.82	13382.72	10306.1	0	4559.77	7674.42	7086.04	15250.75	11225.5
Os03g0245100	1.17.1.8	RXN-14014	Os03g0245100	4668.32	2565.21	12478.81	2370.46	10140.19	6661.31	2277.92	4559.77	0	1739.94	1913.53	14103.25	4613.81
Os02g0436400	1.17.1.8	RXN-14014	Os02g0436400	6871.98	942.9	7230.75	588.57	1559.32	605.74	1085.88	7674.42	1739.94	0	52.69	11393.05	143.27
Os03g0299900	2.6.1.83	RXN-7737	Os03g0299900	1965.91	890.95	4668.32	315.69	2738.2	1197.87	1116.82	7086.04	1913.53	52.69	0	16039.96	81.95
Os03g0195100	2.6.1.83	RXN-7737	Os03g0195100	13219.74	16314.71	14023.89	13705.27	10639.58	12146.14	12395.8	15250.75	14103.25	11393.05	16039.96	0	10057.43
Os02g0440000	4.1.1.20	DIAMINOPIMDECARB-RXN	Os02g0440000	1475.71	1992.58	8285.72	105.24	8598.82	878.4	306.63	11225.5	4613.81	143.27	81.95	10057.43	0

Figure S4: Pathway visualization tools
Exploration of the pathway visualization tools in PMN. (A) L-lysine biosynthesis VI (PWY-5097) in OryzaCyc 7.2.0. (B) co-expression view of the rice genes that code for enzymes which catalyze reactions in the lysine pathway. Genes are in both rows and columns. Each numerical cell shows the co-expression of the two genes as ATTED-II mutual rank (Obayashi et al. 2018). Lower numbers indicate stronger co-expression. Medium gray indicates the pairing is in the top 5% of the mutual rank score; dark gray indicates top 1%. Genes with no co-expression data have been manually removed from the table.

