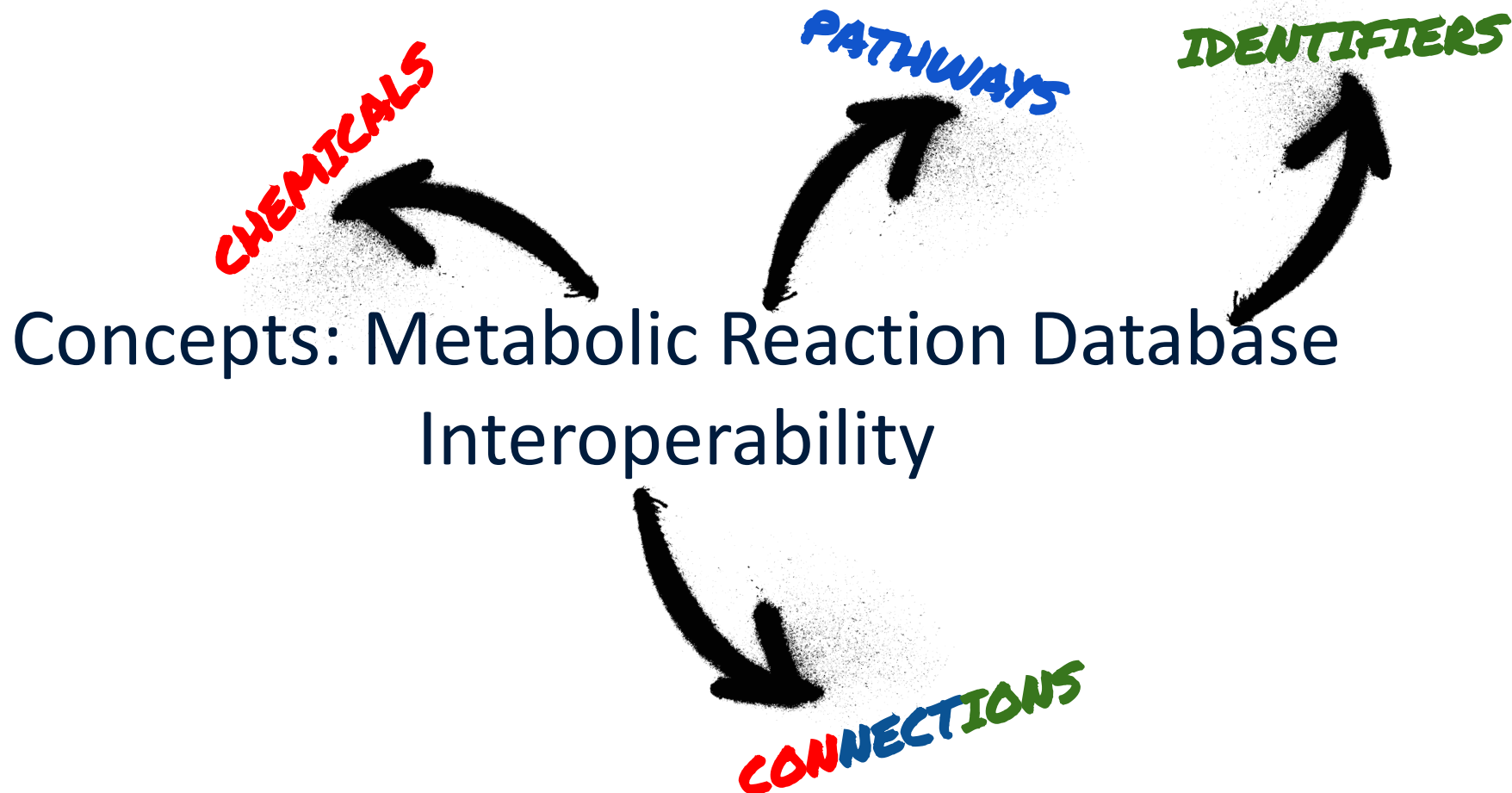# BridgeDb @ Fluxomics workshop

Denise Slenter, Egon Willighagen
0000-0001-8449-1318, 0000-0001-7542-0286

Twitter: @BridgeDbProject
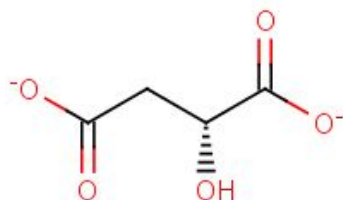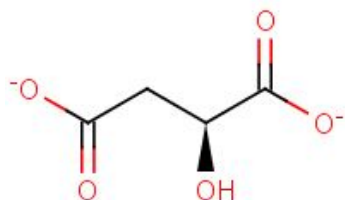API: webservice.bridgedb.org
website: bridgedb.github.io

2021-10-05

Maastricht University

L12. Standardization of metabolic reaction databases and use in fluxomics

CHEMICALS

PATHWAYS

IDENTIFIERS

Concepts: Metabolic Reaction Database Interoperability

CONNECTIONS

Maastricht University

CHEMICALS

s-malate(2-) vs r-malate(2-)

Computers have issues reading chemical structures in Figures [1]

What about reactions/pathways?

[1] Rajan, K., Brinkhaus, H.O., Zielesny, A. *et al.* A review of optical chemical structure recognition tools. *J Cheminform* 12, 60 (2020). https://doi.org/10.1186/s13321-020-00465-0

