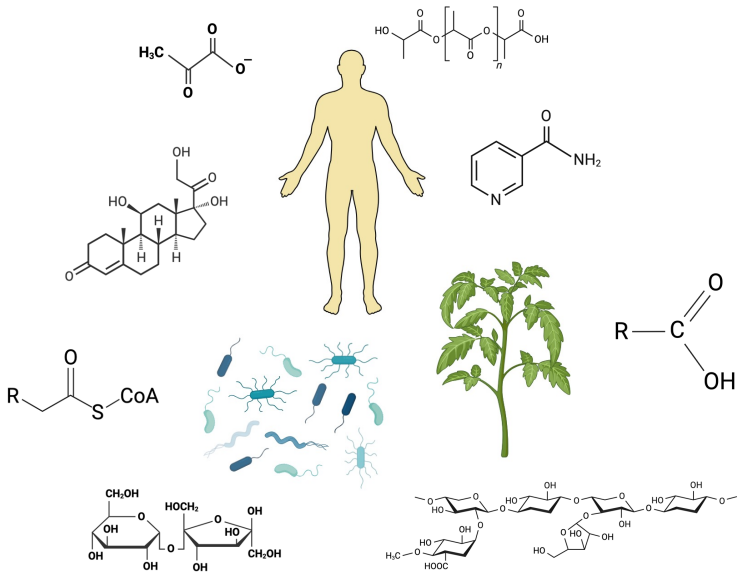
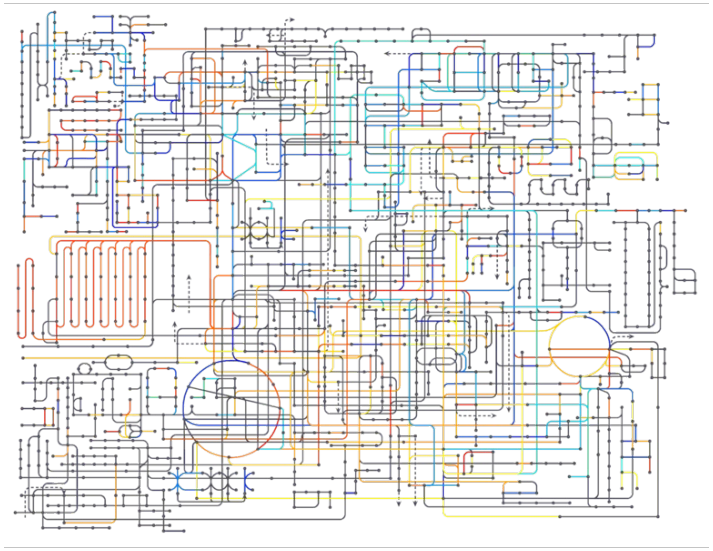
創発的再解析のための
メタボローム統合データベース

早川英介

九州工業大学 情報工学研究院
理化学研究所 環境資源科学研究センター

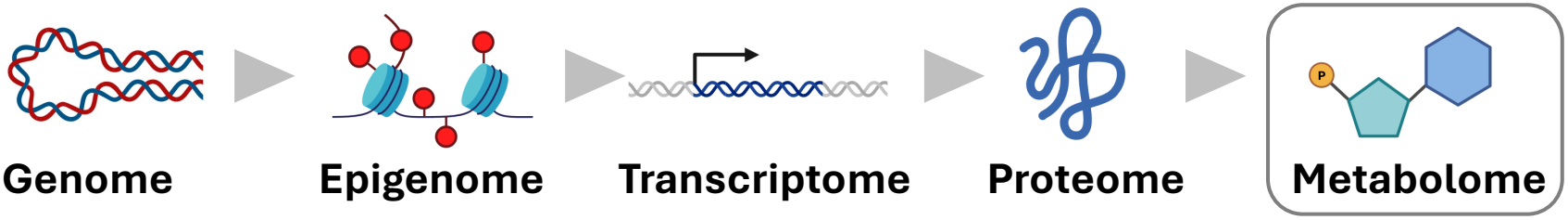
Metabolomics: A Comprehensive Overview of Metabolism





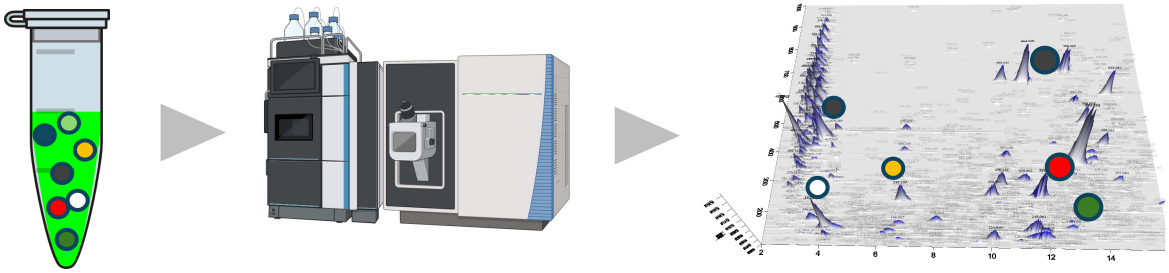
"Metabolic Network" by BruceMcAdam, CC BY-SA 3.0.

