

# MetaDraft

## 1. Welcome

This is the MetaDraft user guide. MetaDraft is part of the CBMPy MetaToolkit project and is developed by the Systems Bioinformatics group at the VU University Amsterdam. For more information on MetaDraft please visit the MetaDraft pages at <https://systemsbioinformatics.github.io/cbmpy-metadraft/#>. MetaDraft is primarily written in [Python](#) and uses Qt for its GUI, as well as [CBMPy](#) for its object model and [SBML](#) support.

For information on installing MetaDraft and its dependencies please see the **readme.md** file included in your [download](#) or on [GitHub](#), you will also find information on the `systemtest.py` utility that will check whether your system is ready to run MetaDraft.

The screenshot displays the MetaDraft GUI. The main window is titled "MetaDraft 0.9.0 result: (test)-(bigg2-eco)-(opt)". It features a menu bar with "File", "Build options", "Model options", "Sessions", and "Help". Below the menu is a toolbar with "Build", "Genes", "Reactions", and "Metabolites". The central area contains a table with the following columns: "source", "match", "score", a checkbox, and "org". The table lists 14 rows of gene matches. Row 6 is selected, showing a match of "D5949" with a score of 1.0. To the right of the table is an "Information" panel for the selected gene, displaying details such as "function", "GO\_component", "gene\_synonym", "name", "gene", "product", "protein\_id", and "db\_xref".

source	match	score		org	
1	in_stu0160	b3057	1.0	<input checked="" type="checkbox"/>	bigg2-eco
2	in_stu0163	b3784	1.0	<input checked="" type="checkbox"/>	bigg2-eco
3	in_stu0246	b0968	1.0	<input checked="" type="checkbox"/>	bigg2-eco
4	in_stu0262			<input type="checkbox"/>	
5	in_stu0264	b1363	0.219	<input type="checkbox"/>	bigg2-eco
6	in_stu0264	D5949	1.0	<input checked="" type="checkbox"/>	bigg2-eco
7	in_stu0265	b3290	1.0	<input checked="" type="checkbox"/>	bigg2-eco
8	in_stu0505	b3560	1.0	<input checked="" type="checkbox"/>	bigg2-eco
9	in_stu0590	b3770	1.0	<input checked="" type="checkbox"/>	bigg2-eco
10	in_stu0596			<input type="checkbox"/>	
11	in_stu0603	b4254	1.0	<input checked="" type="checkbox"/>	bigg2-eco
12	in_stu0603	b0273	0.001	<input type="checkbox"/>	bigg2-eco
13	in_stu0605	b0809	1.0	<input checked="" type="checkbox"/>	bigg2-eco
14	in_stu0692	b4129	0.82	<input type="checkbox"/>	bigg2-eco

**Information**

GeneMatch: stu0264 --> b3849

**Matching Gene Information: b3849**

function	transport; Transport of small molecules; Cations
GO_component	GO:0019886 - organelle inner membrane; GO:0092274 - peptidoglycan-based cell wall
gene_synonym	ECK3841; JW5576; sapI
name	potassium uptake, requires TrkE
gene	trkH
product	potassium transporter
protein_id	YP_026273.1
db_xref	CE:49176432 ASAP:ABE-0012574 UniProtKB/Swiss-Prot:P0AFZ7 ECOCYC:EG1.1021 BioGene:EG1.1021 GeneID:948333

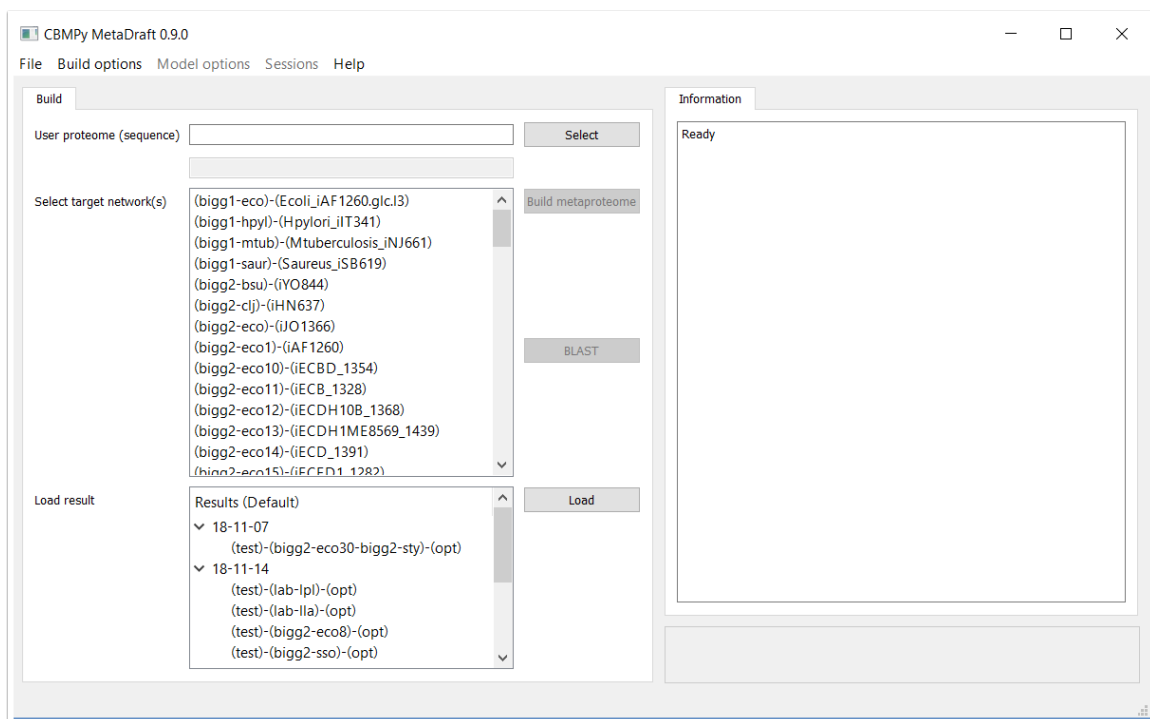
**Associated reaction(s)**

R\_Kt2pp (bigg2-eco) Potassium transport in via

**Python GUI based genome-scale model reconstruction**

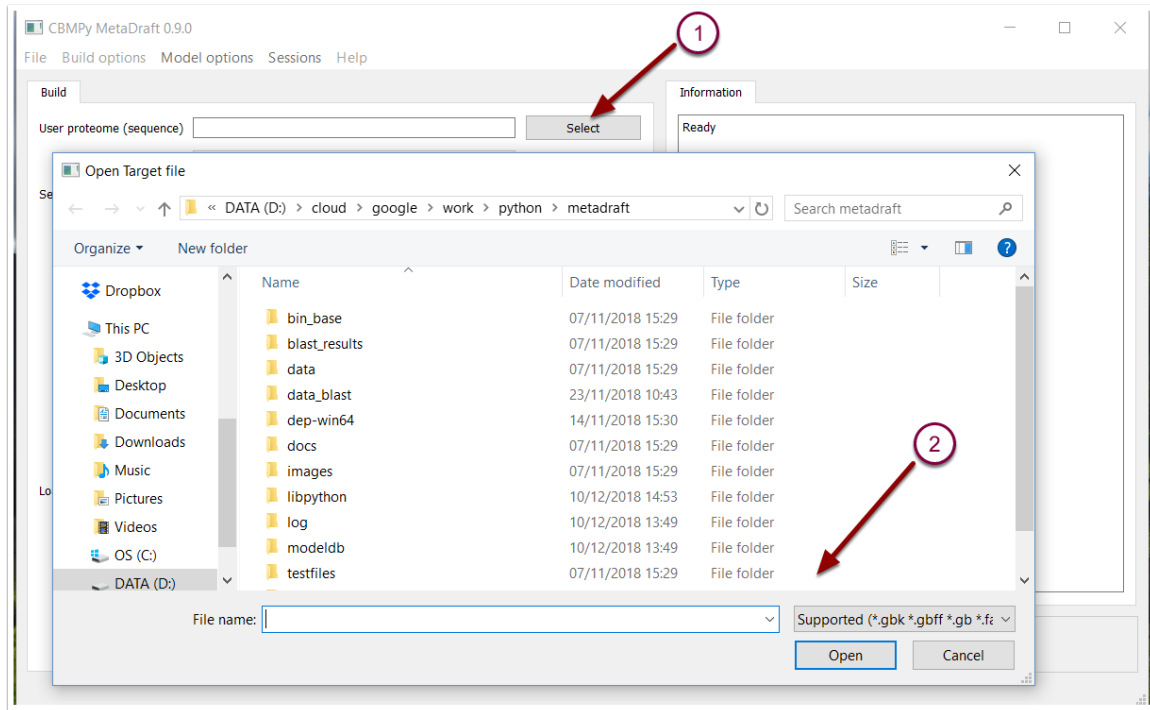
## 2. Main reconstruction screen

When MetaDraft starts up you will find yourself at the reconstruction screen. Here you are able to input the protein sequence file (either a protein FASTA or GenBank format file with CDS annotations), create a template database that you will BLAST your sequence against. Using the "Build Option" menu you can create your own, user-defined, template models and create custom sets of template models. You can also reload previously generated draft models in the "Load result" section.