Maastricht University

PATHWAYS

- Are useful to show relationships between metabolites and enzymes
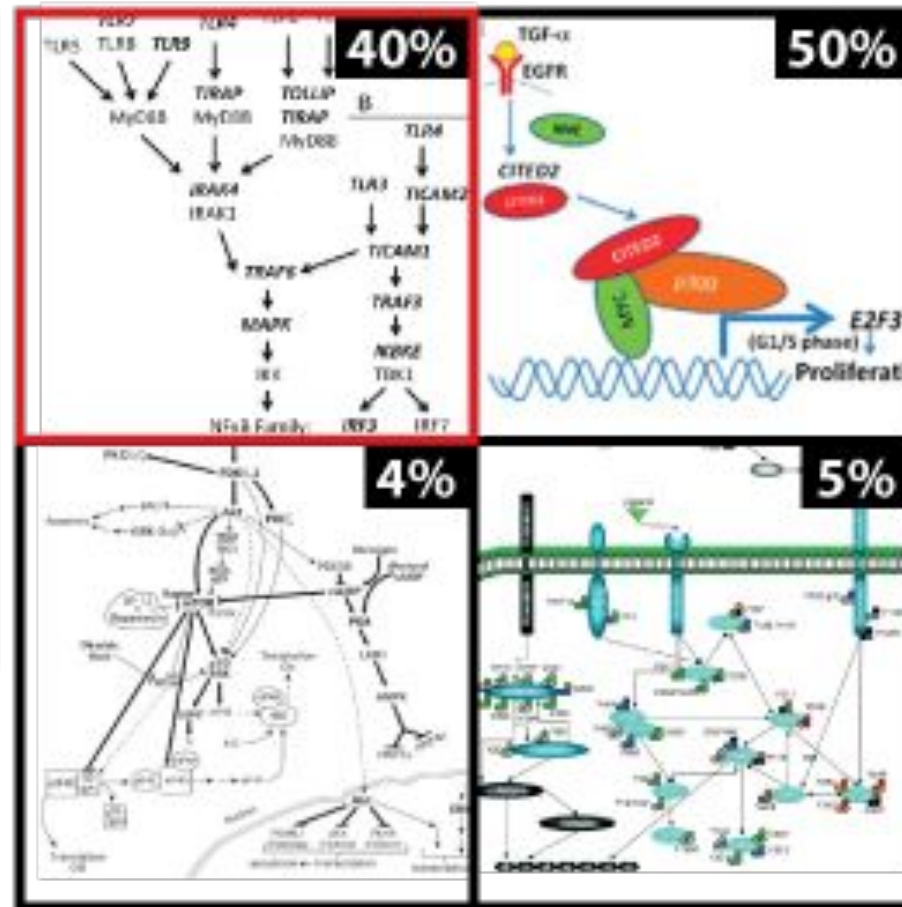
PATHWAYS

- Are useful to show relationships between metabolites and enzymes

  However

- Often captured in Figures as static images
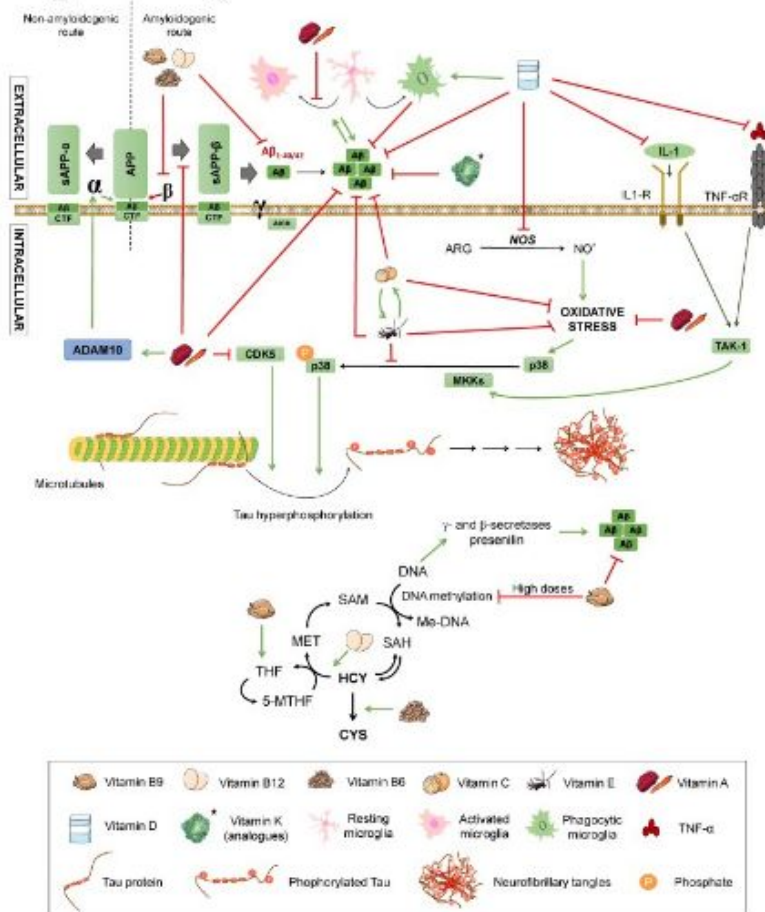- Can be difficult to interpret
- Use non-standard names, abbreviations

Maastricht University

Slide courtesy of San Francisco Team: <u>Alex Pico</u>, Anders Riutta, Kristina Hanspers

Maastricht University

Example of extracted OCR content

Slide courtesy of San Francisco Team: Alex Pico, Anders Riutta, Kristina Hanspers

Maastricht University

Fig. 5 adapted from Fernández-Sanz, Paloma, Daniel Ruiz-Gabarre, and Vega García-Escudero. 2019. "Modulating Effect of Diet on Alzheimer's Disease" Diseases 7, no. 1: 12. https://doi.org/10.3390/diseases7010012
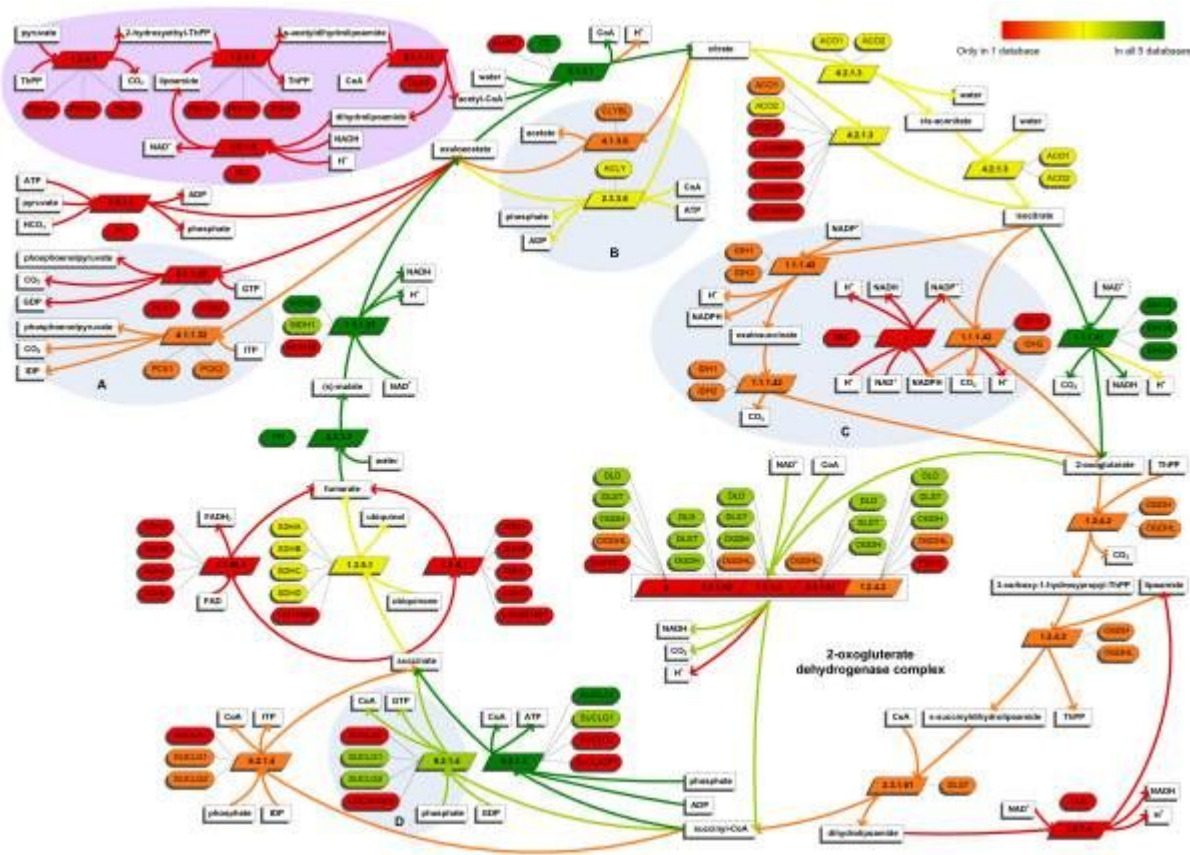
PATHWAYS

- Are useful to show relationships between metabolites and enzymes

PATHWAYS

- Are useful to show relationships between metabolites and enzymes

  However

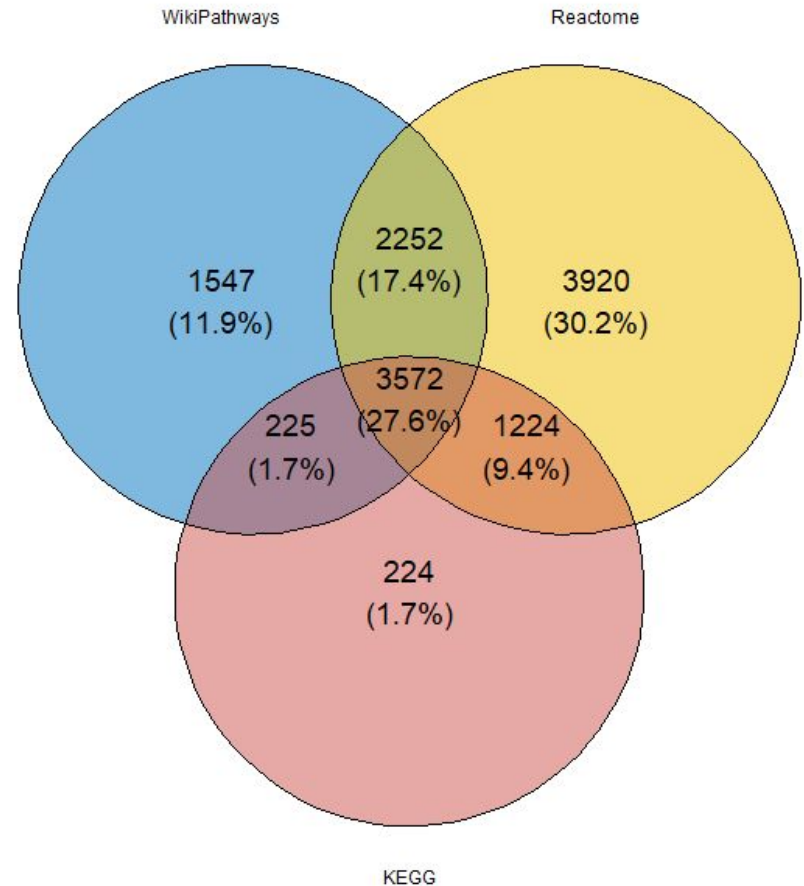- Pathway databases are incomplete

Comparison of the TCA cycle in five metabolic pathway databases. (BiGG, EHMN, HumanCyc, KEGG, Reactome) Map illustrating the (lack of) consensus for the TCA cycle. Metabolites are represented by rectangles, genes by rounded rectangles, and EC numbers by parallelograms. Color indicates how many of the five databases agree on a specific entity (gene, EC number, reaction). We first matched reactions based on their metabolites. Genes and EC numbers were matched within matching reactions. Color of an arrow indicates the number of databases that agree upon an entire reaction, i.e., all its metabolites (except H+ which was matched separately). 'x' denotes a missing EC number.

Maastricht University

Stobbe MD, Houten SM, Jansen GA, van Kampen AH, Moerland PD. Critical assessment of human metabolic pathway databases: a stepping stone for future integration. BMC Syst Biol. 2011 Oct 14;5:165. doi: 10.1186/1752-0509-5-165.

A more recent comparison

Gene set content overlap analysis between KEGG, Reactome and WikiPathways; data obtained from MSigDb (release 7.4 from April 2021).

