Biomass Reaction Operations On Metabolic Models: an application to aid the curation of genome-scale metabolic models.

V. Fairclough & Dr J. W. Pinney

Theoretical Systems Biology group, Imperial College London

December 8, 2013

Contents

Chapter 1

Overview

1.1 Introduction to BROOMM

Genome-scale metabolic models can now be generated semi-automatically using information in certain online databases, such as KEGG, Model SEED and MetaCyc [1, 2, 3, 4]. However, they contain only a generic biomass reaction with incorrect stoichiometries (discussed further in chapter 4), which causes quantitative analyses of them to be inaccurate. Therefore, the user must curate this reaction, based on current knowledge of the cell type being modelled. This process of curating metabolic models is time consuming [5] and challenging because of the inevitable errors that are often introduced when manually editing files in XML format. The key motivation for developing the application is to allow experts in the field of molecular and cellular biology to curate existing metabolic models without needing to directly access the XML files. Biomass Reaction Operations on Metabolic Models (BROOMM) is a desktop application providing a graphical user interface that allows easy curation of biomass reactions within genome-scale metabolic models. By aiding collaboration between systems biologists and wet-lab scientists, we hope that this software will allow the development of more accurate genome-scale metabolic models.

The application has been implemented entirely in Python and requires the installation of the libSBML, wxPython, NumPy and matplotlib packages [6, 7, 8, 9, 10]. libSBML was used for parsing the XML files and manipulating the models, wxPython was used to create the graphical user interface and matplotlib and NumPy were used to generate the graphics.

1.2 Functionality and uses

The software application carries out two main functions, referred to as the *Transfer mode* and *Edit mode*.

Transfer mode

This mode allows the user to transfer information from a curated model (henceforth referred to as the ***template model***) to the model being curated (henceforth referred to as the ***target model***). The template model should be a curated model of a cell that is expected to have a similar biomass composition to that of the target model (e.g. a closely related cell).

It is highly unlikely that the biomass reaction of the target cell will be exactly the same as that of the template cell; therefore, the user has the option to select specific subsets of metabolites to transfer from the ***template biomass reaction***. The process of transferring metabolite information from the template model to the target model is helped by organising the reactants into seven different categories (Protein, Carbohydrates, Lipids, Nucleic acids, Inorganic ions, Growth-associated maintenance/Non-growth-associated maintenance (GAM/NGAM) and Other) according to the ChEBI ontology for classification of chemical entities [11]. The broad composition of each biomass reaction (***template and target reactions***) can be compared by referring to stacked bar plots (***biomass composition diagrams***) that were produced to show the relative contribution of each of these categories to the biomass.

Edit mode

This mode allows the user to edit the stoichiometries of metabolites within the biomass reaction of their target model. It should be noted that the user cannot add metabolites to the reaction, unlike in the transfer mode. The broad composition of each biomass reaction (***template and target reactions***) can be compared by referring to their stacked bar plots, as in the transfer mode.

Summary of terminology

Target model – the model requiring curation.
Target (biomass) reaction – the selected biomass reaction within the target model.
Template model – a curated model with an accurate biomass reaction.
Template (biomass) reaction – the selected biomass reaction within the template model.
ID type – the ID type used to specify metabolites within the model (KEGG, Model SEED or Other).
Biomass composition diagram – stacked bar plots produced to show the relative mass contributions of the seven ChEBI metabolite categories to the biomass.

Chapter 2

Installation

The BROOMM software can be downloaded from www.theosysbio.bio.ac.uk/broomm. Its contents should be extracted in the location desired by the user. Formal installation is not required – the program is run without installation from the command line (see section 3.1 for more information).

All of the required Python packages must be installed in order for the software to function. All packages used are shown in table 2.1

Non-standard requirements: libSBML, wxPython, NumPy, matplotlib

BROOMM has been tested and works with the following versions on the following platforms:

- **Linux:** Python v2.7.2, libSBML v5.0.0, wxPython v2.8.10.1, NumPy v1.6.1, matplotlib v1.1.0 on the Linux system.
- **Microsoft Windows:** Enthought distribution of Python (Enthought Canopy v1.1.0.1371) which also contains the Numpy library. This was used alongside libSBML v5.9.0 for Windows, matplotlib v1.3.1 for Windows and wxPython v2.8 for Windows.

Table 2.1: *Summaries of the packages used to implement the graphical user interface.*

Package	**Contribution to GUI functionality**
libSBML	Used for parsing the XML files, manipulating the models and saving them.
wxPython	Used to create the graphical user interface.
matplotlib	Used to generate the mass distribution graphics.
os	Used for file navigation.
re	Used for identifying database IDs using regular expressions based upon ID formats.
NumPy	Used in the production of the mass distribution graphics.

