Systems Biology & BiGCaT

Egon Willighagen
BiGCaT, NUTRIM, Maastricht University
2019-07-17, Luxembourg
ORCID: 0000-0001-7542-0286
@egonwillighagen

Maastricht University

BiGCaT

Prof. dr. Chris Evelo

Dr. Friederike Ehrhart

Dr. Lars Eijssen

Dr. Egon Willighagen

Dr. Martina Summer-Kutmon

Dr. Susan Steinbusch-Coort

www.BiGCaT.unimaas.nl

BiGCaT

department of bioinformatics

NUTRIM

Postdoctoral researchers

Serena Bonaretti, PhD

Lauren Dupuis, PhD.

PhD students

Nasim B. Sangani, MSc.

Marvin Martens, MSc

Denise Slenter, MSc

Amadeo Muñoz García, MSc

Ryan Miller, MSc

Mirella Kalafati, Msc

Technical staff

Nuno Nunes

Funded projects

NanoCommons
Nano-Knowledge Community

EUTOXRISK

EuroDISH

OpenRiskNet
RISK ASSESSMENT E-INFRASTRUCTURE

RISK GONE

eNM
eNanoMapper

NanoSolveIT

HORIZON 2020

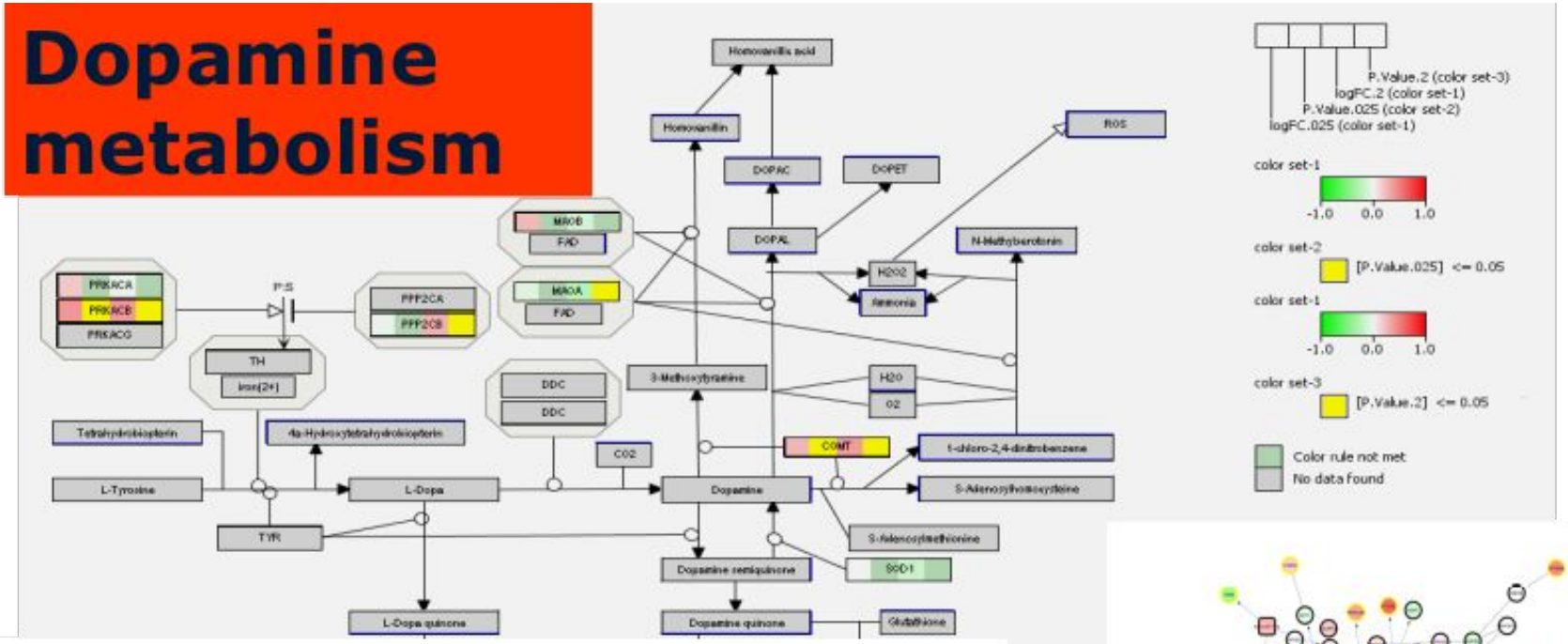
Helis Academy

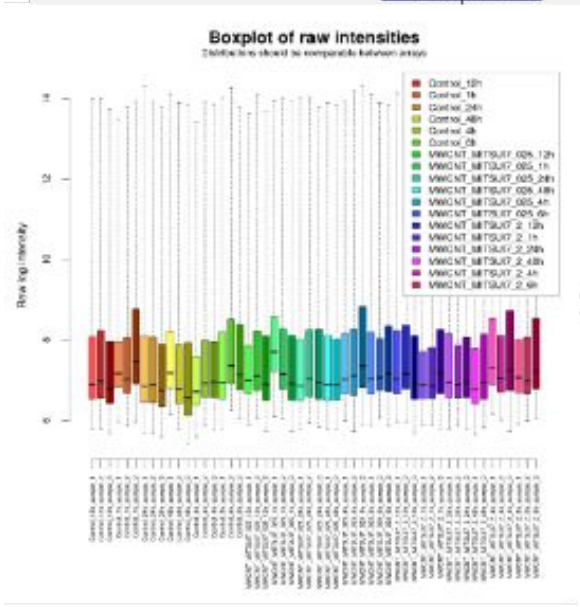
Maastricht University

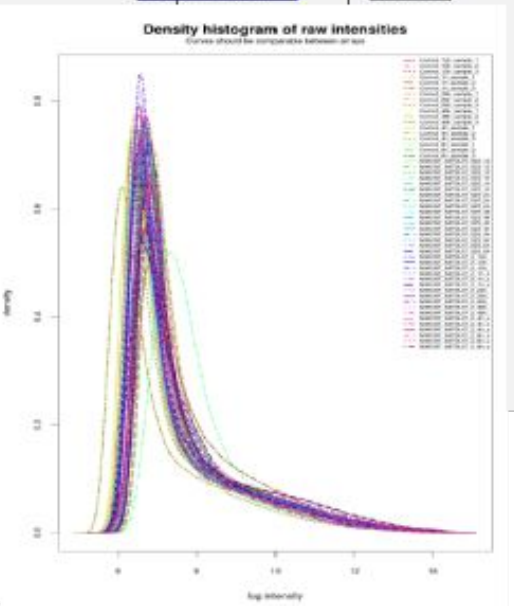
FAIRplus

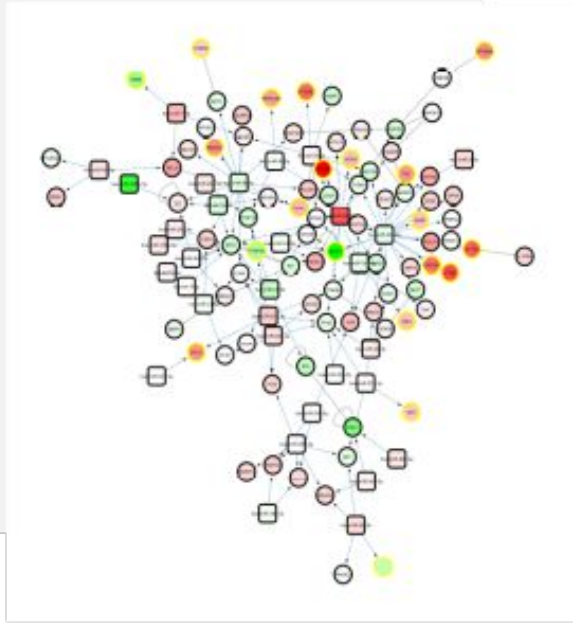
EUROPEAN JOINT PROGRAMME
RARE DISEASES

Dopamine metabolism







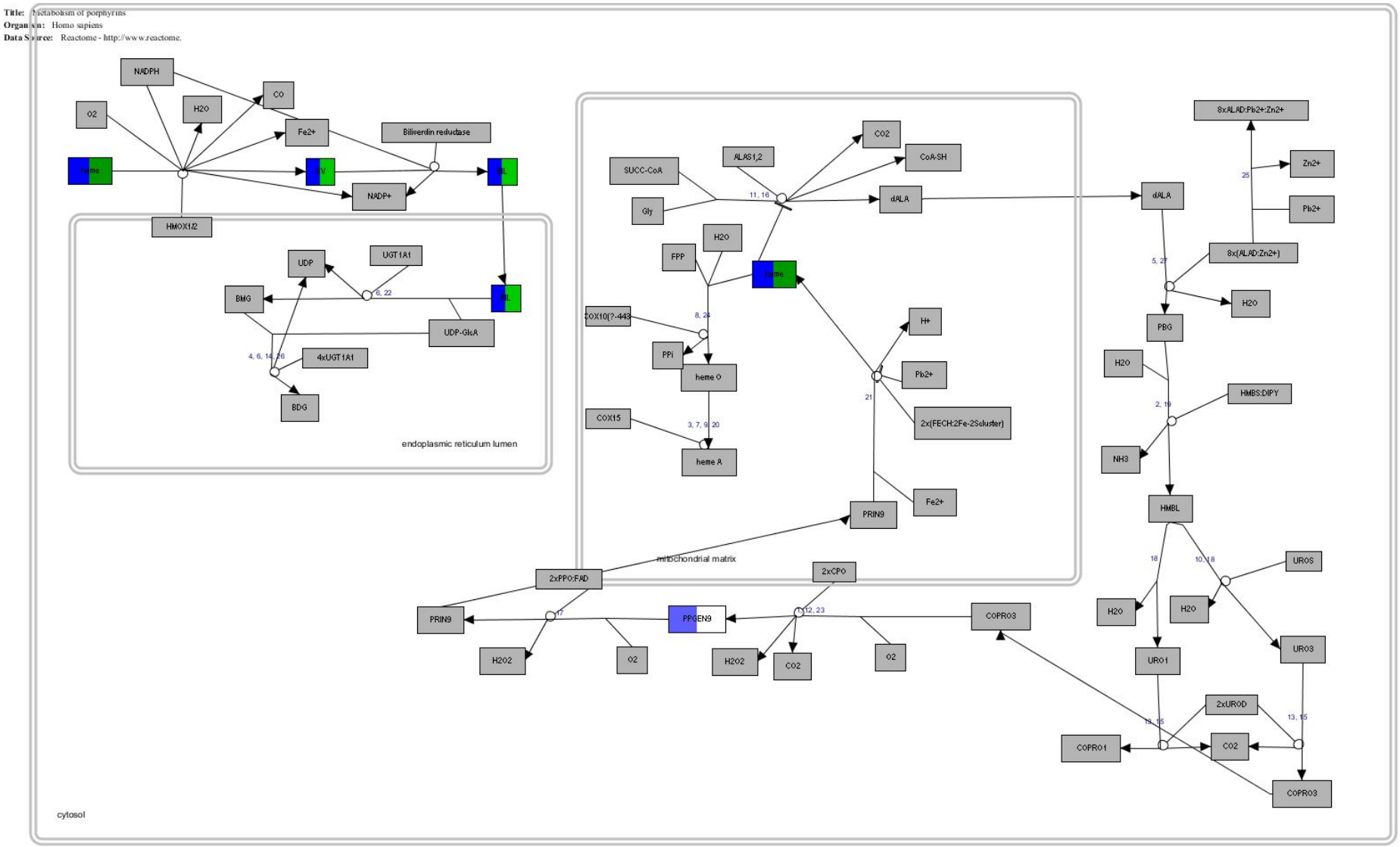


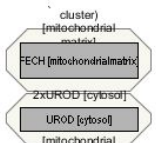
Marloes Poort

Maastricht University

The effect of Multi-walled carbon nanotubes on gene expression in bronchial epithelial BEAS 2B cells, B.Sc. Thesis, 2015

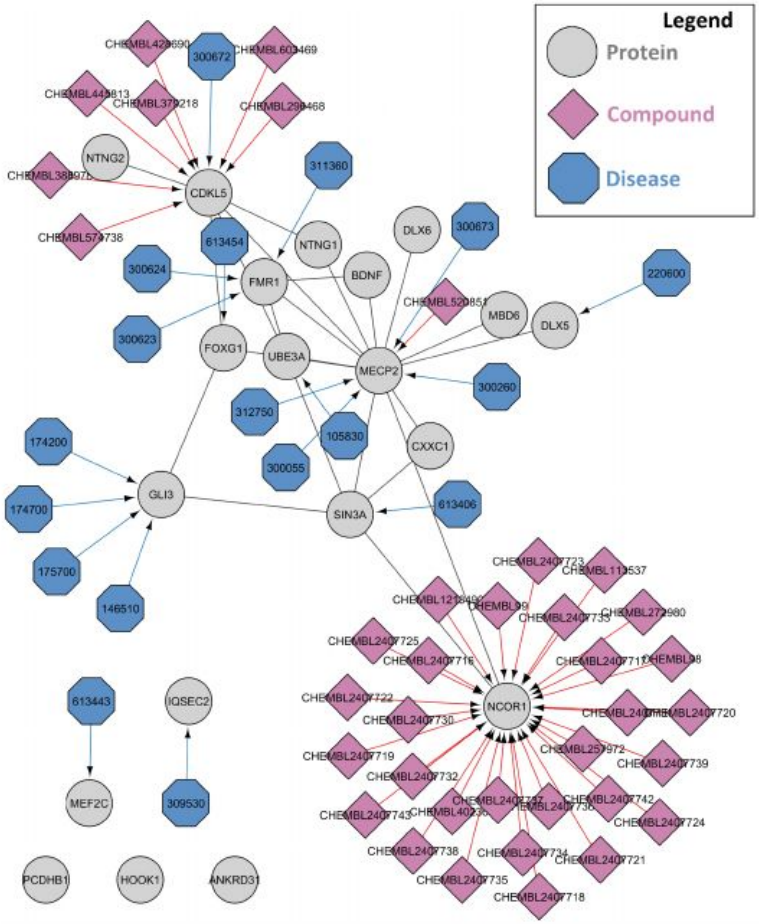
The effect of troglitazone on heme biosynthesis

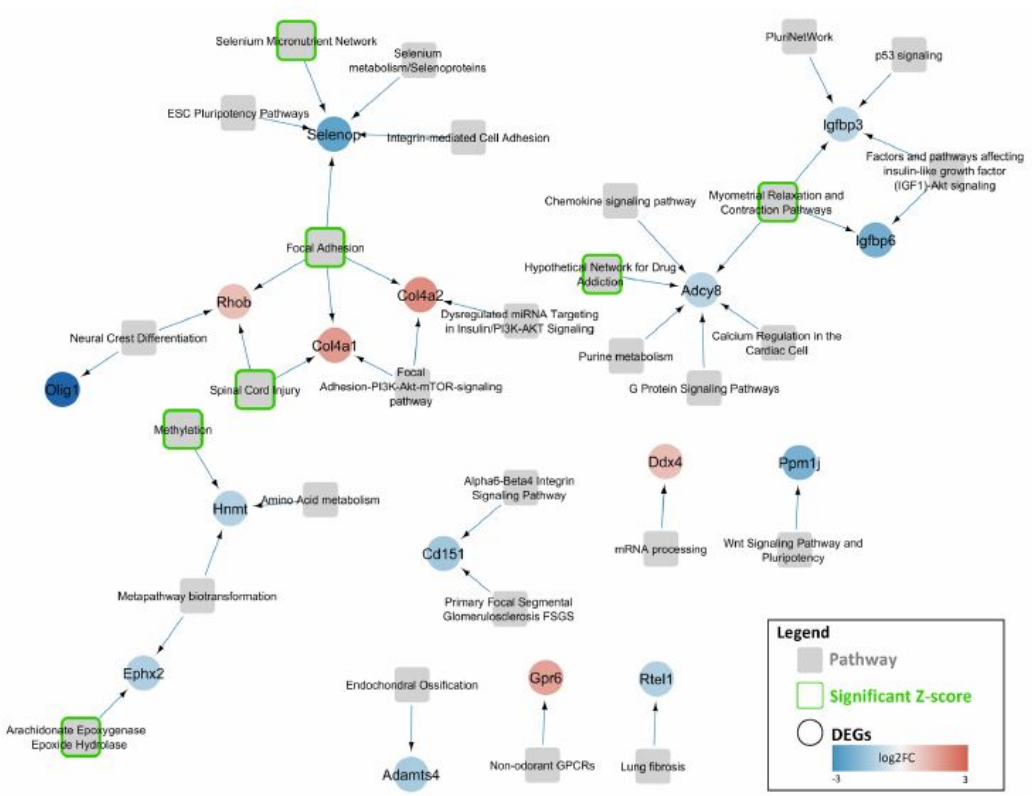




Metabolism of porphyrins (Homo sapiens), 2018, Reactome Team, Anwesha Bohler, Egon Willighagen, Martijn van Iersel, http://identifiers.org/wikipathways/WP1852

