MrBac: A web server for draft metabolic network reconstructions for bacteria

Yu-Chieh Liao, Ph.D.
Assistant Investigator
National Health Research Institutes

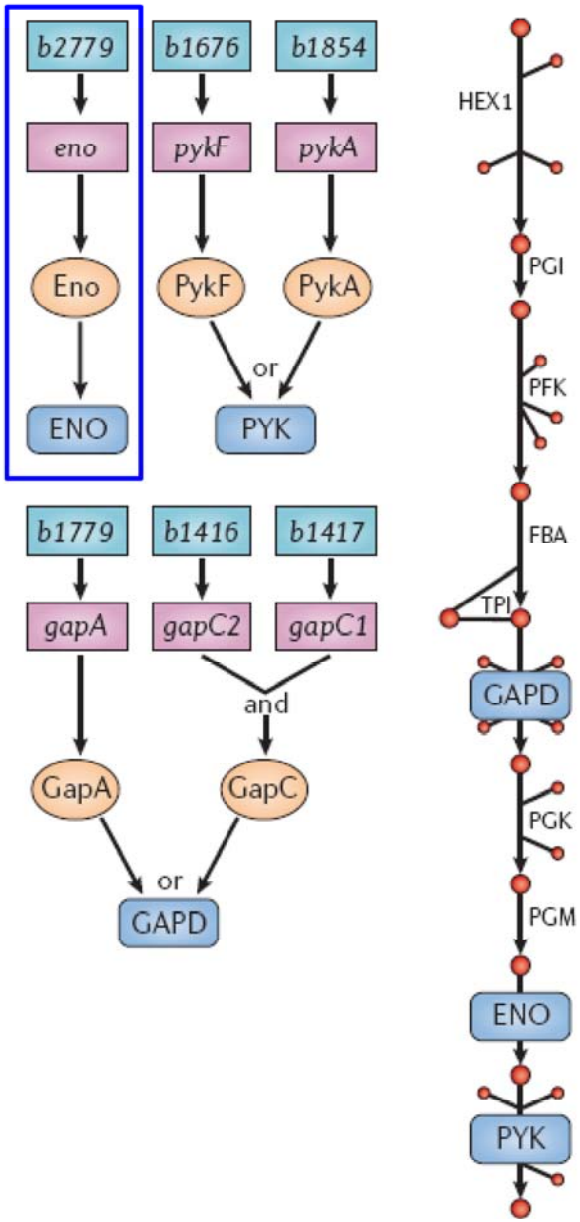
http://sb.nhri.org.tw/MrBac

Box 3 | **Assembly and representation**

A list of charge and elementally balanced metabolic reactions can be represented in a stoichiometric matrix (S), where rows and columns correspond to metabolites and reactions and the elements are the stoichiometric coefficients. In genome-scale metabolic networks these stoichiometric matrices contain few non-zero elements, as relatively few metabolites participate in a given reaction. Connections between genes and reactions can be represented as gene–protein–reaction (GPR) associations by using Boolean rules or visualized using graphic images. In the GPR scheme, the first level (teal) corresponds to genetic loci, the second level (pink) to transcripts, the third level (orange) to functional proteins, and the fourth level (blue) to reactions. [c], cytoplasmic reactions.

Abbreviation	Glycolytic reactions	Genes
HEX1	[c]GLC + ATP → G6P + ADP + H	*glk*
PGI	[c]G6P ↔ F6P	*pgi*
PFK	[c]ATP + F6P → ADP + FDP + H	*pfkA*, *pfkB*
FBA	[c]FDP ↔ DHAP + G3P	*fbaA*, *fbaB*
TPI	[c]DHAP ↔ G3P	*tpiA*
GAPD	[c]G3P + NAD + PI ↔ 13DPG + H + NADH	*gapA*, *gapC1*, *gapC2*
PGK	[c]13DPG + ADP ↔ 3PG + ATP	*pgk*
PGM	[c]3PG ↔ 2PG	*gpmA*, *gpmB*
ENO	[c]2PG ↔ H_2O + PEP	*eno*
PYK	[c]ADP + H + PEP → ATP + PYR	*pykA*, *pykF*

