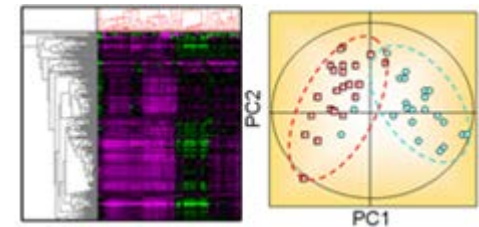


植物代謝物プロファイリングデータベース AtMetExpressの開発とオミックスデータ 統合化の推進

福島 敦史

理化学研究所・環境資源科学研究センター (CSRS)



内容

1. 研究背景
2. 研究開発の目的・全体像
3. 結果
4. まとめ

モデル植物シロイヌナズナの メタボローム

植物界全体では推定
~200,000代謝物が存在

メタボロミクス

~5,000 metabolites?

The AraCyc 8.0 release from April 2011 contains 446 pathways, 5520 enzymes, 2689 reactions, and 2825 compounds

プロテオミクス

(トランスクリプトミクス)

~4,000-5,000 enzymes

ゲノミクス

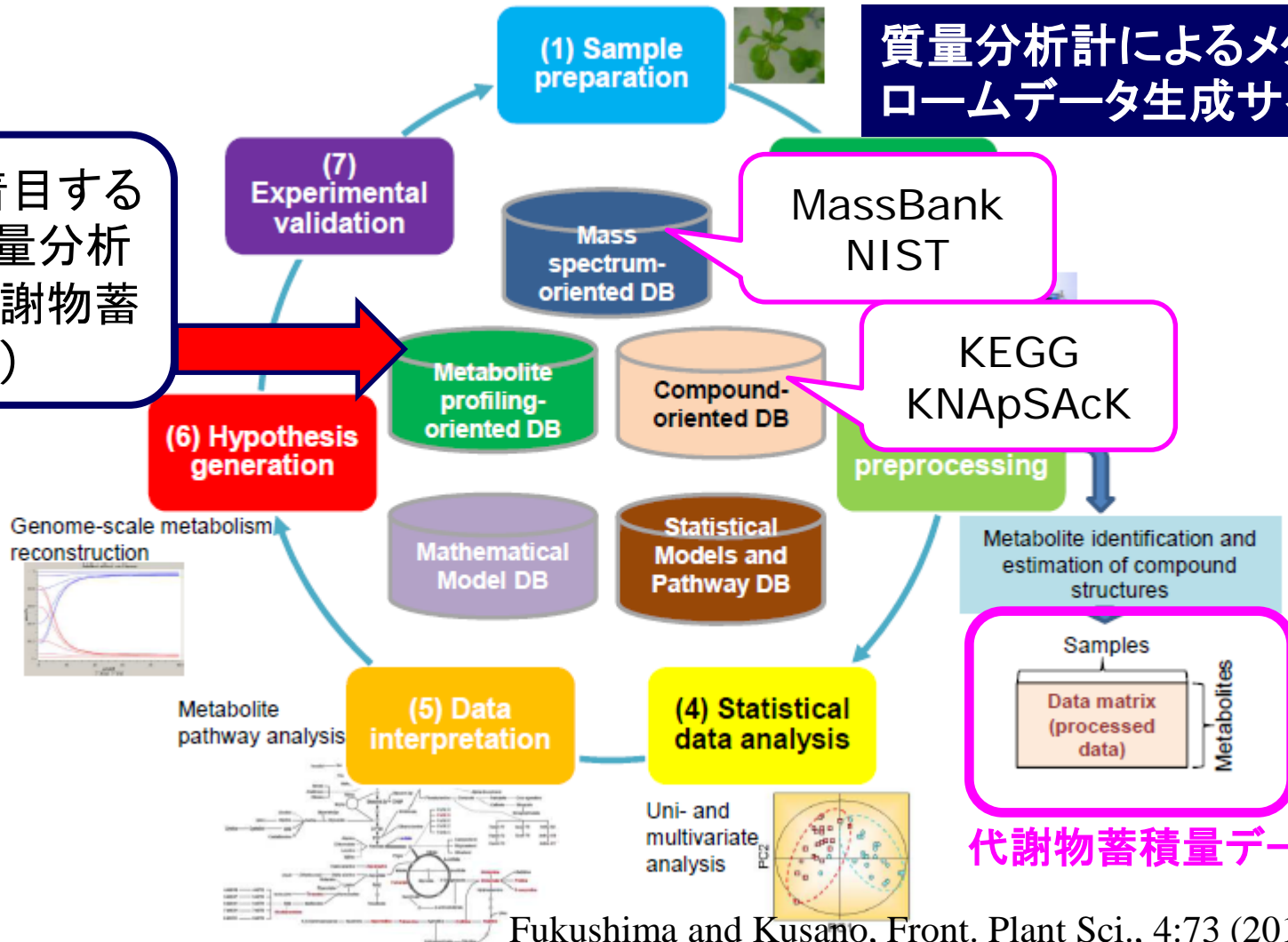
~28,000 genes

メタボロミクスは細胞内代謝物の総体を 化学分析する学問

生命の根本である「情報」と「物質」とが交差する特異的な分野の一つ

質量分析計によるメタボロームデータ生成サイクル

本課題で着目するデータ (質量分析計による代謝物蓄積量データ)



代謝物蓄積量データ行列

例. PlantMetabolomics.org

Arabidopsis Metabolomics Consortium

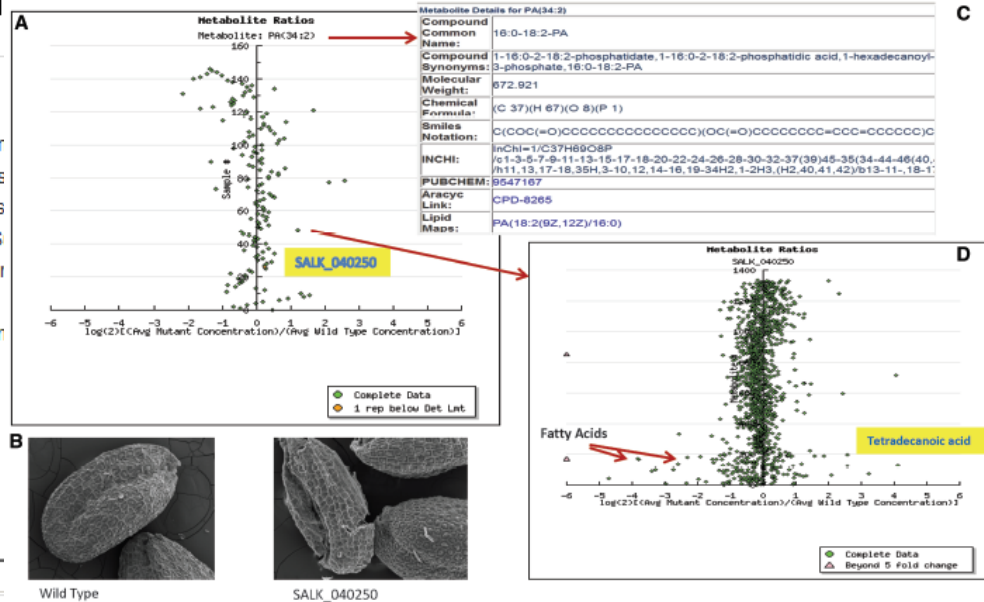
Home	About	Datasets	Tools	Tutorials	Gene Targets	Search
------	-------	----------	-------	-----------	--------------	--------

About PlantMetabolomics

This is an NSF-funded multi-institutional project that is developing metabolite Arabidopsis genes. The Consortium has established metabolomic platforms chemically defined. Initial studies have focused on investigating the robustness minimize the environmental and developmental effects on the metabolome. S Arabidopsis mutant lines in genes whose functions are currently not fully understood. T-DNA knockout alleles for these targeted genes, some of which have been integrated with phenotypic data and data concerning protein function, transcription concerning the functions of the targeted genes.

What Can You Do?

Search for your Metabolite:



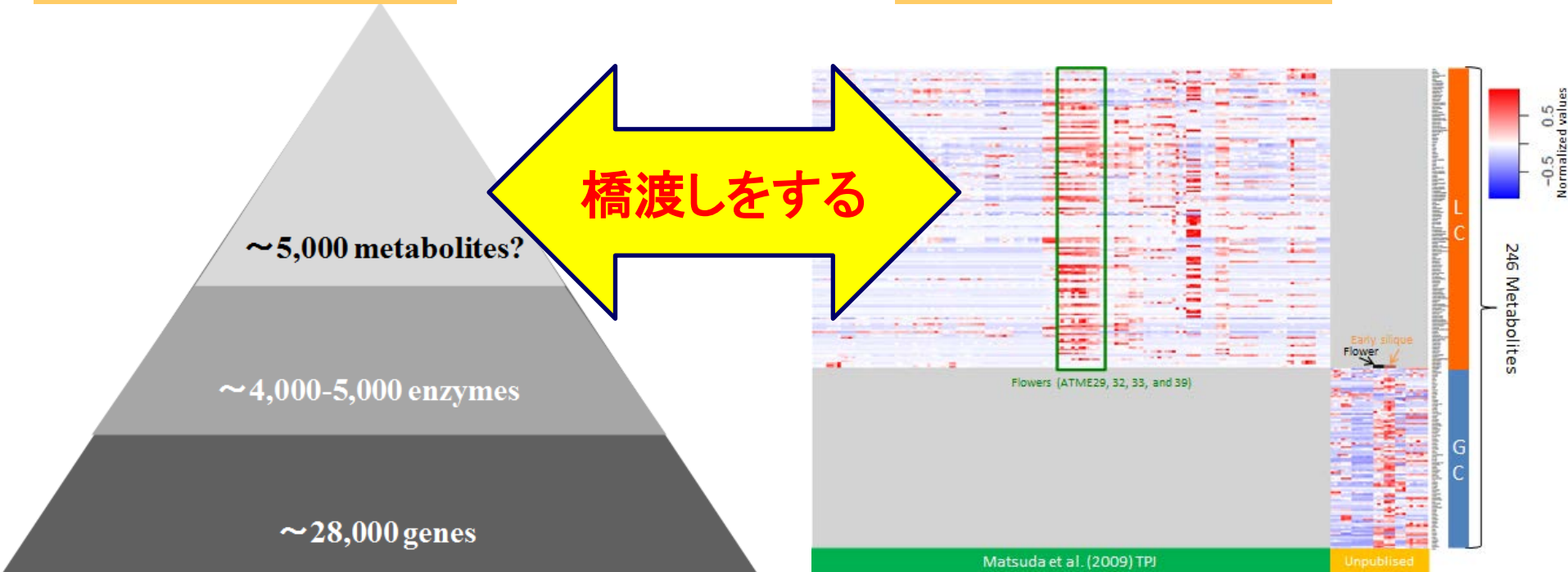
Bais et al., Plant Physiol (2010); Bais et al., NAR (2012);
<http://plantmetabolomics.vrac.iastate.edu/ver2/>

See also, Plant Metabolomics Resource (PMR)

研究開発の目的

予測されるメタボローム

測定可能なメタボローム

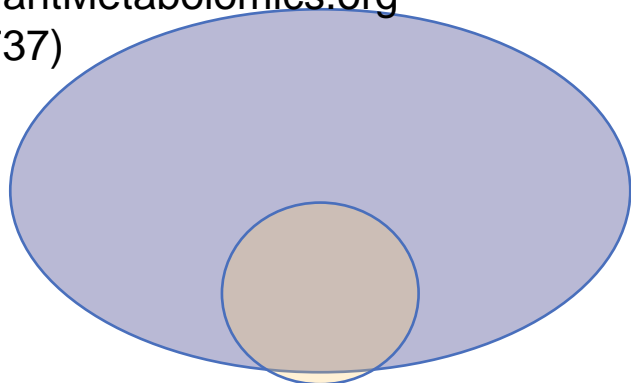


- モデル植物シロイヌナズナのメタボローム把握するための文献整理・データ再解析
- シロイヌナズナの代謝物プロファイルデータを閲覧・加工・共有するフレームワークの開発・情報基盤づくり

研究開発の概要

化合物データセット

PlantMetabolomics.org
(737)



KNApSAcK (625)



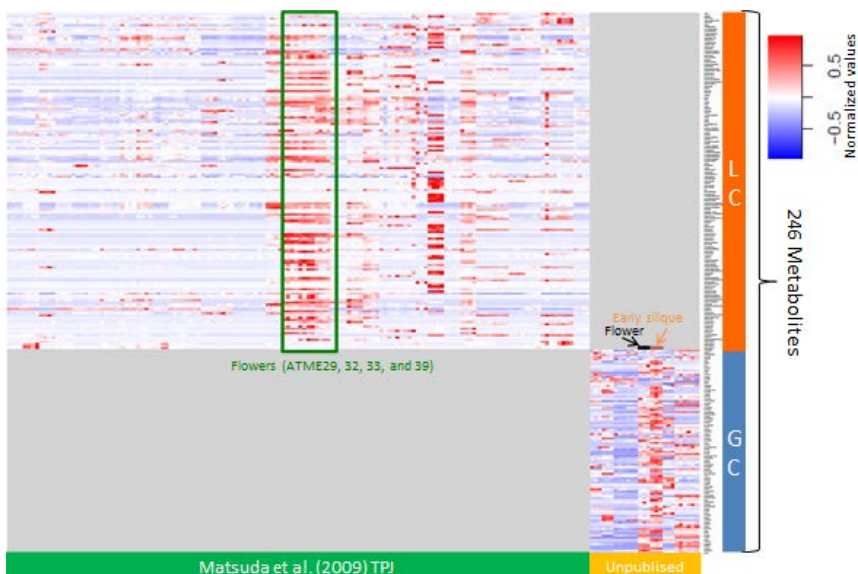
RIKEN

代謝物プロファイル測定データ

- GC-TOF-MS (83*, 92, 93)
- CE-TOF-MS (74, 60)
- UPLC-q-TOF-MS (106, 289, 99)
- LC-IT-TOF-MS (55, 96)

*括弧内は、測定できた代謝物の数

AtMetExpressの開発



Wikiスタイルページ



AtMet Discussion Read View source View history Go Search

AtMet:Malate::1

Malate	
Suzukichack	Pmdetect
PmdetectAracyclID	malic acid
Pmdetectname	Malate
Pmdetectkegg	malic acid
KNApSAcK-name	l-malic acid,636-61-3
Molecular_formula	C4H6O5
MW_calculate_by_MWC	134.0215226
Organism	Arabidopsis thaliana
datasetname	17128_Nikolau_DataSheet2
each_dataset_Uniq_ID	PM714

IDs and Links	
PUBCHEM	222656
Aracycl_ID	MAL
CAS_ID	97-67-6
KEGG	C00149
KNApSAcK-ID	C00001192

Navigation: Main Page, Recent changes, Random page, Help

metabolites

Flavonoid

Basic metabolites

Crude Drug

Plant taxa

Toolbox

What links here

Related changes

Special pages

Tweet Like Be the first of your friends to like this.

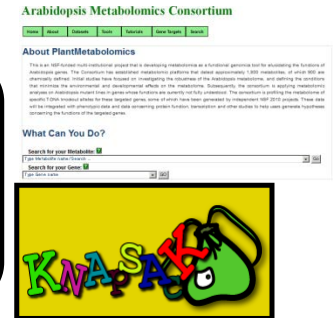
This page was last modified on 7 February 2014, at 11:00.

収集済みのデータセットは14

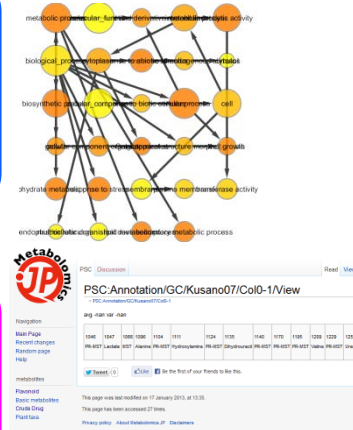
データセット名	PMID	検出した既知代謝物数	
Matsuda et al. (2010) Plant Physiol - AtMetExpress Development	20023150	289	Published data
Matsuda et al. (2009) Plant J – MS2T paper	18939963	99	
Matsuda et al. (2011) Front in Plant Sci – AtMetExpress Ecotype	22645535	106	
Hirai et al. (2010) JPR – LC, ecotype	20369372	38	
Ide et al. (2011) JXB – GC and CE data	21131548	74	
Kusano et al. (2007) BMC Syst Biol – GC data	18028551	93	
Kusano et al. (2010) Amino Acids – GC data	20354740	87	
Kusano et al. (2011) Plant J – GC and LC data	21466600	92	
Okazaki et al. (2012) Metabolomics – LC-IT data	23463370	96	
Watanabe et al. (2008) Plant Physiol – GC and CE data	18024555	60	
Fukushima et al. (2009) PNAS – GC data	19359492	83	
AtMetExpress Development (GC data), unpublished	NA	93	Unpublished data
Time course GC data - unpublished	NA	78	
Development of MeKO database (GC data)	NA	99	

実施した研究項目

1) 既存の化合物情報の収集・統合 (H25年10月)



2) 代謝物プロファイルデータの収集・再解析 (H25年9-10月)



3) Wiki化・データ視覚化・データ共有 (H25年11月-H26年1月)



→ シロイヌナズナのメタボローム中、質量分析計で測定可能な代謝物数はおおよそ1,200個

Metabolite information in Wiki



Navigation

Main Page

Recent changes

Random page

Help

metabolites

Flavonoid

Basic metabolites

Crude Drug

Plant taxa

Toolbox

What links here

Related changes

Special pages

Log in

AtMet

Discussion

Read

View source

View history

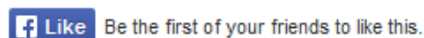
Go

Search

AtMet:Malate::1

Malate	
Suzukicheck	Pmdetect
PmdetectAracycID	malic acid
Pmdetectname	Malate
Pmdetectkegg	malic acid
KNApSAcK-name	l-malic acid,636-61-3
Molecular_formula	C4H6O5
MW_calculate_by_MWC	134.0215226
Organism	Arabidopsis thaliana
datasetname	17128_Nikolau_DataSheet2
each_dataset_Uniq_ID	PM714

IDs and Links	
PUBCHEM	222656
Aracyc_ID	MAL
CAS_ID	97-67-6
KEGG	C00149
KNApSAcK-ID	C00001192



This page was last modified on 7 February 2014, at 11:00.