CyTargetLinker

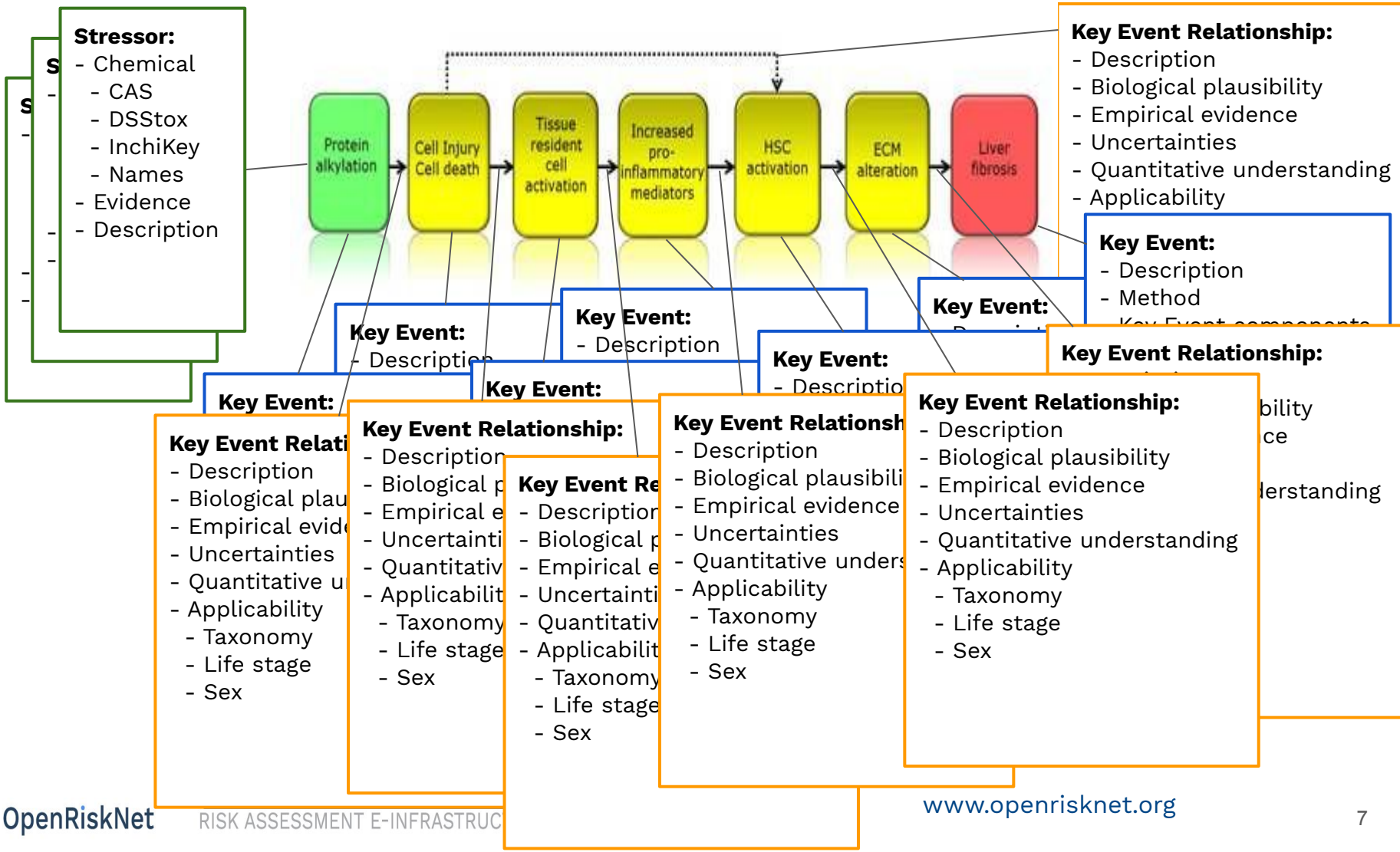




CyTargetLinker app update: A flexible solution for network extension in Cytoscape. 2018. F1000Research. 7. 743. 10.12688/f1000research.14613.1

Maastricht University

AOPs in AOP-Wiki contain diverse types of info



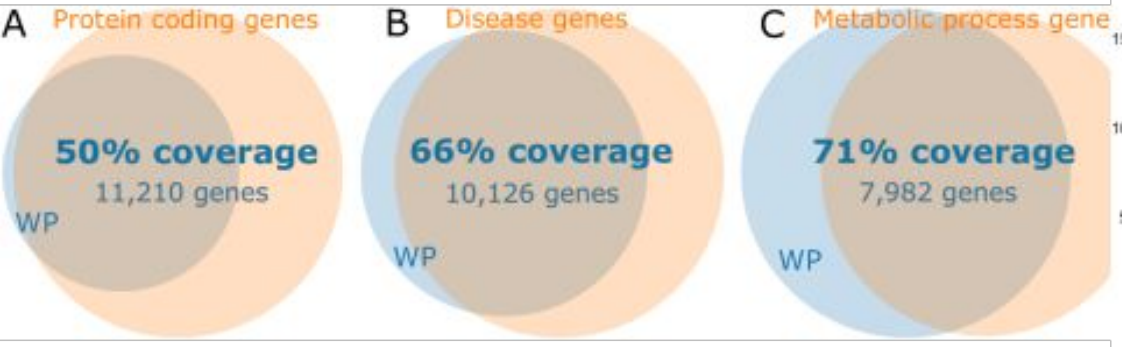
WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research

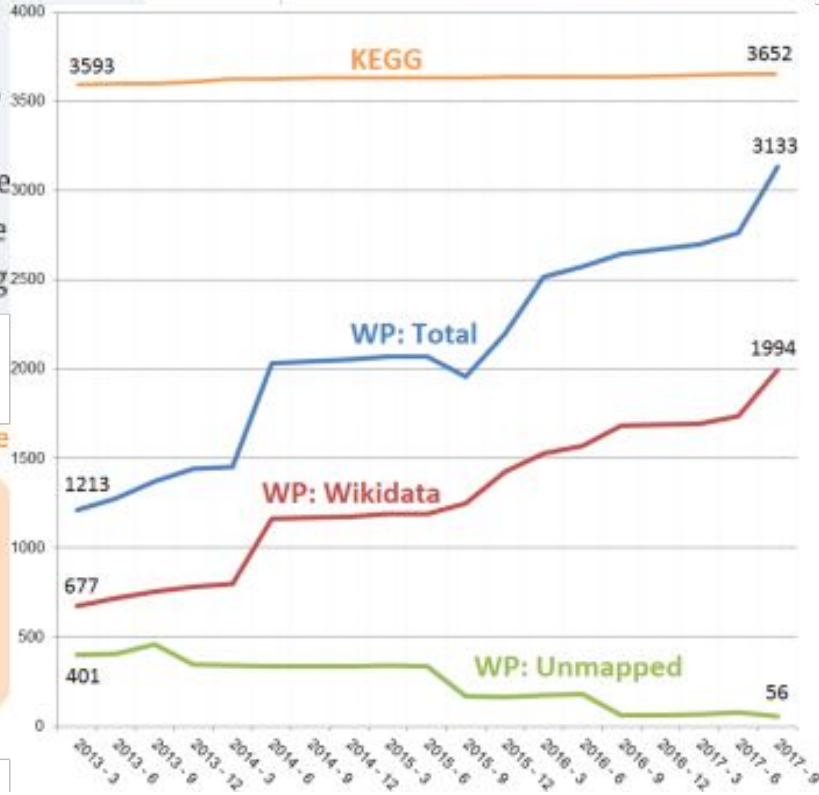
Denise N Slenter, Martina Kutmon, Kristina Hanspers, Anders Riutta, Jacob Windsor, Nuno Nunes, Jonathan Mélius, Elisa Cirillo, Susan L Coort, Daniela Digles ... Show more

Nucleic Acids Research, Volume 46, Issue D1, 4 January 2018, Pages D661–D667,
https://doi.org/10.1093/nar/gkx1064
Published: 10 November 2017 **Article history** ▾

Abstract

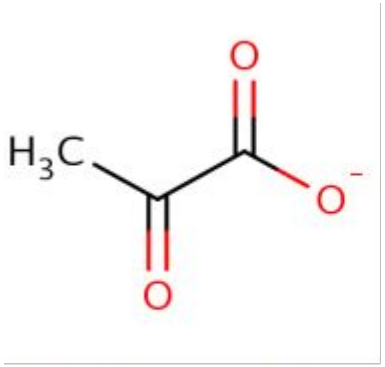
WikiPathways (wikipathways.org) captures the collective knowledge represented in biological pathways. By providing a database in a curated, machine readable way, omics data analysis and visualization is enabled. WikiPathways and other pathway databases are used to analyze experime data by research groups in many fields. Due to the open and collaborative nature of the WikiPathways platform, our content keeps growing and is g



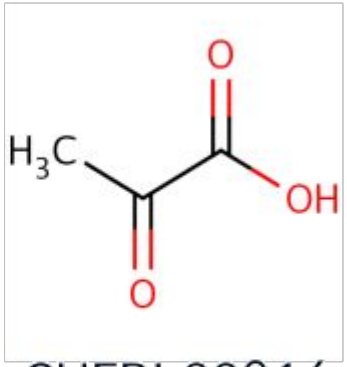


Chemistry in metabolic pathways

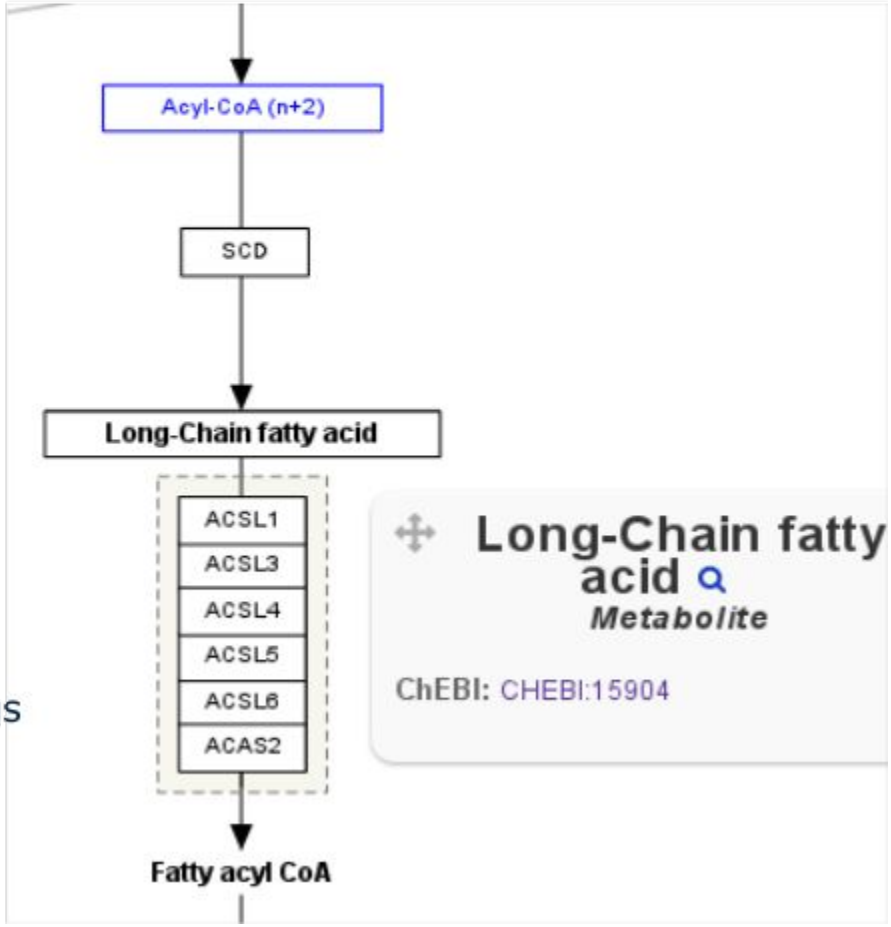
CHEBI:15361 (Pyruvate) -> Ce:CHEBI:32816 (conjugate) -> Ck:C00022 -> [WP2456 HIF1A and PPARG regulation of glycolysis, WP2453 TCA Cycle and PDHc]



CHEBI:15361



CHEBI:32816



Brenninkmeijer, CYA, et al. "Scientific Lenses over Linked Data: An approach to support task specific views of the data. A vision." Proceedings of 2nd International Workshop on Linked Science. 2012.

So, what IDs are used in WikiPathways?

2017

datasource	numberEntries
ChEBI	1923
HMDB	623
CAS	299
KEGG Compound	251
PubChem-compound	245
Chemspider	174
PubChem-substance	33
LIPID MAPS	10
Reactome	4
Wikidata	3
ChEMBL compound	2

+ Reactome

2015

source	co
HMDB	56
ChEBI	49
KEGG Compound	40
CAS	29
PubChem-compound	21
Chemspider	15
PubChem-substance	24
LIPID MAPS	11
Wikipedia	9
ChemIDplus	7
Reactome	4
ChEMBL compound	2
Other	1
CTD Chemical	1
ChemSpider	1

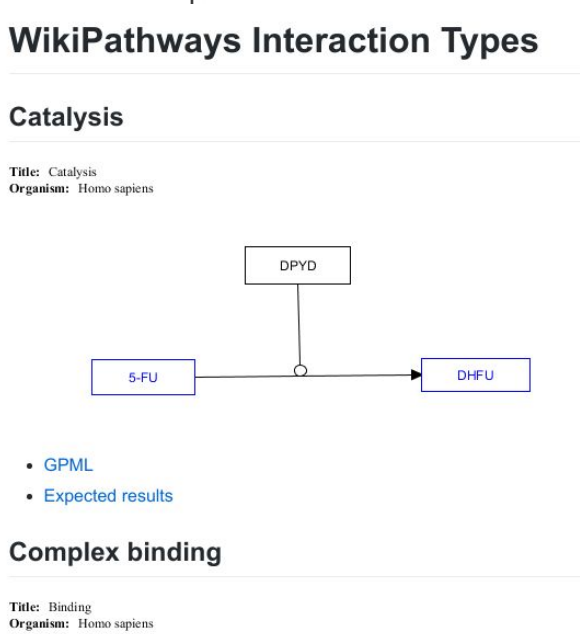
2019

datasource	numberEntries
ChEBI	2428
HMDB	931
KEGG Compound	397
PubChem-compound	357
CAS	293
Wikidata	273
Chemspider	202
LIPID MAPS	148
KNApSAcK	112
ChEMBL compound	7
Reactome	2
KEGG Glycan	2
PubChem-substance	2
TTD Drug	1

(What about reactions (- identifiers)? In progress

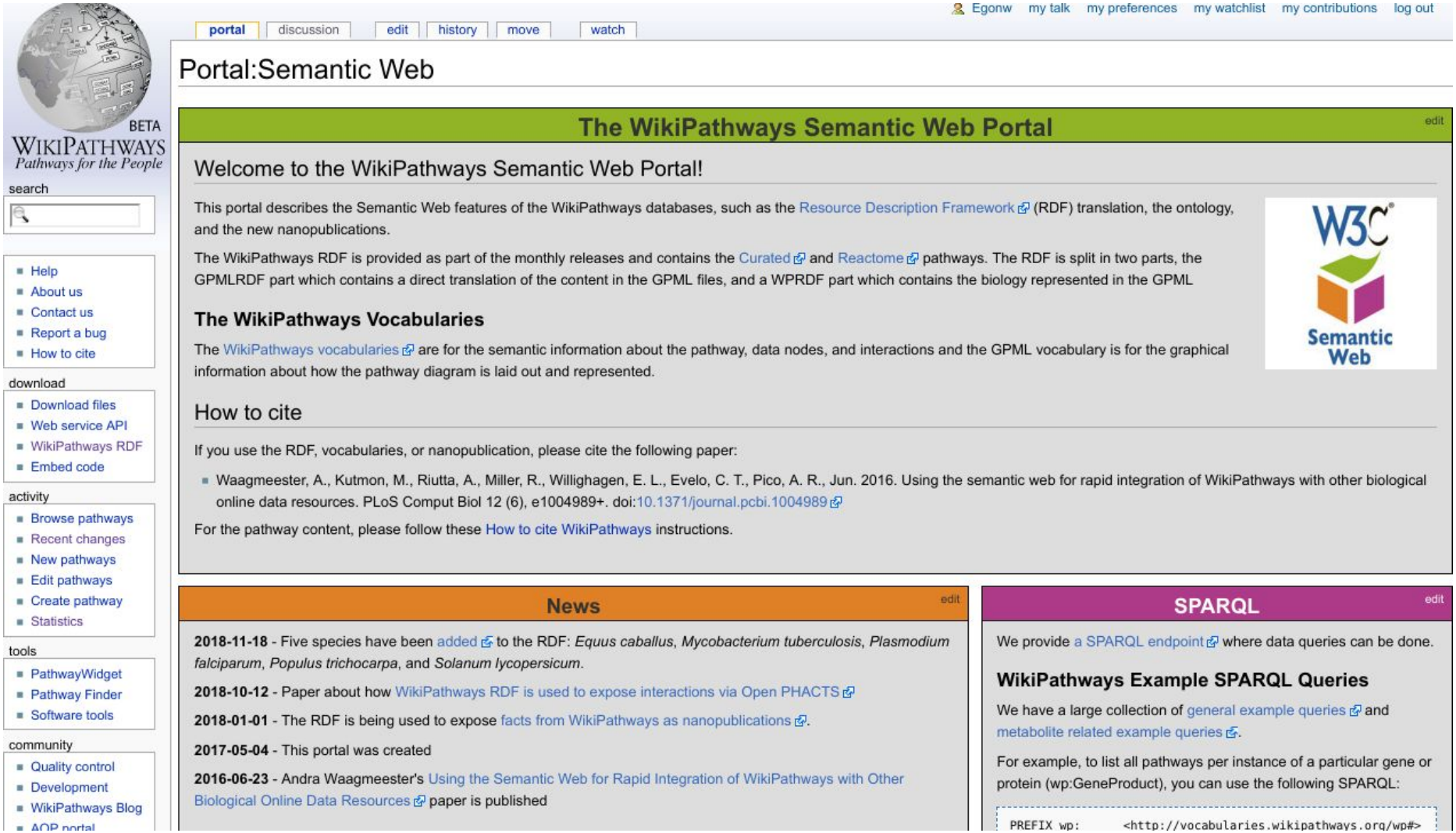
source	count
Reactome	11320
KEGG Reaction	61
Rhea	59
WikiPathways	13
KEGG Pathway	8
XMetDB	2
SPIKE	1
WormBase	1

	Interaction Type	Interaction Count
0	Interaction	15228
1	DirectedInteraction	11586
2	Conversion	1362
3	Catalysis	1169
4	Inhibition	1081
5	ComplexBinding	889
6	Binding	889
7	Stimulation	721
8	TranscriptionTranslation	256



Maastricht University

R. Miller et al, Understanding signaling and metabolic paths using semantified and harmonized information about biological interactions, 2019, in preparation.


Egonw my talk my preferences my watchlist my contributions log out
portal discussion edit history move watch
BETA
WIKIPATHWAYS
Pathways for the People
search
Help
About us
Contact us
Report a bug
How to cite
download
Download files
Web service API
WikiPathways RDF
Embed code
activity
Browse pathways
Recent changes
New pathways
Edit pathways
Create pathway
Statistics
tools
PathwayWidget
Pathway Finder
Software tools
community
Quality control
Development
WikiPathways Blog
Portal:Semantic Web
The WikiPathways Semantic Web Portal
edit
Welcome to the WikiPathways Semantic Web Portal!
This portal describes the Semantic Web features of the WikiPathways databases, such as the Resource Description Framework (RDF) translation, the ontology, and the new nanopublications.
The WikiPathways RDF is provided as part of the monthly releases and contains the Curated and Reactome pathways. The RDF is split in two parts, the GPMLRDF part which contains a direct translation of the content in the GPML files, and a WPRDF part which contains the biology represented in the GPML
W3C
Semantic Web
The WikiPathways Vocabularies
The WikiPathways vocabularies are for the semantic information about the pathway, data nodes, and interactions and the GPML vocabulary is for the graphical information about how the pathway diagram is laid out and represented.
How to cite
If you use the RDF, vocabularies, or nanopublication, please cite the following paper:
Waagmeester, A., Kutmon, M., Riutta, A., Miller, R., Willighagen, E. L., Evelo, C. T., Pico, A. R., Jun. 2016. Using the semantic web for rapid integration of WikiPathways with other biological online data resources. PLoS Comput Biol 12 (6), e1004989+. doi:10.1371/journal.pcbi.1004989
For the pathway content, please follow these How to cite WikiPathways instructions.
News
edit
2018-11-18 - Five species have been added to the RDF: Equus caballus, Mycobacterium tuberculosis, Plasmodium falciparum, Populus trichocarpa, and Solanum lycopersicum.
2018-10-12 - Paper about how WikiPathways RDF is used to expose interactions via Open PHACTS
2018-01-01 - The RDF is being used to expose facts from WikiPathways as nanopublications.
2017-05-04 - This portal was created
2016-06-23 - Andra Waagmeester's Using the Semantic Web for Rapid Integration of WikiPathways with Other Biological Online Data Resources paper is published
SPARQL
edit
We provide a SPARQL endpoint where data queries can be done.
WikiPathways Example SPARQL Queries
We have a large collection of general example queries and metabolite related example queries.
For example, to list all pathways per instance of a particular gene or protein (wp:GeneProduct), you can use the following SPARQL:
PREFIX wp: <http://vocabularies.wikipathways.org/wp#>

SPARQL endpoint

Virtuoso SPARQL Query Editor

Default Data Set Name (Graph IRI)

Query Text

```
select ?KeyEventID ?KeyEventName ?AssayText
where {
 ?KeLook a aopo:KeyEvent ;
 rdfs:label ?KeyEventID ;
 dc:title ?KeyEventName ;
 dcterms:isPartOf ?aop1 ;
 mmo:0000000 ?AssayText .
 ?aop1 dc:identifier ?AopAssoc  .
 values ?aop1 {aop:38}
}
```

(Security restrictions of this server do not allow you to retrieve remote RDF data, see details.)

Results Format: HTML

Execution timeout: 0 milliseconds *(values less than 1000 are ignored)*

Options:
- [x] Strict checking of void variables
- [ ] Log debug info at the end of output (has no effect on some queries and output formats)
- [ ] Generate SPARQL compilation report (instead of executing the query)

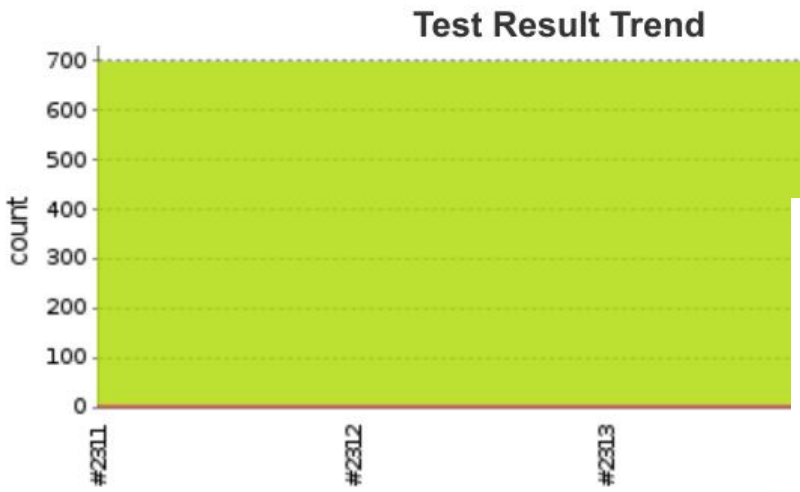
(The result can only be sent back to browser, not saved on the server, see details)

Run Query | Reset

grlc REST API

Computer assisted Curation (SPARQL++)

Wikidata Checks for Metabolomics	6 hr 29 min - #63
WikiPathways Curation (SPARQL endpoint)	6 days 13 hr - #5510
WikiPathways Curation (WP2RDF Curation)	33 min - #5605
WikiPathways Curation (WP2RDF Curation) (all)	5 hr 32 min - #2315
WikiPathways Curation (WP2RDF Curation) (all, dev)	5 hr 10 min - #1483
WikiPathways Curation Events	5 hr 10 min - #710



Test Result

2 failures (+1)

693 tests (±0)

Took 8 min 42 sec.

All Failed Tests

Test Name	Duration	Age
nl.unimaas.bigcat.wikipathways.curator.ChEBIMetabolites.secondaryChEBIIdentifiers	0.26 sec	1
nl.unimaas.bigcat.wikipathways.curator.Metabolites.ChEBIIDsNotMarkedAsMetabolite	43 ms	1

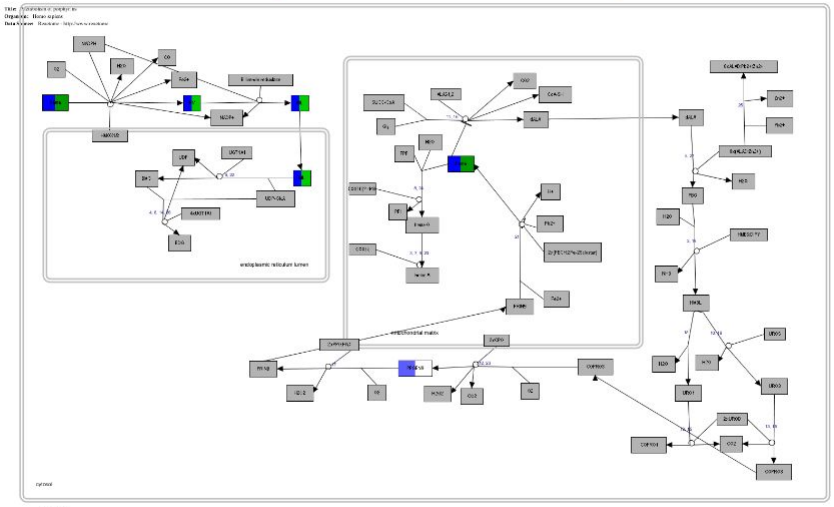
All Tests

Package	Duration	Fail	(diff)	Skip	(diff)	Pass	(diff)	Total	(diff)
nl.unimaas.bigcat.wikipathways.curator	1 min 13 sec	2	+1	0		691	-1	693	

Maastricht University

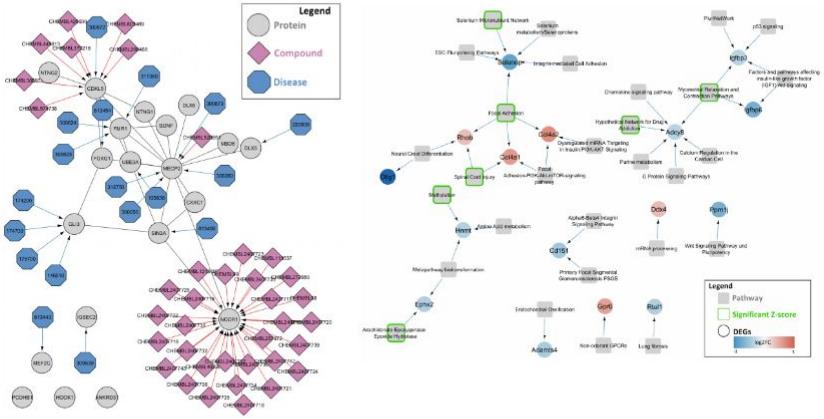
From individual entities to huge networks

The effect of troglitazone on heme biosynthesis



Metabolism of porphyrins (Homo sapiens), 2018, Reactome Team, Anwesha Bohler, Egon Willighagen, Martijn van Iersel, http://identifiers.org/wikipathways/WP1852

CyTargetLinker



CyTargetLinker app update: A flexible solution for network extension in Cytoscape. 2018. F1000Research. 7. 743. 10.12688/f1000research.14613.1

Maastricht University

Wikidata / Scholia

Redirecting

If you know the identifier then Scholia can make a lookup based on the identifier:

cas/50-00-0

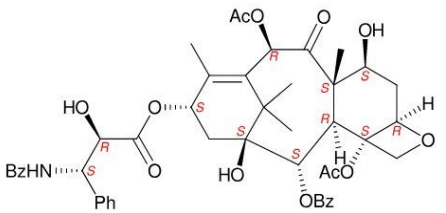
Lookup CAS 50-00-0. This will identify formaldehyde and redirect to its Scholia page.

inchikey/QTBSBXVTEAMEQO-UHFFFAOYSA-N

Redirect also works for InChIKeys, here for acetic acid.

paclitaxel (Q423762)

Paclitaxel (PTX), sold under the brand name Taxol among others, is a chemotherapy medication used to treat a number of types of cancer. This includes ovarian cancer, breast cancer, lung cancer, Kaposi sarcoma, cervical cancer, and pancreatic cancer. It is given by injection into a vein. ... (from the English Wikipedia)



Identifiers

Show 10 entries Search:

IDpred	Id
ATC code	L01CD01
CAS Registry Number	33069-62-4

2019: 10.3897/rio.5.e35820
2017: 10.6084/m9.figshare.6356027.v1

Maastricht University

Wikidata / Scholia

DSSTOX substance identifier	DTXSID30678817
	1 reference
	stated in: Mapping file of InChIStrings, InChIKeys and DTXSIDs for the EPA CompTox Dashboard

IDpred	IDpredLabel	count
wd:P235	InChIKey	152393
wd:P233	canonical SMILES	152233
wd:P234	InChI	149944
wd:P662	PubChem CID	145798
wd:P661	ChemSpider ID	125510
wd:P2017	isomeric SMILES	84844
wd:P683	ChEBI ID	84011
wd:P231	CAS Registry Number	72475
wd:P652	UNII	59293
wd:P592	ChEMBL ID	49622
wd:P3117	DSSTOX substance identifier	36373
wd:P232	EC ID	20335
wd:P1579	Beilstein Registry Number	19083
wd:P665	KEGG ID	15065
wd:P2566	ECHA InfoCard ID	12362
wd:P715	Drugbank ID	7786
wd:P595	Guide to Pharmacology Ligand ID	5950
	HMDB ID	5705
	KNApSAcK ID	4272

Show 10 entries Search:

Mol	InChIKey	CAS	ChemSpider	PubChem CID
acetic acid	QTBSBXVTEAMEQO-UHFFFAOYSA-N	64-19-7	171	176
deuterated acetic acid	QTBSBXVTEAMEQO-GUEYOVJQSA-N	1186-52-3	2006083	2723903
acetic acid c-14	QTBSBXVTEAMEQO-HQMMCQRPSA-N	2845-03-6	144444	164769
acetic acid c-13	QTBSBXVTEAMEQO-VQEHIDDOSA-N	1563-79-7	8329490	10153982
acetic acid c-11	QTBSBXVTEAMEQO-JVVVGQRLSA-N	78887-71-5	396653	450349
acetate ion	QTBSBXVTEAMEQO-UHFFFAOYSA-M	71-50-1	170	175

Edit on query.Wikidata.org

Showing 1 to 6 of 6 entries Previous 1 Next

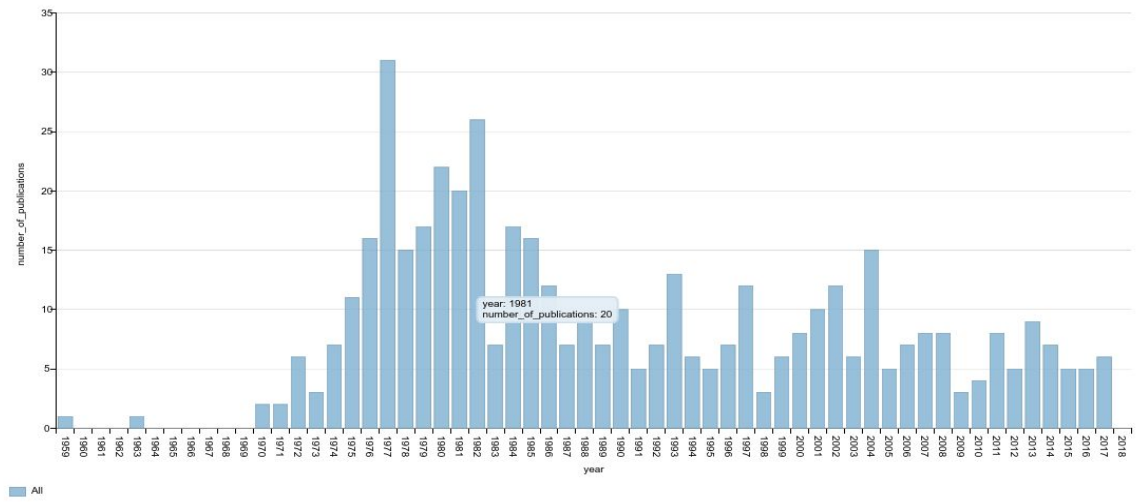
Wikidata / Scholia

Physchem Properties

Show 10 entries Search:

PropEntity	Value	Units	Qualifiers	Source	Doi
acid dissociation constant	4.74	1		Small Scale Determination of the pKa Values for Organic Acids	10.1021/ED071PA6
mass	60.021129	atomic mass unit		PubChem	
acid dissociation constant	4.756	1	temperature: 25	CRC Handbook of Chemistry and Physics (95th edition)	
boiling point	117.9	degrees Celsius	pressure: 101325	CRC Handbook of Chemistry and Physics (95th edition)	
density	1.0446	gram per cubic centimetre	temperature: 25	CRC Handbook of Chemistry and Physics (95th edition)	

Publications per year

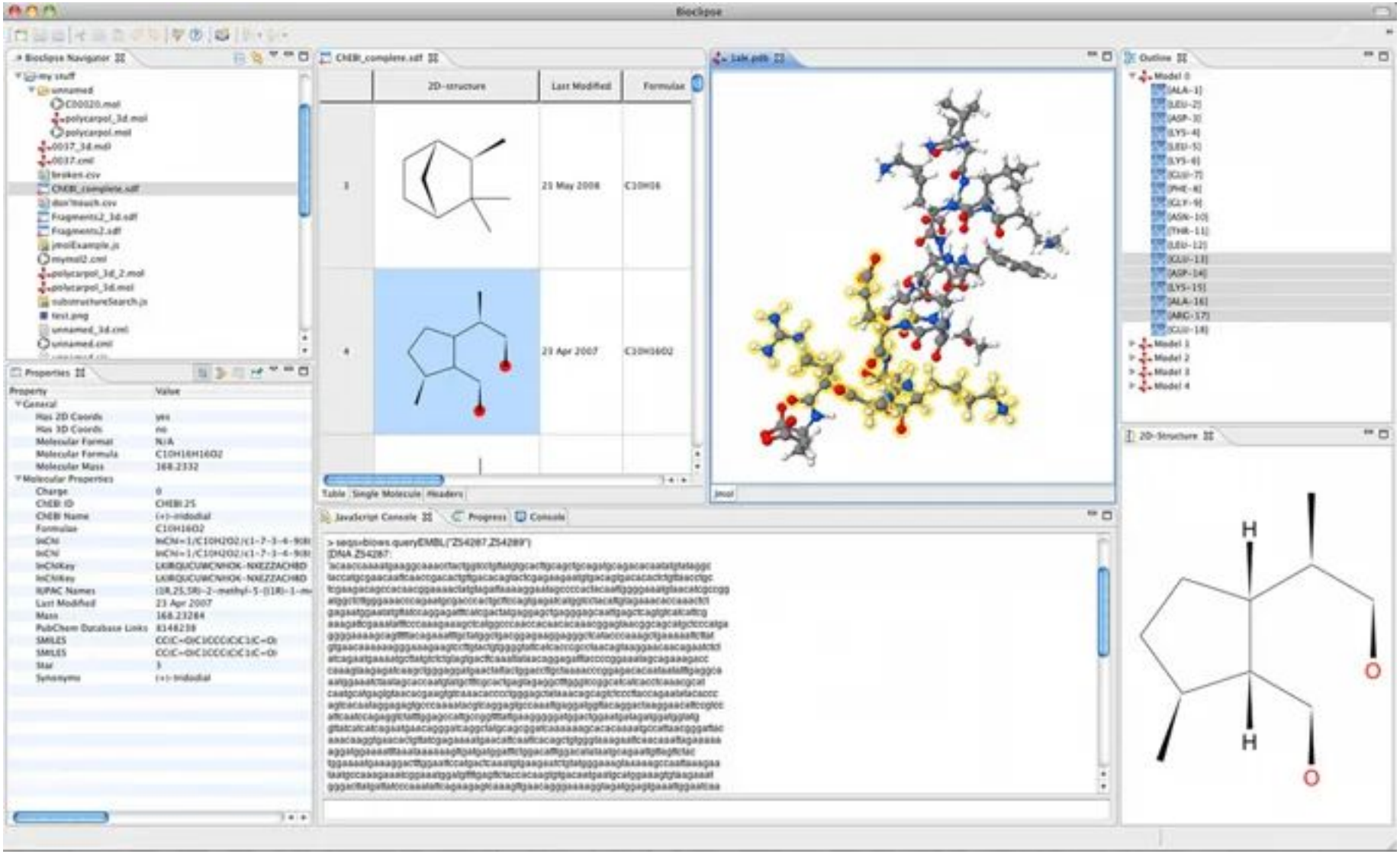


Recently published works on the chemical

Show 10 entries

Date	Work	Type	Topics
2017-08-09	In vitro human skin permeation of benzene in gasoline: effects of concentration, multiple dosing and skin preparation	scholarly article	oil and gas extraction // benzene
2017-04-27	Nicotine, aerosol particles, carbonyls and volatile organic compounds in tobacco- and menthol-flavored e-cigarettes	scholarly article	toluene // benzene

Automation / repeating



Bioclipse

topic

Bioclipse (Q1769726)

The Bioclipse project is a Java-based, open-source, visual platform for chemo- and bi
scripting functionality in 2009.Like any RCP application, Bioclipse uses a plugin archit
Eclipse, such as help system, software updates, preferences, cross-platform deploym
bioinformatics, and extension points that easily can be extended by other, possibly pr
Wikipedia)

Recently published works on the topic RSS

Show 10 entries

Date	Work
2013-01-15	Bioclipse-R: integrating management and visualization of life science data with statistical analysis
2013-01-01	Applications of the InChI in cheminformatics with the CDK and Bioclipse
2012-01-01	Interactive predictive toxicology with Bioclipse and OpenTox
2012-01-01	Open source drug discovery with bioclipse.
2011-01-01	Computational toxicology using the OpenTox application programming interface and Bioclipse
2010-01-01	Use of historic metabolic biotransformation data as a means of anticipating metabolic sites using MetaPrint2D and Bioclipse
2009-01-01	Bioclipse 2: a scriptable integration platform for the life sciences
2007-	Bioclipse: an open source workbench for chemo- and bioinformatics

BMC Part of Springer Nature Explore Journals Get Published About BMC

BMC Bioinformatics

Home About Articles Submission Guidelines

Software | OPEN | Published: 03 December 2009

Bioclipse 2: A scriptable integration platform for the life sciences

Ola Spjuth, Jonathan Alvarsson, Arvid Berg, Martin Eklund, Stefan Kuhn, Carl Mäsak, Gilleain Torrance, Johannes Wagener, Egon L Willighagen, Christoph Steinbeck & Jarl ES Wikberg

BMC Bioinformatics **10**, Article number: 397 (2009) | Download Citation

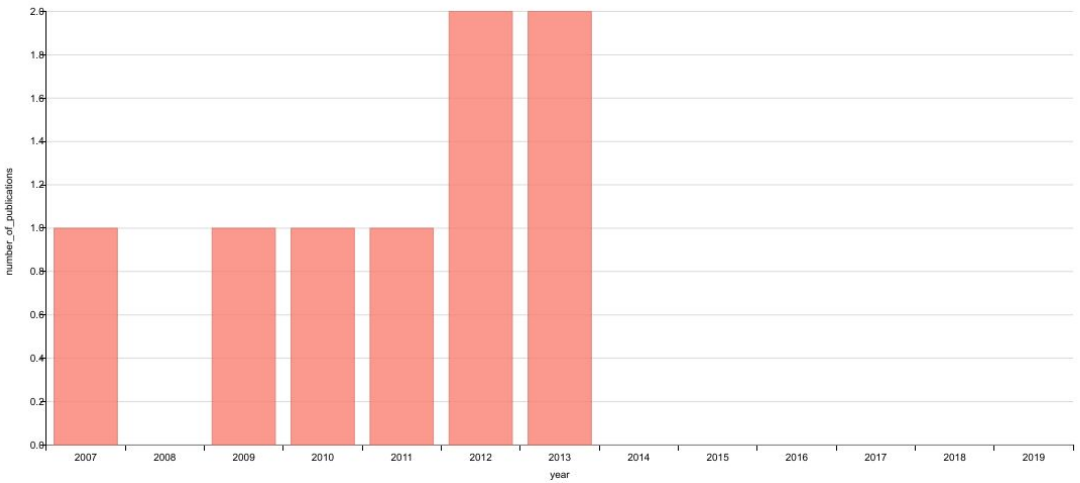
Abstract

Background

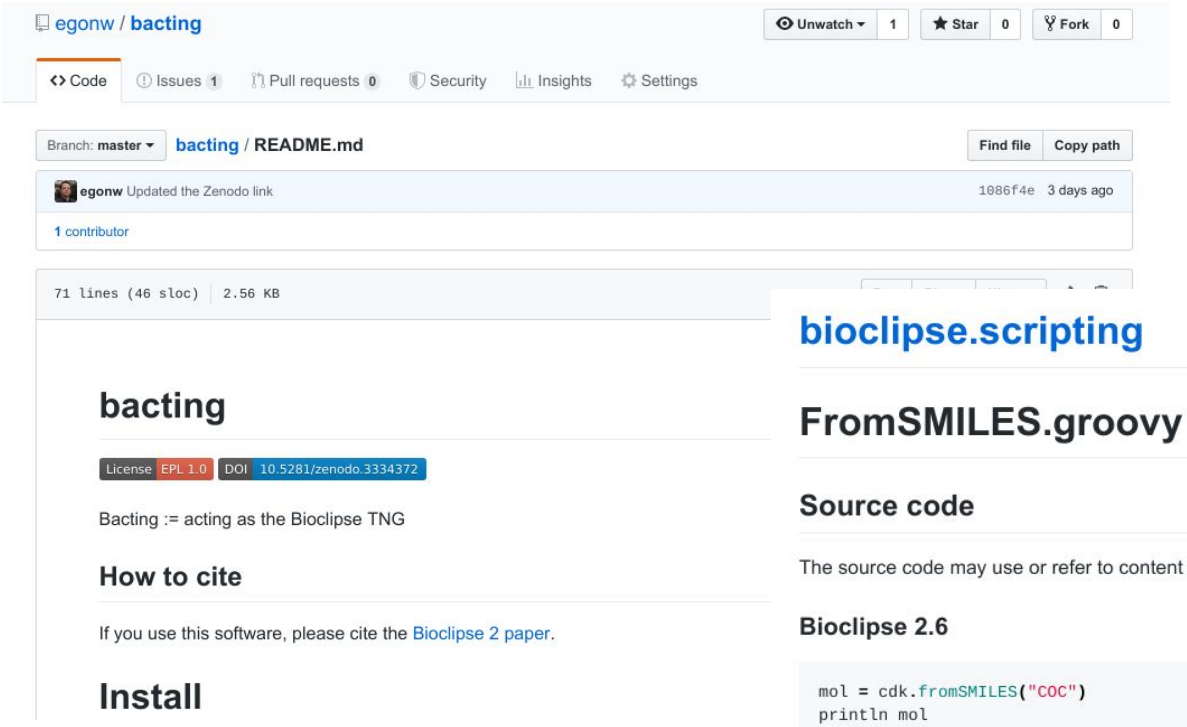
Contemporary biological research integrates neighboring scientific domains to answer complex questions in fields such as systems biology and drug discovery. This calls for tools that are intuitive to use, yet flexible to adapt to new tasks.

