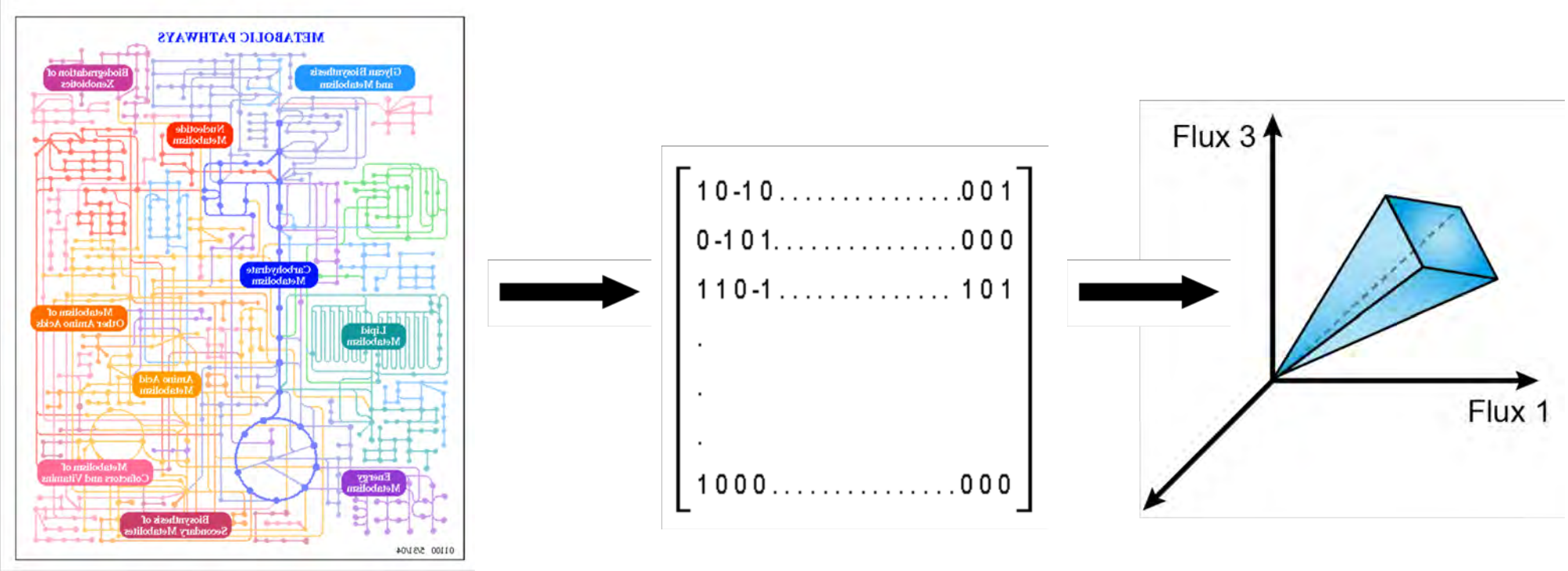
The RAVEN Toolbox 2.0: a versatile platform for metabolic network reconstruction and simulation

Hao Wang Ph.D.
hao.wang@chalmers.se

Division of Systems and Synthetic Biology
Department of Biology and Biological Engineering
Chalmers University of Technology, Sweden

Genome-scale models (GEMs) for Systems biology research



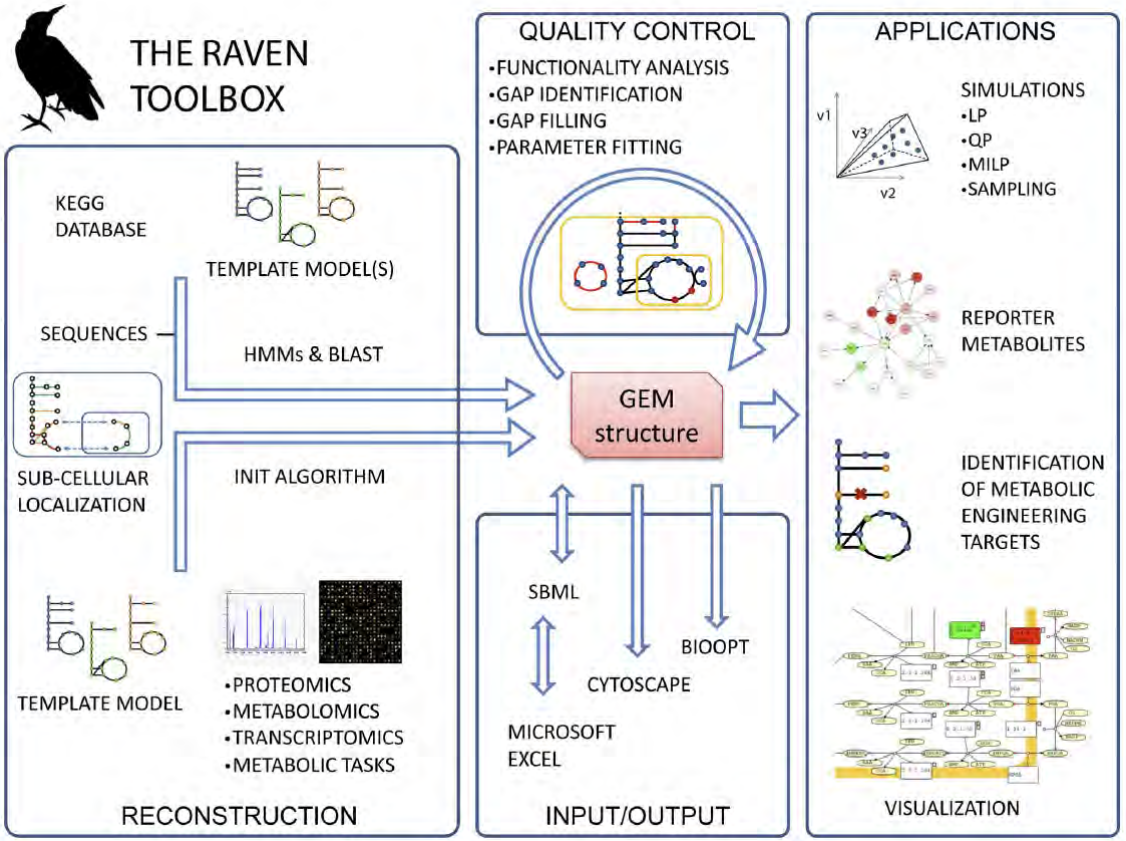
Genome-scale metabolic model reconstruction

Simulation of metabolic capabilities

Automated draft GEM reconstruction

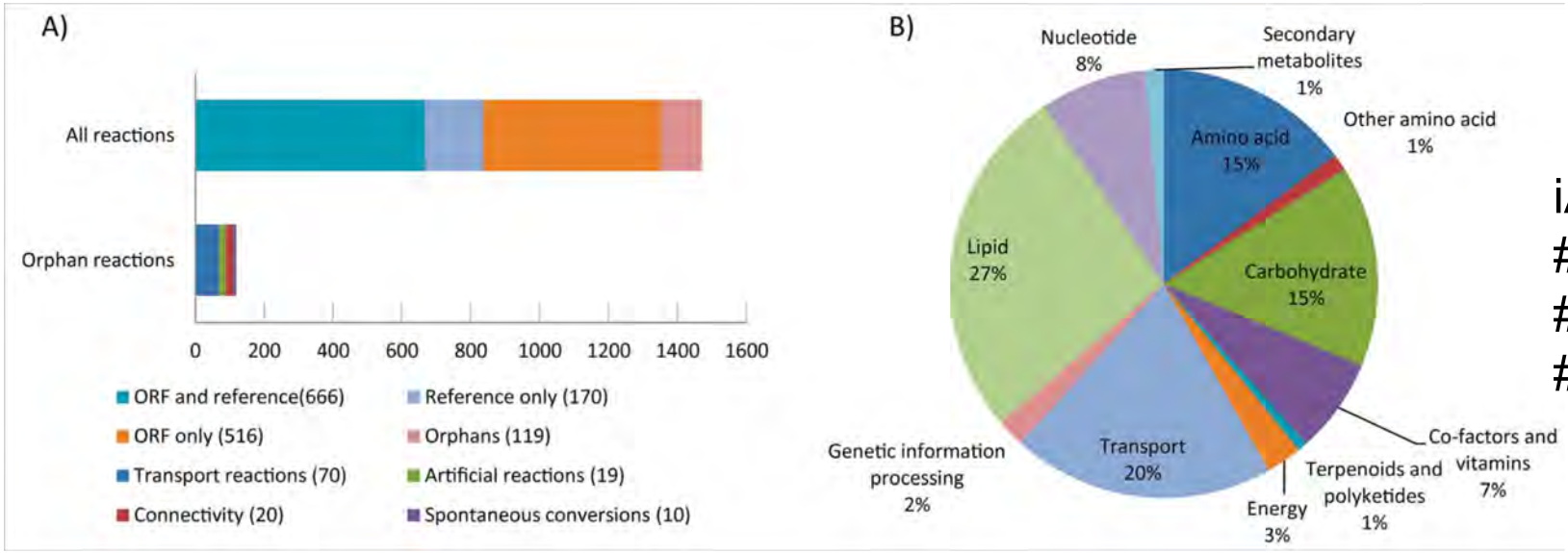
- Various packages/solutions to aid reconstruction
  - Kbase
  - CoReCo
  - ModelSEED
  - *merlin*

- **RAVEN** (Reconstruction, Analysis and Visualization of Metabolic Networks) Matlab-based Toolbox



GEM reconstruction for *Penicillium chrysogenum*

Features	***ANi***	***AO***	***AN***	***PC***
Genome size (Mb)	30.1	37.2	34.9	32.2
Number of chromosomes/supercontigs	8	8	8	49
Number of total protein sequences	10 560	12 074	11 197	12 811
Functional assignments				
Pairwise comparison	***ANi* and *PC***	***AO* and *PC***	***AN* and *PC***	***PC***
Number of protein sequence orthologues[a]	5749	5614	5632	-
Number of metabolic orthologues based on COG[b]	1316	1471	1313	2330
Number of metabolic orthologues based on GEMs[c]	576	915	563	1143