再解析ツール (AtMetExpress): a heatmap

PRIME Visualization Tools

RIKEN PRIME provides web-based data analysis and visualization tools for public access. Users may analyze datasets from both AtMetExpress and MeKO, as well as upload custom datasets.

Analysis parameters

Data source: *AtMetExpress*

Dataset: *Kusano10AminoAcids*

Downloads

None

Edit parameters

Summary (MVA)

Hierarchical Cluster

Heat Map

Principal Component Analysis

hclust Method

Complete

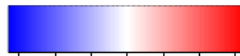
dist Method

Euclidean

scaling Method

auto

Color Key



0 2 4 6

Select data source

AtMetExpress

MeKO

GC-MS (RIKEN format)

Generic

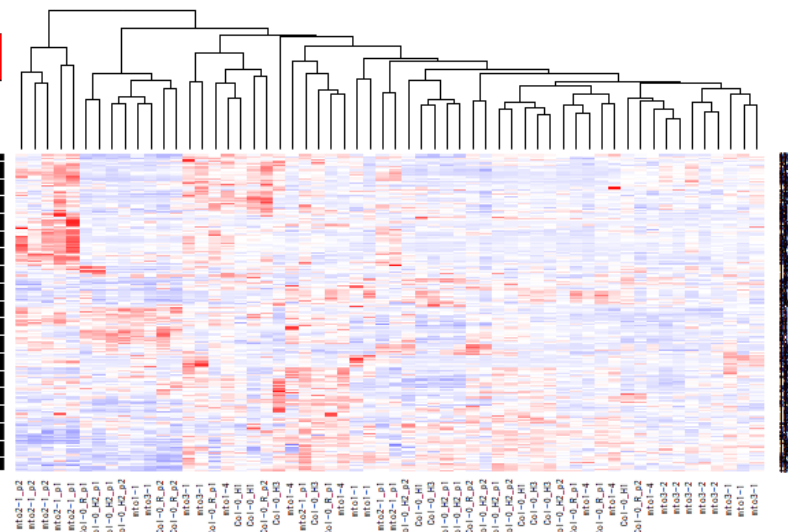
We integrated multiple metabolome datasets in Arabidopsis and constructed our database, called AtMetExpress, to store the information. The integrated data analyses showed that Arabidopsis has ~1,200 metabolites, which we can detect using mass spectrometry-based metabolite profiling.

Dataset

Kusano10AminoAcids

Run analysis

Cancel



再解析ツール (AtMetExpress): HCA

PRIME Visualization Tools

RIKEN PRIME provides web-based data analysis and visualization tools for public access. Users may analyze datasets from both AtMetExpress and MeKO, as well as upload custom datasets.

Analysis parameters

Data source: *AtMetExpress*

Dataset: *Kusano10AminoAcids*

Downloads

None

[Edit parameters](#)

Summary (MVA)

Hierarchical Cluster

[Heat Map](#)

[Principal Component Analysis](#)

hclust Method

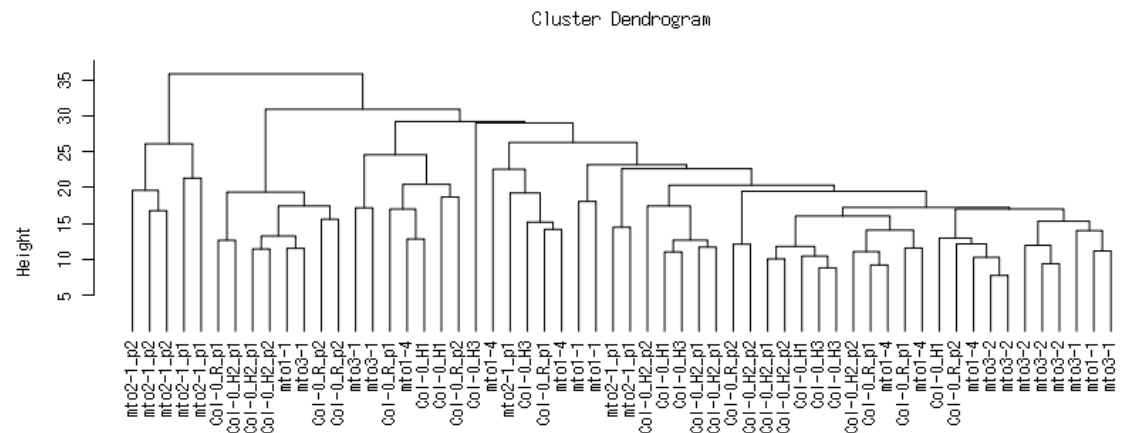
Complete

dist Method

Euclidean

scaling Method

auto



COMPLETE & EUCLIDEAN
hclust (*, "complete")