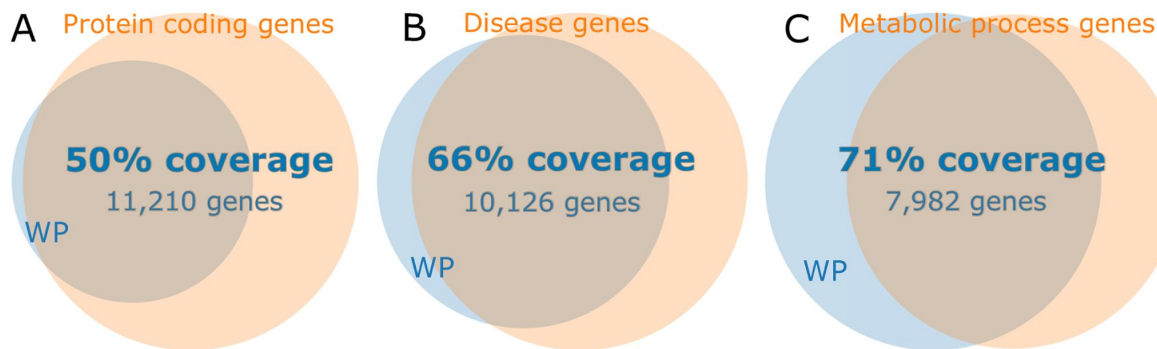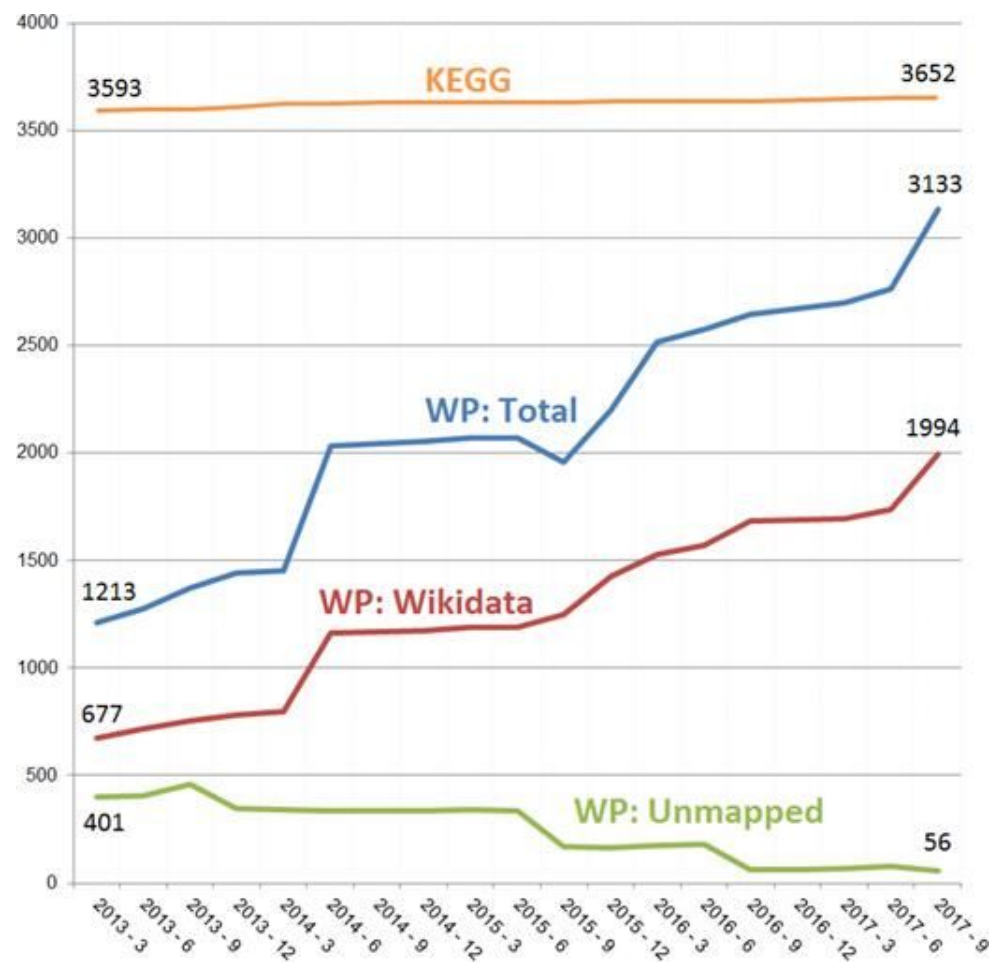And metabolite/reaction content is even harder to compare

Maastricht University

A — Protein coding genes

50% coverage
11,210 genes

WP

B — Disease genes

66% coverage
10,126 genes

WP

C — Metabolic process genes

71% coverage
7,982 genes

WP

KEGG — 3593 ... 3652

WP: Total — 1213 ... 3133

WP: Wikidata — 677 ... 1994

WP: Unmapped — 401 ... 56

Maastricht University

Identifiers

ORCID
DOI
PMID
ISBN

Identifiers

ORCID

DOI

PMID

ISBN

but also:

CAS

ChEBI
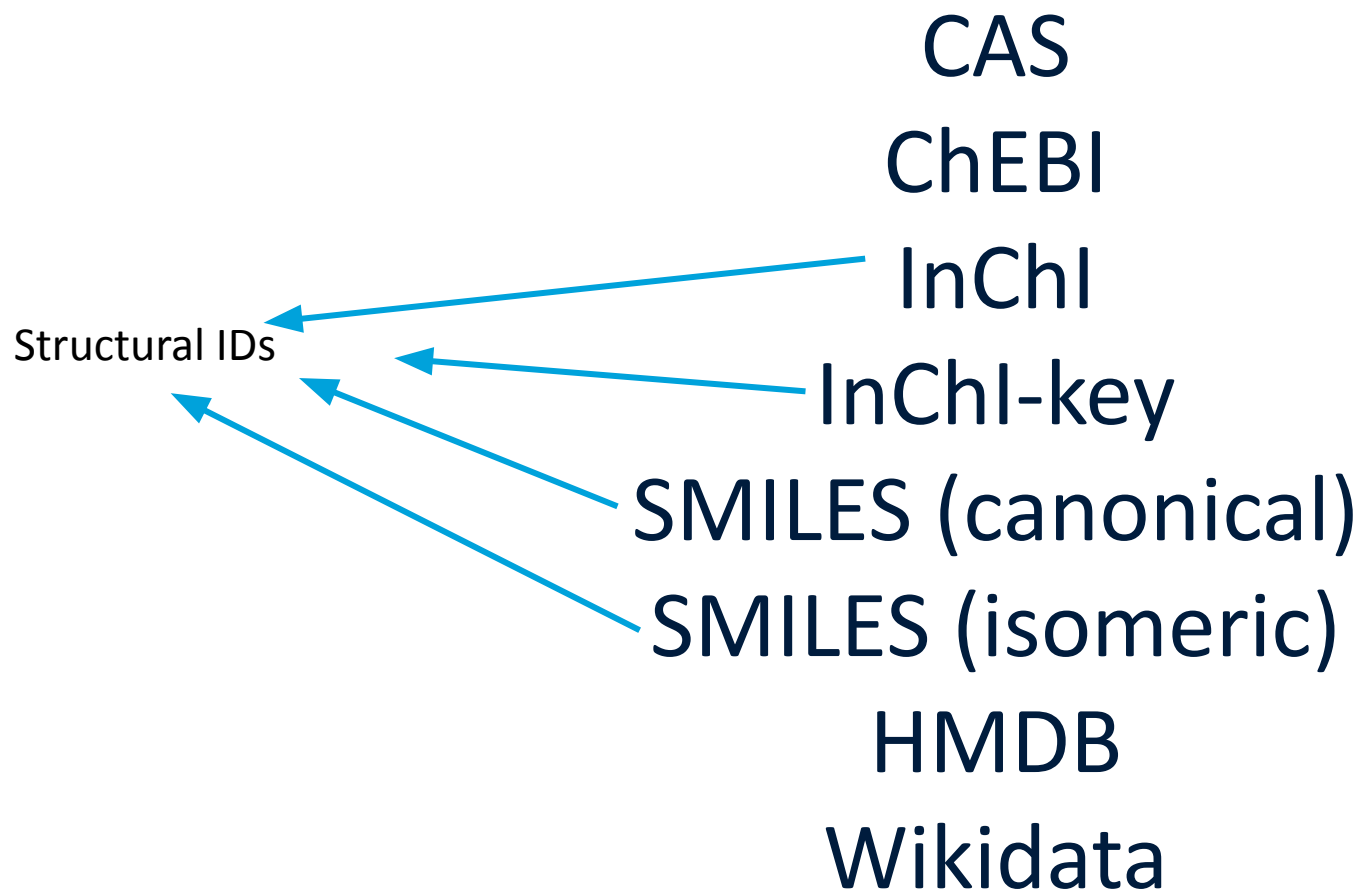
InChI

InChI-key

SMILES (canonical)
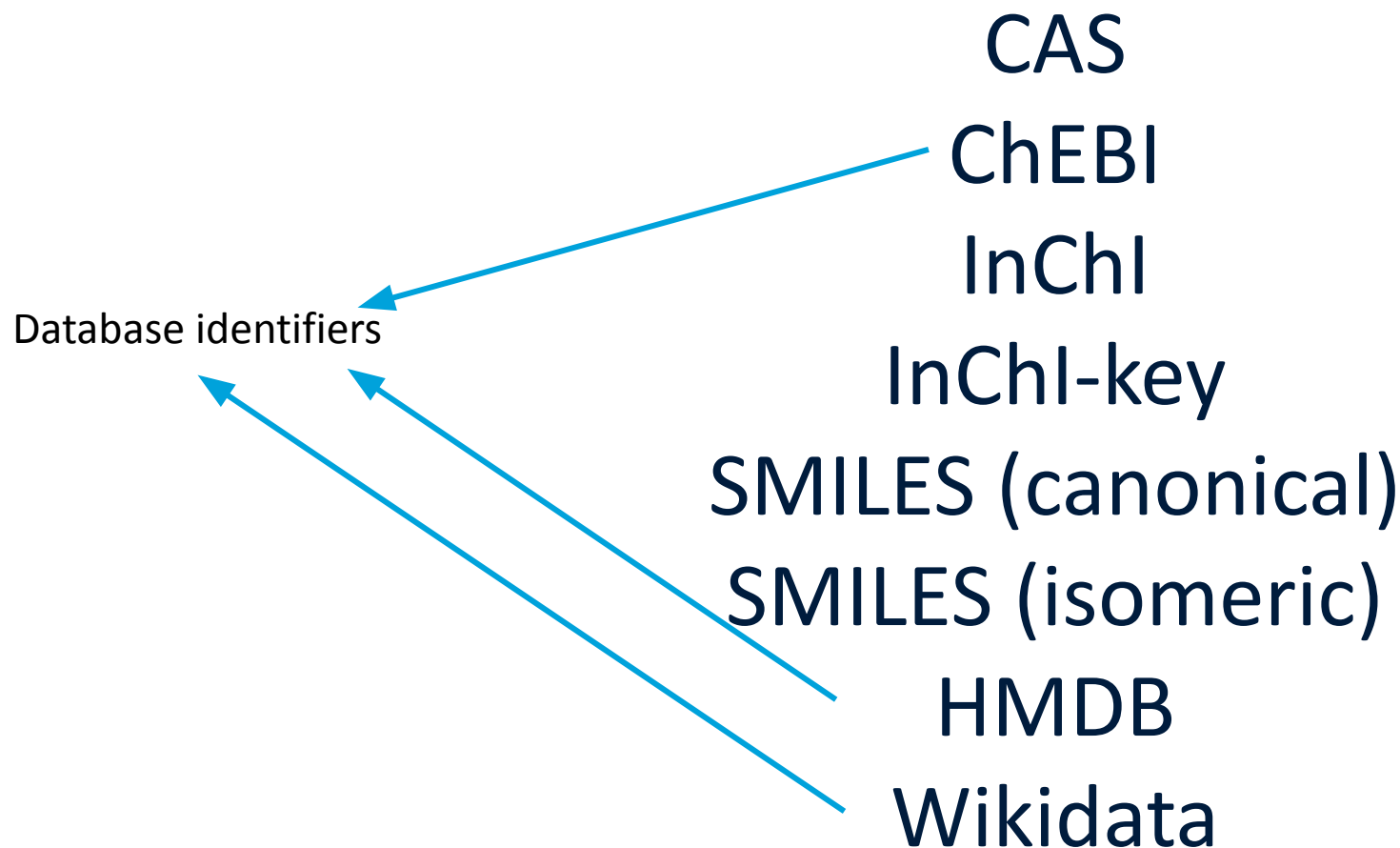
SMILES (isomeric)

HMDB

Wikidata

Different use for IDs

CAS

ChEBI

InChI

InChI-key

SMILES (canonical)

SMILES (isomeric)

HMDB

Wikidata

Ordering chemicals for wetlab

Different use for IDs

CAS

ChEBI

InChI

Structural IDs

InChI-key

SMILES (canonical)

SMILES (isomeric)