Chapter 3

Using the application

3.1 Overview

Running the software

The program is run from the command line; the user must navigate via the command line to the expanded BROOMM directory and then to the 'src' directory contained within it. Once within the 'src' directory, the following command should be entered to open and display the graphical user interface:

```
python broomm.py
```

Selecting the Transfer or Edit mode

Once the window has opened, the user should select the 'Tabs' button in the menu bar and then select either the Transfer or Edit option.

Loading the models

Each input file (.xml file) must contain only one metabolic model and its contents must follow the standard formatting required for SBML files. To load a model file in BROOMM, click on the button provided to select the file using a pop-up file navigation window.

Selecting format of metabolite IDs

The ID type used within the model must be selected (i.e. KEGG IDs, Model SEED IDs or Other). When using models that use non-standard IDs (i.e. Other, not KEGG or Model SEED IDs), the user must supply a tab-delimited flat file for converting the model IDs to ChEBI IDs. The files must be formatted in two columns separated by a single tab, as follows:

```
ModelID1    ChEBI
ID2     01234
ID3     12345
ID4     23456
...      ...
```

The tab-delimited flat-files must be saved under the following names:

- `other_id_type_Template.txt` Conversion file for 'Template' model (Transfer mode only)
- `other_id_type_Target.txt` Conversion file for 'Target' model (Transfer **and** Edit mode)

An example ID conversion file can be found in the `\example` file provided with the software with the filename `other_id_type_Template_example.txt`.

Selecting and displaying reaction information

Once the model is loaded, the reaction IDs are parsed from the model and displayed within a list box where the user can select the biomass reaction. Alternatively, they can type it into the text control box provided. The user must verify this selection using the select button provided.

Reactant categories

The reactants of the biomass reaction are classified according to the ChEBI ontology and are assigned to one of the seven groups described in section 1.2 above. This is done by utilising the ChEBI ontology flat files to classify the reactants according to their 'ancestors' and/or 'descendants' within the hierarchical classification system.

Reactants within a given category may be visualised by selecting the relevant button on the left hand side of the panel. A new window is displayed showing the following information:

- *Model IDs* – metabolite IDs, as seen within the model
- *ChEBI IDs* – ChEBI IDs, identified by using an ID conversion dictionary (see below)
- *Metabolite name* – metabolite names from within the model
- *Stoichiometry* – The metabolite stoichiometries within the selected reaction.
- The template model within the transfer mode will also contain a column of tick boxes to allow the user to select which metabolites should be transferred from the template to the target model.

Biomass composition diagram

A stacked bar plot is used to show the composition of biomass according to the relative masses of the ChEBI categories (e.g. Proteins, Carbohydrates). This diagram is updated whenever any changes are made or a new reaction is selected. An example can be seen in figure 4.1.

Converting IDs

ID conversion dictionaries taken from BioDB are included and used within the software to match up the ID types within the template and target models [12]. Conversion dictionaries are included for the standard IDs (KEGG, Model SEED and ChEBI).

Saving the output

Once the user has finished editing/altering a target model, the altered model can be saved to a new file in XML format by selecting the 'Save altered model' button in the bottom right hand side of the screen. This brings up a pop-up file navigation window for saving the file with a filename of their choice. Note that the filename must also include the '.xml' file extension.

Creating a new reaction (Transfer mode only)

If the target model does not already contain a biomass reaction, the user can create one. There are boxes for the user to enter a unique reaction ID, its compartment within the model, the flux upper bound, flux lower bound and the flux value. There is also a drop-box to select the units for these flux variables as well as a tick-box to allow the user to set the objective coefficient of the new biomass reaction to 1. If this box is selected, the objective coefficient of all other reactions within the model will be set to 0.

Transferring metabolites (Transfer mode only)

Once the target and template models have been loaded into the two panels, the user can transfer reactant/product information uni-directionally from the template model to the target model. To transfer reactants, the user must select the relevant category (e.g. Protein, Carbohydrate or Lipid etc.) to view a pop-up window containing all metabolites within that category. Within this window, the user can select and transfer the metabolite(s) with their stiochiometric values. Similarly, product metabolites can be transferred directly below the reactant categories by selecting the required tick boxes.

Editing metabolite stoichiometries (Edit mode only)

Once the target model has been loaded, the user can select the biomass reaction and edit its reactant/product stoichiometries. To edit reactant stoichiometries, the user must select the relevant category (e.g. