

mMAP: mouse Metabolomics Analysis Platform

Preeti Bais

09/09/2014

Outline

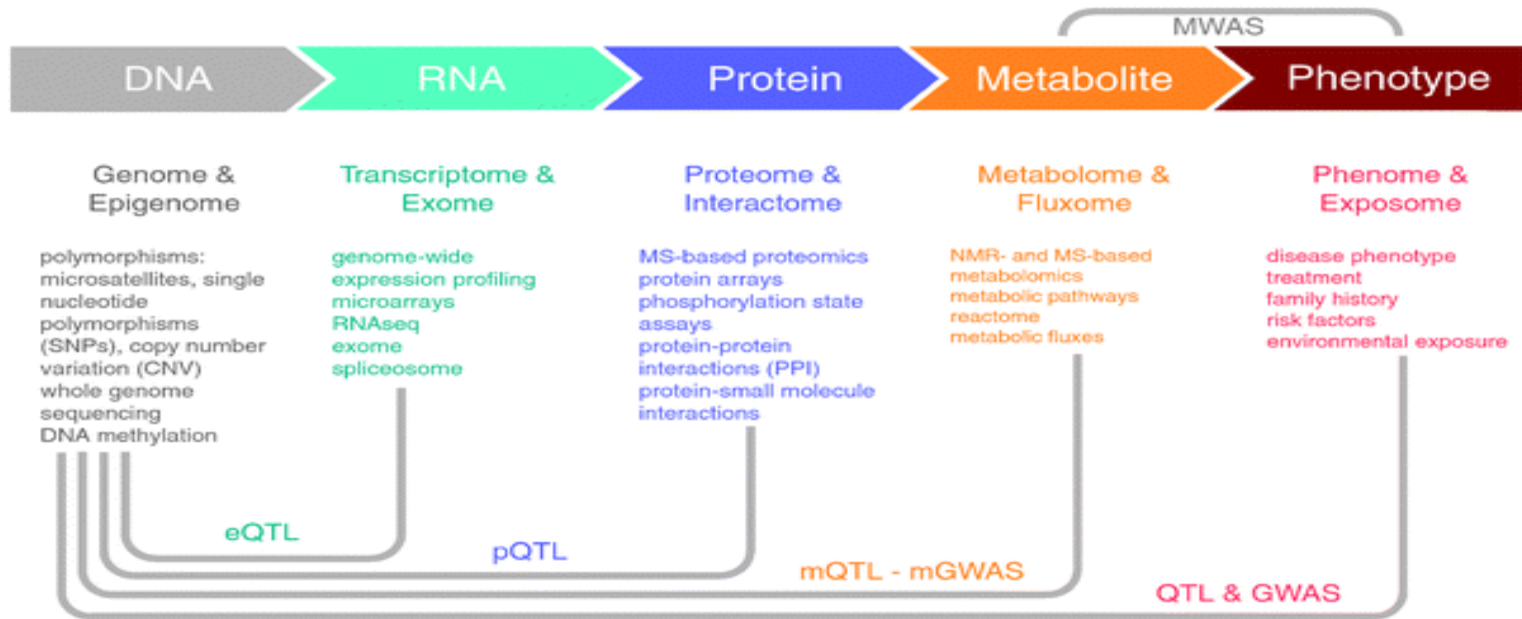
1. Introduction: What is Metabolomics ?

2. Problem: Need for Bioinformatics Infrastructure

3. Solution: mMap Database and Pipeline



What is Metabolomics?