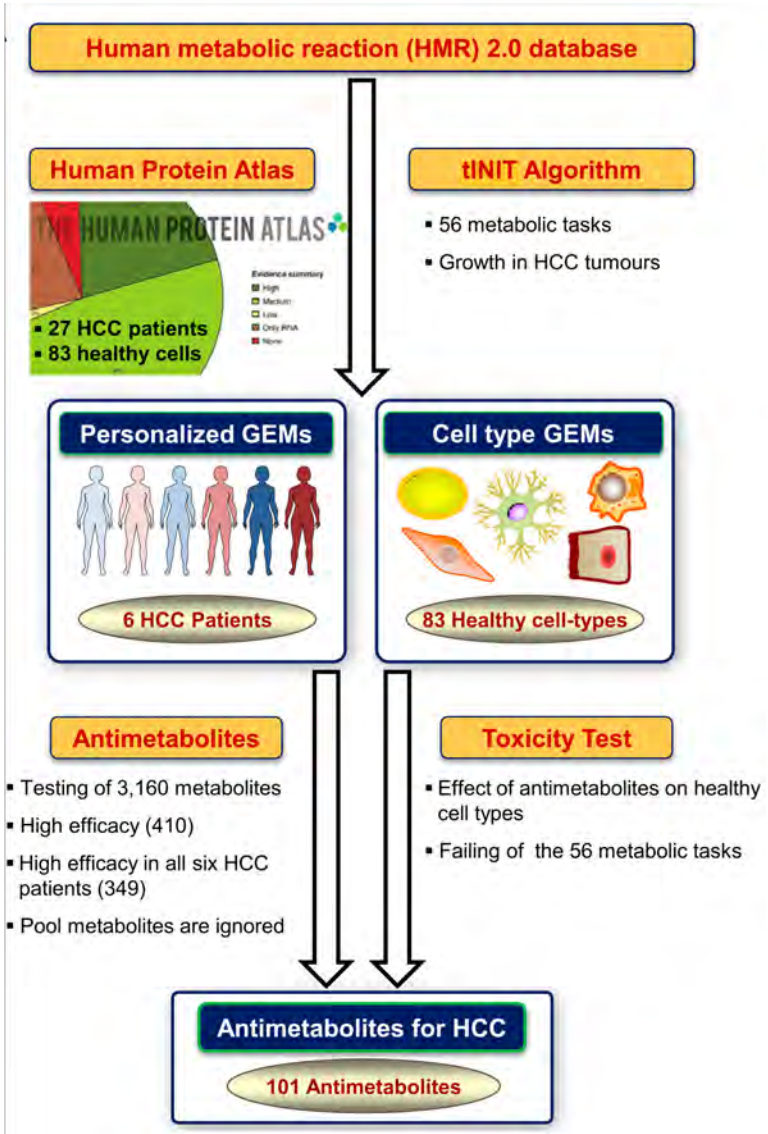


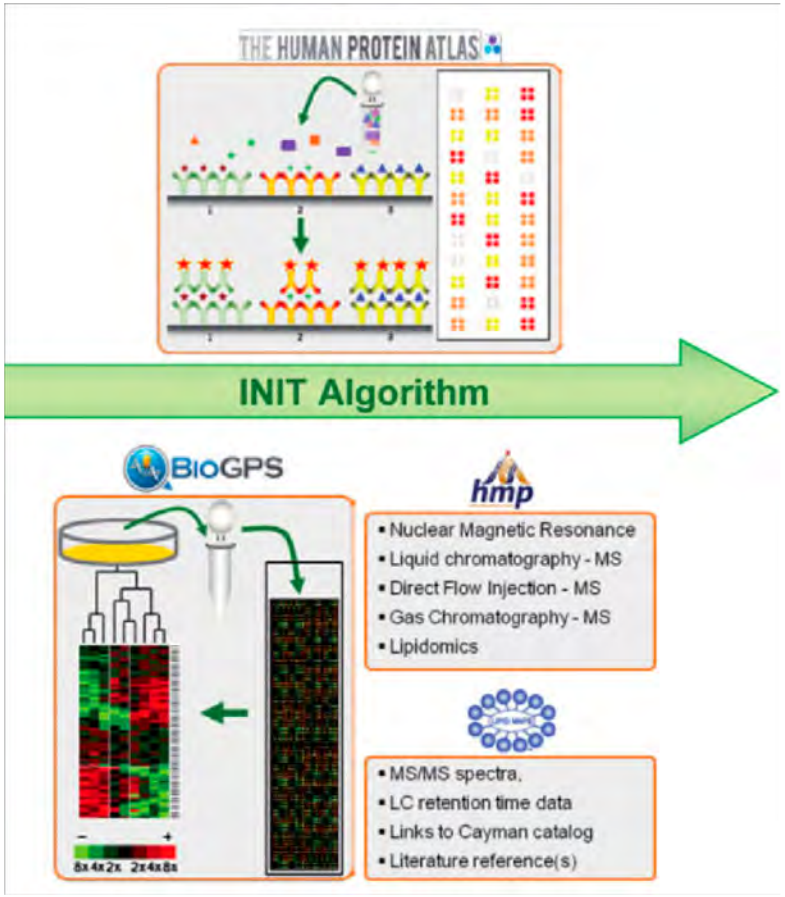
iAL1006
Rxn: 1471
Met: 1235
Gene: 1006

PLOS Comp Biol. 2013, 9:e1002980

Task-driven Model Reconstruction Algorithm (tINIT)

metabolicatlas.org





THE RAVEN TOOLBOX 2

RAVEN Toolbox ver 2.0

sys² bio www.sysbio.se

A wide range of biological pathway databases

Template models

MetaCyc

KEGG Kyoto Encyclopedia of Genes and Genomes

HMR database

HMMs & BLAST

tINIT algorithm

Import/Export

Matlab

RAVEN

COBRA

XML

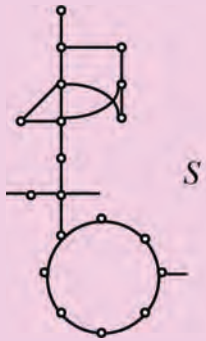
SBML FBCv2

SBML level2

YAML

Excel sheet

GEMs



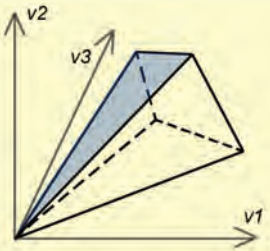
$$S = \begin{array}{c} \begin{matrix} r1 & r2 & \dots & ri \end{matrix} \\ \begin{pmatrix} 1 & 0 & \dots & 0 \\ -1 & 2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & \dots & 1 \end{pmatrix} \end{array} \begin{matrix} m1 \\ m2 \\ \\ mj \end{matrix}$$