- **Metabolomics enables the detailed analysis of metabolites, providing insights into metabolic changes across biological systems.**

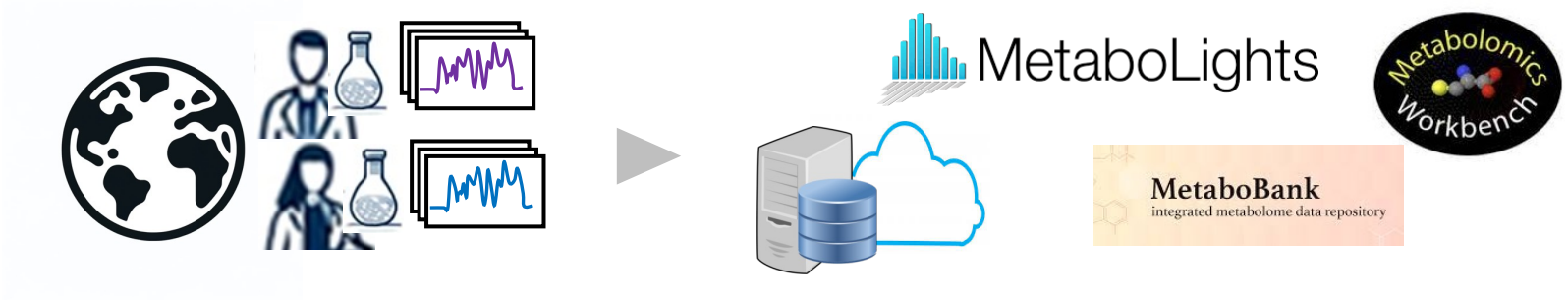


- **Metabolomics lies downstream of the genome, providing insights that are closer to the phenotype**

Metabolomics Data Repositories

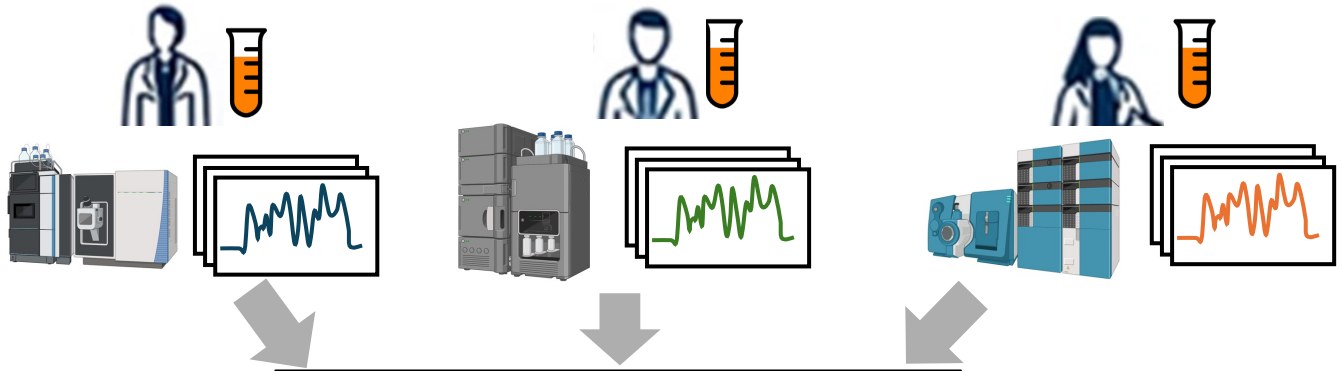


- **Metabolomics generates vast amounts of data from numerous samples through high-throughput mass spectrometry (MS) instruments.**



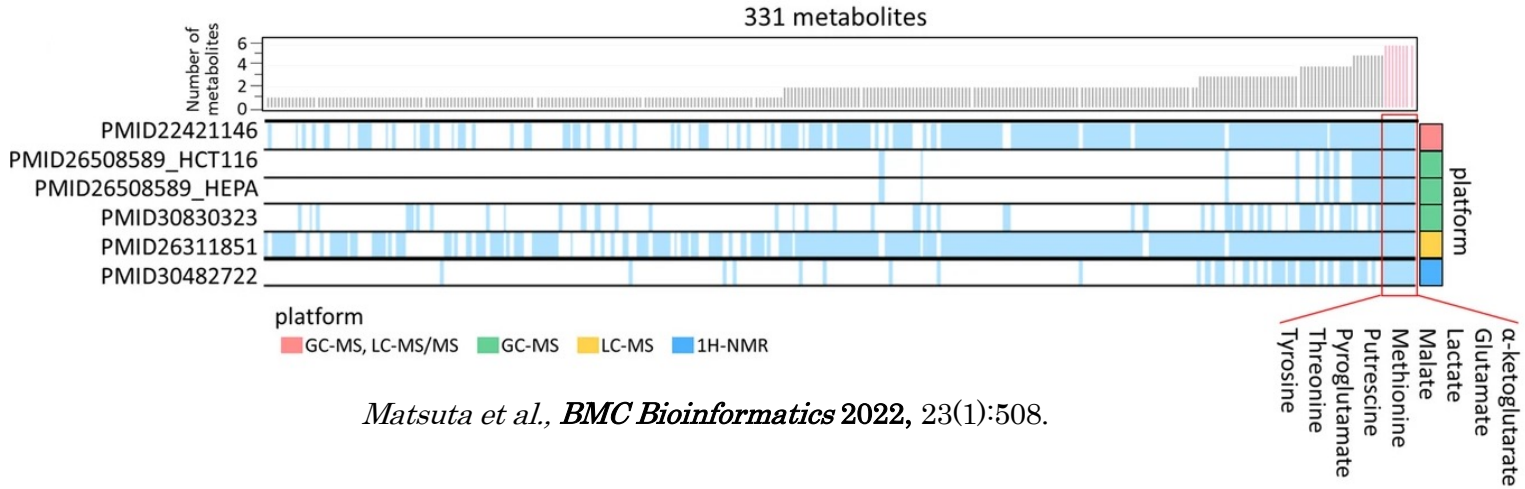
- **Metabolomics data repositories are extensively utilized for the storage and management of metabolomics data.**
- **The volume and diversity of metabolomics data are experiencing rapid growth.**
- **Despite advancements, the reanalysis and integration of metabolomics data remain challenging.**