	HEX1	PGI	PFK	FBA	TPI	GAPD	PGK	PGM	ENO	PYK
ATP	−1	0	−1	0	0	0	1	0	0	1
GLC	−1	0	0	0	0	0	0	0	0	0
ADP	1	0	1	0	0	0	−1	0	0	−1
G6P	1	−1	0	0	0	0	0	0	0	0
H	1	0	1	0	0	1	0	0	0	−1
F6P	0	1	−1	0	0	0	0	0	0	0
FDP	0	0	1	−1	0	0	0	0	0	0
DHAP	0	0	0	1	−1	0	0	0	0	0
G3P	0	0	0	1	1	−1	0	0	0	0
NAD	0	0	0	0	0	−1	0	0	0	0
PI	0	0	0	0	0	−1	0	0	0	0
13DPG	0	0	0	0	0	1	−1	0	0	0
NADH	0	0	0	0	0	1	0	0	0	0
3PG	0	0	0	0	0	0	1	−1	0	0
2PG	0	0	0	0	0	0	0	1	−1	0
PEP	0	0	0	0	0	0	0	0	1	−1
H_2O	0	0	0	0	0	0	0	0	1	0
PYR	0	0	0	0	0	0	0	0	0	1



J. L. Reed et al., Nature Reviews Genetics, Vol. 7, pp. 130-141 (2006)

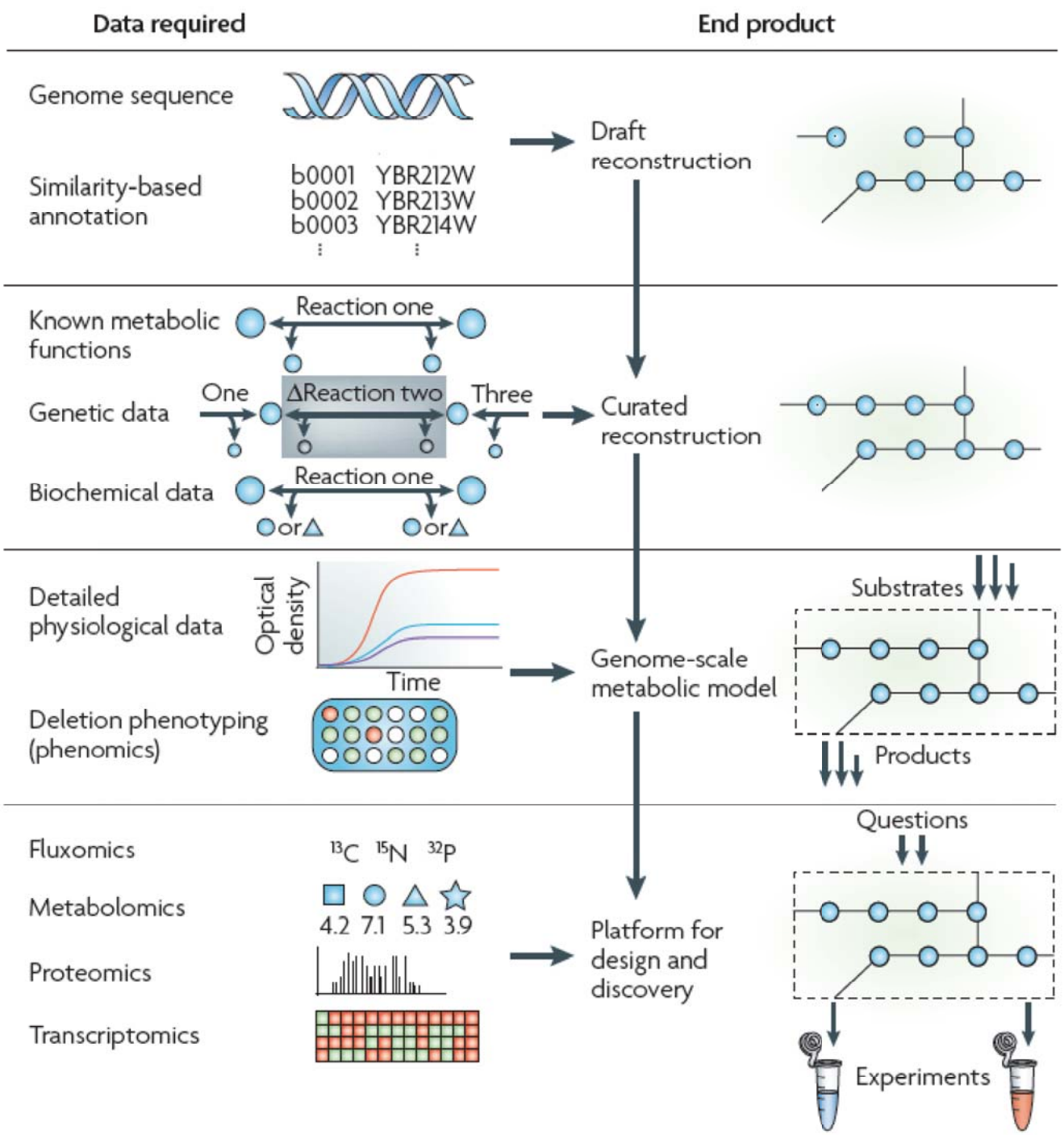


Figure 1 | **Phases and data used to generate a metabolic reconstruction.**

A. M. Feist et al., Nature Reviews Microbiology, Vol. 7, pp. 129-143 (2009)

MrBac (http://sb.nhri.org.tw/MrBac)

A web server to draft

Metabolic network reconstructions for Bacteria

Mr Bac

Home | Draft reconstruction | About | How-to | Contact | Supplementary

*Select a query and a reference species: [Example]

Query Species:	Reference Species:
Acaryochloris marina MBIC11017	Acaryochloris marina MBIC11017
Acholeplasma laidlawii PG-8A	Acholeplasma laidlawii PG-8A
Acidiphilium cryptum JF-5	Acidiphilium cryptum JF-5
Acidithiobacillus ferrooxidans ATCC 23270	Acidithiobacillus ferrooxidans ATCC 23270
Acidithiobacillus ferrooxidans ATCC 53993	Acidithiobacillus ferrooxidans ATCC 53993
Acidobacteria bacterium Ellin345	Acidobacteria bacterium Ellin345
Acidobacterium capsulatum ATCC 51196	Acidobacterium capsulatum ATCC 51196
Acidothermus cellulolyticus 11B	Acidothermus cellulolyticus 11B
Acidovorax citrulli AAC00-1	Acidovorax citrulli AAC00-1
Acidovorax sp. JS42	Acidovorax sp. JS42

*Choose Blast Settings: [Default / Select All]