## 2.1 Load your sequences

Let's start by loading the sequence file that you would like to turn into a draft model reconstruction. MetaDraft is able to process protein FASTA (\*.fasta) or GenBank format files (that include CDS annotation), to load your file mouse over to the "User proteome (sequence)" field and hit the associated "Select button" (1). This will trigger a FileOpen window where you may select the appropriate, supported file (2).

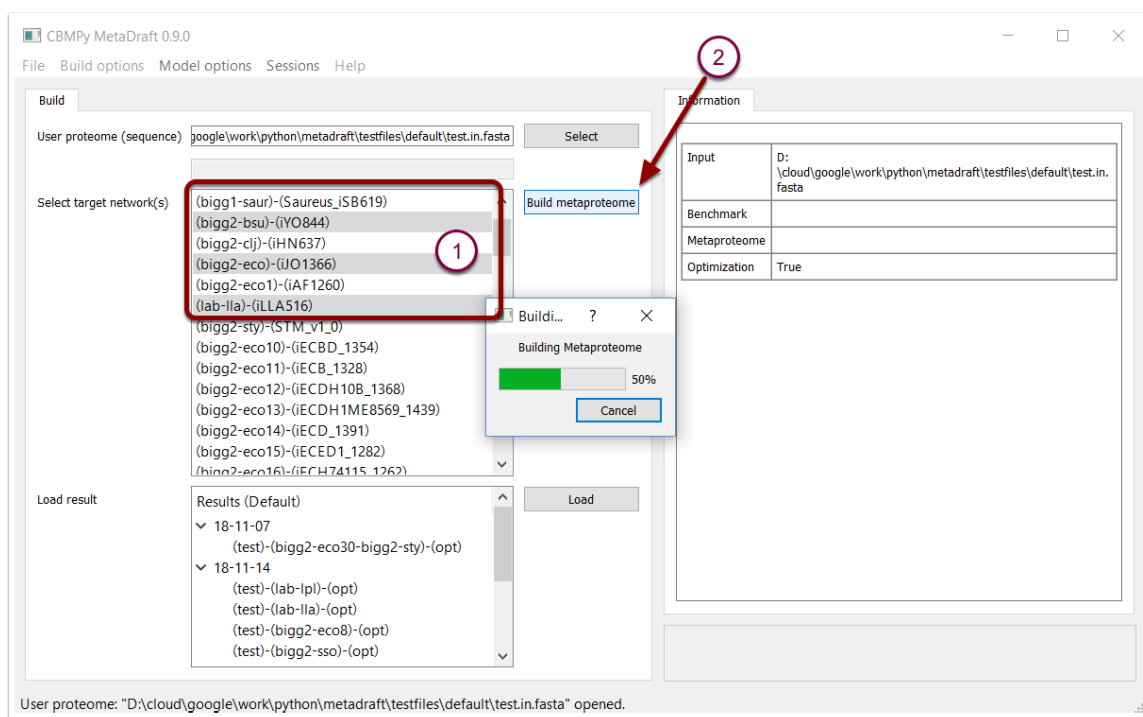


## 2.2 Build a metaproteome from template models

MetaDraft works by comparing your input sequence to a database of known models and sequences or "template models". However, the modeller has control over the number and priority, of the template models that are assembled into what MetaDraft calls a metaproteome. In this release template models are created from the BiGG database and thus use a consistent set of reaction and metabolite identifiers, this is not a prerequisite for user-defined templates.

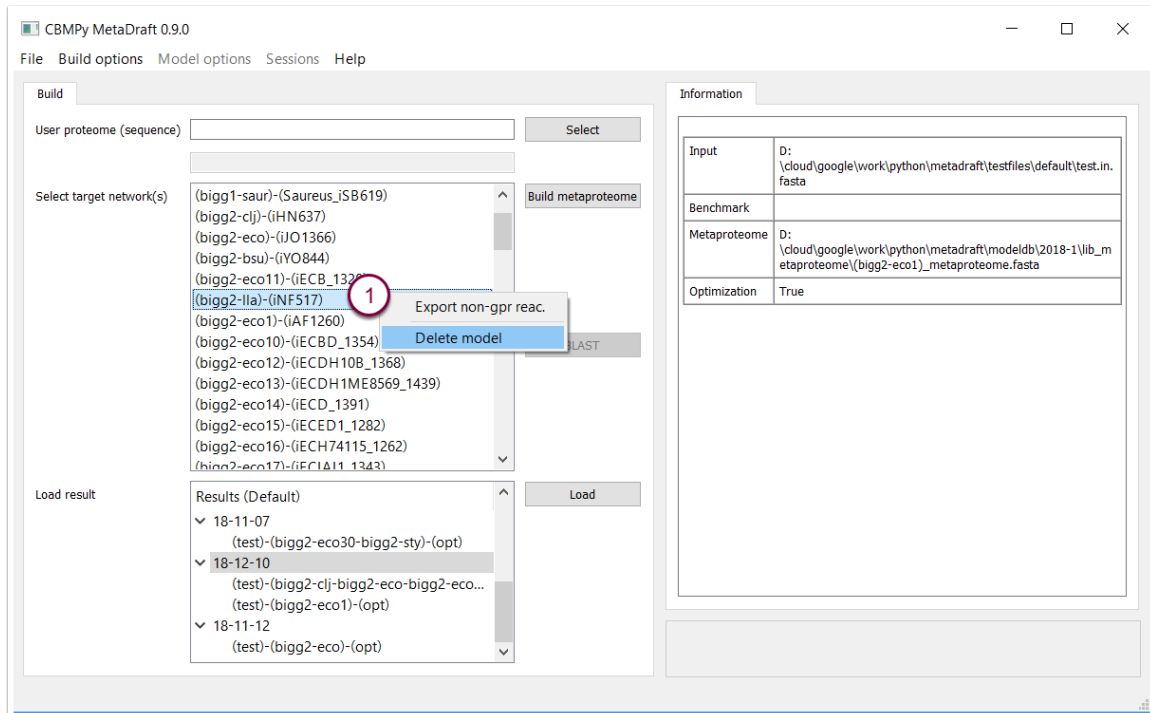
Once a user sequence has been input (previous step) one or more template models in the "Target network" list can be selected to form the basis of the metaproteome (1). Note that order does matter, template models higher in the list have priority when "ID optimization" is selected. Template model priority can include factors such as, phylogenetic distance from your source model, physiological factors or template models quality/completeness.

Once a selection has been made the metaproteome can be constructed by pressing the "Build metaproteome" button (2).



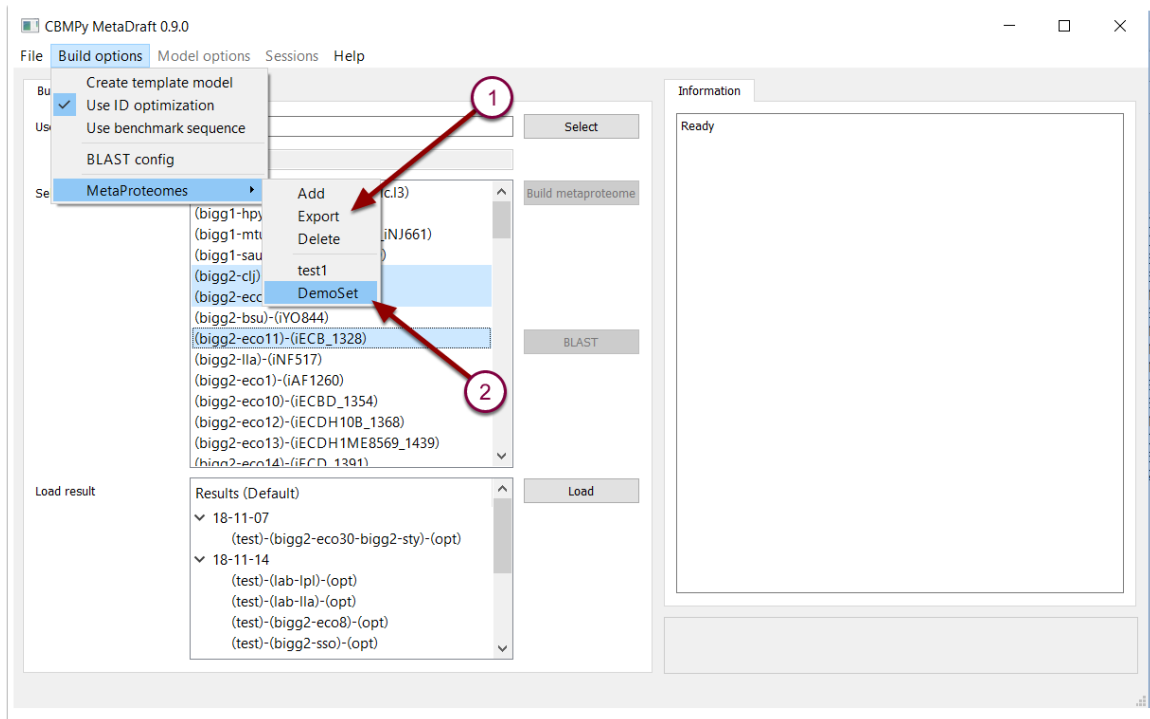
## 2.3 Template model options

Right-clicking on a template model allows you to delete it from the model list or export the model reactions that are not associated with genes via a GPR association (1). Non-GPR reactions will be exported as an SBML Level 3 FBC file that can be used later in the reconstruction process.