HMDB

Wikidata

Maastricht University

Different use for IDs

CAS

ChEBI

InChI

InChI-key

SMILES (canonical)

SMILES (isomeric)

HMDB

Wikidata

Database identifiers

Maastricht University
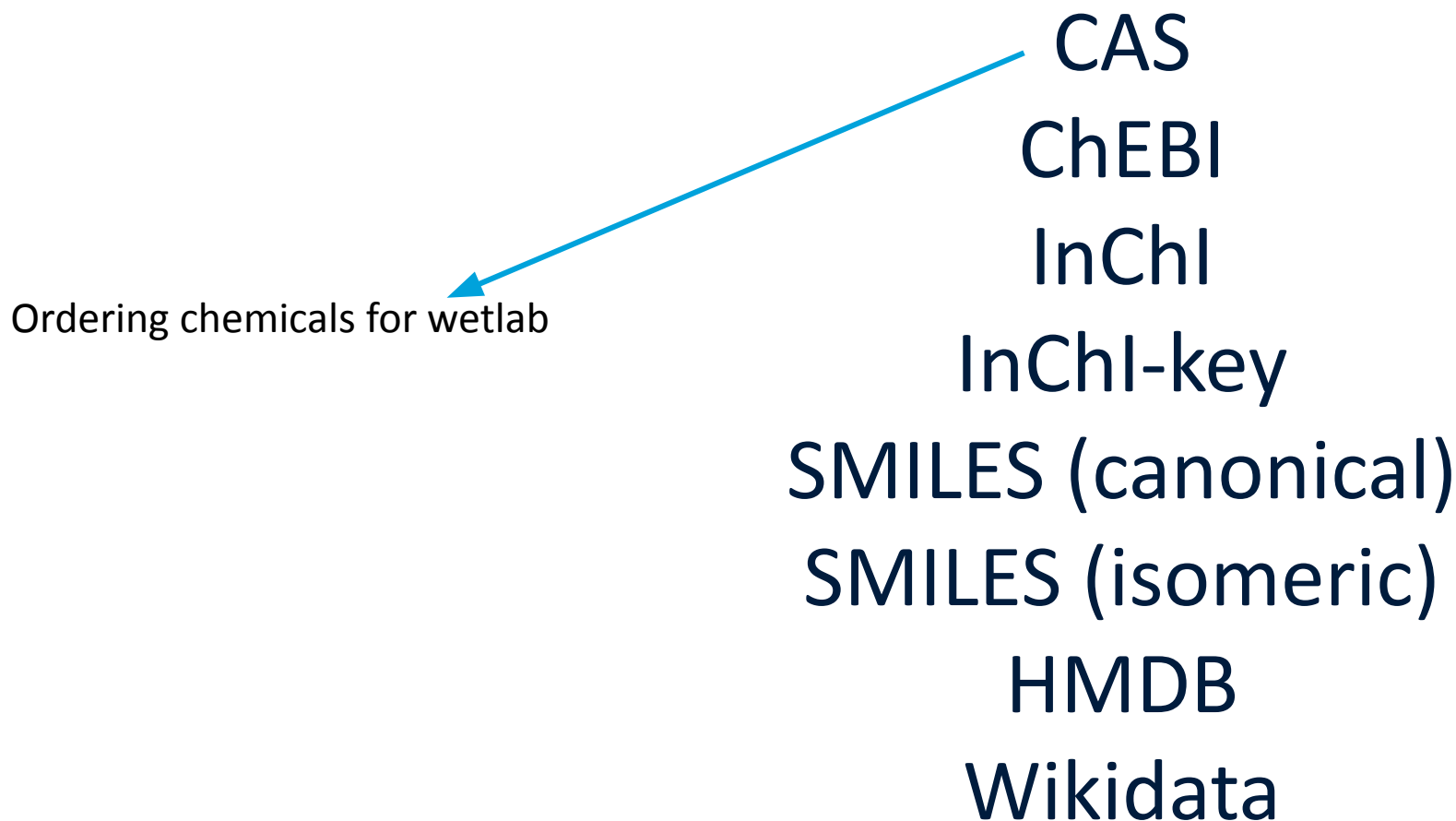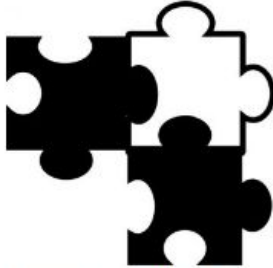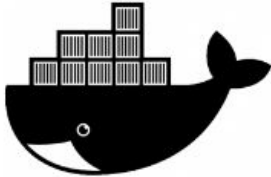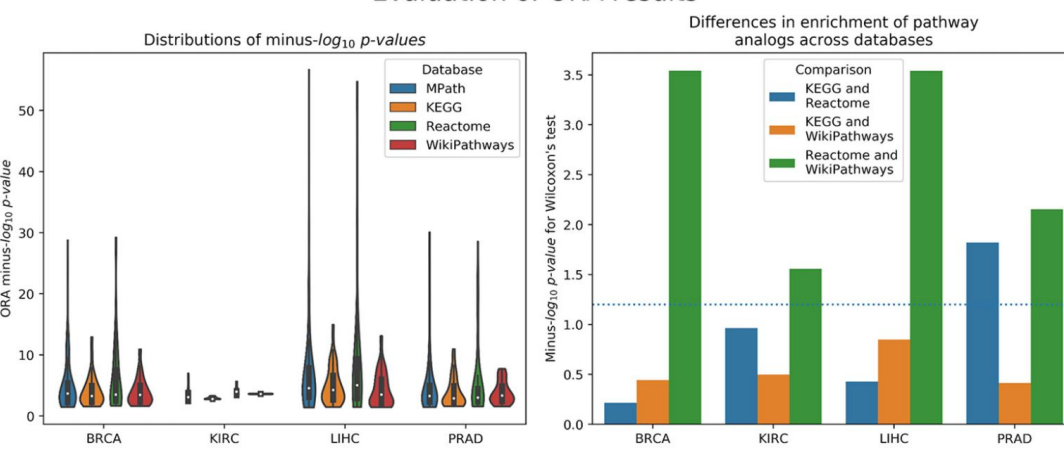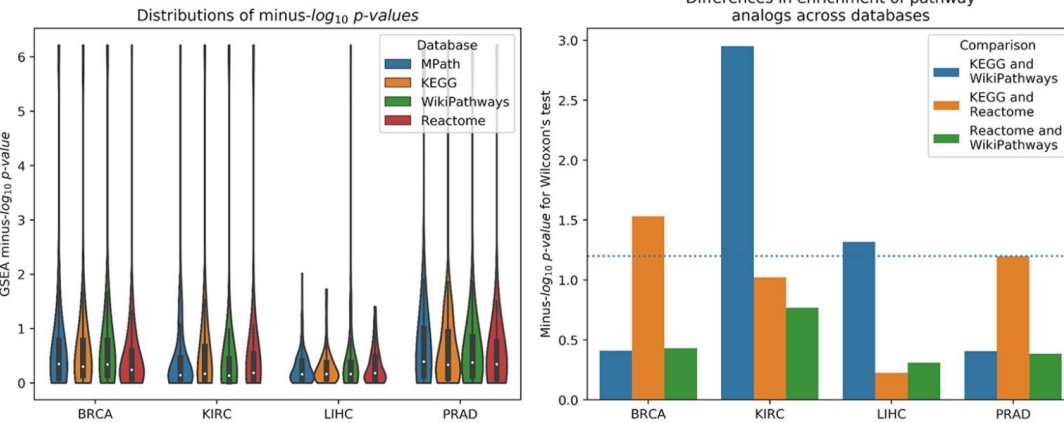
Interoperability creates CONNECTIONS