Technical Challenges in Data Integration : MS



Data integration: infeasible

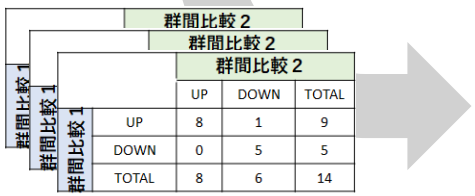
- **Variability in instrumentation creates standardization challenges.**
- **Limited absolute quantification and instrument-dependent signal intensities impact data comparability.**
- **Metabolite detection varies with instrumental settings, affecting data consistency across platforms.**



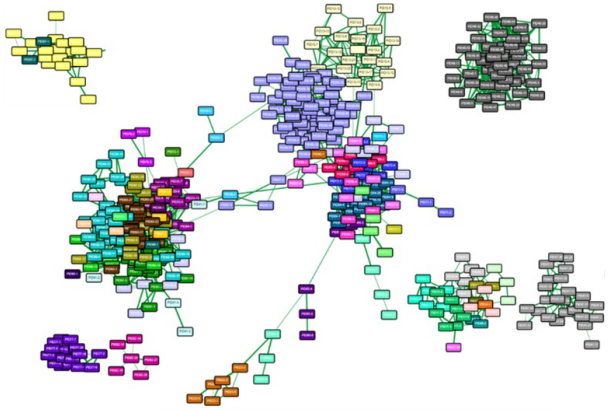
Matsuta et al., ***BMC Bioinformatics*** **2022,** 23(1):508.

Objective: Advancing Metabolomics Data Integration

Metabolomics datasets

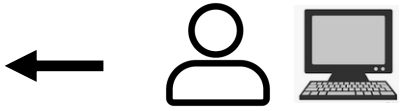


	UP	DOWN	TOTAL
UP	8	1	9
DOWN	0	5	5
TOTAL	8	6	14



Network-based integrated DB

Metabolomics researchers



Data scientists

1.Metabolite differential profile-based Integration:

- **Employing a differential profile-based approach to integrate metabolomics data, enabling more precise comparative analysis across different datasets.**

2.Metadata-Based Integration:

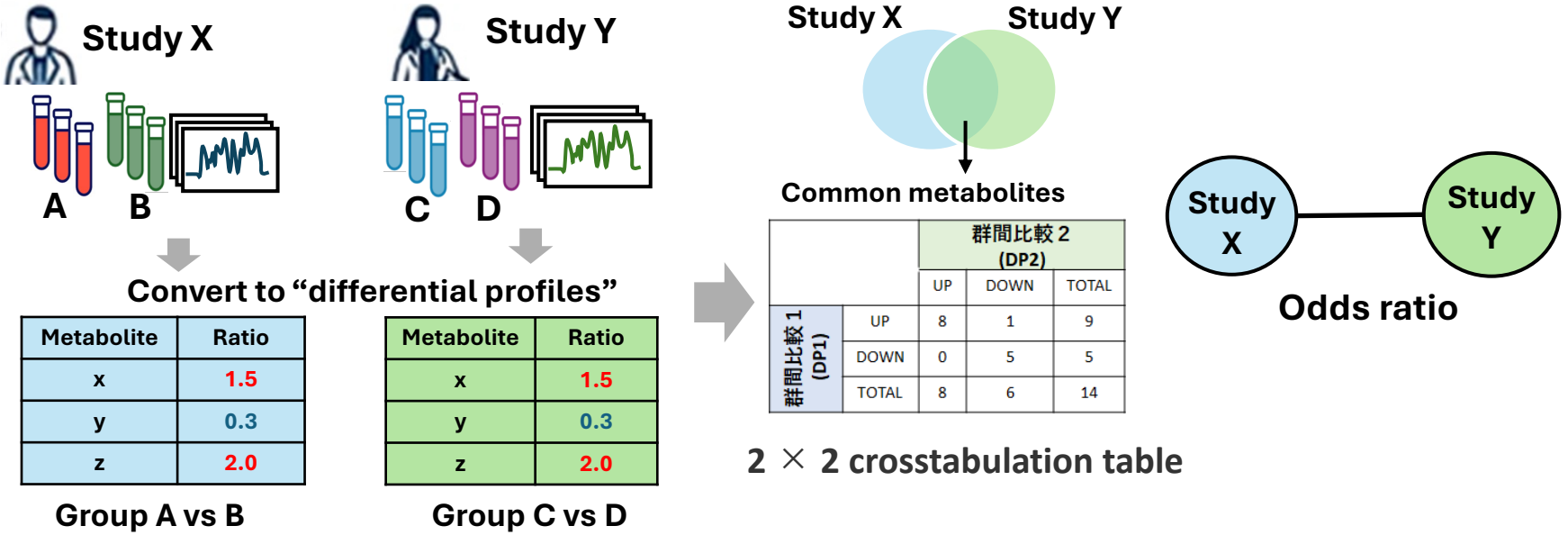
- **Keyword-based data integration to support interpretation of complex metabolome data sets.**

3.Unique Integrated Network for Data Exploration:

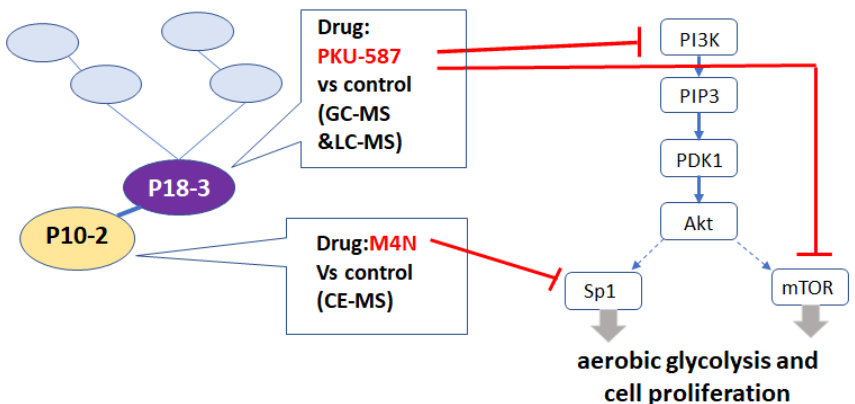
- **Developing an integrated network platform that facilitates the exploration of vast metabolomics data, paving the way for novel discoveries and insights.**

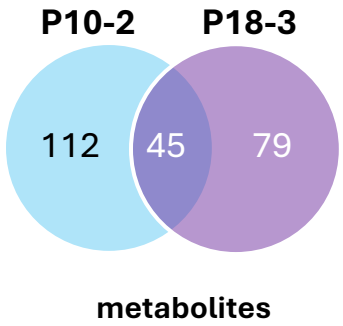
Differential metabolomic profile-based integration

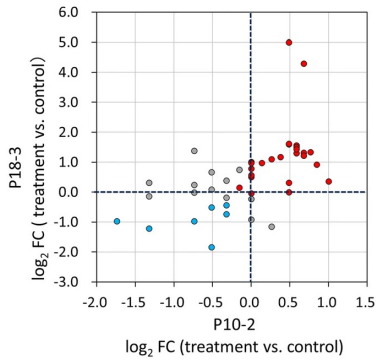
- **iDMET:** network-based approach for integrating metabolomics differential analysis



- **Connect studies (differential profiles) based on similarity of metabolic change**







Matsuta et al., BMC Bioinformatics 2022, 23(1):508.

Data Structure of Metabolomics Repository

MetaboLights

Study 1

S_MTBLS**.txt**

Sample	Characteristic 3
C1	**Control (GroupA)**
C2	**Control (GroupA)**
N1	**Drug Treated (Group B)**
N2	**Drug Treated (Group B)**

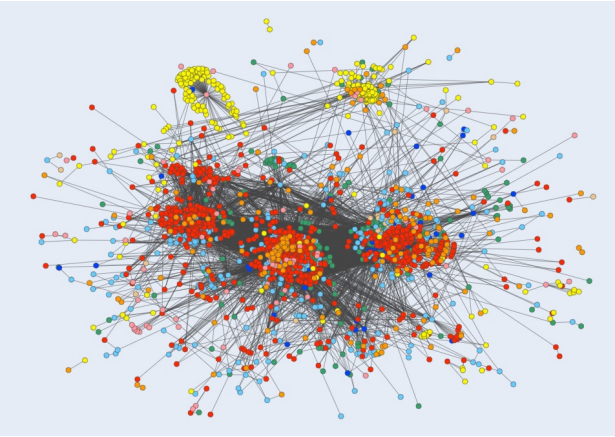
maf.tsv

Metabolite	C1	C2	N1	N2
Estrone	1.69E+08	7.0E+08	5.7E+08	7.2E+08
Dihydrocortisol	1.52E+09	1.2E+09	3.6E+09	1.4E+09
Adipate semialdehyde	3.26E+09	1.9E+09	3.2E+09	1.4E+09
3-Aminopentanedioate	1.69E+08	7.0E+08	3.2E+09	3.6E+06
N-[(2S)-2-Amino-2-c...	1.69E+08	1.4E+09	7.1E+08	3.6E+09

