INFORMATICS PROFESSIONALS. LEADING THE WAY.

RaMP-DB 3.0 Relational Database of Metabolomics Pathways

Workshop: **An Introduction to NCATS Public Resources and Analytics for Rare Diseases, Targets, Drug Substances, and Analytes**

Ewy Mathé, Tim Sheils, Keith Kelleher

IFX, Division of Preclinical Innovation

National Center for Advancing Translational Sciences/National Institutes of Health

Disclosure and Learning Objectives

- I have no relevant relationships with commercial interests to disclose.

- Provide details on the data provenance for our resources, so that our users understand how the data could be useful for their work.

- Provide hands on experience interacting with our resources in various ways (e.g. web applications, APIs, or other packages).

- Interact and discuss how the data can be used, augmented, and further incorporated with other efforts. We anticipate new collaborative opportunities through the workshop.

RaMP-DB: Comprehensive Metabolite/Gene Annotations

Up-to-date and comprehensive biological, chemical, ontology annotations from multiple sources

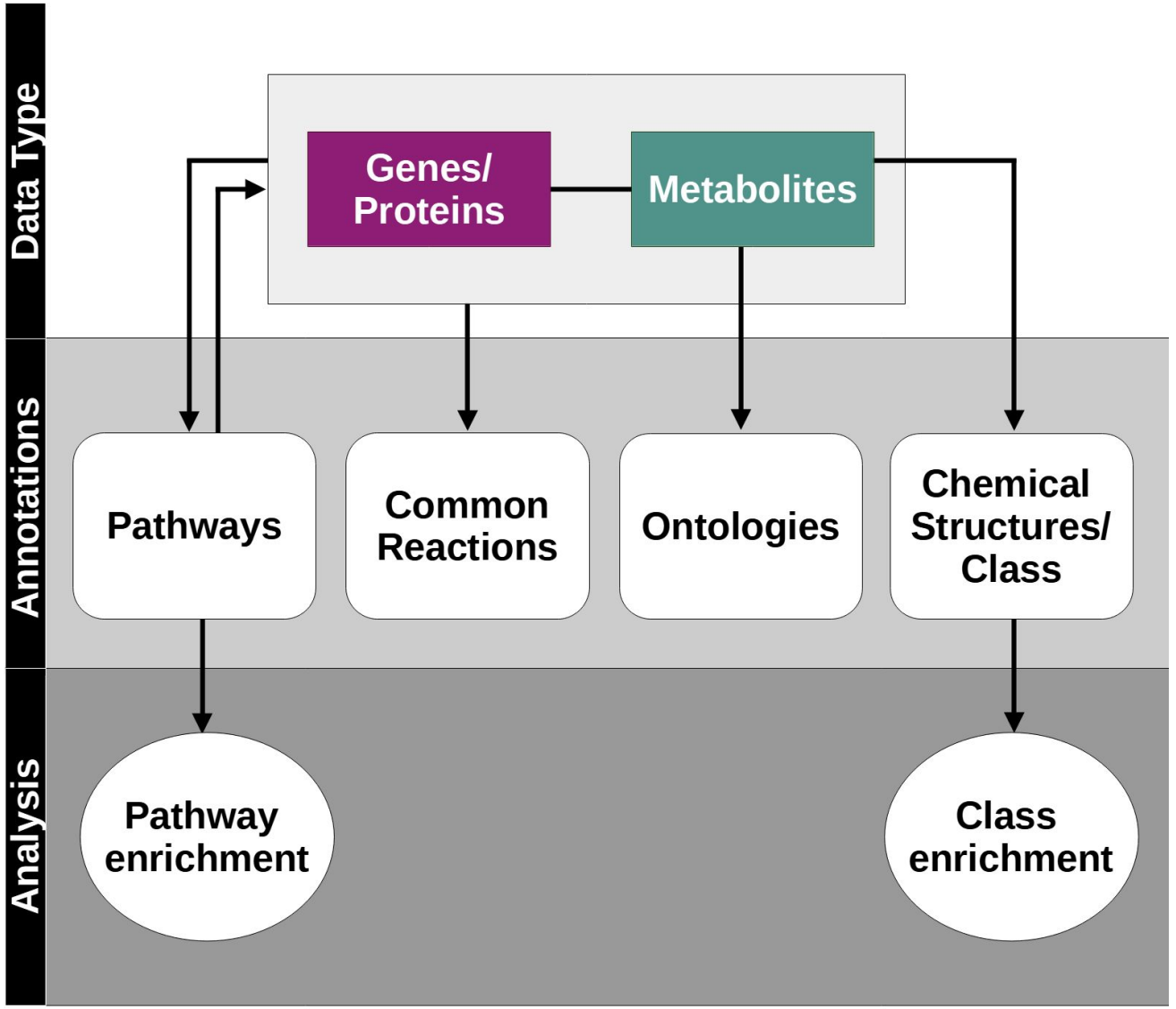
Open-source, stand-alone usable database