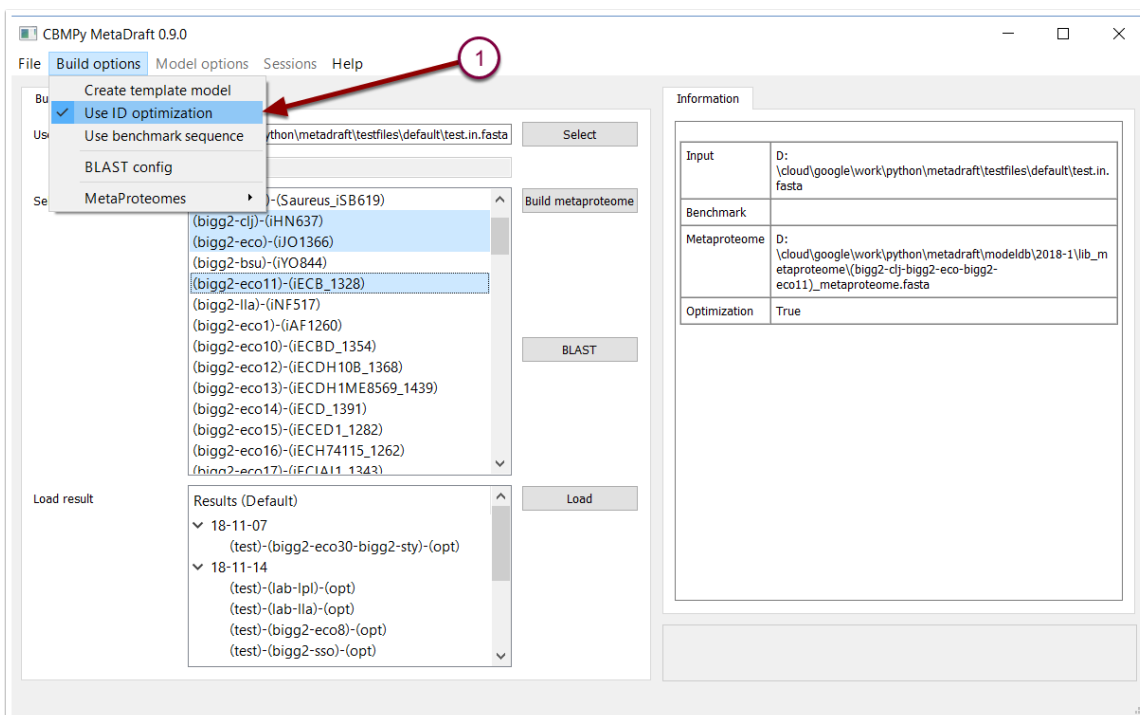
## 2.4 User defined metaproteome selections

MetaDraft allows users to define their own metaproteomes which, once defined, can easily be selected from the "Build options" menu. Selecting the "MetaProteomes" sub-menu (1) you can "Add" or "Delete" metaproteomes or "Export" a metaproteome in FASTA format. This allows metaproteomes to be exported and used for other purposes outside of MetaDraft. Once defined metaproteomes become available for selection in the lower half of the "Metaproteomes" menu (2).



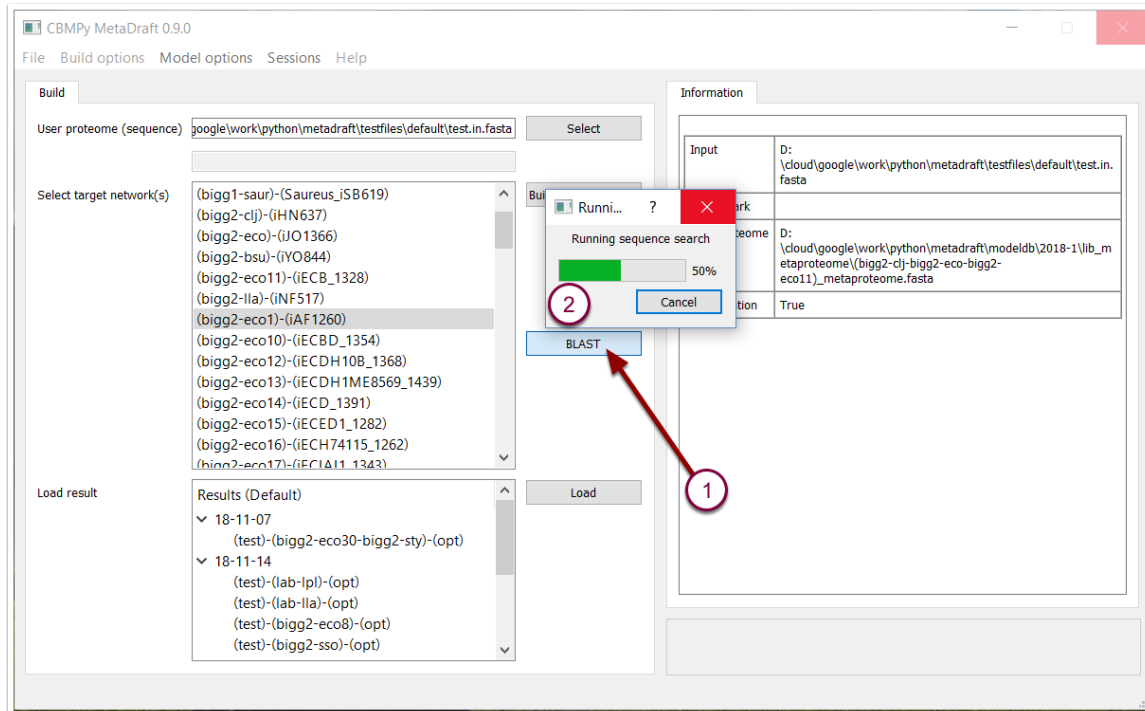
## 2.5 Metaproteome optimization

By default MetaDraft uses ID optimization to reduce the number of related reactions present in the metaproteome. ID optimization works best when all the template models in the metaproteome use a common reaction and (less important) metabolite set. In this optimization the highest ranking model (top selection in the template model list) is used as a starting point, for any additional template only the reactions which are not present in the base template are added to the metaproteome. This option may easily be toggled on/off by selecting the "Use ID optimization" item in the "Build options" menu (1).



## 2.6 Run sequence search

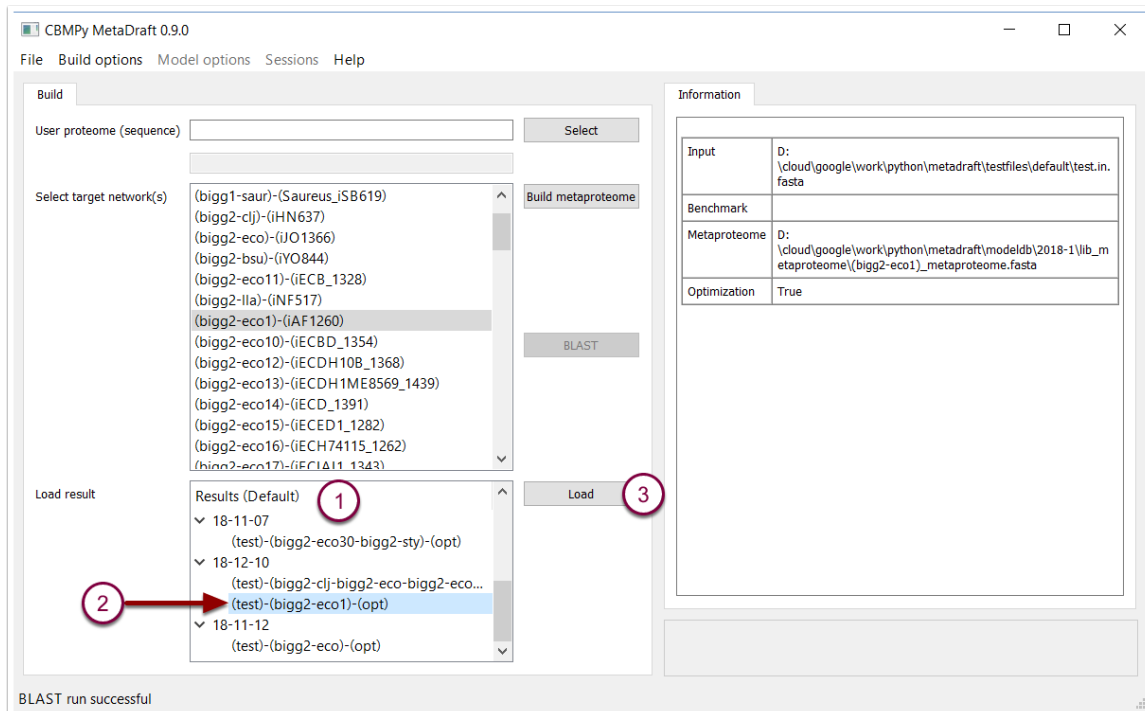
Once a user proteome has been loaded and a metaproteome defined the "BLAST" button becomes active (1). Pressing this will initiate the orthology mapping, a process which might take minutes or hours depending on the size of the user proteome and metaproteomes. On certain operating systems (some flavours of Linux)