$$S \cdot v = 0$$

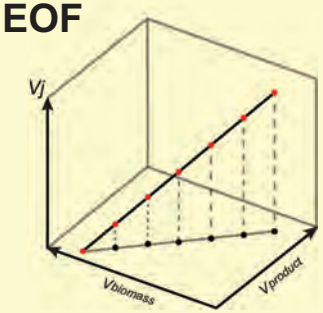
COBRA toolbox

Model simulation

FBA



FSEOF

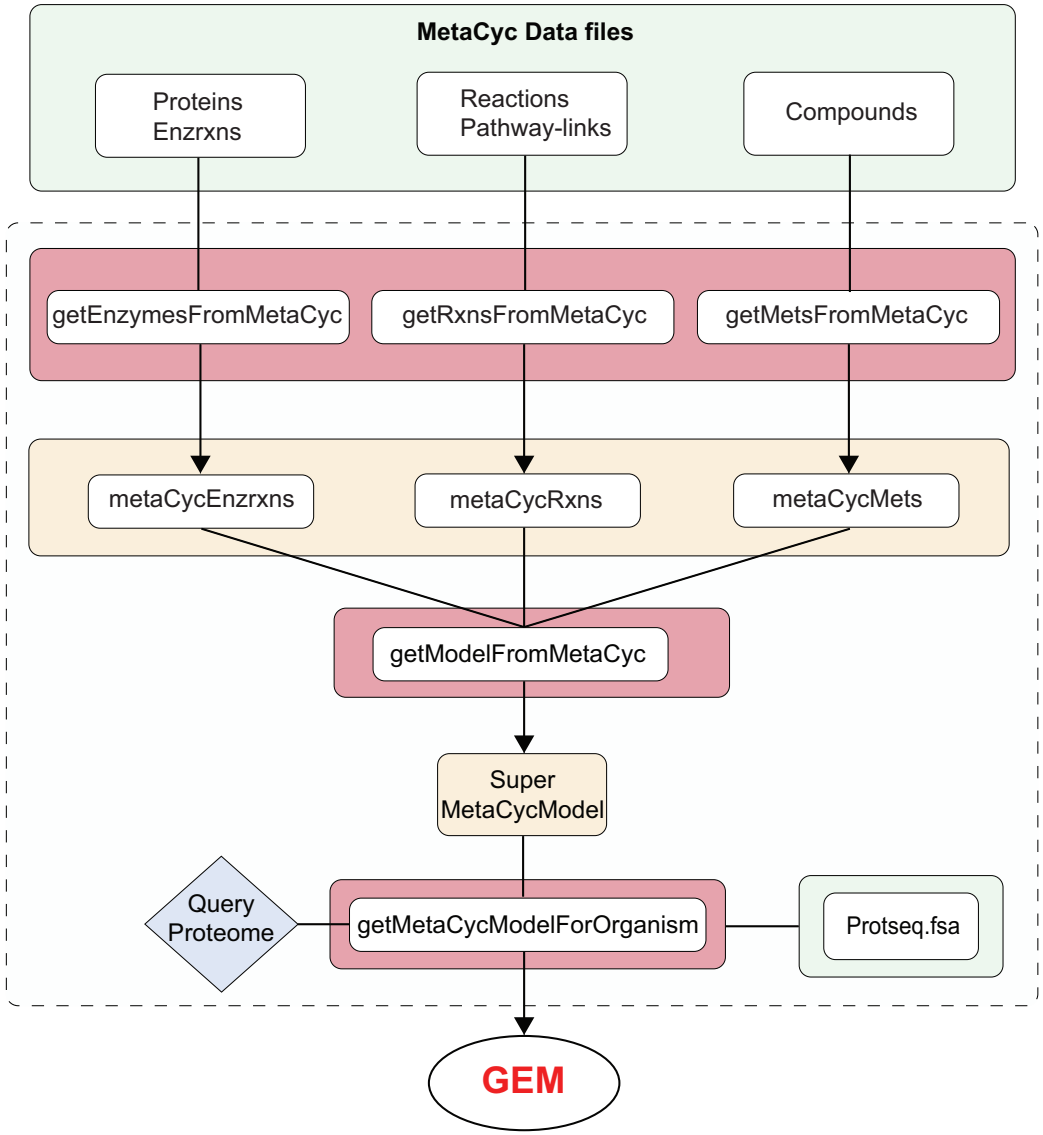


CHALMERS

Features of RAVEN Toolbox ver 2.0

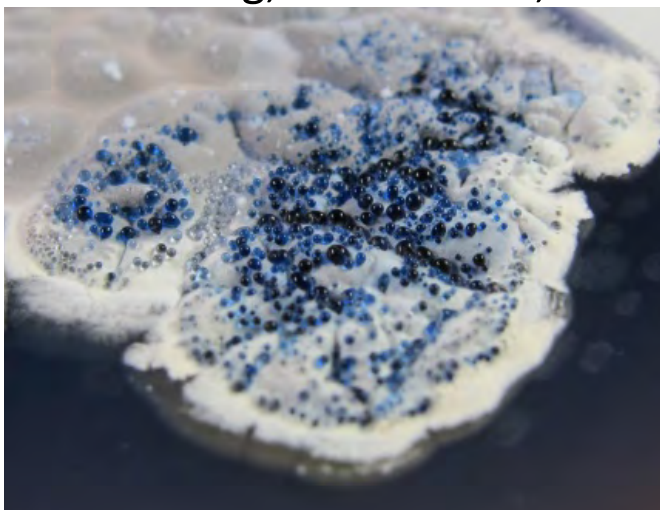
	MEMOSys	FAME	Microbes Flux	CoReCo	Pathway Tools	RAVEN 1.0	COBRA 2.0	Model SEED	merlin	RAVEN 2.0
Prokaryotic models										
Tissues/organ models										
Eukaryotic models										
Uses KEGG										
Uses MetaCyc										
Uses template model										
Mass and charge balance										
Assign sub-cellular localization										
Transporters annotation										
Spontaneoues reaction										
Simulation analysis										
Pathways visualisation										
Export to SBML										
Export to Excel										
OS independent										

MetaCyc reconstruction module in RAVEN 2.0

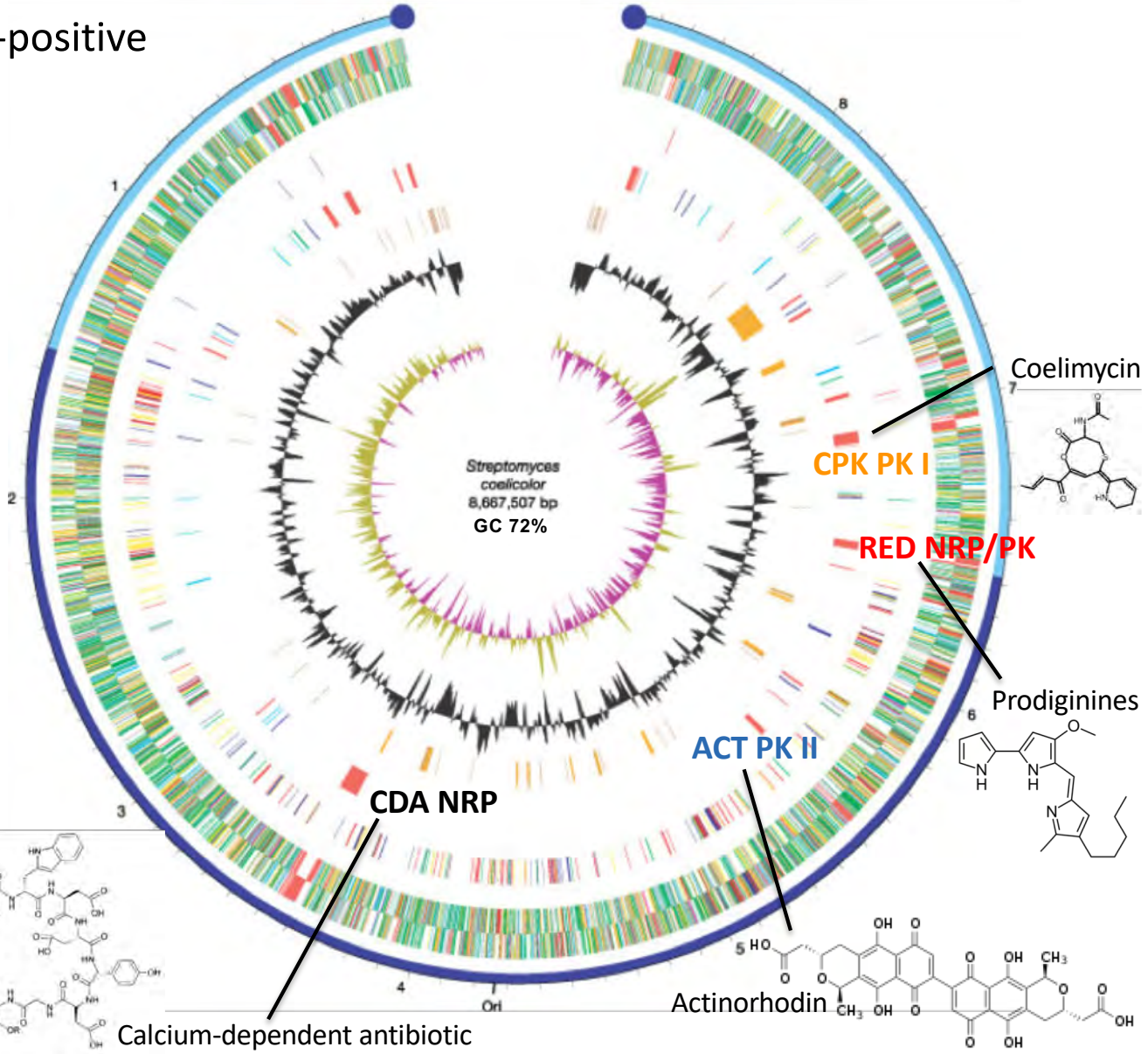


Streptomyces coelicolor

Soil-dwelling, filamentous, Gram-positive



- Model organism of Actinomycete for antibiotic production
- Produce a number of different types of compounds with antibiotic properties
- A good candidate of cell factory for novel antibiotic production



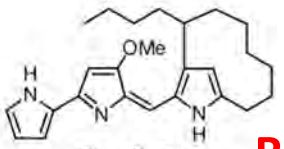
Nature. 2002 417:141-7.

GEMs of *Streptomyces coelicolor* A3(2)

Metabolic model	iIB711	iMA789	iMK1208
GENES INCLUDED	711	789	1208
REACTIONS	700	896	1643
Metabolic reactions	561	745	1443
Transport reactions	139	151	200
EXCHANGE REACTIONS	–	–	216
GENE-PROTEIN-REACTION ASSOCIATIONS			
Gene associated or spontaneous	495	669	1405
No gene associated	205	227	238
METABOLITES			
Unique metabolites	500	643	1246
Cytoplasmic	–	–	1236
Extracellular	–	–	200

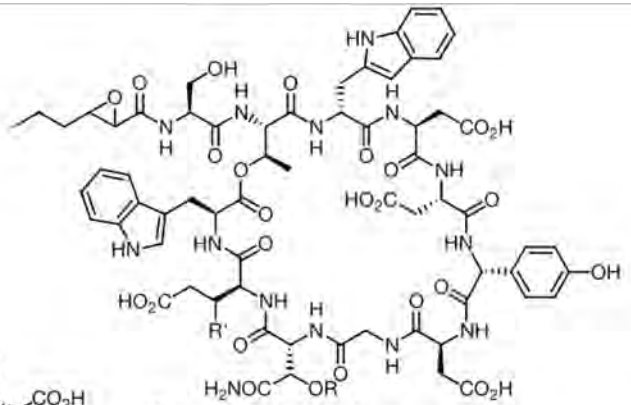
Biotechnol J. 2014, 9:1185-94

Secondary metabolites produced by *S. coelicolor*

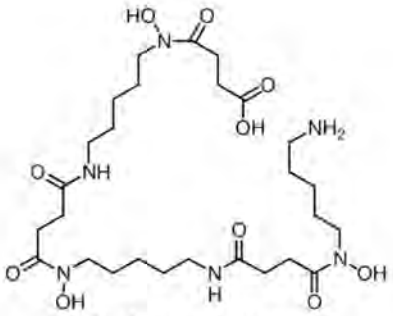


Prodiginines **RED**

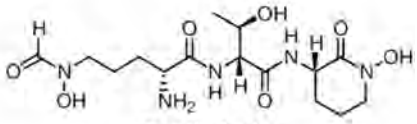
Actinorhodin **ACT**



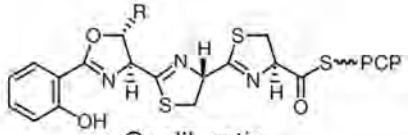
Calcium-dependent antibiotic **CDA**



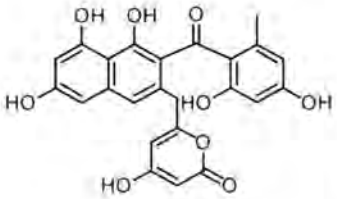
Desferrioxamines **SCO2782-2785**



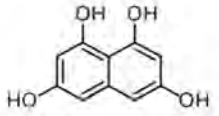
Coelichelin **SCO0489-0499**



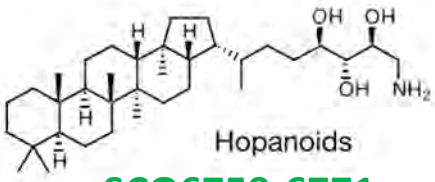
Coelibactin **SCO7681-7691**



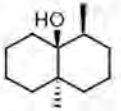
TW95a **SCO5314-5321**



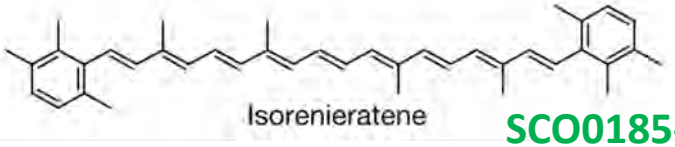
Tetrahydroxynaphthalene **SCO1206-1208**