Multiple access points through the R package, SQLite database, APIs, interactive web app to perform queries and pathway analyses

Transparent and reproducible process, including semi-automated curation, code, and web app

Basis for improving pathway analysis methods

rampdb.nih.gov/



RaMP-DB 3.0 - Major Improvements

- **Expanded content**
  - Source database updates
  - Increased focus on lipid annotations
  - Incorporation of curated reactions from Rhea

- **Improved semi-automated integration** scripts

- **SQLite version** (replacing MySQL)
  Greatly facilitates R package installation and integration with other tools (COMETS Analytics, MetaboAnalyst, Bioconductor Computational Metabolomics Ecosystem)

- **New queries of biochemical reactions**

rampdb.nih.gov/

RaMP-DB 3.0 – Increased Annotations and Pathway Mappings

>230K metabolites, ~16K genes/enzymes, >53K pathways

	Total	HMDB	KEGG*	Reactome	WikiPathways	LIPIDMAPS	Rhea
# Distinct Metabolites	234,543	196,426	5,578	2,239	3,569	39,368	12,155
# Distinct Genes/Enzymes	16,443	7,073		11,229	13,841		3,498
# Distinct Pathways	53,952	49,613	363	2,627	1,349		
# Metabolite-Pathway Mappings	412,990	367,370	1,673	30,410	13,537		
Gene-Pathway Mappings	406,132	208,211	8,459	127,077	62,385		

* From HMDB

RaMP-DB 3.0 – Chemical Descriptions and Reactions (New)

> 63K reactions, > 250K chemical property records

	Rhea
Reactions	63,396 (15,849)
Metabolite-Reaction Mappings	299,272 (74,818)
Protein-Reaction Mappings	16,436 (4,109)

	Total Distinct Compounds	HMDB	ChEBI	LIPIDMAPS
Chemical Properties	256,537 (distinct InChIKeys)	217,776	18,471	45,227
Chemical Class Information	155,293	126,101		39,368
Metabolite Ontology	134,653	134,653		

New SQLite greatly improves accessibility and facilitates usage

Original MySQL Installation Process

- Install the RaMP R Package
- Download the MySQL RaMP Data File
- Download the MySQL Software and Install
- Start the MySQL Server
- Load the RaMP Data into MySQL Server
- Load the RaMP Package
- Establish RaMP to MySQL Connection
- Run RaMP Queries

New Simplified Process

- Install the RaMP R Package
- Load the RaMP Package
- Run RaMP Queries

RaMP 3.0 automatically downloads the SQLite Database file and stores it in a local file cache.

A RaMP function allows users to see and select RaMP DB versions on their computer OR in our RaMP-DB GitHub Large File Storage System.