E value: & Percent Identity: %

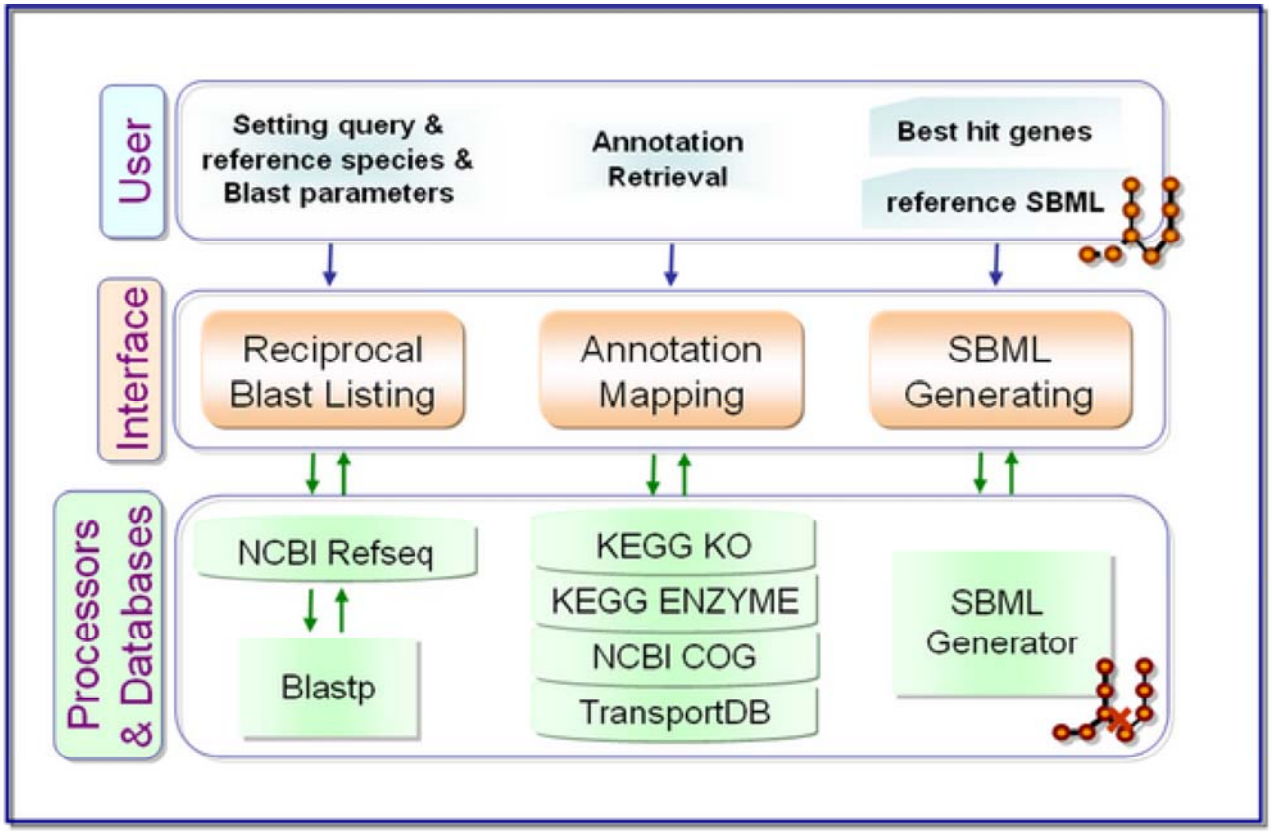
Choose the blast columns to be shown:

☑Ref ☑Query ☑PercIden ☐MLen ☐mismatches ☐gap openings ☐q.start ☐q.end ☐r.start ☐r.end ☑EVal ☐BitScore