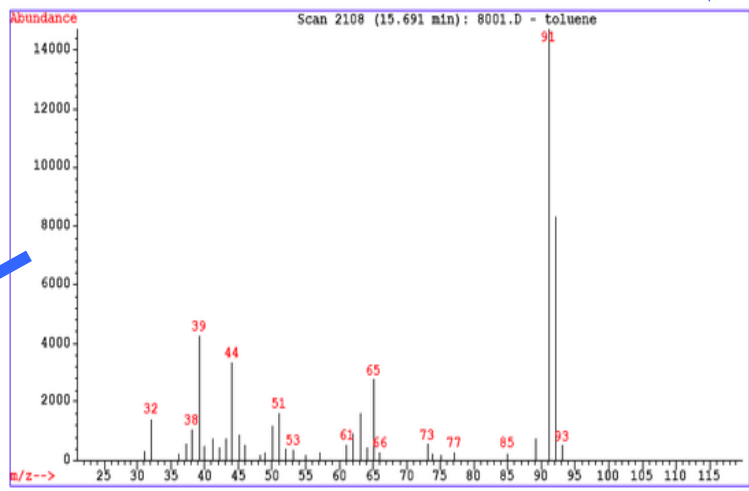
Mol. BioSyst., 2012,8, 2494-2502



Metabolomics Workflow



	Methionine	Serine	Glu	Metabolite 10	Metabolite 11	Metabolite 12	Metabolite 13
MT1	1.0	23.3	2000				..
MT2	1.5	23	2200				..
MT3	1.3	22	2300				..
WT1	100.2	22	2500				..
WT2	111.1	22	2500				..
WT3	100	22.1	2200				..



Project Goals

- To identify metabolic pathways related to pathogenesis in mouse models of 3 neuromuscular degeneration mutants
- To develop informatics tools to assign metabolomics data to curated biochemical pathways and functional annotations for mouse genes and proteins

Data Generation -> Data Analysis -> Data Interpretation

Metabolites->Biochemical Pathways->Genes->Annotations



Bioinformatics infrastructure is lacking for metabolomics data analysis and interpretation

- Metabolite names are variable and are used inconsistently (e.g. B-Alanine, Beta-Alanine)
- Very few metabolomics data analysis pipelines are designed for the biologists (e.g HMDB, plantmetabolomics.org)
- Need to combine the metabolomics data with genomics data and biological annotations for interpretation



mMAP Features: Data Base and Data Analysis Pipeline

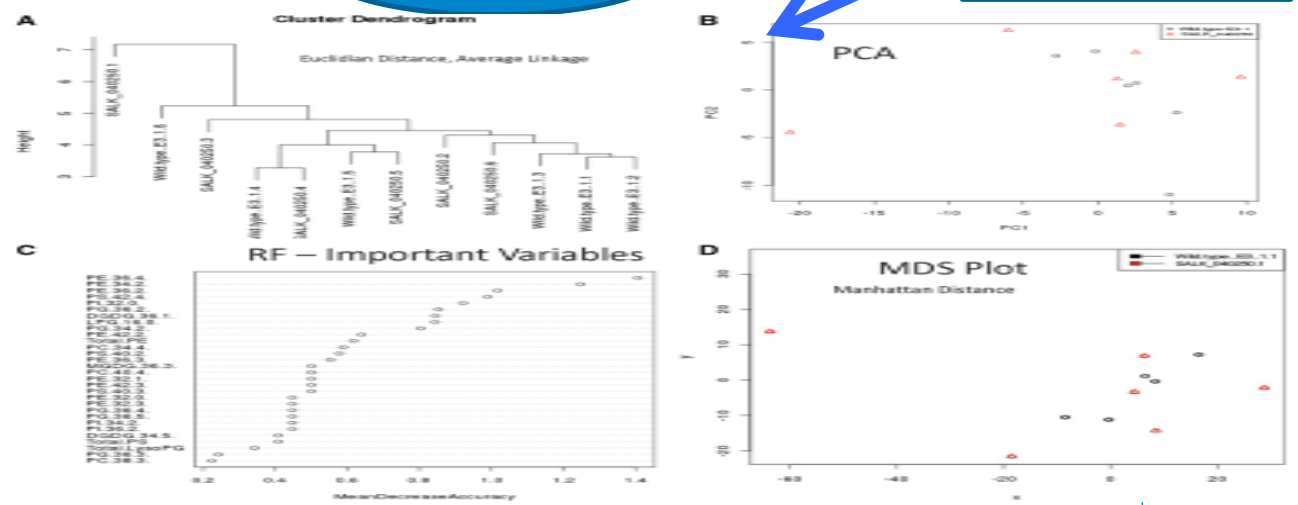
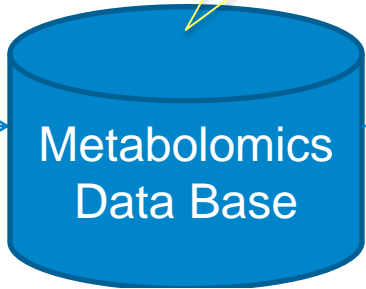


Vs



	Methionine	Serine	Glu	Methionine 10	Methionine 11	Methionine 12	Methionine 13
MT1	1.0	23.3	2000				..
MT2	1.5	23	2200				..
MT3	1.3	22	2300				..
WT1	100.2	22	2500				..
WT2	111.1	22	2500				..
WT3	100	22.1	2200				..

Metabolomics Data Reporting Standards



mMap Features: Advanced Data Analysis Tools

Data Analysis

Missing Value Imputation	Step 1
Data Normalization	Step 2
Hierarchical cluster analysis	Step 3- A

?

Distance Measure
Euclidean

Linkage Method
Complete

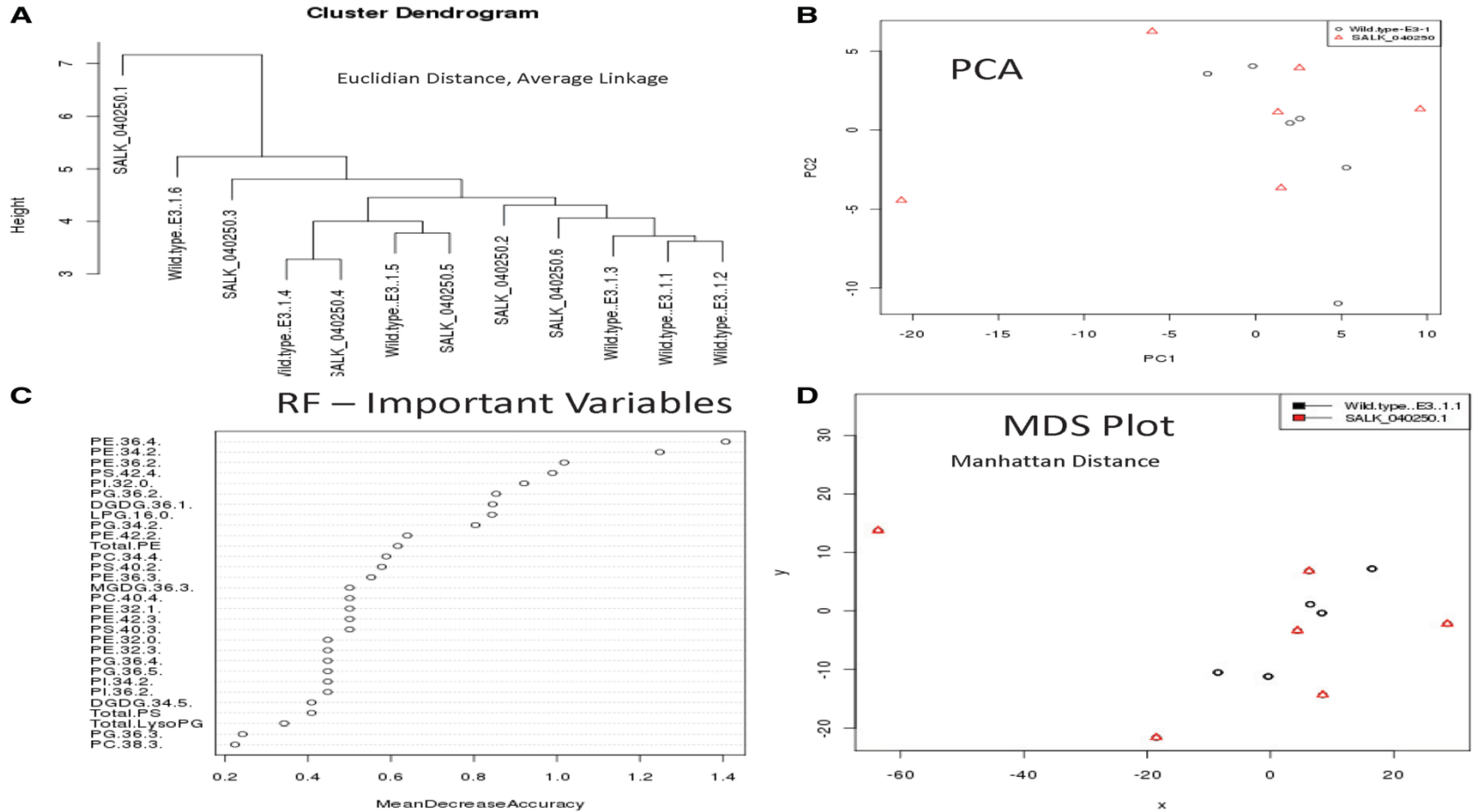
Submit

MDS Plot analysis	Step 3 -B
PCA Analysis	Step 3 -C
Random Forest Analysis	Step 3 -D

- Menu Driven automated data analysis pipeline
- Best practices in metabolomics data analysis
- Connected to genomics information
- Interactive plots : Download/Print
- Flexible pipeline
- Help in interpretation of plots with examples



mMap Features: Data Analysis Plots