Newly incorporate Biochemical Reactions

Input: list of metabolite and gene/protein ids

Output: Rhea reactions, including

- Associated mappings to input metabolites
- Associated mappings to input genes/proteins

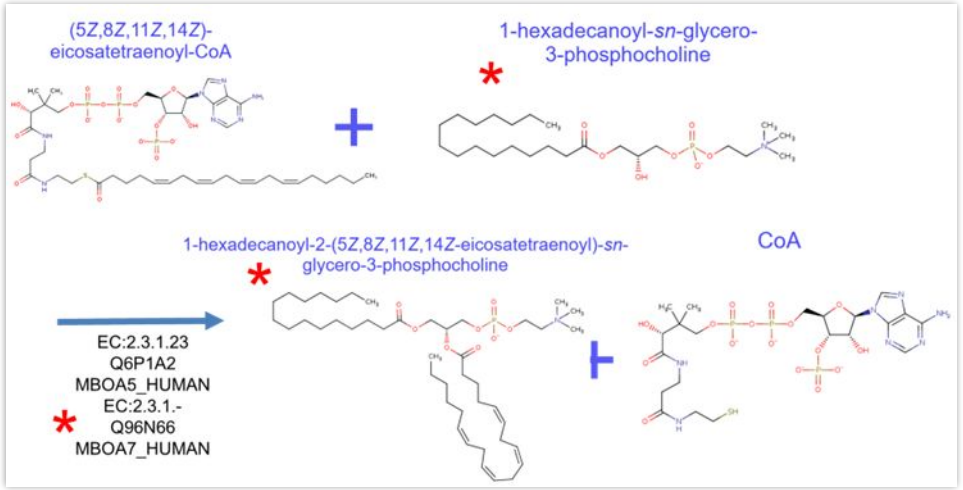
```
getReactionsForAnalytes(
  db = RaMP(),
  analytes,
  analyteType = "metabolites",
  namesOrIds = "ids",
  onlyHumanMets = F,
  humanProtein = F,
  includeTransportRxns = F,
  rxnDirs = c("UN")
)
```

Options

- Exclusion of cofactors and transport reactions.
- Return of reactions involving human proteins or human metabolites only (as defined by ChEBI).
- Further filtering can be performed to report on reactions having more than one participant represented in the input query.