再解析ツール (AtMetExpress): User's data

Meta-analysis selection (GC-MS RIKEN format or GENERIC format)

User's Data selection

Analysis Parameters

Select data source

AtMetExpress MeKO **GC-MS (RIKEN format)** **Generic**

We integrated multiple metabolome datasets in Arabidopsis into a single matrix and constructed our database, called AtMetExpress, to store the information. The integrated data analyses showed that Arabidopsis has xxx metabolites, which we can detect using mass spectrometry-based metabolite profiling.

Dataset: Fukushima11BMC_Syst_Biol

Run analysis Cancel

Users can select data format.



Analysis Parameters

Select data source

AtMetExpress MeKO GC-MS (RIKEN format) Generic

Raw data (TSV): ファイルが選択されていません

Phenodata (CSV): ファイルが選択されていません

Internal standard: Hexadecanoate_13C4

Factors (to select multiple factors, hold Ctrl while clicking selections)

No. of components: 2

Meta-analysis tools

HCA, Heatmap, and PCA

RKN PRIDE provides web-based data analysis and visualization tools for public access. Users may analyze datasets from both AtMetExpress and MeKO, as well as upload custom datasets.

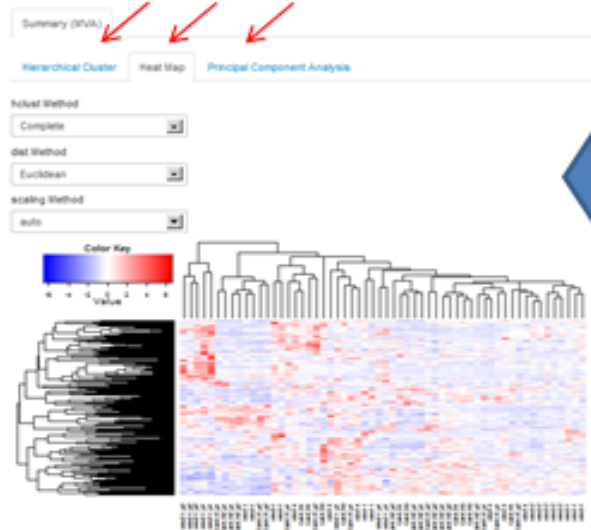
Analysis parameters

Data source: AtMetExpress

Dataset: Kusano10AminoAcids

Downloads: None

Edit parameters



Uploading user's Data

Analysis Parameters

Select data source

AtMetExpress MeKO GC-MS (RIKEN format) Generic

Raw data (TSV): Kusano10AminoAcids_Data.txt
Upload complete

Phenodata (CSV): Kusano10AminoAcids_phenodata.csv
Upload complete

Internal standard: Hexadecanoate_13C4

Factors (to select multiple factors, hold Ctrl while clicking selections)

Genotype: goid
onlygeno
sampleOrg

No. of components: 2

Raw data matrix (TSV format)