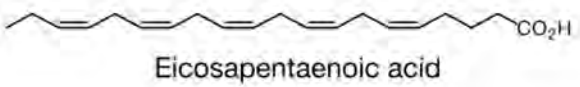
Hopanoids **SCO6759-6771**



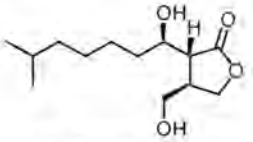
Geosmin **SCO6073**



Isorenieratene **SCO0185-0191**



Eicosapentaenoic acid

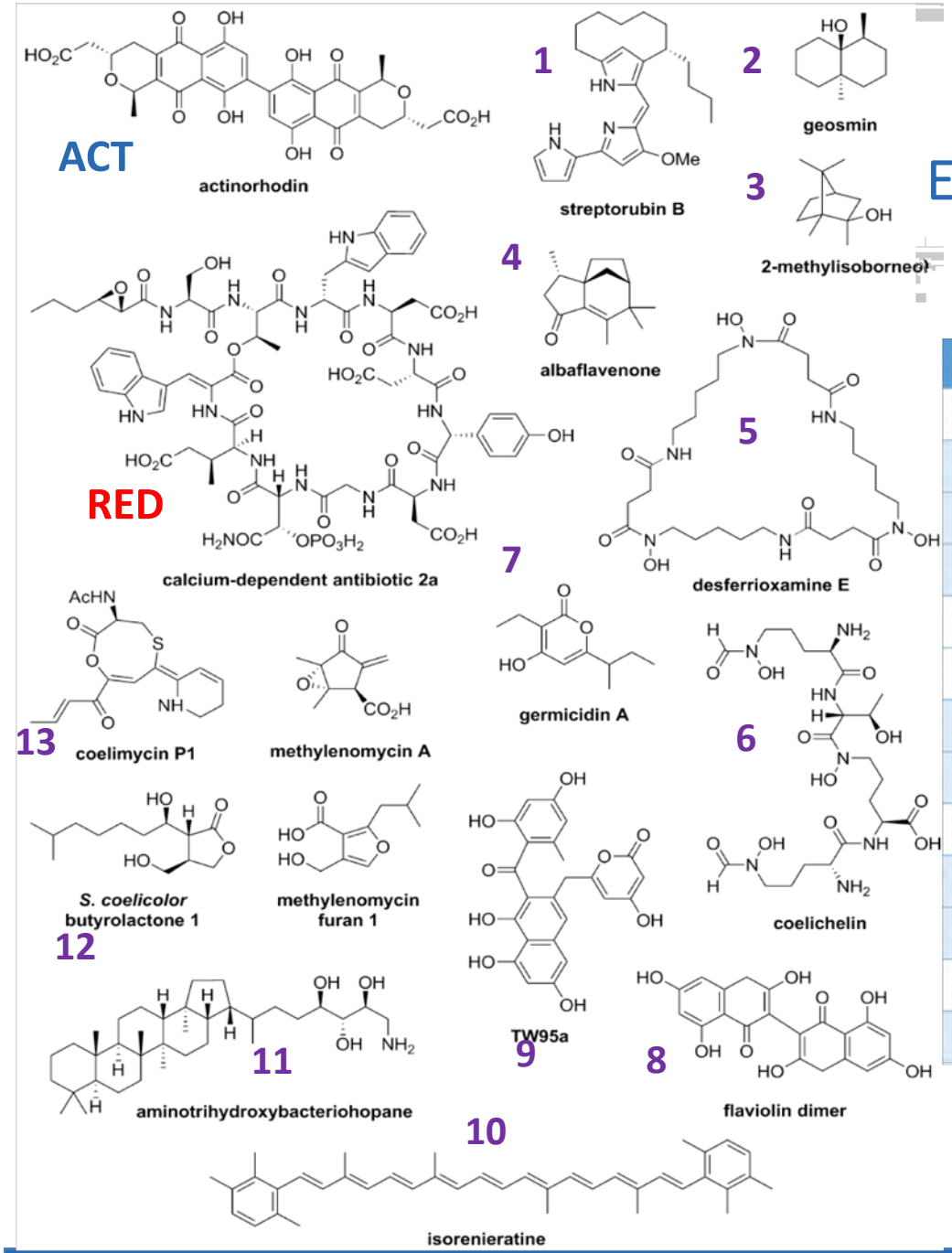


Butyrolactones **SCO6264-6267**

iIB711: 25 (12 + **8** + **3**)
iMA789: 33 (10 + **10** + **13**)
iMK1208: 83 (22 + **19** + **42**)

Nature. 2002 417:141-7.

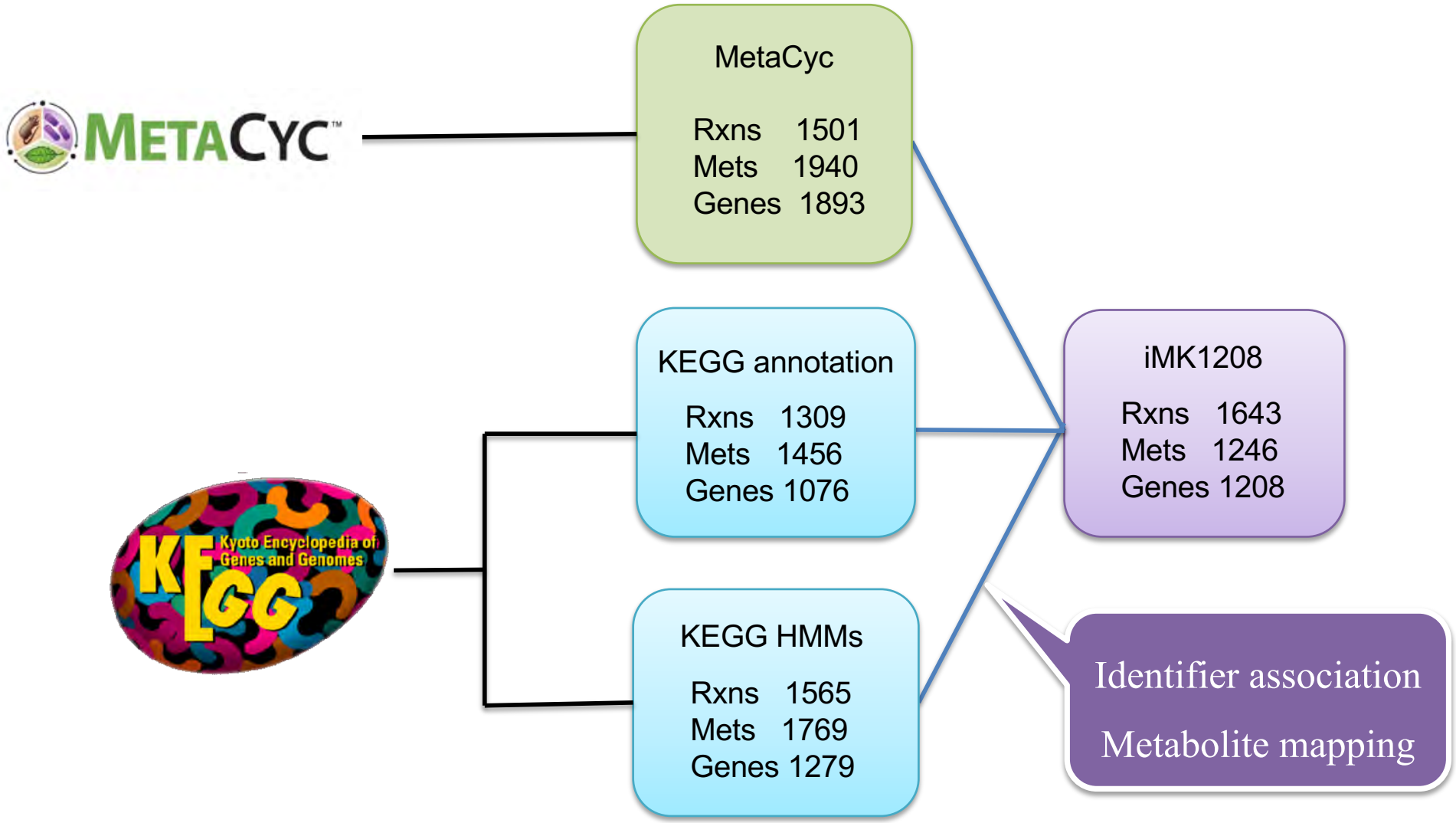
Experimentally elucidated secondary metabolic pathways



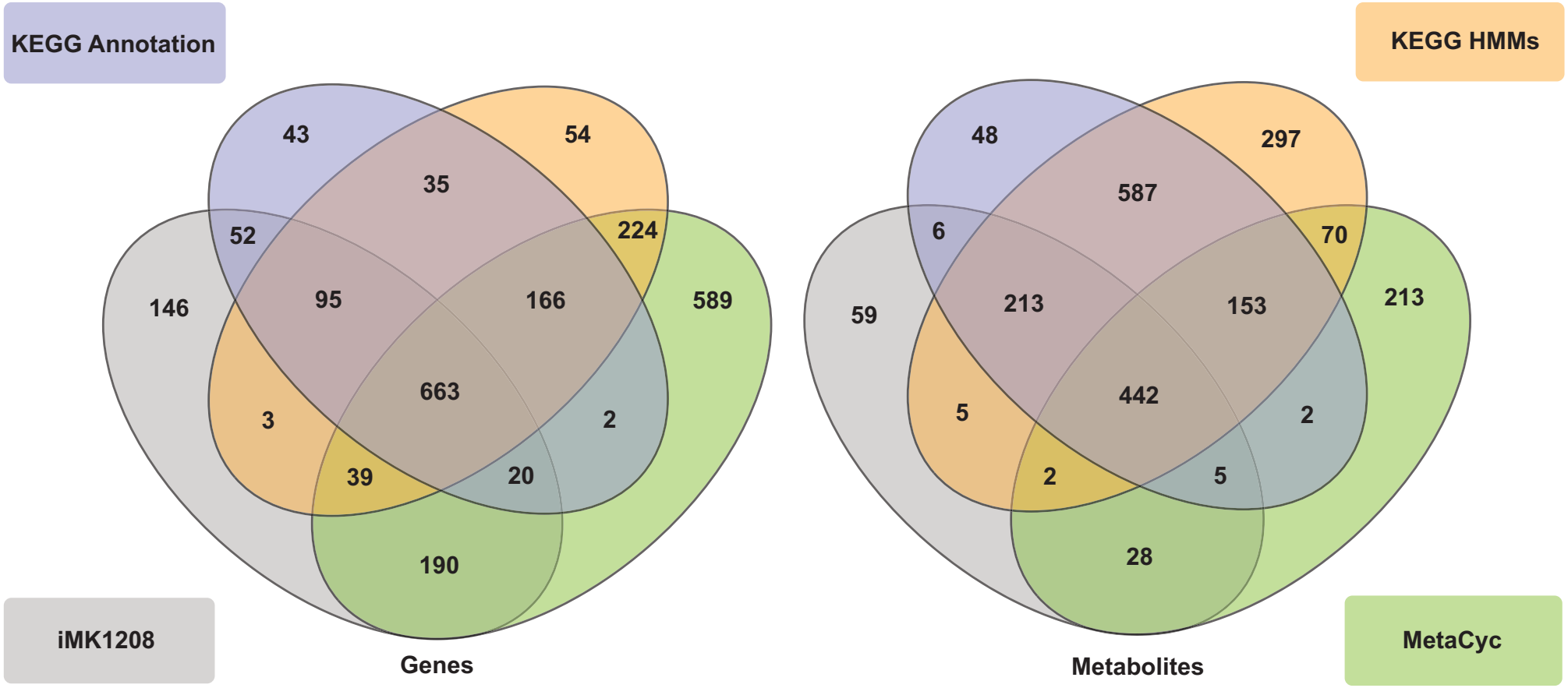
	Name	# met	# rxn	# gene
1	Streptorubin B	1	1	1
2	Geosmin	4	3	1
3	2-Methylisoborneol	2	2	2
4	Albaflavenone	3	3	2
5	Desferrioxamine E	4	5	4
6	Coelichelin	4	3	3
7	Germicidin A	3	3	1
8	THN/flaviolin	5	5	4
9	TW95a	4	3	7
10	Isoprenoid	12	10	4
11	Hopanoid	9	9	7
12	Butyrolactone	5	3	2
13	Coelimycin P1	6	6	8

J Ind Microbiol Biotechnol 2014, 41:219–32

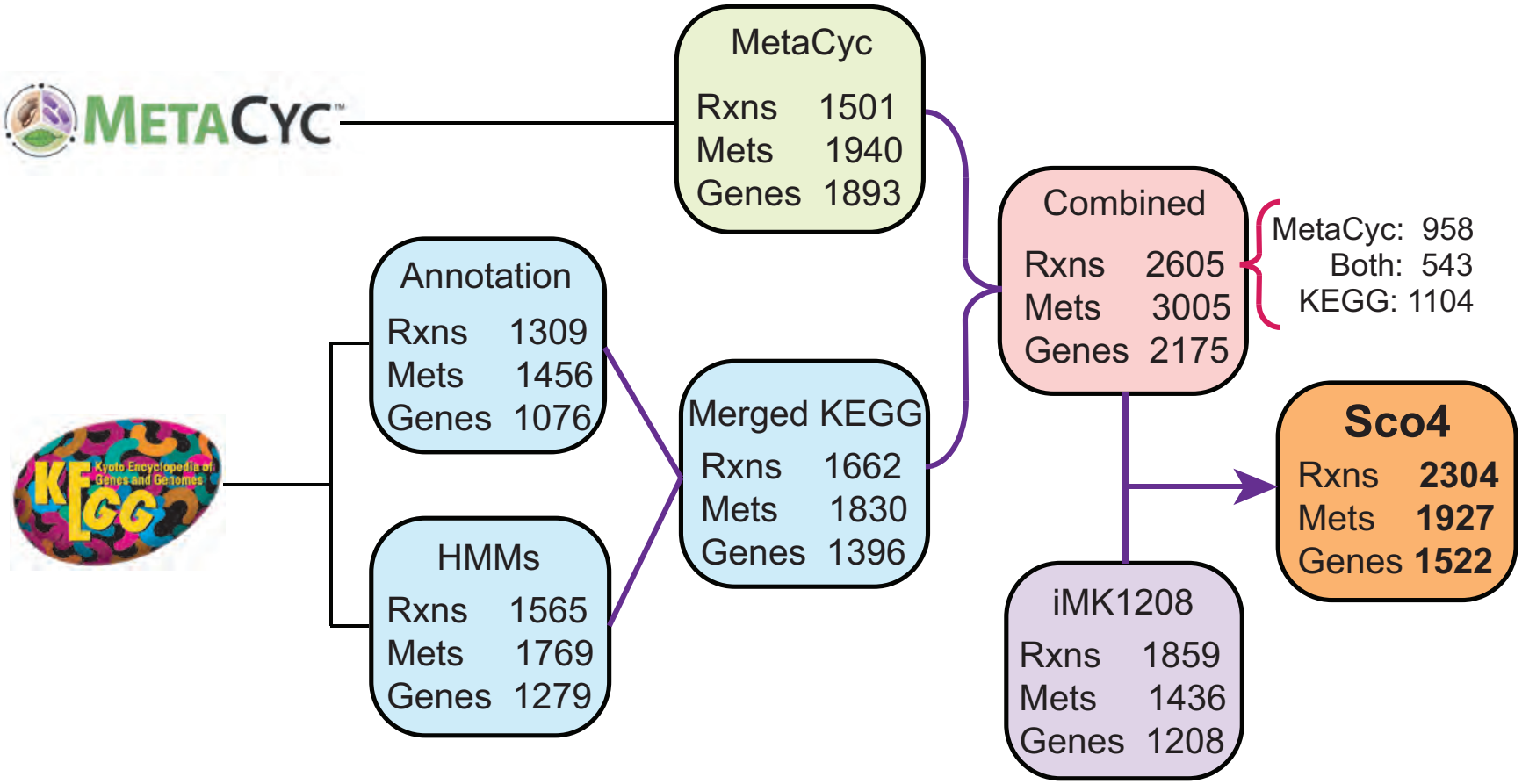
Draft GEM reconstruction for *S. coelicolor* by RAVEN 2.0