mMap Features : Help in Interpreting the Plots

Browse Feedback Help

Data Analysis

Missing Value Imputation

Step 1

Threshold % Of Missing values for a metabolite

50

Hint: The metabolites with more missing values than the given threshold will be discarded from the next analysis steps. Metabolites with less than given percentage of missing values will be imputed by the mean of the concentration for that metabolite.

The default is 50%.

Data Normalization

Step 2



- Range Scaling (Default)
- Log Transformation
- Pareto Scaling
- Auto Scaling

Hierarchical cluster analysis

Step 3 - A

MDS Plot analysis

Step 3 - B

PCA Analysis

Step 3 - C

Random Forest Analysis

Step 3 - D

Data file created at [DataFile](#)

Please provide your choices for Missing Value imputation (Step 1)
[Contact Us](#)

Cluster Analysis

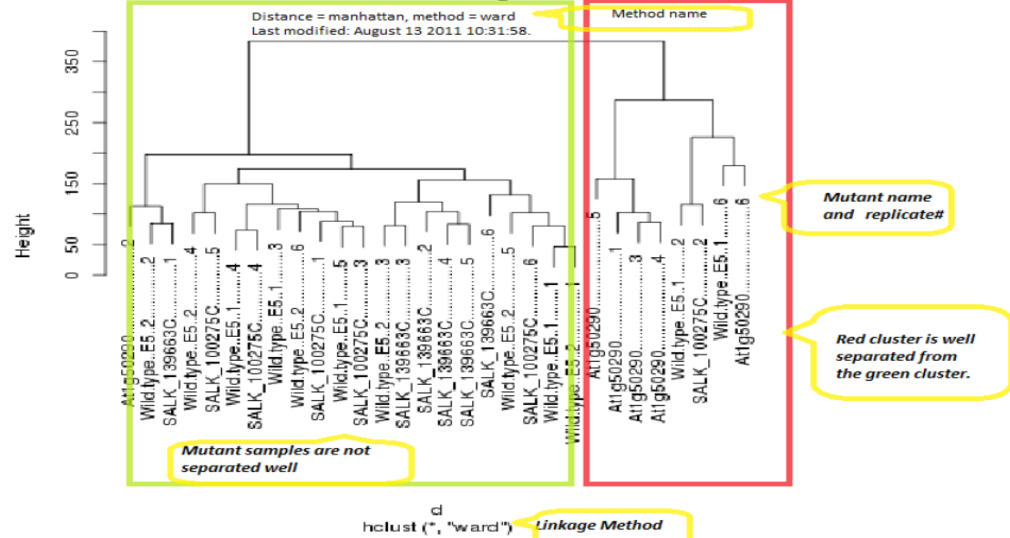
How to interpret this plot?