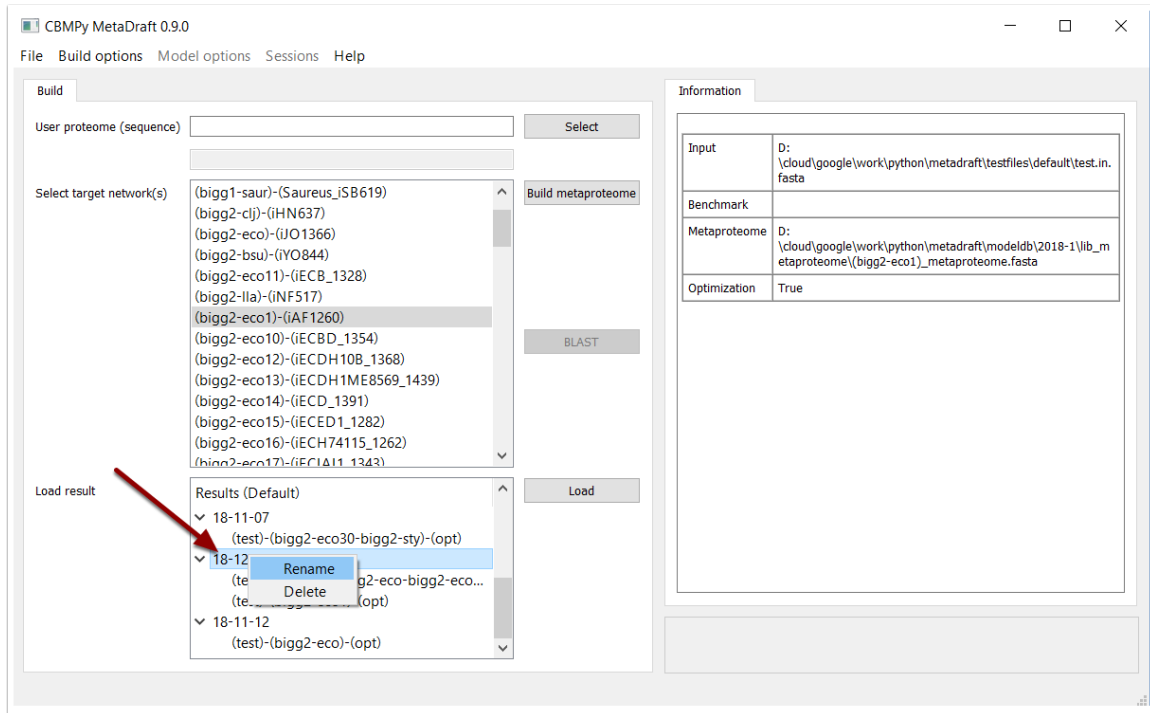
### 3. BLAST search results

When a BLAST search is complete or when MetaDraft is started up, the results of all previously run BLAST searches are shown in the results panel (1). By default these results are grouped by creation data with the results listed as (user proteome)-(metaproteome)-(optimization). To load a results, select the results (2) and press the "Load button" (3).



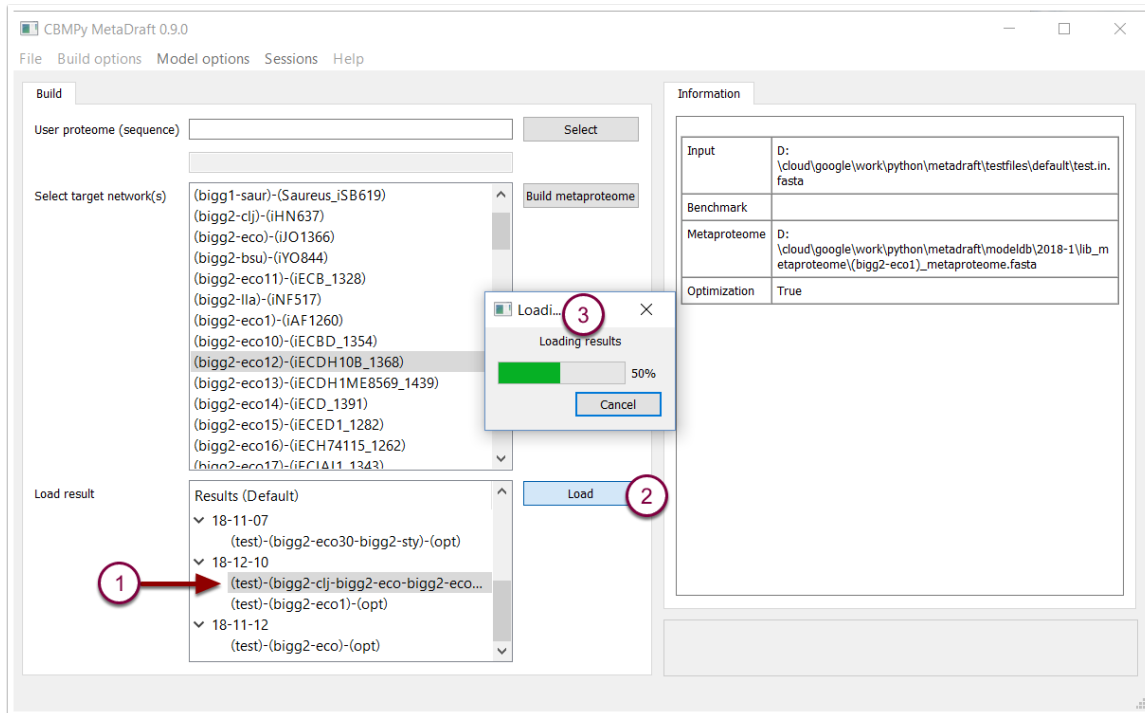
### 3.1 Delete or rename collections of results

Right-clicking on the result group allows one to either delete the entire group or rename it to something more suitable. In addition, right-clicking on a result allows one to delete it, note there is currently no undelete function and deletion is therefore permanent.



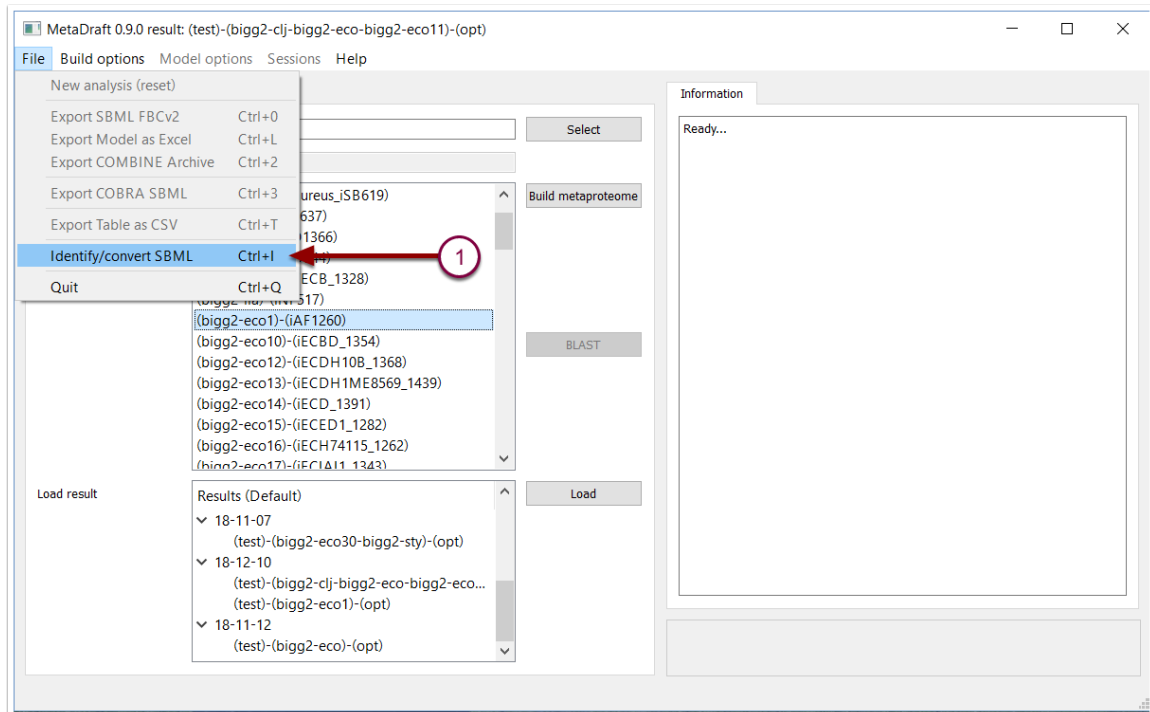
### 3.2 Loading results for further reconstruction

Once the results have been generated, they can be loaded for further interpretation and manipulation. This is done by selecting a results set (1), pressing the "Load" button (2) and waiting for the model to load into the edit screen (3).