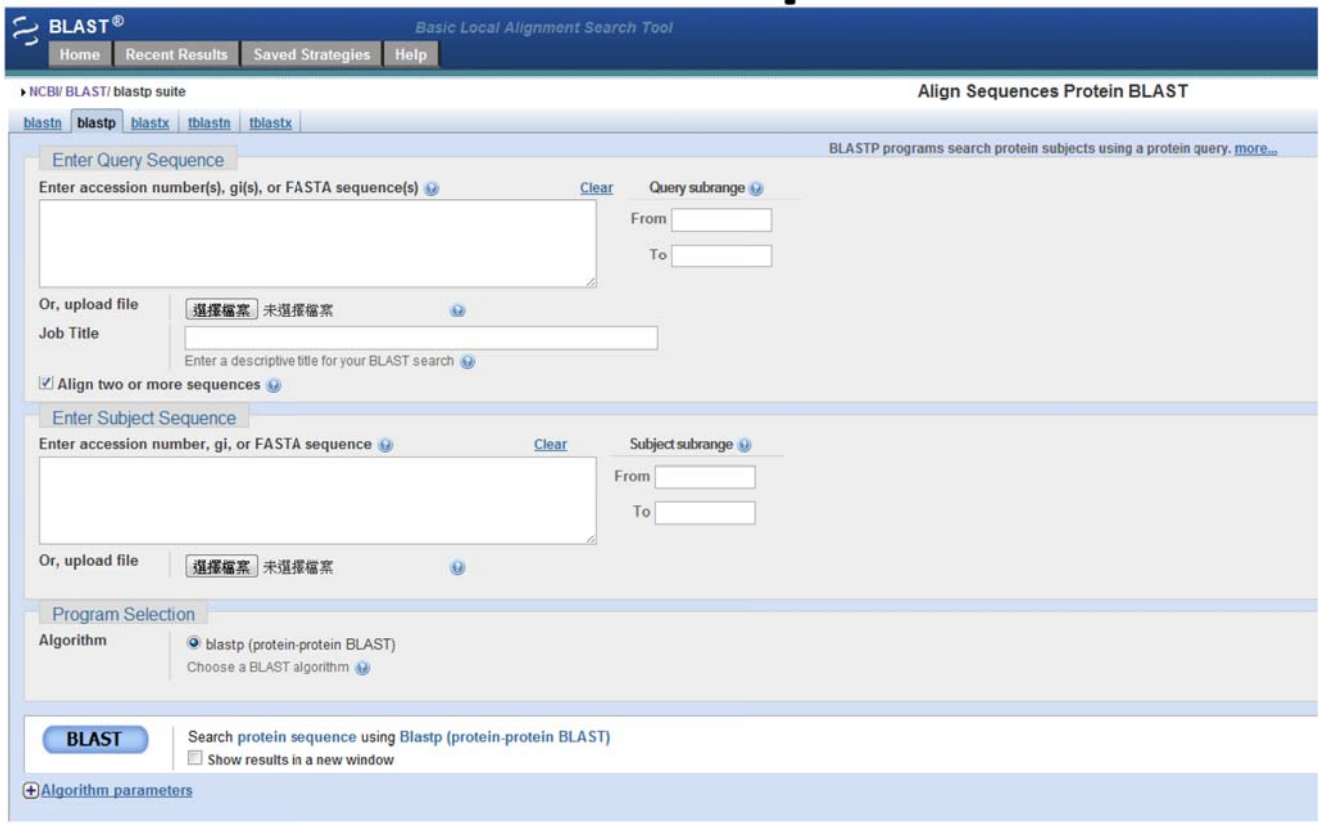
* indicates required input.

Submit

Flow Chart


User
Setting query & reference species & Blast parameters
Annotation Retrieval
Best hit genes
reference SBML
Interface
Reciprocal Blast Listing
Annotation Mapping
SBML Generating
Processors & Databases
NCBI Refseq
Blastp
KEGG KO
KEGG ENZYME
NCBI COG
TransportDB
SBML Generator

Reciprocal Blast



http://blast.ncbi.nlm.nih.gov/

http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastDocs&DOC_TYPE=Download

Query	Ref	PercIden	MLen	mismatches	gap openings	q.start	q.end	r.start	r.end	EVal	PercIden2	MLen2	mismatches2	gap openings2	q.start2	q.end2	r.start2	r.end2	EVal2
STM0002	b0002	94.51	820	45	0	1	820	1	820	0.0	94.51	820	45	0	1	820	1	820	0.0
STM0003	b0003	93.51	308	20	0	1	308	1	308	2e-165	93.51	308	20	0	1	308	1	308	6e-154
STM0004	b0004	93.46	428	28	0	1	428	1	428	0.0	93.46	428	28	0	1	428	1	428	0.0
STM0005	b0006	86.38	257	35	0	1	257	1	257	6e-132	86.38	257	35	0	1	257	1	257	6e-132
STM0006	b0007	76.26	476	113	0	1	476	1	476	0.0	76.26	476	113	0	1	476	1	476	0.0
STM0007	b0008	94.64	317	17	0	1	317	1	317	0.0	94.64	317	17	0	1	317	1	317	0.0

Salmonella enterica subsp. enterica serovar Typhimurium str. LT2 versus Escherichia coli str. K-12 substr. MG1655