Please click on the picture to see an example.



Distance = euclidean, method = complete
Last modified: August 27 2014 11:20:54.

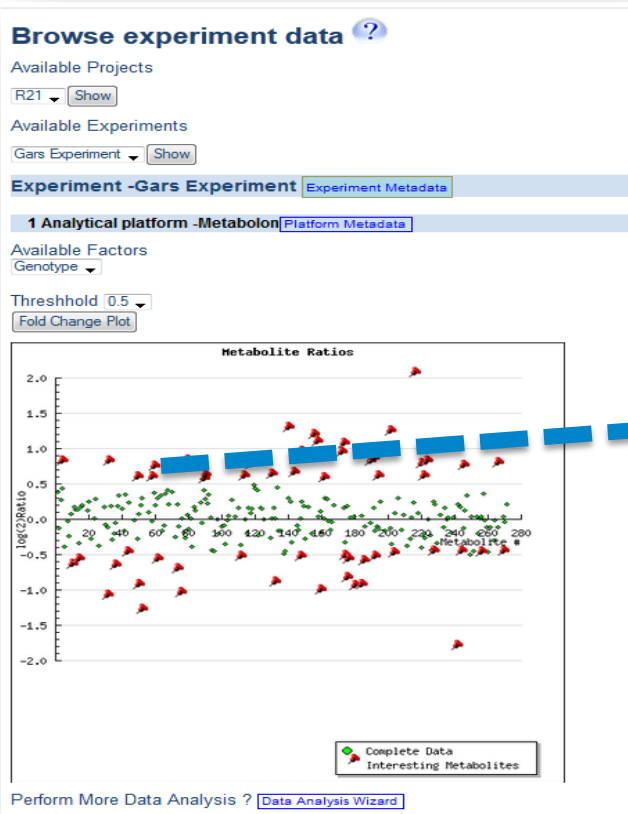
Cluster Dendrogram



This cluster dendrogram shows that the samples of the mutant at11g50290 are clearly separable from the rest of the samples. Manhattan distance measure and Ward linkage methods were used for this plot. One can choose different methods from the drop down menus on the left side and plot. Please click on the help icon to get more information about the choice of methods.