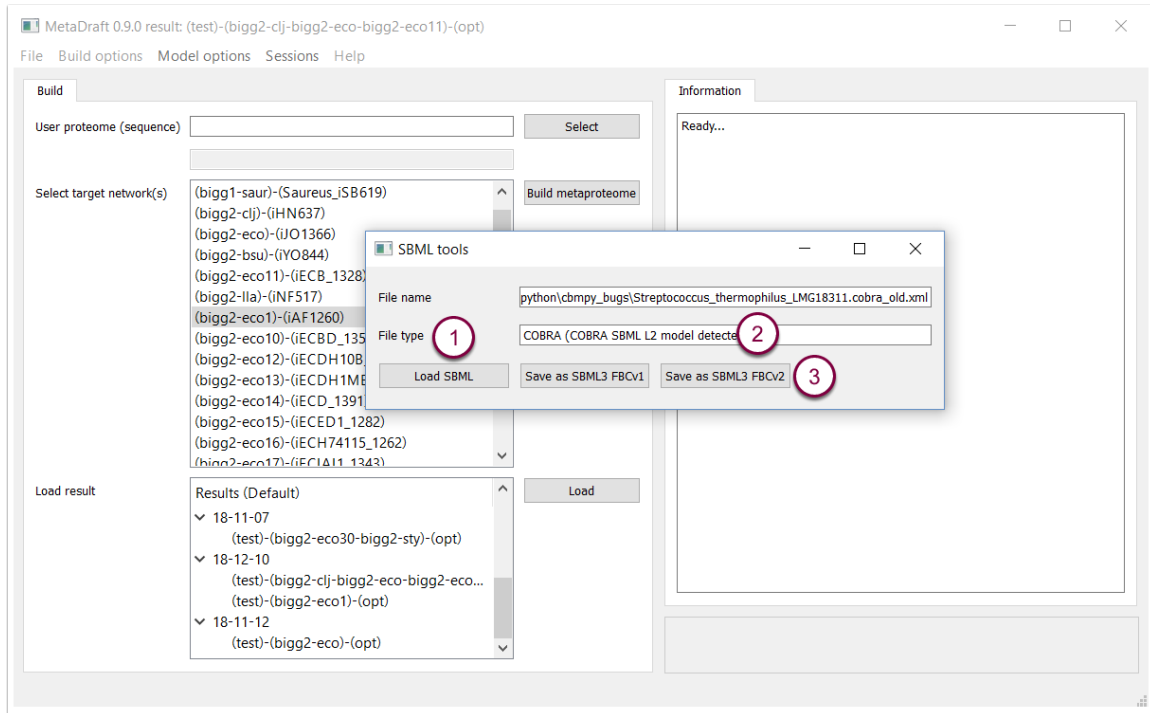
## 4. Utility function: analyse and convert an SBML model

MetaDraft template models must be encoded in the latest version of the open, community, standard SBML L3V1 FBCv2. However, many potential template models are encoded in older versions of SBML or SBML dialects, therefore, Metadraft contains functions that assist the modeller in the identification of model type and the creation of template models. To analyse, identify and optionally convert an SBML to the format used by the MetaDraft template generator select "Identify/convert" (1) from the "File" menu.



### 4.1 Utility function: analyze and convert an SBML model

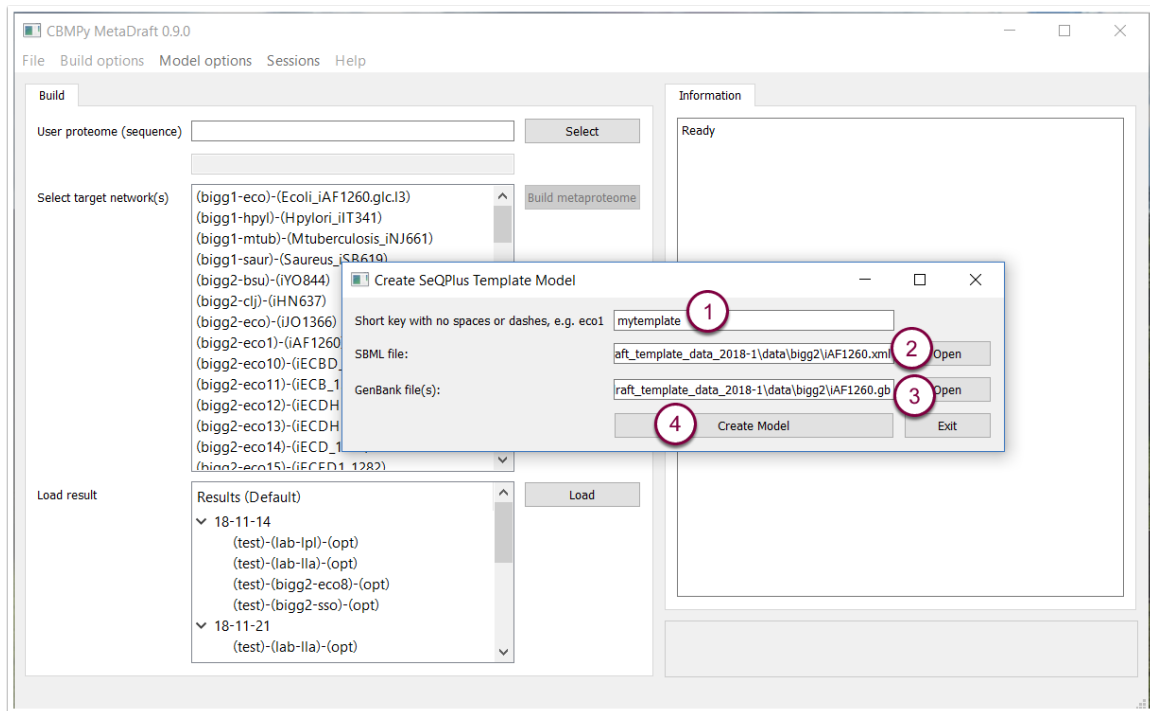
To use the SBML identification/conversion tool, simply load an SBML model (1). If the model encoding is identified this will be displayed in the "File type" field (2) and can then be converted to SBML3 FBCv2 by pressing the "Save as" button (3).



## 5. Utility function: create a user-defined, template model

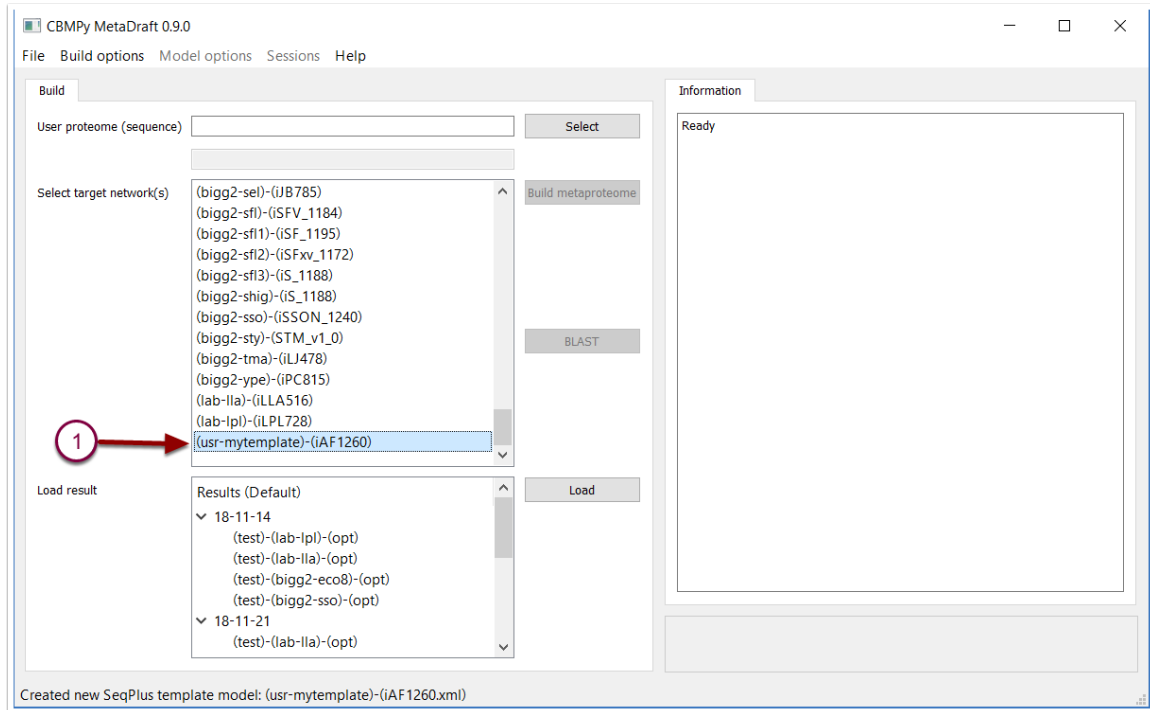
MetaDraft's template database can be extended with the addition of customised SeqPlus templates. These templates are the result of the combination of a genome-scale model (encoded using SBML L3V1 FBCv2 with GPR information) and the associated genome (in GenBank format). Please note that currently, the source SBML and GenBank file needs to be located in the same directory. The "Create template model" function is available in the "Build options" menu.

Once selected the template builder window allows you to input a unique model identifier (1), a genome-scale model in SBML format (2) and an associated genbank file that contains CDS annotations for genes present in the model (3). Once all of the previous information has been entered the template can be created by pressing the "Create Model" button (4). The final step may take a few minutes.



## 5.1 Utility function: create a user-defined, template model

User defined templates are accessible as "(usr-template\_id)" in the "Target networks" list.



## 6. Primary result editing screen: genes

When the results of a BLAST search are loaded they are displayed in MetaDrafts result editing screen. The results are displayed on three interlinked tabs on the left hand side of the panel (1) while any selected components annotation is displayed on the right panel.

Source genes (genes in your input proteome) are displayed in the first column (3), while, where relevant, matches from the metaproteome are displayed in the "match column" (4). The "score" column provides the match score which ranges from 0 to 1 (5) and the "org" column displays the matching template model (6). Finally, the selection column (7) allows genes to be included in the final reconstruction or not, for convenience those genes with a 100% match are automatically selected.