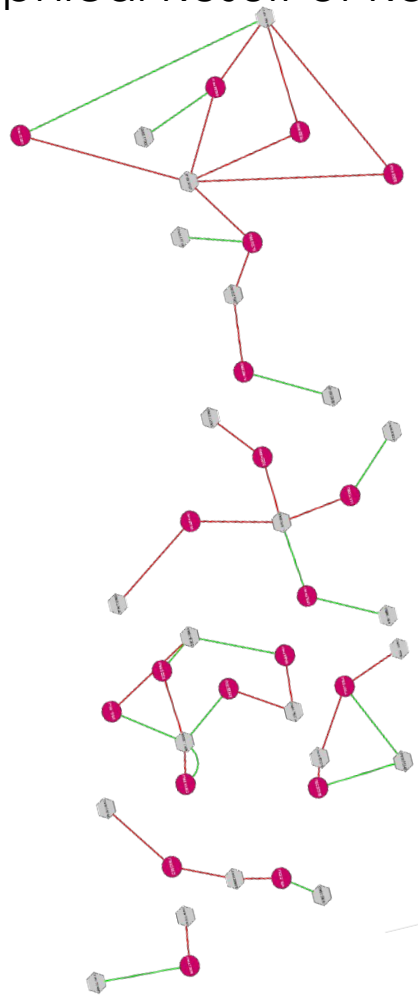
Visualizing Returned RaMP-DB Reactions

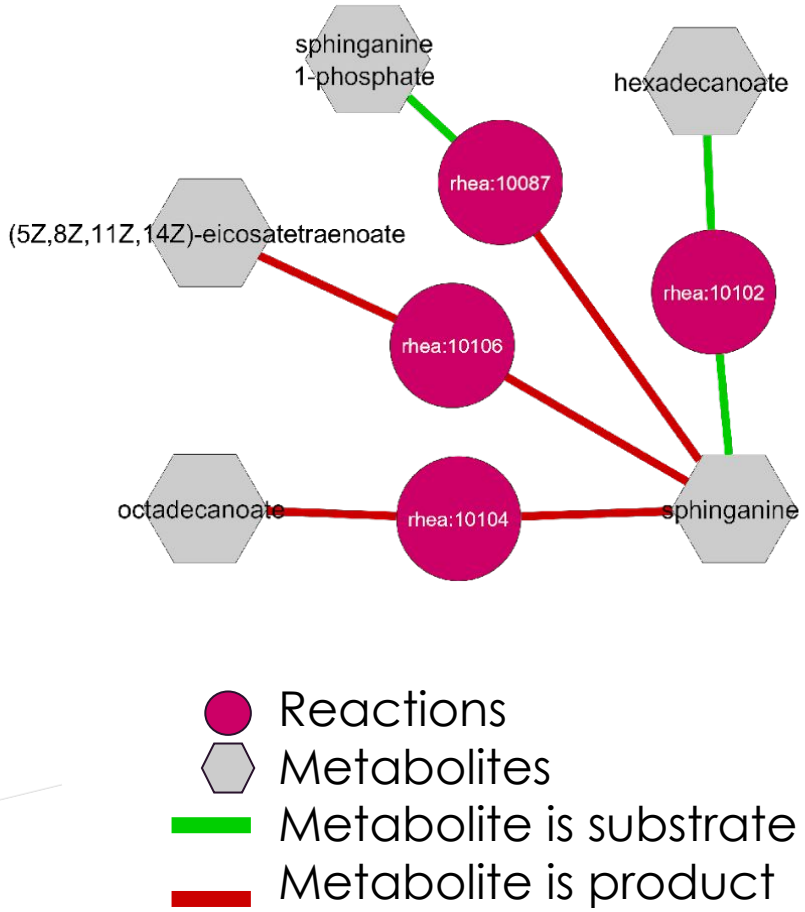
chebi:72998
uniprot:Q9N66
chebi:73003
…



Example reaction (RHEA:35999)

Graphical Result of Reactions





RaMP-DB – Many Uses and Accessibility

User-friendly web app:
https:rampdb.nih.gov

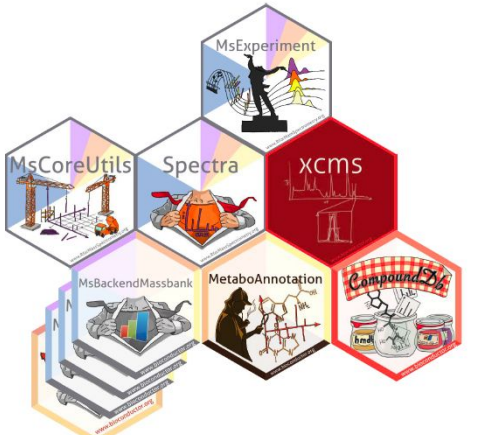
R package and Vignette/Tutorial:
https://github.com/ncats/RaMP-DB
https://ncats.github.io/RaMP-DB/RaMP_Vignette.html

Documented API Access:
https://ramp-api-alpha.ncats.io/__docs__/

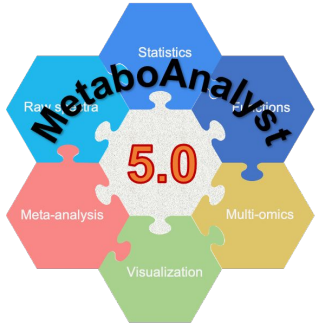
Code for building back-end and front-end:
https://github.com/ncats/RaMP-Client/
https://github.com/ncats/RaMP-Backend

Uses:
Reuse of parsing scripts, incorporate in data analysis workflows, develop new methods using RaMP-DB, etc.