Your current blast settings are
Evalue: 1e-5
PercIden: 50.0

Show all Query Genes: Yes ○ No ◉

***Choose the annotation columns to be shown for both query and reference:** ?

☑ QueryID ☑ gene
☐ 5'Coordinate ☐ 3'Coordinate ☐ Product ☐ ProteinLength
☐ **KO:** ☐ KONum ☐ KODEFIN ☐ KOCLASS
☐ **EC:** ☐ ECNum ☐ ECNAME ☐ ECCLASS ☐ ECSYSNAME
☐ **COG:** ☐ FunID ☐ FUNCategory ☐ FUNDesc
☐ **MT:** ☐ TransporterType ☐ FamilyName ☐ Substrate

☑ ReferenceID ☑ gene
☐ **KO:** ☐ KONum ☐ KODEFIN ☐ KOCLASS
☐ **EC:** ☐ ECNum ☐ ECNAME ☐ ECCLASS ☐ ECSYSNAME
☐ **COG:** ☐ FunID ☐ FUNCategory ☐ FUNDesc
☐ **MT:** ☐ TransporterType ☐ FamilyName ☐ Substrate

Back to blast filtering Proceed to annotation matching...

Annotation retrieval

Salmonella enterica subsp. enterica serovar Typhimurium str. LT2 versus **Escherichia coli str. K-12 substr. MG1655**
Evalue: 1e-5 Percent Identity: 50.0

MyBac

<< Back to blast filtering
< Back to column selection

Save as .xls file

Save Gene Pair File
Metabolic Network Generator>

Query	gene	5'Coordinate	3'Coordinate	Product	ProteinLength	KONum	KODEFIN	ECCLASS	ECNum	ECN
PSLT067	--	54974	55216	putative cytoplasmic protein	80	--	--	--	--	--
STM0002	thrA	337	2799	bifunctional aspartokinase I/homoserine dehydrogenase I	820	K00003;; K00928	homoserine dehydrogenase [EC:1.1.1.3];; aspartate kinase [EC:2.7.2.4]	Metabolism; Amino Acid Metabolism; Glycine, serine and threonine;metabolism [PATH:ko00260];Metabolism; Amino Acid Metabolism; Cysteine and methionine;metabolism [PATH:ko00270];Metabolism; Amino Acid Metabolism; Lysine biosynthesis;[PATH:ko00300]	1.1.1.13;; 1.1.1.3;; 2.7.2.4	L-arabinitol 2-dehydrogenase; L-a forming); L-arabinitol (ribulose-fo homoserine dehydrogenase; HSDI aspartokinase; AK; beta-aspartoki
STM0003	thrB	2801	3730	homoserine kinase	309	K00872	homoserine kinase [EC:2.7.1.39]	Metabolism; Amino Acid Metabolism; Glycine, serine and threonine;metabolism [PATH:ko00260]	2.7.1.39	homoserine kinase; homoserine ki
STM0004	thrC	3734	5020	threonine synthase	428	K01733	threonine synthase [EC:4.2.3.1]	Metabolism; Amino Acid Metabolism; Glycine, serine and threonine;metabolism [PATH:ko00260];Metabolism; Metabolism of Cofactors and Vitamins; Vitamin B6;metabolism [PATH:ko00750]	4.2.3.1	threonine synthase; threonine synt phospho-lyase (adding water)
STM0005	yaaA	5887	5114	hypothetical protein	257	K09861	hypothetical protein	Unclassified; Poorly Characterized; Function unknown	--	--
STM0006	yaaJ	7396	5966	alanine/glycine transport protein	476	K03310	alanine or glycine:cation symporter, AGCS family	Environmental Information Processing; Membrane Transport; Other;ion-coupled transporters	--	--

Metabolic Network Draft Generation

Draft metabolic network generation requires a reference SBML file and the list of reciprocal blast best hits as inputs to proceed.

By taking these two inputs, we are assuming that query ORFs in the best hit list have the same functions as their corresponding reference ORFs. The generator only keeps reactions, in which all associated reference ORFs under "AND" condition and partial reference ORFs under "OR" condition have best hit matches to query ORFs. Only when those two criteria are met, the reference ORF id are replaced by corresponding query id, and the metabolites involved in those reactions are kept in the draft SBML file.