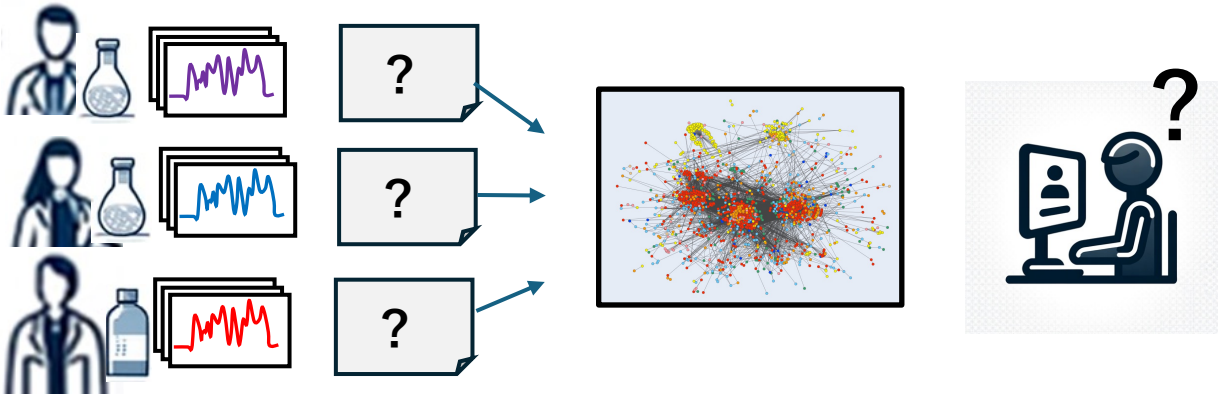
Create Differential Profile

群間比較 1		群間比較 2		
		UP	DOWN	TOTAL
UP		8	1	9
DOWN		0	5	5
TOTAL		8	6	14

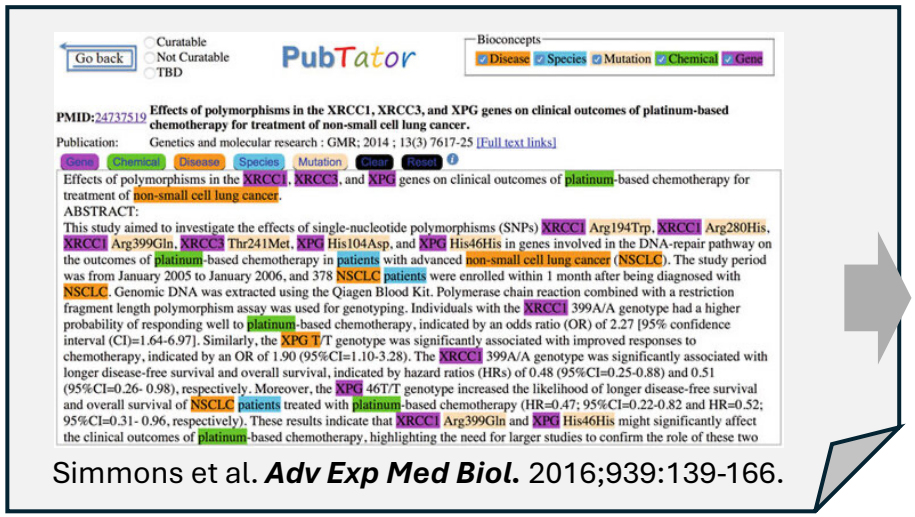
iDMET



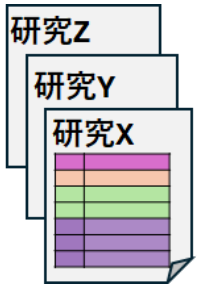
Metadata-based integration/filtering

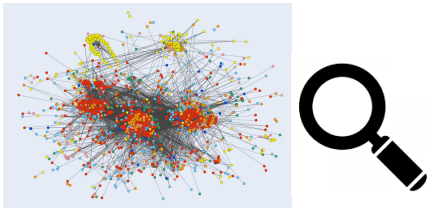


- **Metabolomics data repository hosts diverse studies, each with unique backgrounds. Difficult to find similarities and connections between studies.**
- **We use tools to annotate key biological concepts (diseases, species, genes and chemicals) in studies.**

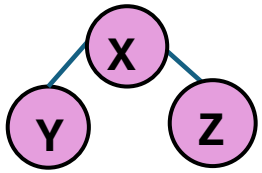


Simmons et al. ***Adv Exp Med Biol.*** 2016;939:139-166.