Bioconductor Computational Metabol0mics ecosystem

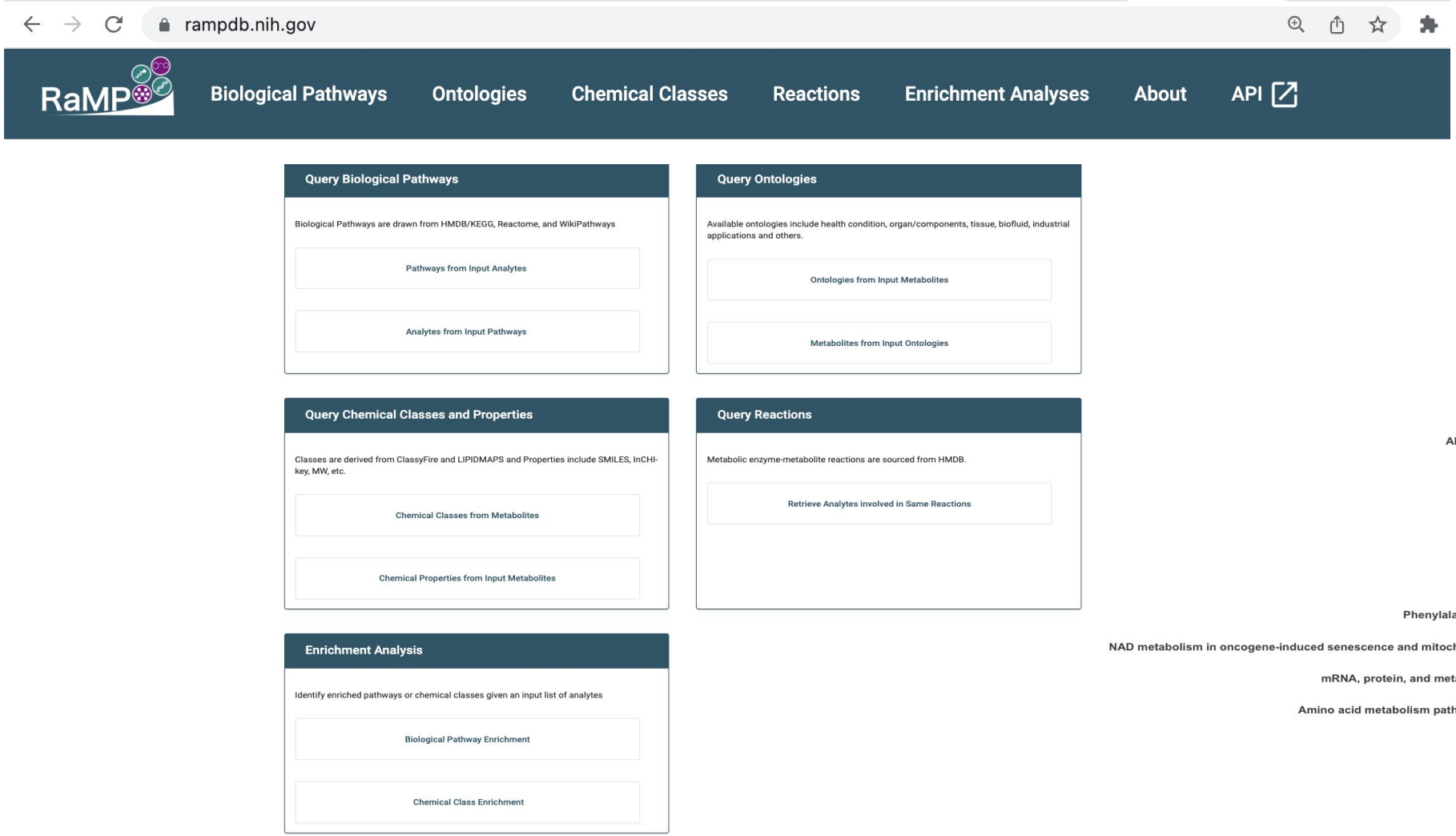


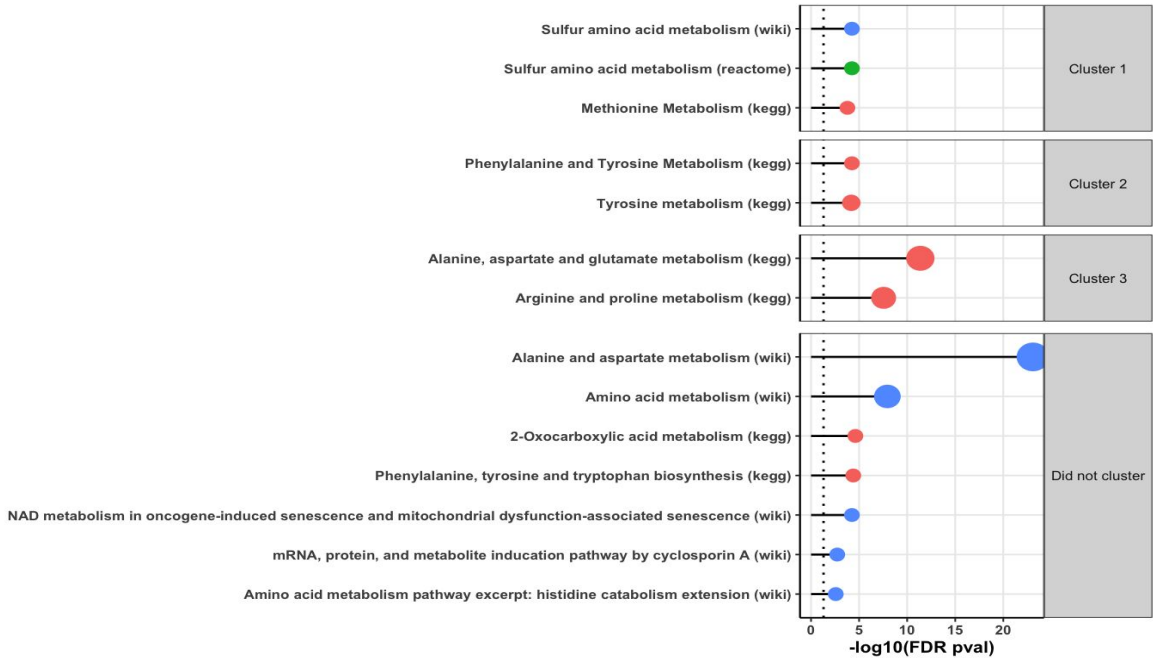
COMETS Analytics



TRANSLATOR
Funded by NIH's National Center for Advancing Translational Sciences

We are always open to feedback!





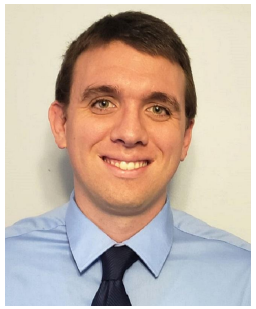
NCATSRaMP@nih.gov

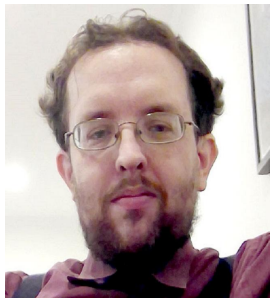
rampdb.nih.gov

It’s a team effort!