The draft in SBML format can be saved as the final output after query is submitted and can be further curated and edited manually in flux balance analysis softwares such as COBRA.

Upload a reference SBML file (sample, iAF1260)

SBML File: 選擇檔案 未選擇檔案

Upload a one to one gene pair file (sample, SLT2vsECO)

Gene Pair File: 選擇檔案 未選擇檔案

Type in a model name

New Model Name:

Submit Query

```
<?xml version="1.0" encoding="UTF-8"?>
<sbml level="2" version="1" >
    <model id="Ec_iAF1260" name="Ec_iAF1260" >
        <listOfUnitDefinitions >
            <unitDefinition id="mmol_per_gDW_per_hr" >
                <listOfUnits >
                    <unit kind="mole" scale="-3" />
                    <unit exponent="-1" kind="gram" />
                    <unit exponent="-1" kind="second" multiplier=".00027777" />
                </listOfUnits>
            </unitDefinition>
        </listOfUnitDefinitions>
        <listOfCompartments >
            <compartment id="Extra_organism" />
            <compartment id="Periplasm" outside="Extra_organism" />
            <compartment id="Cytosol" outside="Periplasm" />
        </listOfCompartments>
        <listOfSpecies >
            <species boundaryCondition="false" charge="-2" compartment="Cytosol" id="M_10fthf_c" name="M_10_Formyltetrahydrofolate_C2
            <species boundaryCondition="false" charge="0" compartment="Cytosol" id="M_12ppd_R_c" name="M__R__Propane_1_2_diol_C3H8O2"
            <species boundaryCondition="false" charge="0" compartment="Cytosol" id="M_12ppd_S_c" name="M__S__Propane_1_2_diol_C3H8O2"
            <species boundaryCondition="false" charge="-4" compartment="Cytosol" id="M_13dpg_c" name="M_3_Phospho_D_glyceroyl_phospha
            <species boundaryCondition="false" charge="0" compartment="Cytosol" id="M_14glucan_c" name="M_1_4_alpha_D_glucan_C36H62O3
            <species boundaryCondition="false" charge="0" compartment="Periplasm" id="M_14glucan_p" name="M_1_4_alpha_D_glucan_C36H62
            <species boundaryCondition="false" charge="2" compartment="Cytosol" id="M_15dap_c" name="M_1_5_Diaminopentane_C5H16N2" />
            <species boundaryCondition="false" charge="2" compartment="Periplasm" id="M_15dap_p" name="M_1_5_Diaminopentane_C5H16N2"
            <species boundaryCondition="false" charge="0" compartment="Periplasm" id="M_1agpe120_p" name="M_1_Acyl_sn_glycero_3_phospl
            <species boundaryCondition="false" charge="0" compartment="Periplasm" id="M_1agpe140_p" name="M_1_Acyl_sn_glycero_3_phospl
            <species boundaryCondition="false" charge="0" compartment="Periplasm" id="M_1agpe141_p" name="M_1_Acyl_sn_glycero_3_phospl
            <species boundaryCondition="false" charge="0" compartment="Periplasm" id="M_1agpe160_p" name="M_1_Acyl_sn_glycero_3_phospl
            <species boundaryCondition="false" charge="0" compartment="Periplasm" id="M_1agpe161_p" name="M_1_Acyl_sn_glycero_3_phospl
            <species boundaryCondition="false" charge="0" compartment="Periplasm" id="M_1agpe180_p" name="M_1_Acyl_sn_glycero_3_phospl
            <species boundaryCondition="false" charge="0" compartment="Periplasm" id="M_1agpe181_p" name="M_1_Acyl_sn_glycero_3_phospl
            <species boundaryCondition="false" charge="-1" compartment="Periplasm" id="M_1agpg120_p" name="M_1_Acyl_sn_glycero_3_phosp
            <species boundaryCondition="false" charge="-1" compartment="Periplasm" id="M_1agpg140_p" name="M_1_Acyl_sn_glycero_3_phosp
```