Source genes with multiple matches are grouped by colour for easy identification while on the right hand side (2) displays all known annotation about the matching gene and the reactions associated with it. This also includes the GPR association which is colour encoded such that green is a gene that is present and selected, red is present but not selected and black has no match in the target metaproteome (8).

MetaDraft 0.9.0 result: (test)-(bigg2-cj)-bigg2-eco-big2-eco11)-(opt)

	source	match	score		org
31	in_stu1013	b34	1.0	<input checked="" type="checkbox"/>	bigg2-eco
32	in_stu1013	b3416	1.0	<input checked="" type="checkbox"/>	bigg2-eco
33	in_stu1264	b2416	1.0	<input checked="" type="checkbox"/>	bigg2-eco
34	in_stu1264	ECB_02316	1.0	<input checked="" type="checkbox"/>	bigg2-eco11
35	in_stu1265	b2415	1.0	<input checked="" type="checkbox"/>	bigg2-eco
36	in_stu1265	ECB_02315	1.0	<input checked="" type="checkbox"/>	bigg2-eco11
37	in_stu1266	b1136	1.0	<input checked="" type="checkbox"/>	bigg2-eco
38	in_stu1267	b0333	1.0	<input checked="" type="checkbox"/>	bigg2-eco
39	in_stu1268	CLJU_RS03...	1.0	<input checked="" type="checkbox"/>	bigg2-cj
40	in_stu1268	CLJU_RS15...	0.475	<input type="checkbox"/>	bigg2-cj
41	in_stu1460	CLJU_RS20...	1.0	<input checked="" type="checkbox"/>	bigg2-cj
42	in_stu1461	ECB_02422	1.0	<input checked="" type="checkbox"/>	bigg2-eco11
43	in_stu1461	b2530	1.0	<input checked="" type="checkbox"/>	bigg2-eco
44	in_stu1492	b0811	1.0	<input checked="" type="checkbox"/>	bigg2-eco
45	in_stu1648	CLJU_RS06...	1.0	<input checked="" type="checkbox"/>	bigg2-cj
46	in_stu1711	CLJU_RS02...	1.0	<input checked="" type="checkbox"/>	bigg2-cj
47	in_stu1869	b3396	1.0	<input checked="" type="checkbox"/>	bigg2-eco
48	in_stu1886			<input type="checkbox"/>	

Information panel (GeneMatch: stu1268 --> CLJU\_RS03255):

**Matching Gene Information: CLJU\_RS03255**

**inference**: EXISTENCE: similar to AA sequence:RefSeq:WP\_013237332.1

**EC\_number**: 4.2.1.3

**note**: Catalyzes the conversion of citrate to isocitrate; Derived by automated computational analysis using gene prediction method: Protein Homology.

**old\_locus\_tag**: CLJU\_c06620

**product**: aconitate hydratase

**protein\_id**: WP\_013237332.1

**db\_xref**: [GI:503002356](http://gi.503002356)

**Associated reactions(s)**

**R\_ACONta** (bigg2-cj): Aconitase (half-reaction A, Citrate hydro-lyase)

**GPR**: ((CLJU\_RS03255 or CLJU\_RS15025))

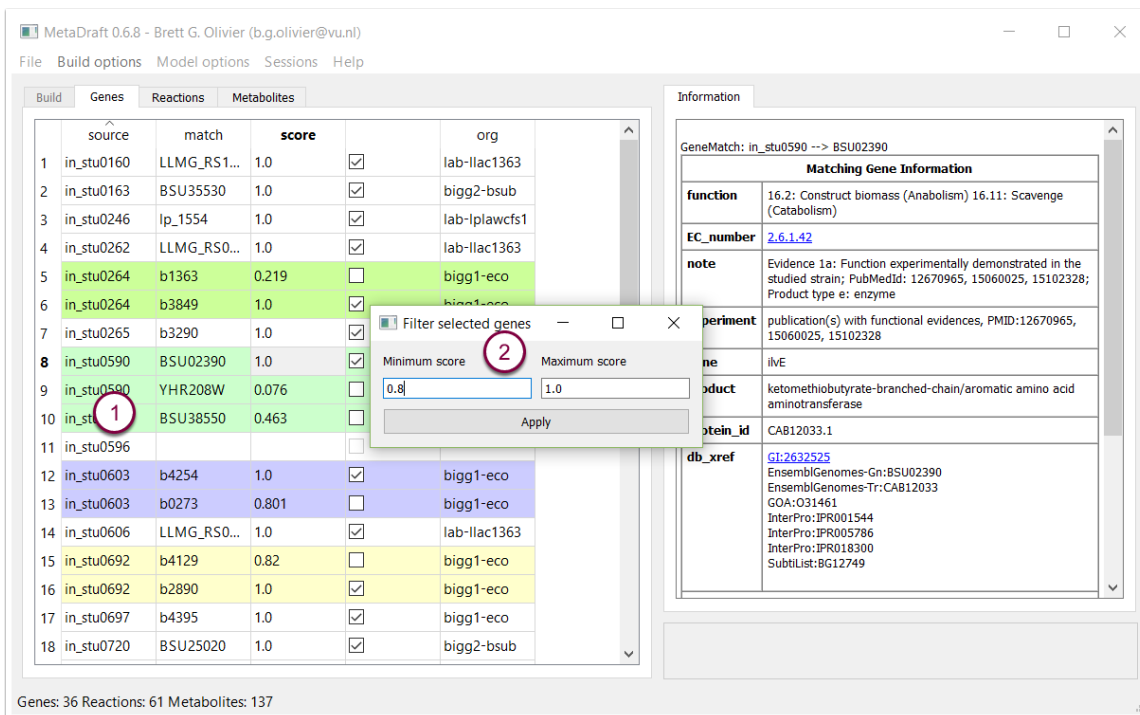
**R\_ACONtb** (bigg2-cj): Aconitase (half-reaction B, Isocitrate hydro-lyase)

**GPR**: ((CLJU\_RS03255 or CLJU\_RS15025))



## 6.1 Gene match score filter

Right clicking on on a row in the "gene" table (1) displays the "Gene Filter" tool where a user defined range matching can be set (between 0 and 1) and applied to quickly select all genes matching the selected score criteria (2).



## 7. Reaction editing screen

As genes are selected in the "Genes table" the corresponding reactions are dynamically displayed in the "Reaction" table (1). Information displayed in the "Reactions" table include the reactions identifier, name and its source model (2). In addition the source gene and its target matches are also provided (3).

On the right hand side any annotation associated with the reaction is displayed including the GPR association given in terms of both the original and newly mapped association (3). Genes are colour encoded to quickly show those that are active, inactive or missing from the original GPR. Hovering over any of the genes in the association displays that genes annotation as a tooltip (4). The model can be further refined by excluding specific reactions from the final reconstruction.

The screenshot shows the MetaDraft 0.9.0 interface. The main window displays a table of reactions with columns for reaction ID, name, org, genes, and src. A red circle (1) highlights the 'Reactions' tab. A second red circle (2) highlights the 'name' column. A third red circle (3) highlights the 'genes' column for reaction R\_ACMUMptspp. On the right, the 'Information' panel shows details for the selected reaction, including organism, association, equation ID, and equation name. A red circle (4) highlights a tooltip for the gene 'ptsH' in the association list.

reaction	name	org	genes	src
1 R_ACGAMT	UDP-N-acet...	bigg2-eco	b3784	stu0163
2 R_ACGAptspp	N-Acetyl-D-...	bigg2-eco	b2416,b2415	stu1264,stu1...
3 R_ACGAptspp	N-Acetyl-D-...	bigg2-eco11	ECB_02316,...	stu1264,stu1...
4 R_ACMANAptspp	N-acetyl-D-...	bigg2-eco11	ECB_02316,...	stu1264,stu1...
5 R_ACMANAptspp	N-acetyl-D-...	bigg2-eco	b2416,b2415	stu1264,stu1...
6 R_ACMUMptspp	N-acetylmur...	bigg2-eco11	ECB_02316,...	stu1264,stu1...
7 R_ACMUMptspp	N-acetylmur...	bigg2-eco	b2416,b2415	stu1264,stu1...
8 R_ACONT	Aconitate hy...	bigg2-clj	CLJU_RS03...	stu1268
9 R_ACONTa	Aconitase (h...	bigg2-clj	CLJU_RS03...	stu1268
10 R_ACONtb	Aconitase (h...	bigg2-clj	CLJU_RS03...	stu1268
11 R_ADOCBLPP	Adenosylco...	bigg2-eco11	ECB_00607	stu0697
12 R_AHSERL2	O acetylho...	bigg2-clj	CLJU_RS03...	stu0987
13 R_AMALT1	Amylomalta...	bigg2-eco	b3416	stu1013
14 R_AMALT2	Amylomalta...	bigg2-eco	b3416	stu1013
15 R_AMALT3	Amylomalta...	bigg2-eco	b3416	stu1013
16 R_AMALT4	Amylomalta...	bigg2-eco	b3416	stu1013
17 R_APH120	Acylphosph...	bigg2-eco	b0968	stu0246
18 R_APH140	Acylphosph...	bigg2-eco	b0968	stu0246

Information panel for reaction R\_ACMUMptspp:

- Organism: bigg2-eco11
- R\_ACMUMptspp N-acetylmuramate transport via PEP:Pyr PTS (irreversible) (periplasm)
- Association: ((ECB\_02329 and ECB\_02315 and ECB\_02316 and ECB\_02317))
- Association new: (stu1264 and stu1265)
- Equation ID: M\_pep\_c + M\_acum...
- Equation Name: Phosphoenolpyruvate acetylmuramate 6-ph...
- Subs: M\_pep\_c Phosphoenolpyruvate charge: -3 C3H2O6P; M\_acum\_p N-Acetylmuramate charge: -1 C11H18NO8
- Prox: M\_acum6p\_c N-acetylmuramate 6-phosphate charge: -3 C11H17NO11P; M\_pyr\_c Pyruvate charge: -1 C3H3O3
- Annotations: locus\_tag: ECB\_02315; raw\_location: 2462490..2462747; codon\_start: 1; EC\_number: 2.7.1.69; transl\_table: 11; note: similar to b2415; gene: ptsH; product: phosphohistidinoprotein-hexo... component of PTS system (Hpr); protein\_id: ACT39970.1

## 7.1 Metabolite viewer

The metabolite viewer is dynamically updated depending on the reactions selected from the "Reaction" panel (1). The lefthand panel contains metabolite id, name and source model or organism (2). On the right hand side of the viewer any selected metabolite's associated annotation (as defined in the model) is displayed (3). Where possible and if present in the source model, MIRIAM URL's are displayed as live links that can be opened in a webbrowser.

MetaDraft 0.9.0 result: (test)-(bigg2-clj-big2-eco-big2-eco11)-(opt)

File Build options Model options Sessions Help

1 Genes Reactions Metabolites

	metabolite	name	org	fixed
49	M_ddcap_c	Dodecanoly-phosphate (...)	bigg2-eco	<input type="checkbox"/>
50	M_dha_c	Dihydroxyacetone	bigg2-eco11	<input type="checkbox"/>
51	M_dhap_c	Dihydroxyacetone phosph...	bigg2-eco11	<input type="checkbox"/>
52	M_dhgly_c	Dehydroglycine	bigg2-eco	<input type="checkbox"/>
53	M_dpcoa_c	Dephospho-CoA	bigg2-clj	<input type="checkbox"/>
54	M_dxyl5p_c	1-deoxy-D-xylulose 5-ph...	bigg2-eco11	<input type="checkbox"/>
55	M_f1p_c	D-Fructose 1-phosphate	bigg2-eco11	<input type="checkbox"/>
56	M_f6p_c	D-Fructose 6-phosphate	bigg2-eco11	<input type="checkbox"/>
57	M_fad_c	Flavin adenine dinucleoti...	bigg2-eco11	<input type="checkbox"/>
58	M_fadh2_c	Flavin adenine dinucleoti...	bigg2-eco11	<input type="checkbox"/>
59	M_fe2_c	Fe2+ mitochondria	bigg2-eco11	<input type="checkbox"/>
60	M_fru_p	D-Fructose	bigg2-eco11	<input type="checkbox"/>
61	M_g1p_c	D-Glucose 1-phosphate	bigg2-eco	<input type="checkbox"/>
62	M_g6p_c	D-Glucose 6-phosphate	bigg2-eco11	<input type="checkbox"/>
63	M_galt_p	Galactitol	bigg2-eco11	<input type="checkbox"/>
64	M_galt1p_c	Galactitol 1-phosphate	bigg2-eco11	<input type="checkbox"/>
65	M_gam_p	D-Glucosamine	bigg2-eco11	<input type="checkbox"/>
66	M_gam6p_c	D-Glucosamine 6-phosph...	bigg2-eco11	<input type="checkbox"/>

Information

<b>Id</b>	M_fe2_c
<b>Name</b>	Fe2+ mitochondria
Charge	2
Formula	Fe
ReagentOf	R_I2FE2SS2, R_I2FE2SS
RDF references (opens in browser)	
is	<a href="http://identifiers.org/big2.metabolite/fe2">http://identifiers.org/big2.metabolite/fe2</a>
is	<a href="http://identifiers.org/chebi/CHEBI:13319">http://identifiers.org/chebi/CHEBI:13319</a>
is	<a href="http://identifiers.org/chebi/CHEBI:13321">http://identifiers.org/chebi/CHEBI:13321</a>
is	<a href="http://identifiers.org/chebi/CHEBI:21129">http://identifiers.org/chebi/CHEBI:21129</a>
is	<a href="http://identifiers.org/chebi/CHEBI:24876">http://identifiers.org/chebi/CHEBI:24876</a>
is	<a href="http://identifiers.org/chebi/CHEBI:29033">http://identifiers.org/chebi/CHEBI:29033</a>
is	<a href="http://identifiers.org/chebi/CHEBI:34754">http://identifiers.org/chebi/CHEBI:34754</a>
is	<a href="http://identifiers.org/chebi/CHEBI:49599">http://identifiers.org/chebi/CHEBI:49599</a>
is	<a href="http://identifiers.org/hmdb/HMDB00692">http://identifiers.org/hmdb/HMDB00692</a>
is	<a href="http://identifiers.org/kegg.compound/C14818">http://identifiers.org/kegg.compound/C14818</a>
is	<a href="http://identifiers.org/metanetx.chemical/MNXM111">http://identifiers.org/metanetx.chemical/MNXM111</a>

Genes: 31 Reactions: 99 Metabolites: 168

## 8. Multi-session support

When a result is loaded MetaDraft loads a default selection set which includes the set of selected genes and reactions. Custom selection sets can be created and used in the "Sessions" menu (1). This menu allows you to load (2) and save (3) sessions or clear the list of saved selections (4).

The screenshot shows the MetaDraft 0.9.0 interface. The 'Sessions' menu is open, showing options: 'Save current session', 'Saved sessions', and 'Clear all sessions'. The 'Saved sessions' dropdown is expanded, showing a list of sessions with 'selection1' selected. The main table displays a list of gene matches with columns for source, match, and org. The 'Information' panel on the right shows details for the gene 'b2415'.

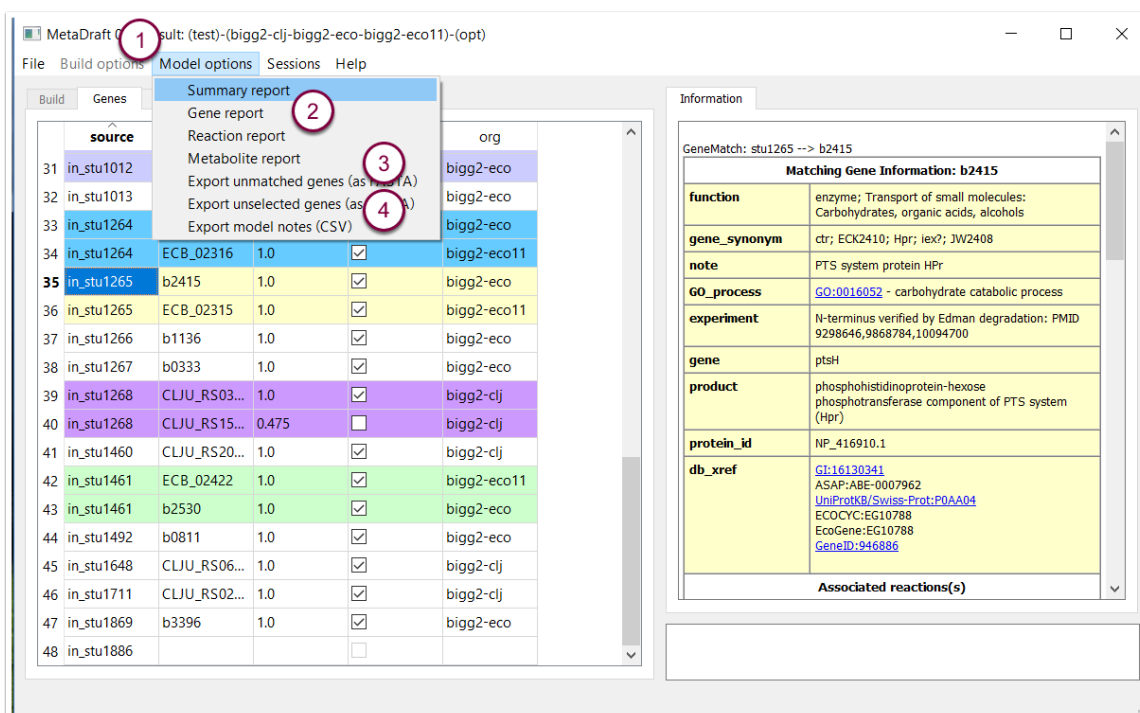
Build	Genes	Reactions	Match	Org
31	in_stu1012	b3428	1.0	bigg2-eco
32	in_stu1013	b3416	1.0	bigg2-eco
33	in_stu1264	b2416	1.0	bigg2-eco
34	in_stu1264	ECB_02316	1.0	bigg2-eco11
35	in_stu1265	b2415	1.0	bigg2-eco
36	in_stu1265	ECB_02315	1.0	bigg2-eco11
37	in_stu1266	b1136	1.0	bigg2-eco
38	in_stu1267	b0333	1.0	bigg2-eco
39	in_stu1268	CLJU_RS03...	1.0	bigg2-clj
40	in_stu1268	CLJU_RS15...	0.475	bigg2-clj
41	in_stu1460	CLJU_RS20...	1.0	bigg2-clj
42	in_stu1461	ECB_02422	1.0	bigg2-eco11
43	in_stu1461	b2530	1.0	bigg2-eco
44	in_stu1492	b0811	1.0	bigg2-eco
45	in_stu1648	CLJU_RS06...	1.0	bigg2-clj
46	in_stu1711	CLJU_RS02...	1.0	bigg2-clj
47	in_stu1869	b3396	1.0	bigg2-eco
48	in_stu1886			

The 'Information' panel shows the following details for GeneMatch: stu1265 --> b2415:

- Matching Gene Information: b2415**
- function:** enzyme; Transport of small molecules: Carbohydrates, organic acids, alcohols
- gene\_synonym:** ctr; ECK2410; Hpr; iex2; JW2408
- note:** PTS system protein HPr
- GO\_process:** [GO:0016052](#) - carbohydrate catabolic process
- experiment:** N-terminus verified by Edman degradation: PMID 9298646,9868784,10094700
- gene:** ptsH
- product:** phosphohistidinoprotein-hexose phosphotransferase component of PTS system (Hpr)
- protein\_id:** NP\_416910.1
- db\_xref:** [GI:16130341](#), [ASAP:ABE-0007962](#), [UniProtKB/Swiss-Prot:P0AA04](#), [ECOCYC:EG10788](#), [EcoGene:EG10788](#), [GeneID:946886](#)

## 9. Model options and reconstruction reports

Metadraft provides the modeller with an extensive set of reports that detail the contents and components of the draft genome-scale model. These reports are available from the "Model options" menu (1) and include gene, reaction and metabolite reports (2). In addition utility functions are provided that allows the modeller to export (as protein FASTA files) the sets of "unmatched genes" that is genes that do not have any match in the template models (3) as well as unselected genes (4).



## 9.1 Report: reconstruction summary

This report shows the source and template models used (1) as well as the overall make up of the new reconstruction. This is provided as a fractional representation of the number of genes, reactions and metabolites, from different sources, that make up the model (2).

The screenshot shows the MetaDraft 0.9.0 interface. The 'MetaDraft Report' window is open, displaying a 'MetaDraft Summary Report'. The report includes the following information:

- Input file:** test.fasta
- Metaproteome:** bigg2-cj-bigg2-eco-bigg2-eco11
- Result file:** D:\cloud\google\work\python\metadraft\blast\_results\default\18-12-10\test-(bigg2-cj-bigg2-eco-bigg2-eco11)-(opt)\_metalink.resplus.json
- Report date:** 2018-12-12 15:48
- MetaDraft version:** 0.9.0

The report is divided into three main statistical sections:

### Gene statistics

Model	Percentage	Genes	Avg. score
bigg2-eco	54.84	17	1.0
bigg2-cj	32.26	10	1.0
bigg2-eco11	12.9	4	1.0
<b>Total</b>		<b>31</b>	

### Reaction statistics

Model	Percentage	Reactions
bigg2-eco	51.52	51
bigg2-eco11	29.29	29
bigg2-cj	19.19	19
<b>Total</b>		<b>99</b>

### Metabolite statistics

Model	Percentage	Metabolites
bigg2-eco	42.86	72
bigg2-eco11	30.95	52
bigg2-cj	26.19	44
<b>Total</b>		<b>168</b>

At the bottom of the report window, there are two buttons: 'View in browser' and 'Export (HTML)'.

## 9.2 Report: genes

The "Gene Report" provides details of the input file, metaproteome composition. Source genes with no orthology to the metaproteome are provided together with a table of genes included in draft reconstruction and their score. In addition the report can be viewed directly in your browser (1) as well as saved as an HTML file (2). The gene id's shown in the report are linked to the gene annotations, show here as they appear in a web browser (3).

The screenshot shows the MetaDraft 0.9.0 interface. The main window displays a table of genes with columns for source, match, and score. A dialog box titled "MetaDraft Gene Report" is open, showing analysis details and a table of genes selected. Two buttons in the dialog are circled: "View in browser" (1) and "Export (HTML)" (2). Below the dialog, a detailed gene annotation for CLJU\_RS03055 is shown, with a circled "3" next to the product name "undecaprenyl-diphosphatase".

source	match	score
31 in_stu1012	b3428	1.0
32 in_stu1013	b3416	1.0
33 in_stu1264	b2416	1.0
34 in_stu1264	ECB_02316	1.0
35 in_stu1265	b2415	1.0
36 in_stu1265	ECB_02315	1.0
37 in_stu1266	b1136	1.0
38 in_stu1267	b0333	1.0
39 in_stu1268	CLJU_RS03...	1.0
40 in_stu1268	CLJU_RS15...	0.475
41 in_stu1460	CLJU_RS20...	1.0
42 in_stu1461	ECB_02422	1.0
43 in_stu1461	b2530	1.0
44 in_stu1492	b0811	1.0
45 in_stu1648	CLJU_RS06...	1.0
46 in_stu1711	CLJU_RS02...	1.0
47 in_stu1869	b3396	1.0
48 in_stu1886		

**MetaDraft Gene Report**

**Analysis**

Input fasta: test.fasta  
 Metaproteome used: bigg2-clj-bigg2-eco-bigg2-eco11  
 Report date: 18-12-12  
 MetaDraft version: 0.9.0

**Genes (28:28)**

Input genes not matched  
 stu0720, stu0596, stu0763, stu0762, stu0785, stu1886, stu0262, stu0117

Genes selected

Source	Target	Score	Model
stu0160	CLJU_RS03055	1.0	bigg2-clj
stu0163	b3784	1.0	bigg2-eco
stu0246	b0968	1.0	bigg2-eco
stu0264	b3849	1.0	bigg2-eco
stu0265	b3290	1.0	bigg2-eco

View in browser (1)      Export (HTML) (2)

GeneMatch: 0160 --> CLJU\_RS03055 (top)

Matching Gene Information: CLJU_RS03055	
inference	EXISTENCE: similar to AA sequence:RefSeq:WP_013237304.1
EC_number	<a href="#">3.6.1.27</a>
note	Derived by automated computational analysis using gene prediction method: Protein Homology.
old_locus_tag	CLJU_c06220
product	undecaprenyl-diphosphatase (3)
protein_id	WP_013237304.1
db_xref	<a href="#">GI:503002328</a>
Associated reactions(s)	
R_UDCPDPex (bigg2-clj)	Undecaprenyl-diphosphatase (periplasm)
((CLJU_RS07805 or CLJU_RS03055))	
R_UDCPDP (bigg2-clj)	Undecaprenyl-diphosphatase
((CLJU_RS07805 or CLJU_RS03055))	

GeneMatch: 0163 --> b3784 (top)

Matching Gene Information: b3784	
function	enzyme; Central intermediary metabolism: Sugar-nucleotide biosynthesis, conversions 1.6.4 metabolism; macromolecules (cellular constituent) biosynthesis; enterobacterial common antigen (surface glycolipid) 6.3 cell structure; surface antigens (ECA, O antigen of LPS)

### 9.3 Report: reactions

The reaction report contains details of the model reconstruction as well as the selected reactions and their annotation.

The screenshot shows the MetaDraft 0.9.0 application window. The 'MetaDraft Report' dialog is open, displaying the following information:

**MetaDraft Reaction Report**

**Analysis** (1)

- Input fasta: test.fasta
- Metaproteome used: bigg2-clj-bigg2-eco-bigg2-eco11
- Report date: 18-12-12
- MetaDraft version: 0.9.0

**Reactions (99)**

**Reactions selected**

Reaction	Name	Model	Source
R_ACGAMT	UDP-N-acetylglucosamine:undecaprenylphosphate N-acetylglucosamine -1-phosphate transferase	bigg2-eco	stu0163
R_ACGAptspp	N-Acetyl-D-glucosamine transport via PEP:Pyr PTS (periplasm)	bigg2-eco	stu1264,stu1265
R_ACGAptspp	N-Acetyl-D-glucosamine transport via PEP:Pyr PTS (periplasm)	bigg2-eco11	stu1264,stu1265
R_ACMANAptspp	N-acetyl-D-mannosamine transport via PTS (periplasm)	bigg2-eco11	stu1264,stu1265
R_ACMANAptspp	N-acetyl-D-mannosamine transport via PTS (periplasm)	bigg2-eco	stu1264,stu1265
R_ACMUMptspp	N-acetylmuramate transport via PEP:Pyr PTS (periplasm)	bigg2-eco11	stu1264,stu1265
R_ACMUMptspp	N-acetylmuramate transport via PEP:Pyr PTS (periplasm)	bigg2-eco	stu1264,stu1265
R_ACONT	Aconitate hydratase	bigg2-clj	stu1268
R_ACONTa	Aconitase (half-reaction A, Citrate hydro-lyase)	bigg2-clj	stu1268
R_ACONTb	Aconitase (half-reaction B, Isocitrate hydro-lyase)	bigg2-clj	stu1268

At the bottom of the report window, there are two buttons: 'View in browser' and 'Export (HTML)'.



## 10. Afterword

Many thanks to all the people that have contributed time in testing, using and providing valuable feedback on earlier versions of MetaDraft.

(C) Systems Bioinformatics, Vrije Universiteit Amsterdam, Amsterdam, 2018.