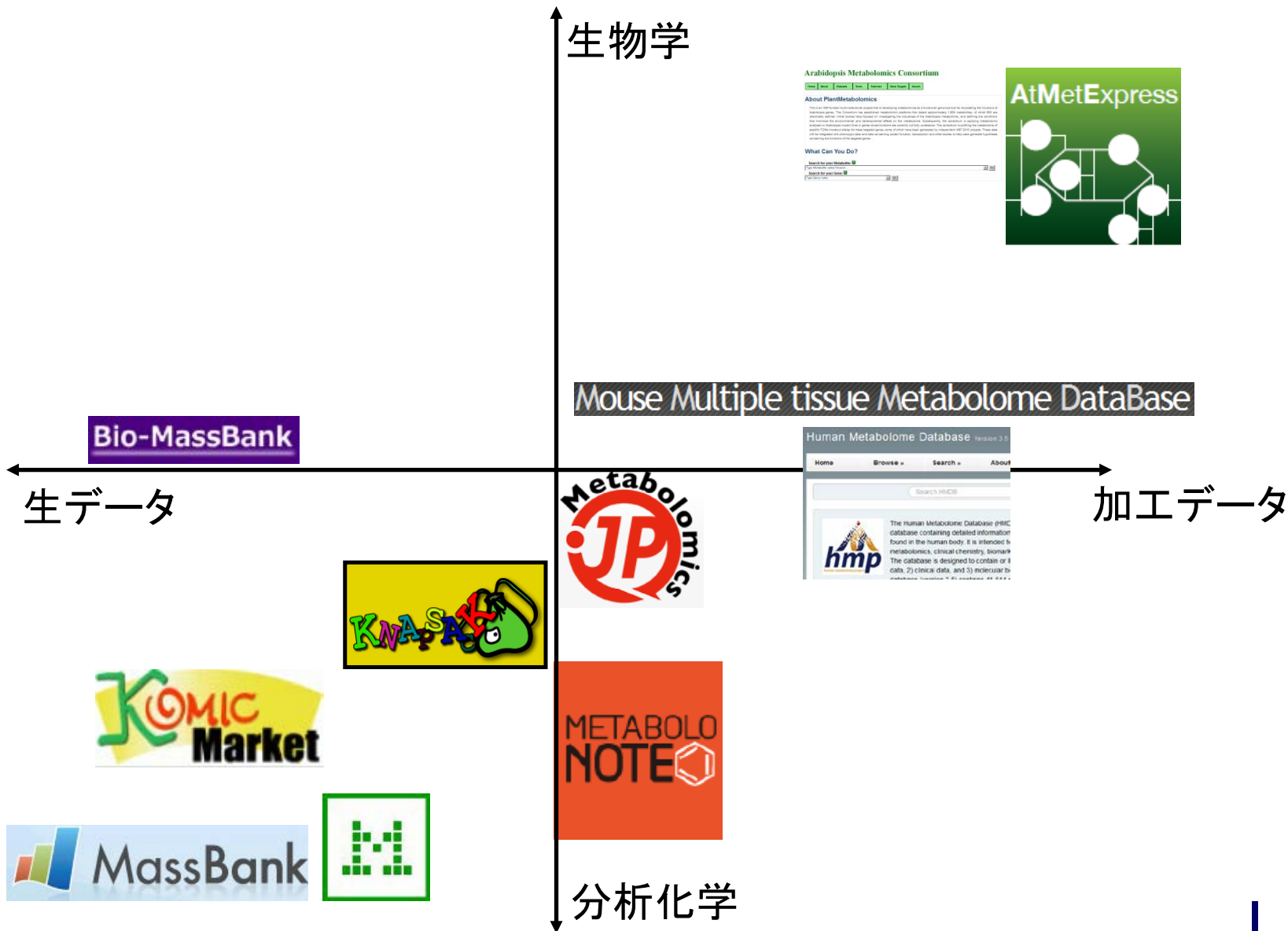
Sample information (CSV format)