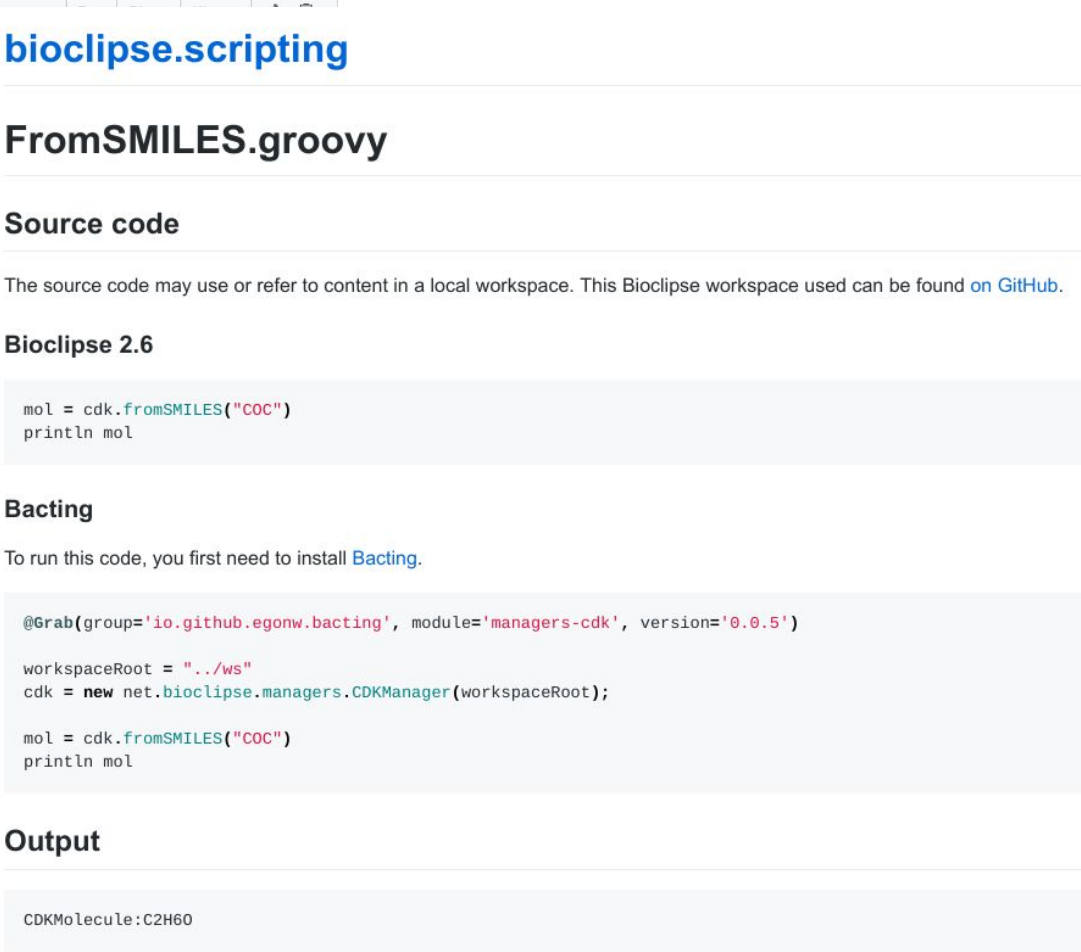
Results

Publications per year



Bacting := acting as the Bioclipse TNG





https://bioclipse.github.io/bioclipse.scripting/

Maastricht University

Bioconductor
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS

Home Install Help Developers About

Search:

Bioschemas

Home » Bioconductor 3.10 » Software Packages » BridgeDbR (development version)

BridgeDbR

platforms all | rank 708 / 1728 | posts 0 | in Bioc 4.5 years
build ok | updated before release

DOI: 10.18129/B9.bioc.BridgeDbR
This is the **development** version of Bioconductor; to use it, please install the devel version of Bioconductor.

Documentation »

Bioconductor

- Package vignettes and manuals.
- Workflows for learning and use.
- Course and conference material.
- Videos.
- Community resources and tutorials.

Code for using BridgeDb identifier mapping framev within R

1 Introduction
2 Concepts
3 Mapping Identifiers
4 Metabolomics
References


```
▼<script type="application/ld+json">
    {
      "@context":"http://schema.org/",
      "@type":"CreativeWork",
      "about":"This tutorial describes how to use the BridgeDbR pa
    mapping.",
      "audience":[
        "http://edamontology.org/topic_3070",
        "http://edamontology.org/topic_3314"
      ],
      "genre":[
        "http://edamontology.org/topic_0605",
        "http://edamontology.org/operation_3282",
        "http://edamontology.org/data_1025",
        "http://edamontology.org/data_0982"
      ],
      "name":"BridgeDbR Tutorial",
      "author":[{
        "@type":"Person",
        "name":"Egon Willighagen",
        "identifier":"0000-0001-7542-0286"
      }],
      "difficultyLevel": "beginner",
      "keywords":"ELIXIR RIR, BridgeDb",
```


BridgeDbR Tutorial

Egon Willighagen

2 May 2019

Package

BridgeDbR 1.19.0

1 Introduction

BridgeDb is a combination of an application programming interface (API), library, and set of data files for mapping identifiers for identical objects [1]. Because BridgeDb is use by projects in bioinformatics, like WikiPathways [2] and PathVisio [3], identifier mapping databases are available for gene products and metabolites.

Questions can be directed to the BridgeDb Google Group.

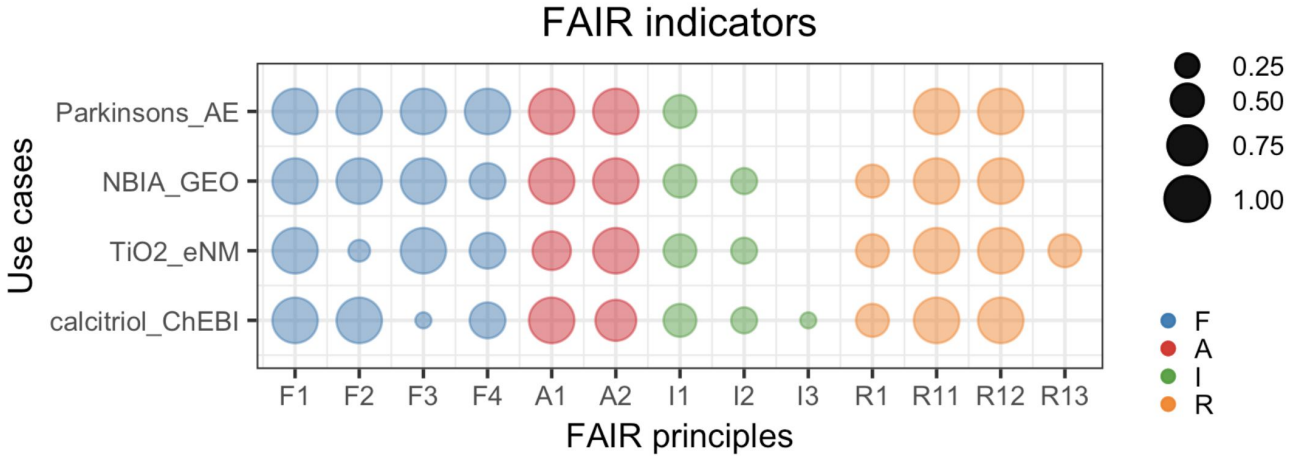
The Bioconductor BridgeDbR package page describes how to install the package. After installation, the library can be loaded with the following command:

```
library(BridgeDbR)
```

```
## Loading required package: rJava
```

Implementation

1. Research question via API Data ID and metadata
2. Calculation of metrics:
   - Findability in a searchable resource : Google Dataset Search
   - About repository (e.g. persistent identifier, license): re3data
   -
3. C



Not final results

Acknowledgments

BiGCaT


Dr. Friederike Ehrhart


Dr. Lars Eijssen


Dr. Egon Willighagen


Dr. Martina Summer-Kutmon


Dr. Susan Steinbusch-Coort

Postdoctoral researchers


Serena Bonaretti, PhD


Lauren Dupuis, PhD.

PhD students

Nasim B. Sangani, MSc.


Marvin Martens, MSc


Denise Slenter, MSc


Amadeo Muñoz García, MSc


Ryan Miller, MSc


Mirella Kalafati, Msc


Prof. dr. Chris Evelo

www.BiGCaT.unimaas.nl

BiGCaT department of bioinformatics

NUTRIM

Technical staff


Nuno Nunes

Funded projects

NanoCommons

EUTOXRISK

EuroDISH

OpenRiskNet RISK ASSESSMENT E-INFRASTRUCTURE

RISK GONE

eNanoMapper

NanoSolveIT

HORIZON 2020

Helis Academy

Maastricht University

FAIRplus

EUROPEAN JOINT PROGRAMME RARE DISEASES

https://www.bigcat.unimaas.nl/

https://tools.wmflabs.org/scholia/organization/Q19845644

Maastricht University

elixir Recommended Interoperability Resources

WIKIPATHWAYS Pathways for the People