Starting materials for manual curation

3	STM1097	--	1184453	1184863	hypothetical protein	136	K07127	--	Unclassifi	--	--	--	--	R	POORLY	General fu	--	--	--	b1970	hiuH
4	STM1098	hpaC	1185708	1185196	4-hydroxyphenylacetate catabolism	170	K00484	4-hydroxy	Metabolis	1.14.13.3	4-hydroxy	Oxidoredu	4-hydroxy	R	POORLY	General fu	--	--	--	--	--
5	STM1099	hpaB	1187288	1185726	4-hydroxyphenylacetate catabolism	520	K00483	4-hydroxy	Metabolis	1.14.13.3	4-hydroxy	Oxidoredu	4-hydroxy	Q	METABC	Secondary	--	--	--	--	--
6	STM1100	hpaR	1187946	1187506	4-hydroxyphenylacetate catabolism	146	--	--	--	--	--	--	--	K	INFORM	Transcript	--	--	--	--	--
7	STM1101	hpaG	1188221	1189510	4-hydroxyphenylacetate catabolism	429	K05921	5-oxopent	Metabolis	4.1.1.68;;	5-oxopent	Lyases; C	5-oxopent	Q	METABC	Secondary	--	--	--	--	--
8	STM1102	hpaE	1189507	1190973	4-hydroxyphenylacetate catabolism	488	K00151	5-carboxy	Metabolis	1.2.1.60	5-carboxy	Oxidoredu	5-carboxy	C	METABC	Energy pr	--	--	--	--	--
9	STM1103	hpaD	1190975	1191826	4-hydroxyphenylacetate catabolism	283	K00455	3,4-dihydı	Metabolis	1.13.11.15	3,4-dihydı	Oxidoredu	3,4-dihydı	S	POORLY	Function u	--	--	--	--	--
10	STM1104	hpaF	1191836	1192216	4-hydroxyphenylacetate catabolism	126	K01826	5-carboxy	Metabolis	5.3.3.10	5-carboxy	Isomerase	5-carboxy	E	METABC	Amino aci	--	--	--	--	--
11	STM1105	hpaH	1192359	1193162	4-hydroxyphenylacetate catabolism	267	K02509	2-oxo-hep	Metabolis	4.2.1.-	--	Lyases; C	--	Q	METABC	Secondary	--	--	--	--	--
12	STM1106	hpaI	1193173	1193964	4-hydroxyphenylacetate catabolism	263	K02510	2,4-dihydı	Metabolis	4.1.2.-	--	Lyases; C	--	G	METABC	Carbohydı	--	--	--	--	--
13	STM1107	hpaX	1194036	1195412	4-hydroxyphenylacetate catabolism	458	K02511	MFS trans	Protein Fa	--	--	--	--	G;; E;; P;;	METABC	Carbohydı	Secondary	The Majo	4-hydroxy	--	--
14	STM1108	hpaA	1195422	1196318	4-hydroxyphenylacetate catabolism	298	K02508	AraC fam	Protein Fa	--	--	--	--	K	INFORM	Transcript	--	--	--	--	--
15	STM1109	--	1196332	1197270	putative periplasmic protein	312	--	--	--	--	--	--	--	R	POORLY	General fu	--	--	--	--	--

- We can identify, for instance, the metabolic genes that are present in SLT2 but absent in ECO.
- Examples include the cob (STM2016-STM2035), hpa (STM1098-STM1108) and pdu (STM2037-STM2058).
- These genes have been reported to confer SLT2 the abilities to synthesize cobalamin and to utilize 4-hydroxylphen-ylacetate and propanediol.

A quick look at GEMSiRV

http://sb.nhri.org.tw/GEMSiRV/en/GEMSiRV

Acknowledgements

- Julie Chih-Yu Chen
- Ming-Hsin Tsai
- Yueh-Hsia Tang

- Dr. Feng-Chi Chen
- Dr. Chao A. Hsiung