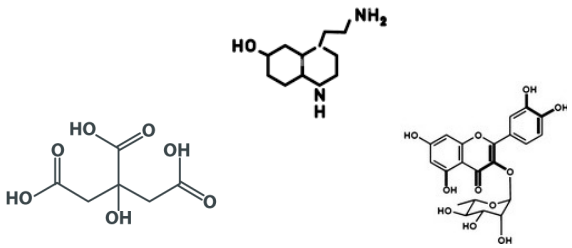
Filtering Integrated Network



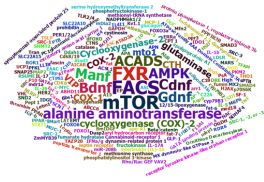
"Metadata" Network

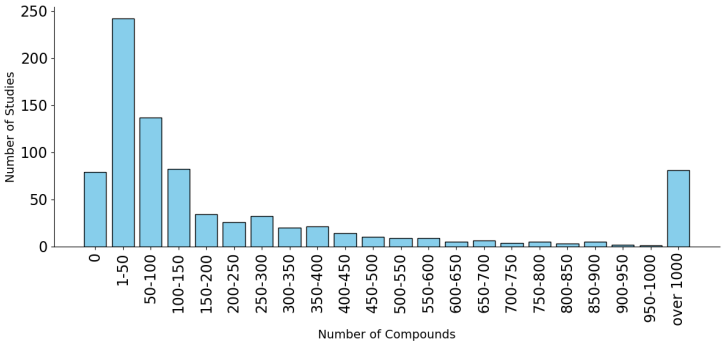
Differential profile-based integration

MetaboLights **748 studies**



Number of metabolites ：35,476

Species	Gene	Disease	Chemical
			
338 species	**198 genes**	**388 diseases**	**696 chemicals**

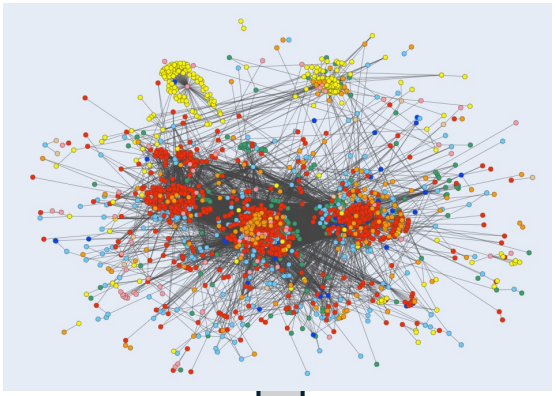


Distribution of the Number of Compounds per Study

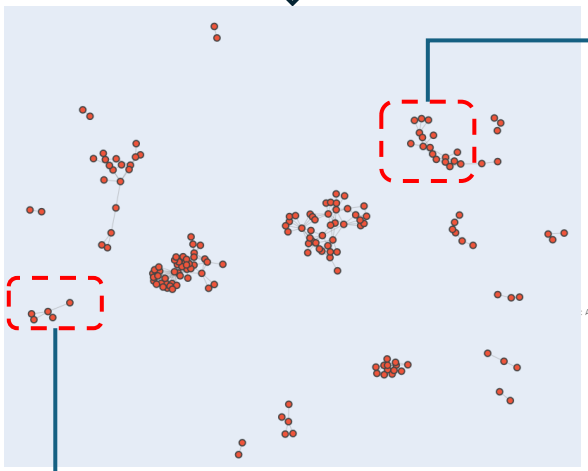
Integrated metabolome study network

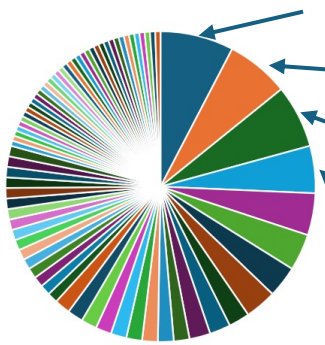
Data Reanalysis with integrated network

Differential profile network



Disease:MESH:D007249
"inflammation"