Model comparison



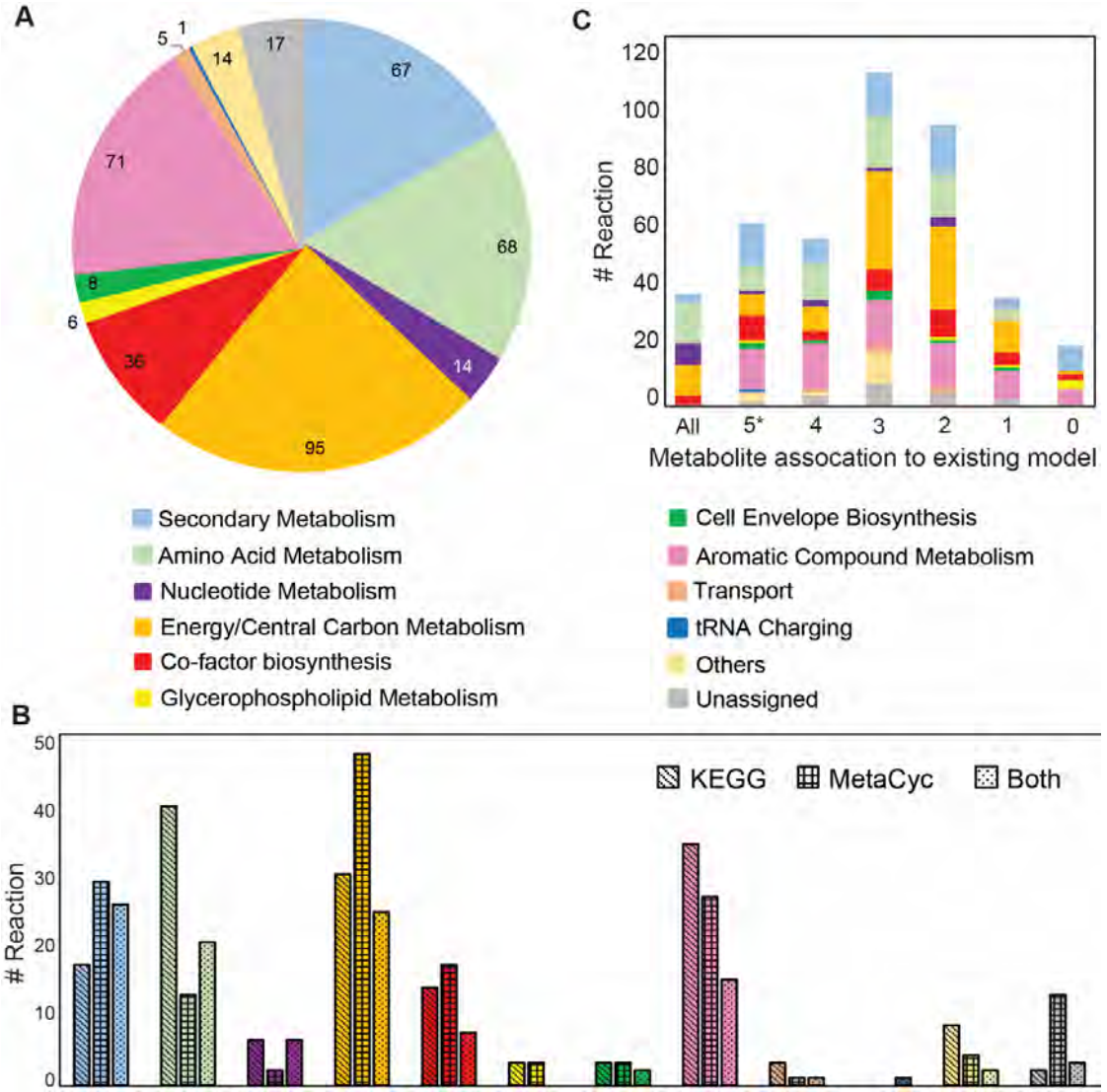
Sco4 model generation by RAVEN 2.0



Gapfilling by RAVEN 2.0

iMK1208			*Combined model*		
SubSystem	**Rxn ID**	**NOTE**	**Gene assocation**	**MetaCyc ID**	**KEGG ID**
Alternate Carbon Metabolism	RMI	Required for growth using L-rhamnose	SCO0812	RHAMNISOM-RXN	R02437
Alternate Carbon Metabolism	RMPA	Required for growth using L-rhamnose	SCO0813	RHAMNULPALDOL-RXN	R02263
Alternate Carbon Metabolism	DABTD	Required for growth using D-arabitol as	SCO1901	RXN-7971	R05604
Arginine and Proline Metabolism	PROD2	Required for growth using L-proline as	SCO5519	R10507	R10507
Cofactor and Prosthetic Group Biosynthesis	THRPD	Required for growth (adenosylcobalam	SCO1859	4.1.1.81-RXN	R06530
Cofactor and Prosthetic Group Biosynthesis	ADCL	Required for growth (tetrahydrofolate	SCO1546	ADCLY-RXN	R05553
Cofactor and Prosthetic Group Biosynthesis	CBLAT	To utilize extracellular cob(1)alamin	SCO1851 or SCO5381	COBALADENOSYLTRANS-RXN	R01492
Cysteine Metabolism	BPNT	Required for growth (acyl-carrier prote	SCO5161	325-BISPHOSPHATE-NUCLEOTIDASE-RXN	R00188
Nucleotide Salvage Pathway	DTMPK	Required for growth (dTTP biosynthesis	SCO3542	DTMPKI-RXN	R02094
Threonine and Lysine Metabolism	OXPTNDH	Required for growth using L-lysine as a car	SCO1204 or SCO1612	GLUTARATE-SEMIALDEHYDE-DEHYDROGE	R02401
tRNA Charging	GLNTRS	tRNA system	SCO5547	GLURS-RXN	R05578

Compositional distribution of the newly included reactions from RAVEN 2.0



Stoichiometric network curation

Identifiers in iMK1208			*Manual curation*		*Updated*
Metabolite	**KEGG**	**Reaction**	**MetaCyc reaction id**	**KEGG reaction id**	**KEGG compound id**
dtdp4d6dg	C00687	TDPDRE	DTDPDEHYDRHAMEPIM-RXN	R06514	C11907
		TDPGDH	DTDPGLUCDEHYDRAT-RXN	R06513	
r1p	C00442	PPM	PPENTOMUT-RXN	R01057	C00620
		PUNP1	ADENPHOSPHOR-RXN	R01561	
		PUNP3	RXN0-5199	R02147	
		PUNP5	INOPHOSPHOR-RXN	R01863	
		PUNP7	XANTHOSINEPHOSPHORY-RX	R02297	
f1p	C02976	FRUK	1PFRUCTPHOSN-RXN	R02071	C01094
		FBA2	RXN-8631	R02568	
		FRUpts	R03232	R03232	
2hymeph	C06191	6PHBG	RXN0-5297	R05134	C02323
		DM_2HYMEPH	*Assumed exchange reaction*		

List of updated metabolite identifiers

Invalid KEGG	*Updated KEGG*	*MetaCycID*
iMK1208	*Sco4*	*MetaCycID*
C00661	C00118	GAP
C00507	C02338	CPD0-1112
C00311	C00451	THREO-DS-ISO-CITRATE
C01602	C00077	L-ORNITHINE
C00348	C17556	CPD-9646
C00023	C14818	FE+2
C00725	C16241	LIPOIC-ACID
C06717	C00051	GLUTATHIONE
C01551	C02350	S-ALLANTOIN
C03340	C20258	CPD-14443
C01188	C06001	CPD-12175
C00034	C19610	MN+2
C00291	C19609	NI+2
C07597		CPD-11957

Spontaneous reaction

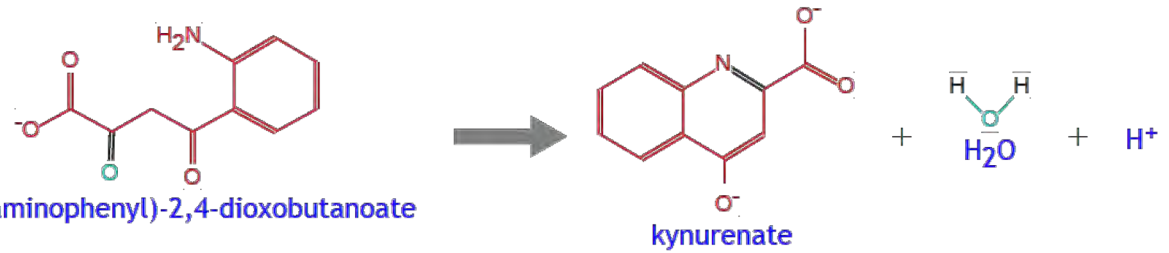
log in to add this to a SmartTable.

MetaCyc **Reaction: [no EC number assigned]**

Superclasses: Reactions Classified By Conversion Type → Simple Reactions → Chemical Reactions
Reactions Classified By Substrate → Small-Molecule Reactions

In Pathway: L-tryptophan degradation XI (mammalian, via kynurenine)

Show Atom Mapping: Coloring? Atom Numbering?



4-(2-aminophenyl)-2,4-dioxobutanoate → kynurenate + H_2O + H^+

The reaction direction shown, that is, A + B ↔ C + D versus C + D ↔ A + B, is in accordance with the direction in which it was curated.

Most BioCyc compounds have been protonated to a reference pH value of 7.3, and some reactions have been computationally balanced for hydrogen by adding free protons. Please see the PGDB Concepts Guide for more information.

This reaction can occur spontaneously -- no enzyme is required.

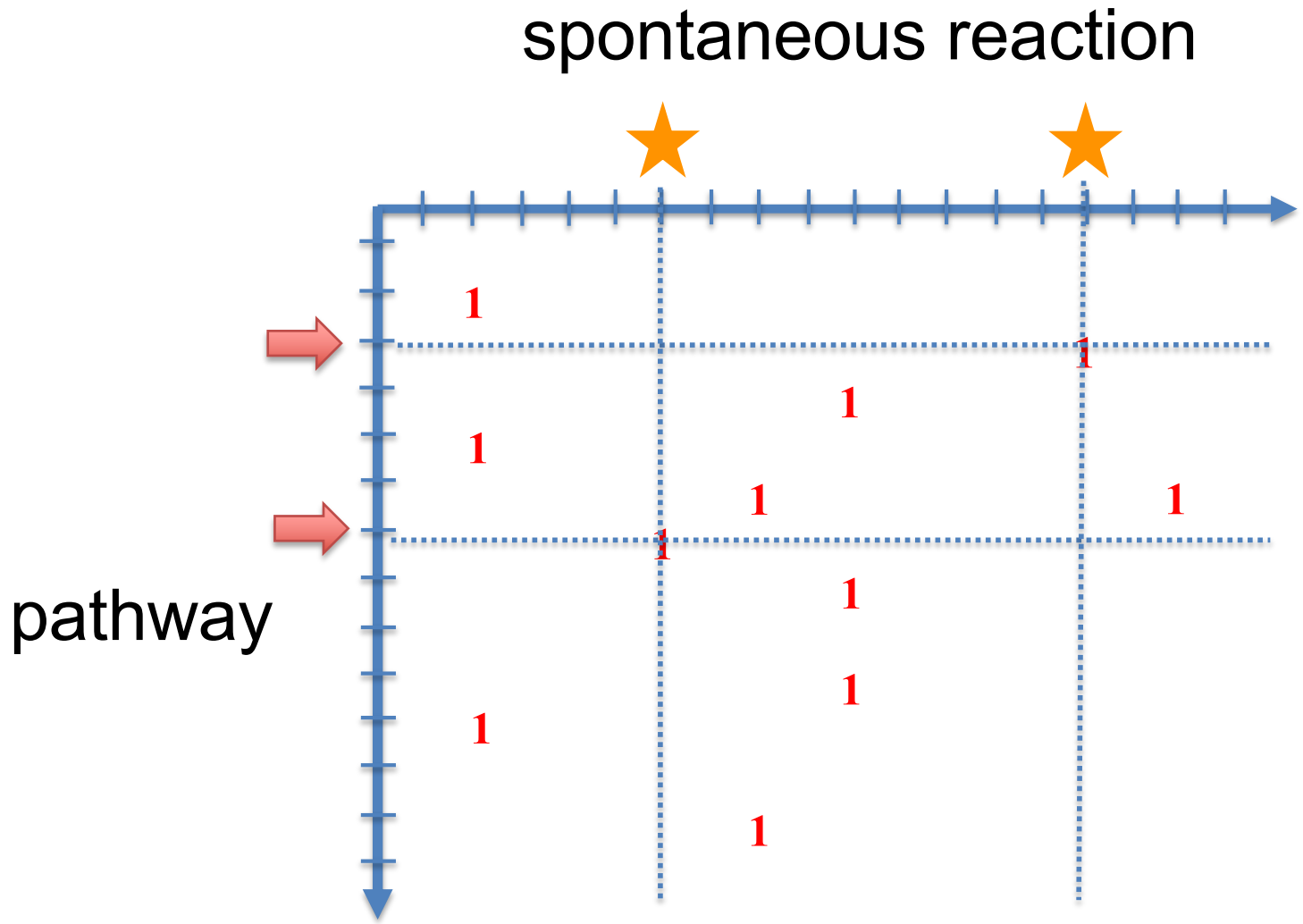
Mass balance status: Balanced.