Framework and libraries

Identifier Mapping Databases

Tutorials and Workflows

Webservice and API

Webservice implementation Docker and semantic web

Documentation

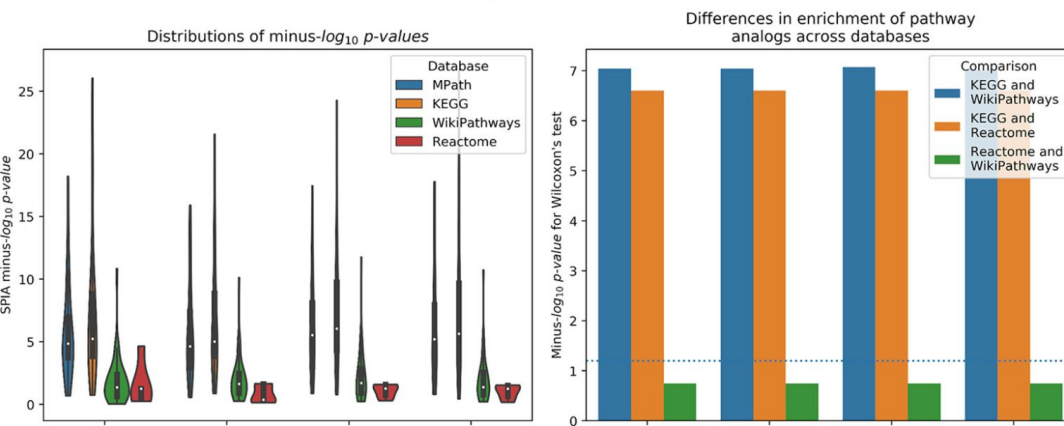Collaborators

Contact us

How to cite

https://bridgedb.github.io/

Left: Distribution of raw p values of pathway analogs across databases [top to bottom: overrepresentation analysis (ORA), gene set enrichment analysis (GSEA), and signaling pathway impact analysis (SPIA)].

Right: Significance of average rank differences of pathway analogs across pairwise database comparisons for the given method.

Fig. 3 adapted from Mubeen S, Hoyt CT, Gemünd A, Hofmann-Apitius M, Fröhlich H and Domingo-Fernández D (2019) The Impact of Pathway Database Choice on Statistical Enrichment Analysis and Predictive Modeling. Front. Genet. 10:1203. doi: 10.3389/fgene.2019.01203
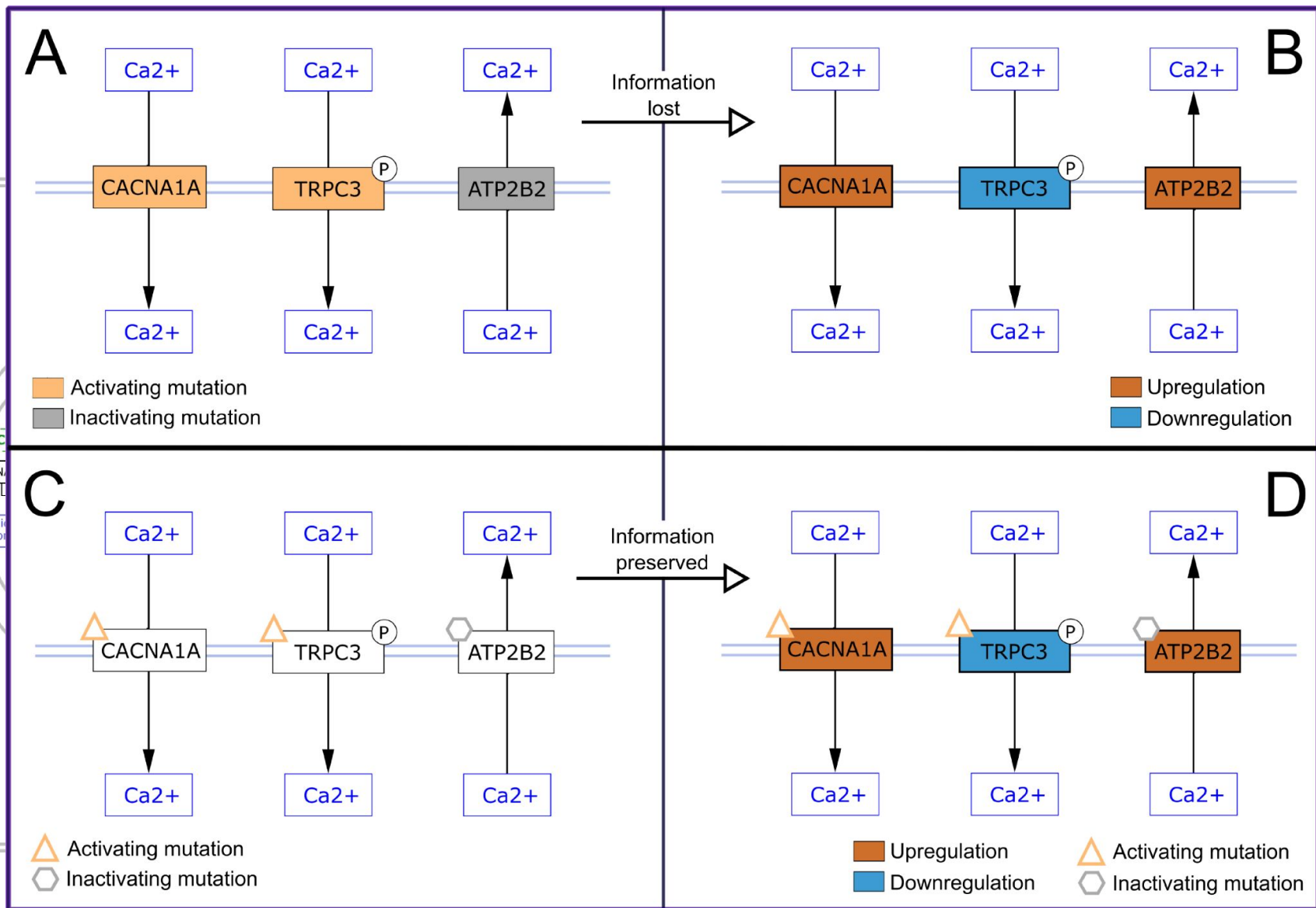
Fig 1-3 adapted from Hanspers K, Kutmon M, Coort SL, Digles D, Dupuis LJ, Ehrhart F, et al. (2021) Ten simple rules for creating reusable pathway models for computational analysis and visualization. PLoS Comput Biol 17(8): e1009226. https://doi.org/10.1371/journal.pcbi.1009226