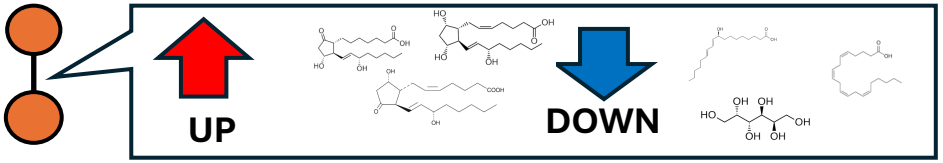
Prostaglandins and related compounds

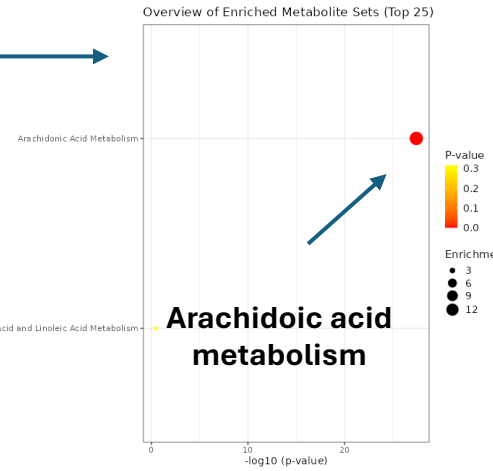
Very long-chainfatty acids

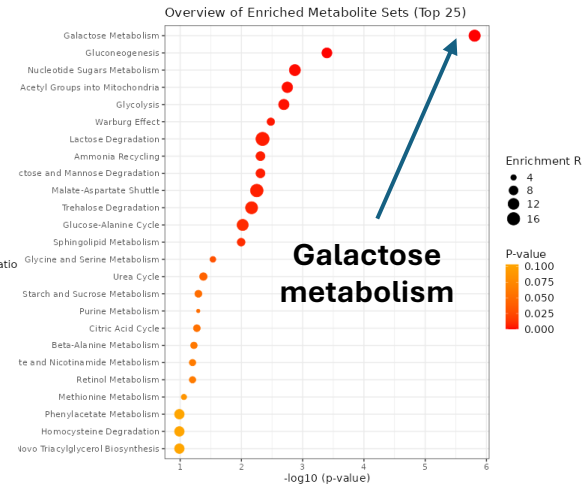
Long-chain fatty acids

Lineolic acids and derivatives

Chemical class (classyfire) composition of "inflammation" network

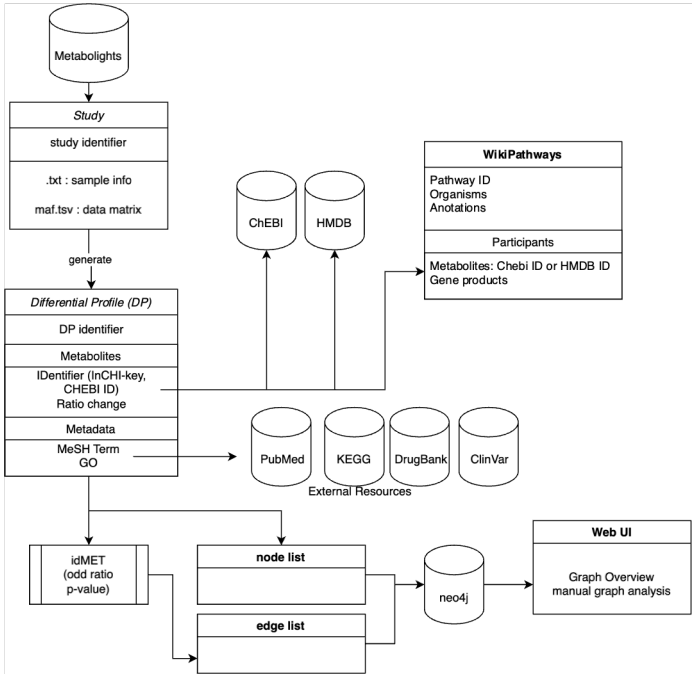






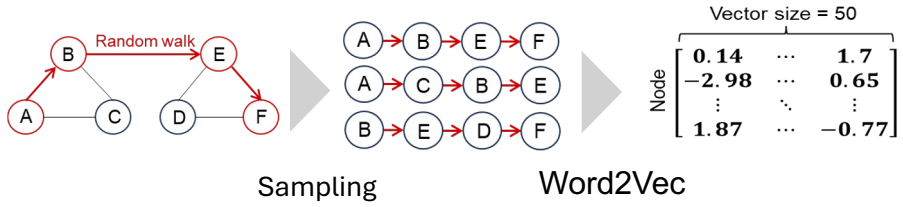
Enhanced Utilization of Integrated Metabolomics Networks

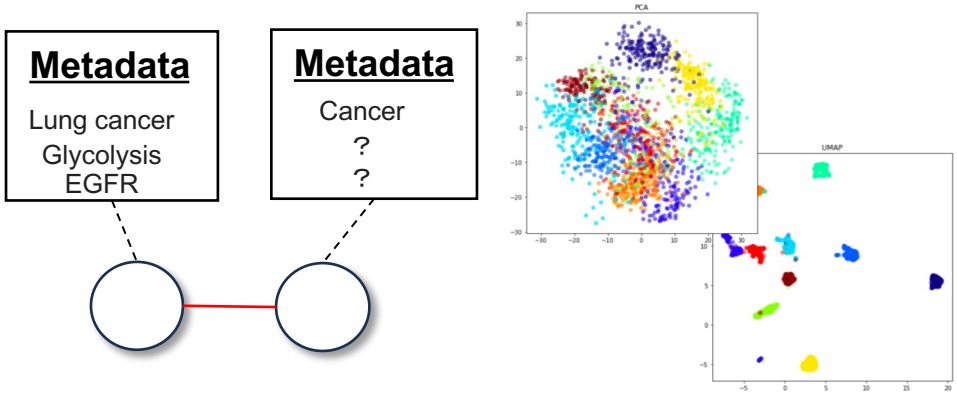
Constructing Knowledge Graphs from Integrated Metabolomics Data