Internal standard information

Information for normalization

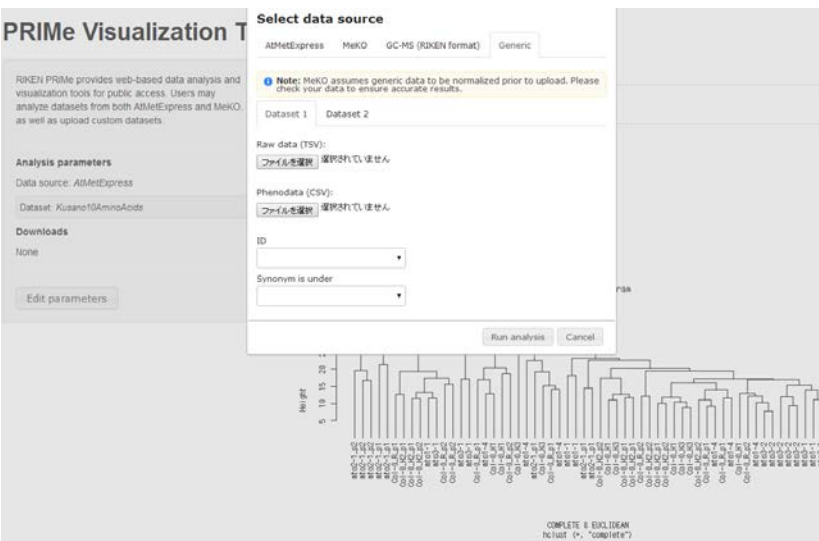
デモ

統合データベースの新たな活用法

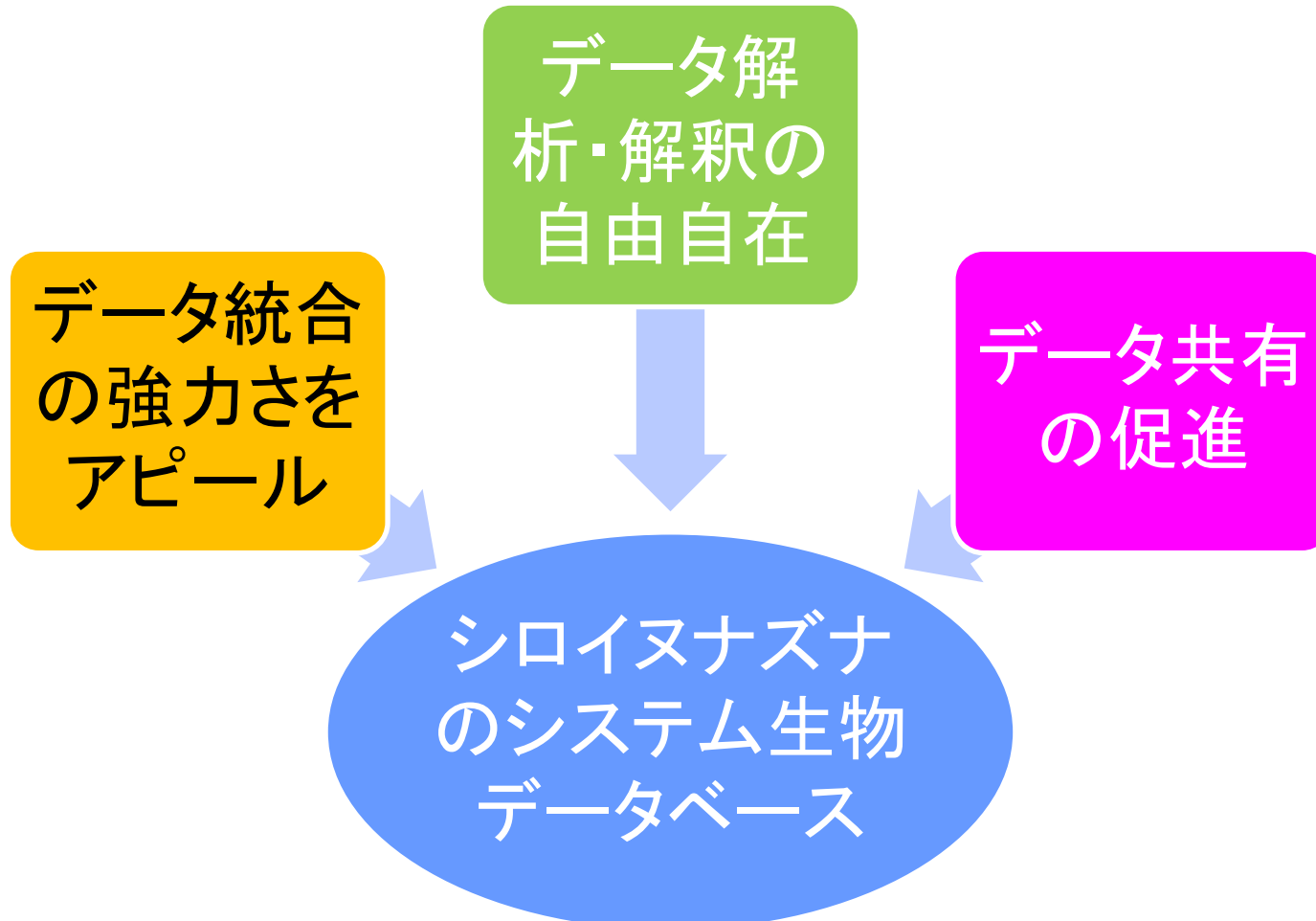


まとめ

- 現状のMSに基づく代謝物プロファイリング→シロイヌナズナの代謝物 ~ 1,200個 が計測可能
- XMetExpress == 汎用的なフレームワーク
- AtMetExpress GUI tool はメタボロームデータの再解析を簡便にする



今後の本研究の将来性



- ・生物学者が活用できるデータ
- ・簡便なデータ共有
- ・研究サイクルの加速

謝辞

理化学研究所CSRS

- Kazuki Saito
- Masanori Arita
- Tetsuya Sakurai
- Miyako Kusano
- Yozo Okazaki
- Ryo Nakabayashi
- Akie Mejia
- Makoto Suzuki
- Makoto Kobayashi
- Yutaka Yamada
- Kenji Akiyama

Others

- Redestig Henning