Standard Gibbs Free Energy ($\Delta_r G'^{\circ}$ in kcal/mol): -38.78524 [Latendresse13]

Summary:
From the comments for EC 2.6.1.7: The product 4-(2-aminophenyl)-2,4-dioxobutanoate is converted into kynurenate by a spontaneous reaction.

- No gene-annotated, still present!

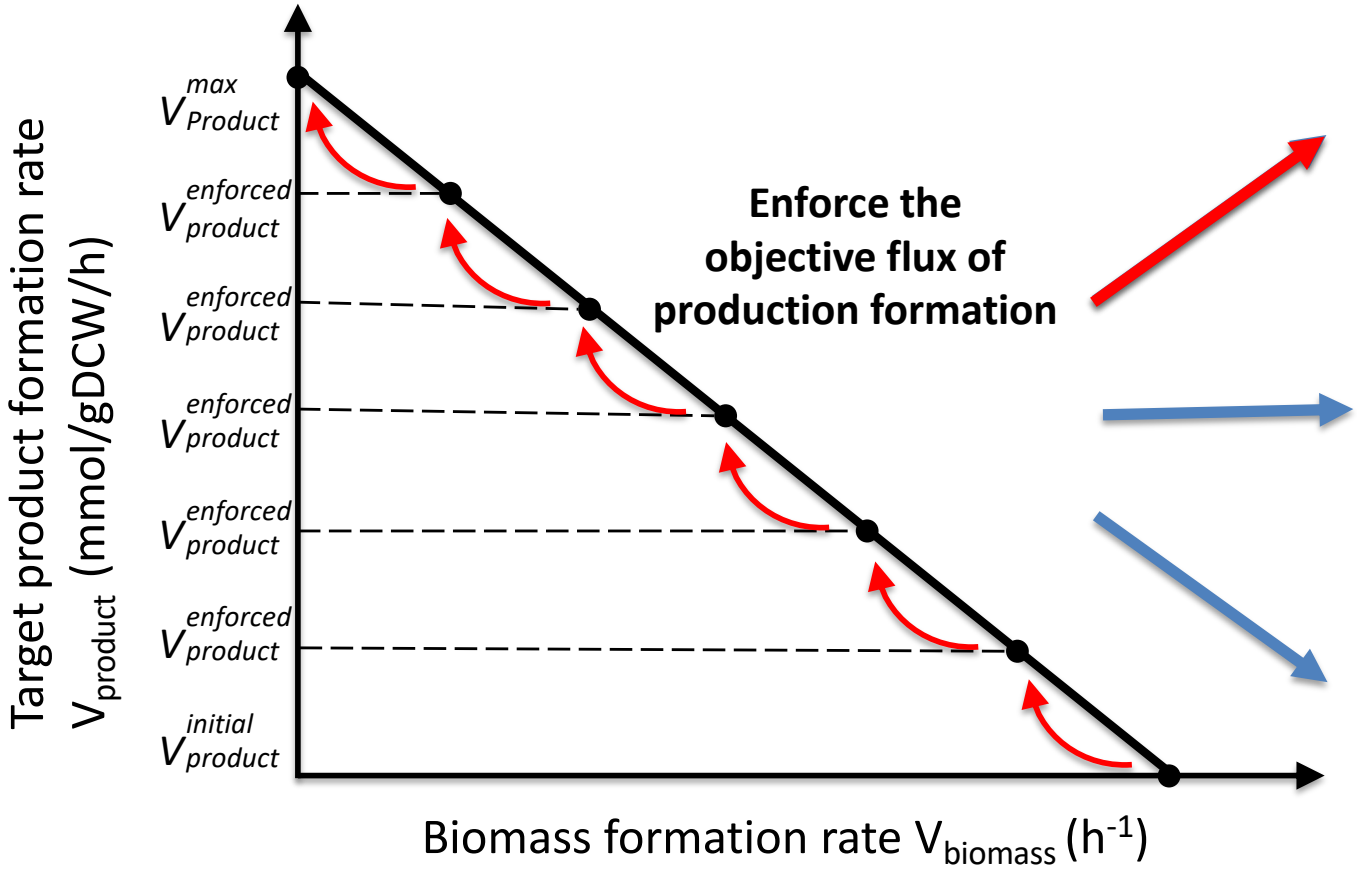
retrieveSpontaneous

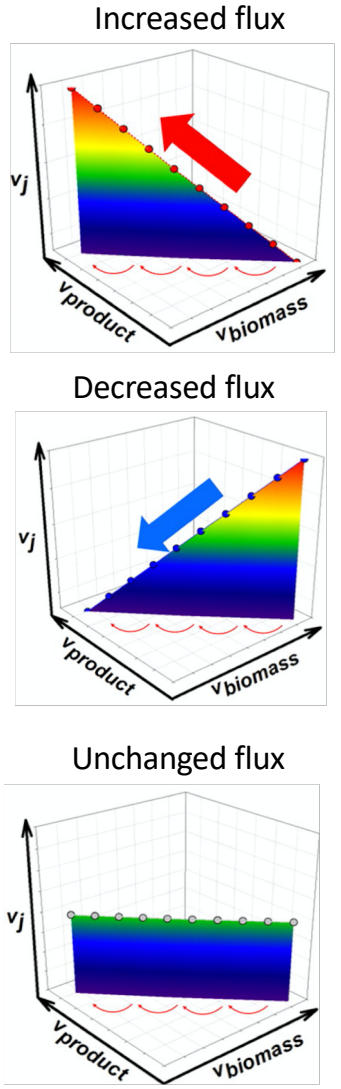


A total of 34 spontaneous Reactions were retrieved into *Sco4*

- Lists of MetaCyc reaction and metabolite identifiers as query

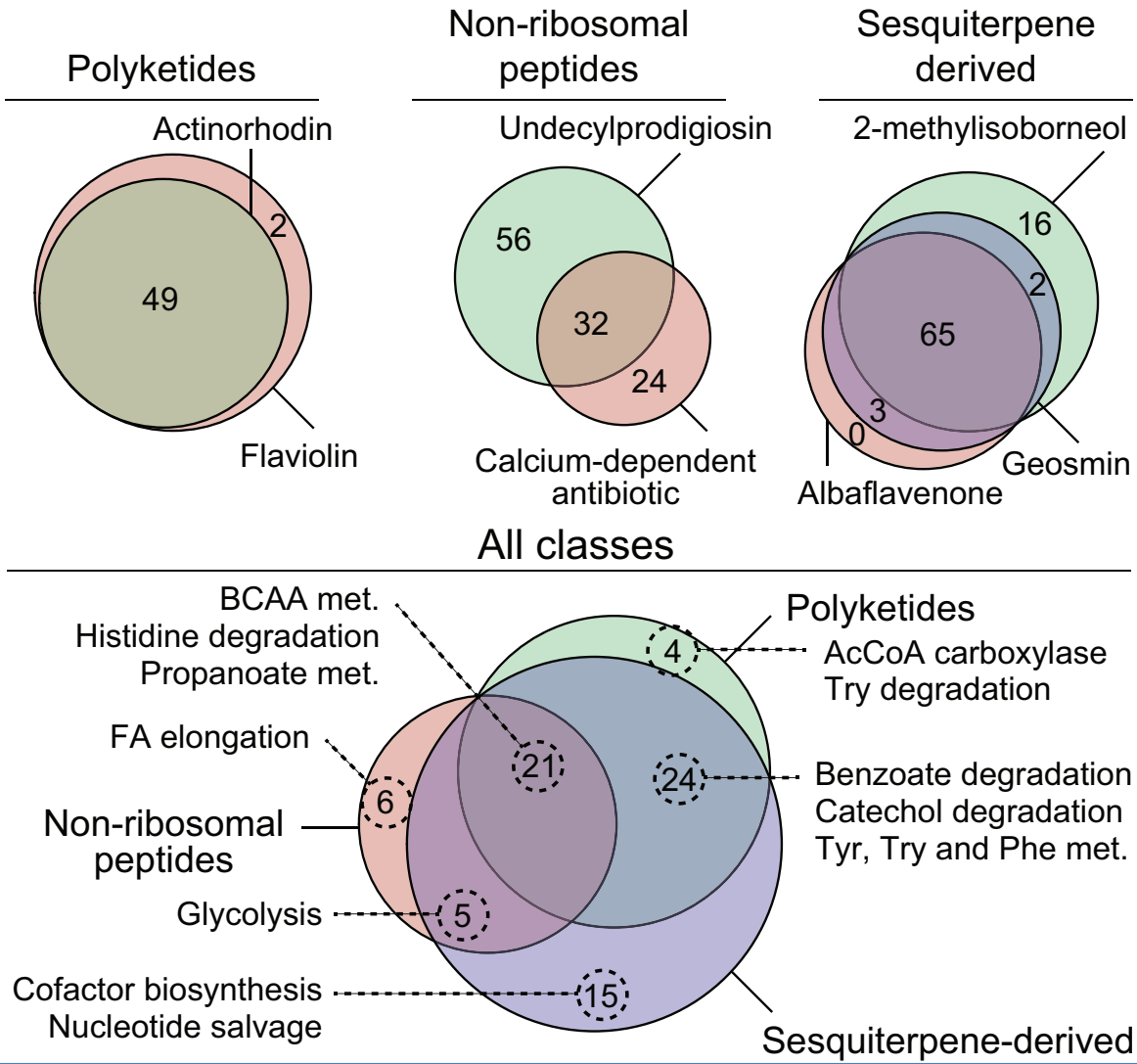
Flux Scanning based on Enforced Objective Flux (FSEOF)