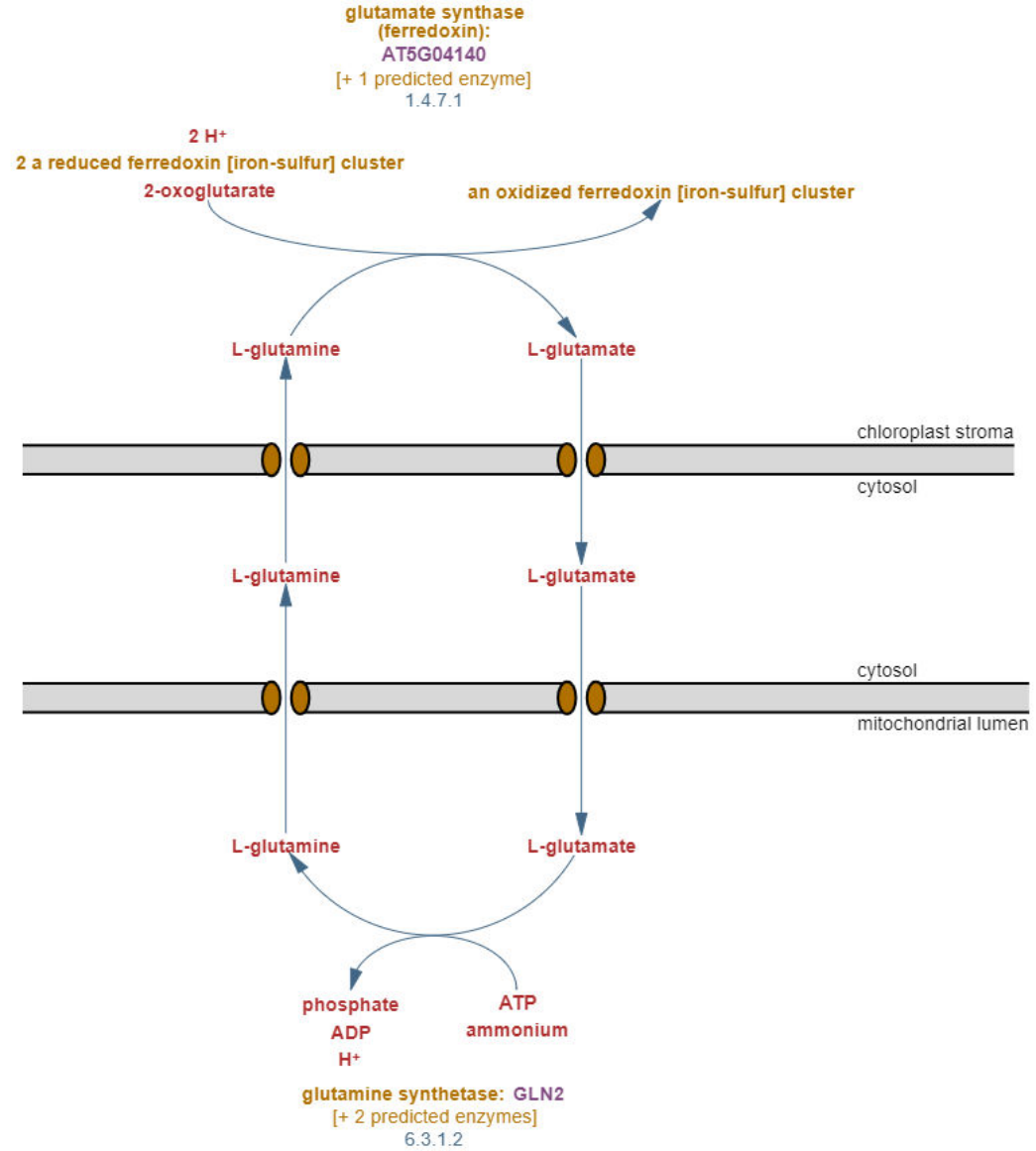

glutamate synthase (ferredoxin):
AT5G04140
[+ 1 predicted enzyme]
1.4.7.1
2 H+
2 a reduced ferredoxin [iron-sulfur] cluster
2-oxoglutarate
an oxidized ferredoxin [iron-sulfur] cluster
L-glutamine
L-glutamate
chloroplast stroma
cytosol
L-glutamine
L-glutamate
cytosol
mitochondrial lumen
L-glutamine
L-glutamate
phosphate
ADP
H+
ATP
ammonium
glutamine synthetase: GLN2
[+ 2 predicted enzymes]
6.3.1.2

Figure S5: Visualization of pathway subcellular localization
Subcellular localization view introduced in PMN 15. The glutamate-glutamine shuttle (PWY-7061) in AraCyc 17.1.1 is shown, with chloroplast and mitochondrial outer membranes displayed on the diagram.

Table S1
Growth of PMN over time, showing the number of pathways, reactions, compounds, enzymes, citations, and species present in each release.

Table S2
Listing of the 127 PMN databases (126 plant-specific databases + the pan-plant database PlantCyc), with database statistics for each.

Table S3, related to Figure S1

Domesticated plants and their wild relatives in PMN. PMN contains databases for 8 domesticated species that have one or more wild relatives for which PMN also contains a database.

Table S4, Related to Figure 2

Summary of the manual pathway assessment results. Number (percentage) of the 120 randomly-selected pathways that were assigned to each of the 5 assessment categories before SAVI (Columns A, B) and how many (what percentage) of them were in each category in the released databases (Columns C, D).

Data S1, Related to Figure 2
Manual pathway assessment results. 120 pathways were selected at random and the taxonomic distribution of each pathway was manually evaluated in the prediction-only and in the released databases. Pathways in CAPP are accepted in a species if they were predicted, and if the species is in the expected range or has predicted enzymes for at least one key reaction of the pathway. Pathways with no SAVI rule are accepted if predicted, as if they were in Accept if Predicted Pathways (AIPP; Schläpfer et al., 2017).

Data S2, Related to Figure 3
Pathway variables and their association with dimensions 1 and 3 of the pathway MCA shown in figure 3A. A pathway variable is the presence or absence of a pathway in a database, which are treated as separate variables in an MCA. Only the 613 pathways that pass filtering (present in at least two species and absent in at least two species) are shown. The Domain Background sheet is the background used for the metabolic domain enrichment (Figure 3B, C). The last three sheets give the first 5 MCA coordinates of each species based on pathways, reactions, and compounds.

Data S3, Related to Figure 5
Datasets used for Index of Cell Identity (ICI) computation and single-cell reference construction. ICI Source Datasets: We used a mixture of RNA-seq and ATH1 microarray datasets published previously to impute putative cell types for each single-cell transcriptome. pSCREEN Datasets: For single-cell RNA-seq datasets, we sourced data from 5 independent studies resulting in a set of 21 individual single-cell datasets.

Data S4-S6, Related to Figure 3
Presence-absence matrices for pathways (S4), reactions (S5), and compounds (S6). The files are tab-delimited text; rows represent the databases, while columns represent the database objects (Pathways, Reactions, or Compounds (PRCs)). “1” indicates the PRC is present in the organism database, while “0” indicates it is absent.

Data S7, Related to Figure 3
Plant group classifications for the 126 PMN species, used to color the MCA plots (Figures 3A, S2). Tab-delimited text. The first column is the database name, and the second is the plant group. The groups were manually chosen to best represent the diversity of species in PMN.

Data S8
Metabolic domain mapping version 2.0. This table updates the pathway domain mapping in Schläpfer et al., 2017’s Supplemental Table S8 with 68 new pathways classified with metabolic domains. The new pathways are highlighted in green.

File S1
User’s guide for PMN 15.

File S2, Related to Figure 3
Lisp code used to generate the PRC matrices (Data S4-S6). Runs within the Pathway Tools software; see the comment at the head the file for instructions.