John Braisted

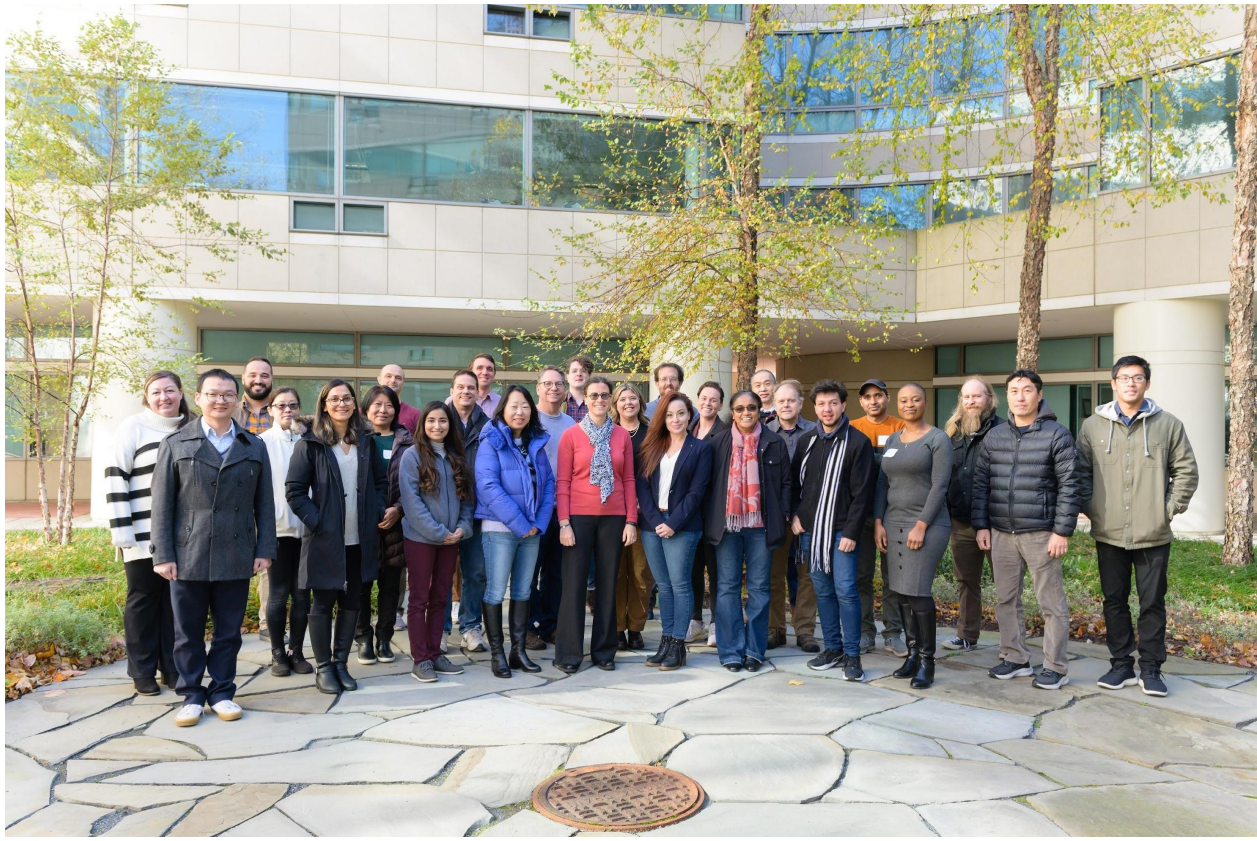

Tim Sheils


Andrew Patt


Keith Kelleher

Also special thanks to

- Tara Eicher, Kyle Spencer, Haley Chatelaine for development/testing.
- Johannes Rainer for help in converting RaMP-DB to SQLite
- Sam Michael, Ke Wang, Betty Li, Amit Virakamath, and NCATS ITRB for computing infrastructure
- Egon Willighagen and WikiPathways team; Jeff ia and MetaboAnalyst team; David Wishart and HMDB team for feedback



https://ncats.nih.gov/preclinical/core/informatics (We’re hiring!)

Let's check it out and get started!

All workshop materials can be found here:

https://shorturl.at/kvzIZ

Some installation prerequisites for RaMP-DB:

https://shorturl.at/bgqrQ

Provide your comments/thoughts here:

https://shorturl.at/sS138