A RAVEN function implementing the FSEOF algorithm has been added

Over-expression targets predicted by FSEOF using *Sco4*



Gene essentiality analysis

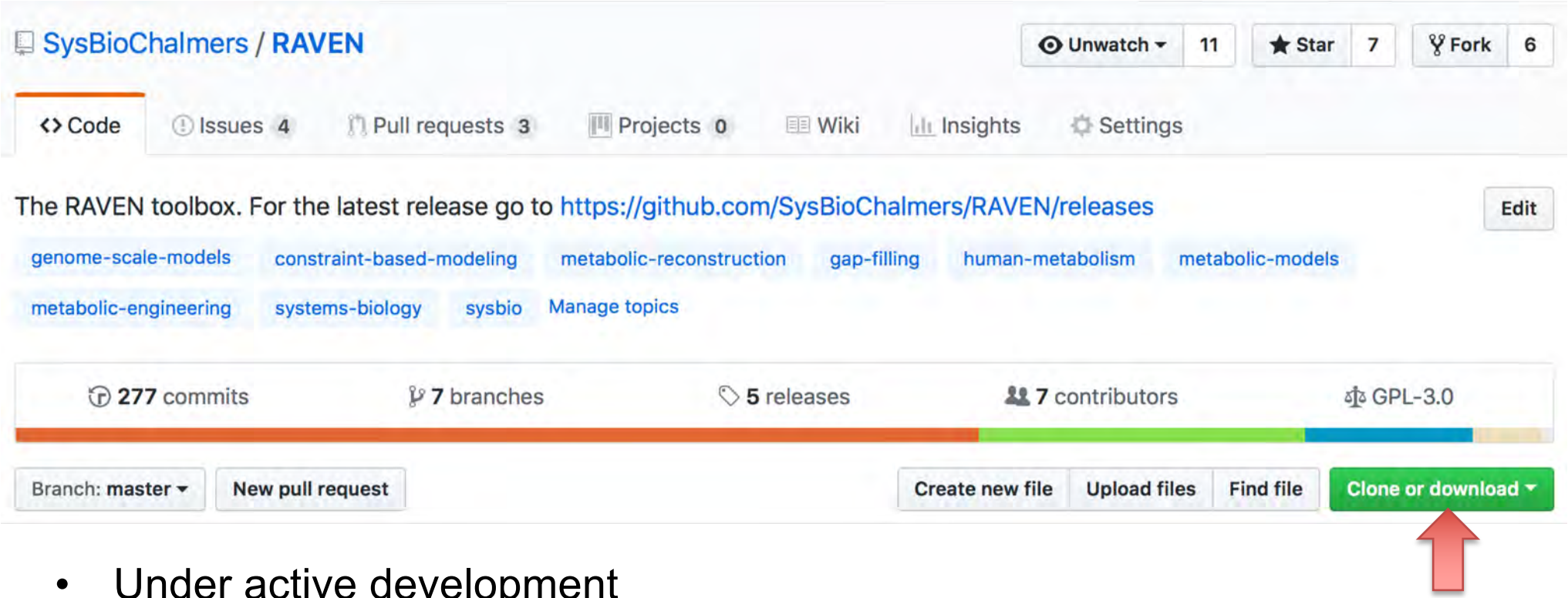
- ***53,900*** mutants identified by global mutagenesis study on *S. coelicolor*

- ***79*** large insertion gaps (>1.9 kb), amount to 203 kb with ***132*** genes. They reflect either cold regions of transposition insertion suggesting putative essential genes

		iMK1208	Sco4
	Reactions	1859	2299
	Metabolites	1436	1924
	Genes	1209	1511
Essentiality prediction	Sensitivity	0.372	0.338
	Specificity	0.880	0.905
	Accuracy	0.855	0.880

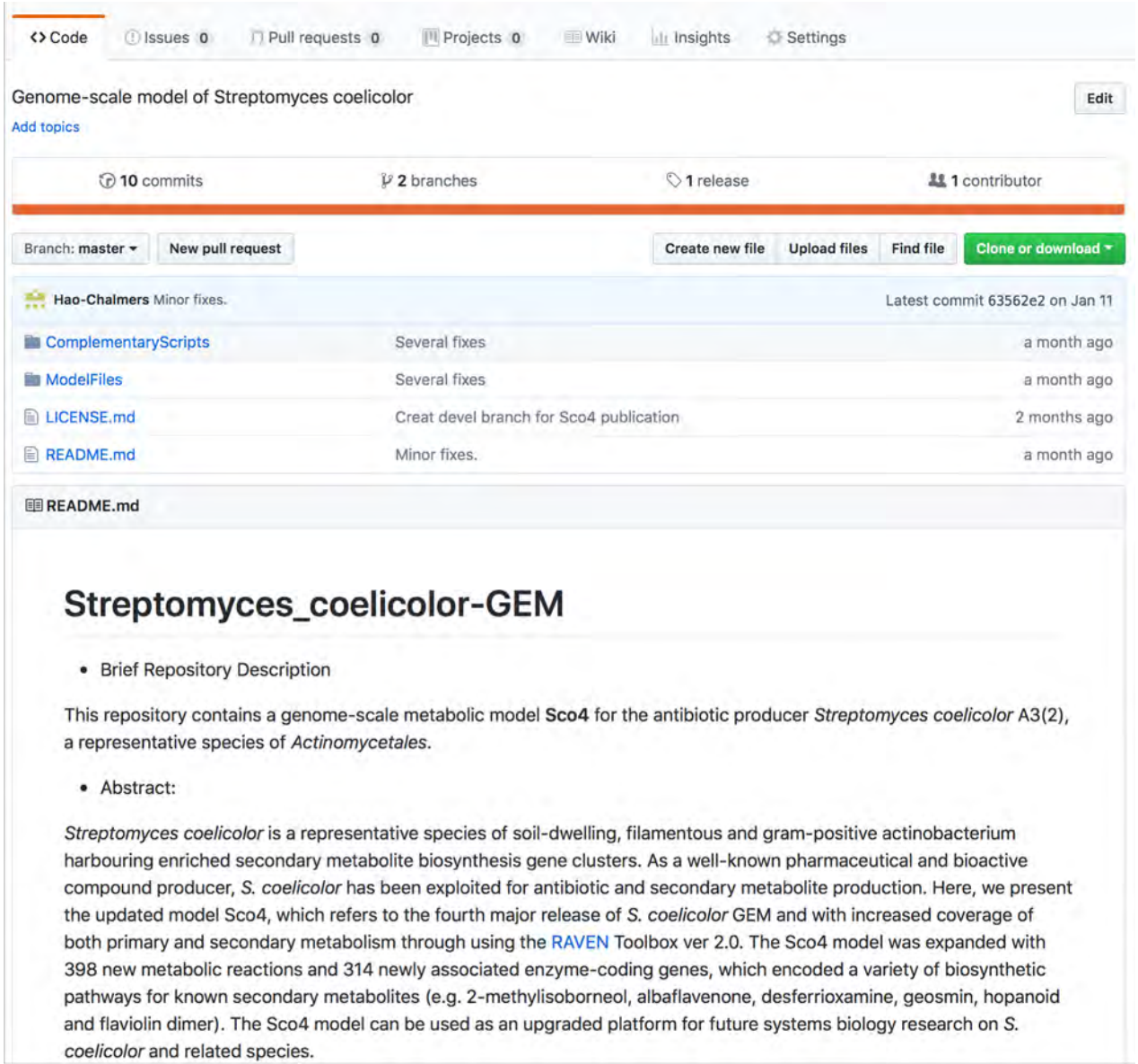
RAVEN on GitHub

1. https://github.com/SysBioChalmers/RAVEN
2. Clone or Download



- Under active development
- Detail help message availabe for all functions
- Comments and bug reports are highly welcome

GEMs on GitHub



https://github.com/SysBioChalmers/Streptomyces_coelicolor-GEM

Summary

- RAVEN Matlab Toolbox on GitHub
- Combines multiple data sources (MetaCyc, KEGG)
- Mass and charge balance
- An efficient platform for high quality draft model reconstruction
- Transport and spontaneous reactions
- A wide range of flux analysis functions
- Improved compatibility (Import/Export, COBRA)
- Useful in updating existing GEMs

Acknowledgements

Sysbio:
Eduard Kerkhoven
Simonas Marcišauskas
Daniel Hermansson
Sylvain Prigent
Rasmus Ågren
Jens Nielsen

SRI:
Ron Caspi
Peter Karp

ERASysAPP ERA-Net for Applied Systems Biology
VÄSTRA GÖTALANDSREGIONEN REGIONUTVECKLINGSNÄMNDEN
Knut and Alice Wallenberg Foundation