Kitazato univ. Kawano Group

Developing Machine Learning-Ready Pipelines

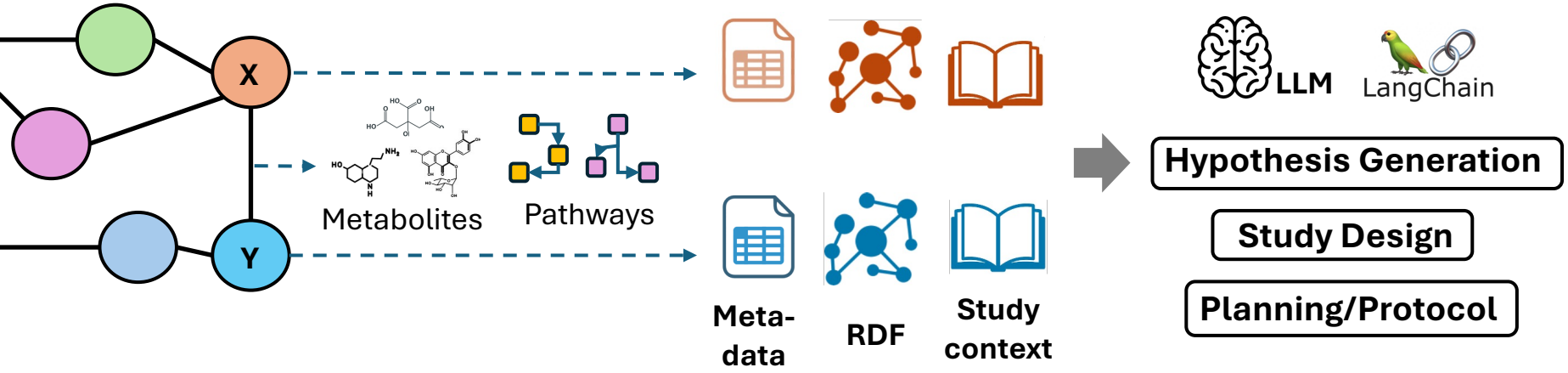




TUAT Tsugawa Group (Oka, Nishida)

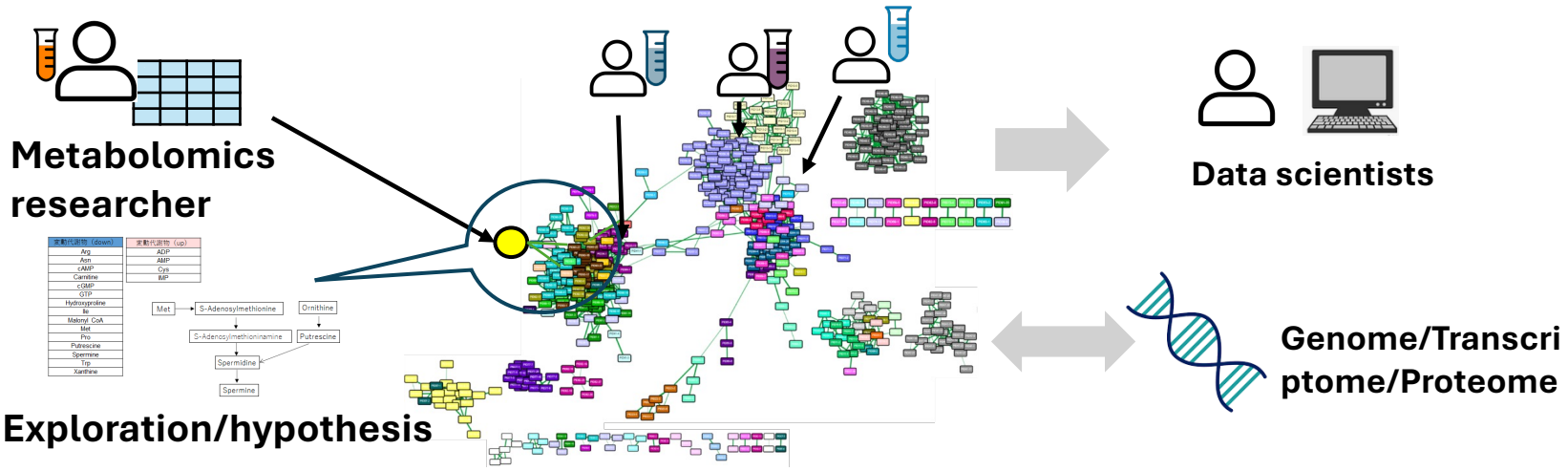
- Leveraging Knowledge Graphs for Advanced Insights and Predictions
- Enabling Efficient Data Integration and Utilization in Metabolomics

Future Outlook:
AI-Driven Study Creation Through Integrated Metabolomics Graph



- **Cross-Study Hypothesis Creation**
  By identifying patterns in metabolic changes between connected studies, we can generate initial hypotheses for further investigation.
- **AI-Driven New Study Design**
  AI-suggests relevant analyses and methodologies, autonomously creating new scientific study designs based on the provided metadata.
- **Accelerating Scientific Discovery**
  AI systematically and autonomously creates scientific study designs, dramatically shortening the time between data analysis and new research conception, enabling rapid scientific progress.

Summary



- **Integrate large metabolome datasets to enables a holistic view of metabolic processes, enhancing our understanding of complex biological phenomena.**

- **This unique network-based platform significantly contributes to the metabolomics/science community by simplifying and enhancing accessibility for data reanalysis**

Acknowledgements

- **Project groups**

RIKEN CSRS Metabolome Informatics Research Team	Eisuke Hayakawa Mikiko Takahashi Yutaka Yamada Masanori Arita
Kyushu Institute of Technology	Eisuke Hayakawa
Kitasato University	Shin Kawano
Tokyo University of Agriculture and Technology	Takaki Oka Kozo Nishida Hiroshi Tsugawa

- **Collaborator**

Human Metabolome Technologies	Rira Matsuta Hiroyuki Yamamoto

- **Funding Support**

統合化推進プログラム

「創発的再解析のためのメタボローム統合データベース」