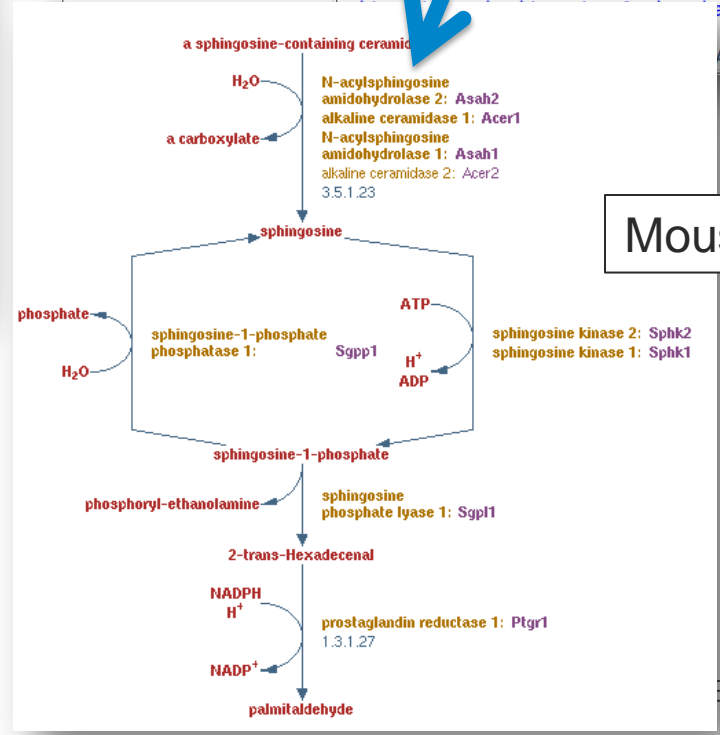
mMAP Features : Interactive Plots



Metabolite Details: sphingosine	
Compound Common Name:	sphingosine
Compound Synonyms:	Sphingenine,sphingosine
Molecular Weight:	299.496
Chemical Formula:	(C 18) (H 38) (N 1) (O 2)
Smiles Notation:	C(CCC)CCCCCCCC=CC(C(CO)[N+])O
PUBCHEM:	25246036
MouseCyc Pathways:	sphingosine and sphingosine-1-phosphate metabolism (As a Product) { MGI } Genes: Sphk1,Sphk2,Sgpl1,Sgpp1,Ptgr1,Asah1,Acer1,Asah2,Acer2 sphingosine and sphingosine-1-phosphate metabolism (As a Product) { MGI } Genes: Sphk1,Sphk2,Sgpl1,Sgpp1,Ptgr1,Asah1,Acer1,Asah2,Acer2 sphingosine and sphingosine-1-phosphate metabolism (As a Reactant) { MGI } Genes: Sphk1,Sphk2,Sgpl1,Sgpp1,Ptgr1,Asah1,Acer1,Asah2,Acer2

mMAP Metabolite Details

MouseCyc Pathway



mMAP Data Analysis Results



mMAP Features: Links to curated pathways, genes, and proteins

Metabolite Details: lanosterol	
Compound Common Name:	lanosterol
Molecular Weight:	426.724
Chemical Formula:	(C 30) (H 50) (O 1)
Smiles Notation:	<chem>CC(C)=CCCC(C)C1(CCC2(C)C3(CCC4(C(C)C)C(O)CCC(C)C(CCC(C)12)=3)4))</chem>
CAS:	79-63-0
PUBCHEM:	25200754
MouseCyc Pathways:	<p>lanosterol biosynthesis (As a Product) { MGI }</p> <p>Genes: Lss (As a Reactant) { MGI }</p> <p>Genes: Sc4mol,Hsd17b7,Ebp,Dhcr7,Sqle,Fdft1,Sc5d,Lss cholesterol biosynthesis I (As a Reactant) { MGI }</p> <p>Genes: Cyp51,Sc4mol,Hsd17b7,Ebp,Dhcr24,Dhcr7,Sqle,Fdft1,Tm7sf2,Sc5d,Nsdhl,Lss cholesterol biosynthesis III (via desmosterol) (As a Reactant) { MGI }</p> <p>Genes: Cyp51,Sc4mol,Hsd17b7,Ebp,Sc5d,Dhcr7,Sqle,Fdft1,Tm7sf2,Nsdhl,Lss</p>

mMAP Metabolite Details

MouseCyc@MGI
MouseMine@MGI

<http://www.informatics.jax.org/>



mMAP Features: Links to curated pathways, genes, and proteins

Metabolite Details: sphingosine
 Compound Common Name: sphingosine
 Compound Synonyms: Sphingenine, sphingosine
 Molecular Weight: 299.496
 Chemical Formula: (C 18) (H 38)

mMAP Metabolite Details

Mousecyc Pathways:

- sphingosine-1-phosphate metabolism (As a Product) [MGI]
- Genes: Sphk1, Sphk2, Sgpl1, Sgpp1, Ptgr1, Asah1, Acer1, Asah2, Acer2
- sphingosine and sphingosine-1-phosphate metabolism (As a Product) [MGI]
- Genes: Sphk1, Sphk2, Sgpl1, Sgpp1, Ptgr1, Asah1, Acer1, Asah2, Acer2
- sphingosine and sphingosine-1-phosphate metabolism (As a Reactant) [MGI]
- Genes: Sphk1, Sphk2, Sgpl1, Sgpp1, Ptgr1, Asah1, Acer1, Asah2, Acer2

ID	Symbol	Name	Type	Chromosome
MGI:1261415	Sgpl1	sphingosine phosphate lyase 1	protein_coding_gene	10
MGI:1277124	Asah1	N-acylsphingosine amidohydrolase 1	protein_coding_gene	8
MGI:1316649	Sphk1	sphingosine kinase 1	protein_coding_gene	10
MGI:1859310	Asah2	N-acylsphingosine amidohydrolase 2	protein_coding_gene	10
MGI:1861380	Sphk2	sphingosine kinase 2	protein_coding_gene	7
MGI:1914353	Ptgr1	prostaglandin reductase 1	protein_coding_gene	4
MGI:1920932	Acer2	alkaline ceramidase 2	protein_coding_gene	4
MGI:214576	Cerap1	ceramide 1-phosphate phosphatase 1	protein_coding_gene	12
MGI:214577	Cerap2	ceramide 2-phosphate phosphatase 2	protein_coding_gene	17

Genes list in MouseMine@MGI

Gene Ontology Enrichment

GO terms enriched for items in this list. All items in your list have been analysed.

Test Correction: Holm-Bonferroni
 Max p-value: 0.05

Background population: Default

View Download

- GO Term**
- sphingosine metabolic process [GO:0006670]
- sphingolipid metabolic process [GO:0006665]
- diol metabolic process [GO:0034311]
- sphingoid metabolic process [GO:0046519]
- membrane lipid metabolic process [GO:0006643]
- polyol metabolic process [GO:0019751]

Anatomy Enrichment

Anatomy (EMAPA) terms enriched for items in this list.

Number of SequenceFeatures in this list not analysed: 0

Test Correction: Holm-Bonferroni
 Max p-value: 0.05

Background population: Default

View Download

- EMAPA Term**
- atrioventricular canal p-Value: 0.010266

Mammalian Phenotype Ontology Enrichment

MP terms enriched for items in this list.

Number of SequenceFeatures in this list not analysed: 0

Test Correction: Holm-Bonferroni
 Max p-value: 0.05

Background population: Default

View Download

- MP Term**
- abnormal sphingolipid level [MP:0012230] 0.036130 2
- abnormal ceramide level [MP:0012232] 0.036130 2

MouseMine@MGI
 Term Enrichment Analysis and
 Associated Disease
 Phenotypes

ID	Symbol	Base Annotations » Subject . Symbol	Subject » Background	Ontology Annotations » Ontology Term . Identifier	Ontology Annotations » Term Name
MGI:1277124	Asah1	Asah1<tm1Esc>/Asah1<+>	involves: 129S1/Sv * C57BL/6	OMIM:228000	Farber Lipogranulomatosis

Summary and Future Goals

- Provides a computational metabolomics platform for JAX researchers to incorporate metabolomics data in their research. Will be available to the community as well
- The automated pipeline is modular and can easily be extended to experiments to include different factors (e.g. Mutation, Drug, Disease etc.)
- Currently the genomic annotations for mice are available but it can be easily extended to other organisms with Biocyc Pathways (e.g. Worm, Fly, Yeast)



Metabolomics of Peripheral Neuropathy Team

- **PI:** Gregory A. Cox



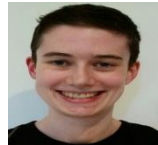
- **PI:** Robert W. Burgess



- **PI:** Carol J. Bult



- Lauren Kuffler



- Preeti Bais



Funding – NIH
R21NS082666