Maastricht University

25 Years of Pathway Figures

Introduction

The WikiPathways team at Gladstone Instiutes has searched the literature over the past 25 years for pathway figures. This interactive tool lets you filter, search and view their findings.
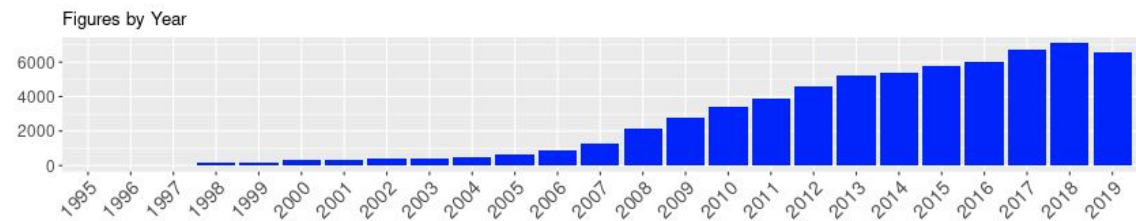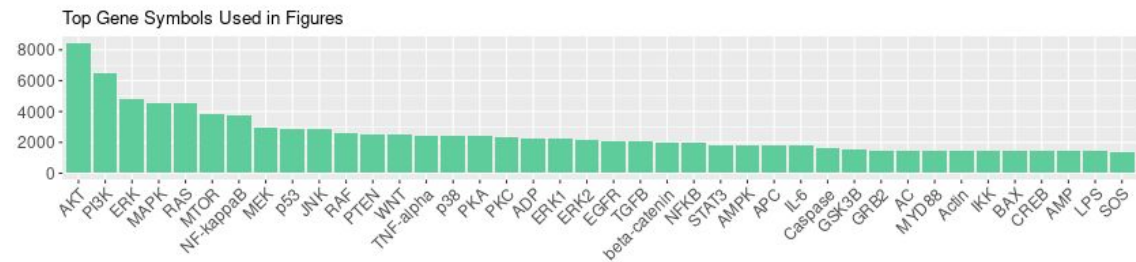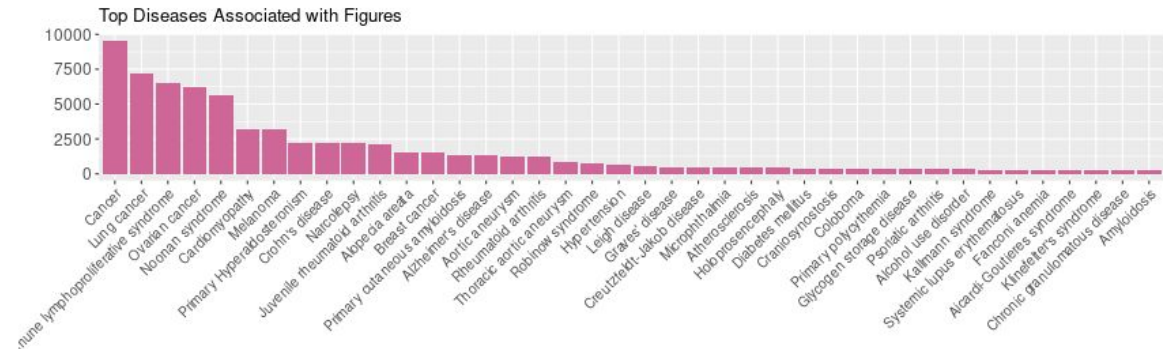
Summary stats

Figures: 64,643
Papers: 56,095
Total genes: 1,112,551
Unique genes: 13,464

1. Filter Figures

Disease Annotations

Gene content

Publication Years

Top Diseases Associated with Figures

Top Gene Symbols Used in Figures

Figures by Year

2. Search Filtered Figures

| Copy | CSV | Excel | PDF |

Search:

pmcid	paper.title	authors	year	number	figure.title

Maastricht University

https://gladstone-bioinformatics.shinyapps.io/shiny-25years/

Acknowledgements

Gladstone Institutes: Alex Pico, Anders Riutta
Heriot Watt University: Alasdair Gray

Maastricht University: Anwesha Bohler, Chris Evelo, Martijn van Iersel, Thomas Kelder, Christ Leemans, Marvin Martens, Luc, Meyer, Jonathan Mélius, Denise Slenter, Lucas Uberti-Bona Marin, Egon Willighagen

University of Manchester: Christian Brenninkmeijer, Ian Dunlop, Carole Goble, Nick Juty, Stian Soiland-Reyes
Others: Manas Awasthi, Jianjiong Gao, Isaac Ho, Kozo Nishida

Time for some exercises!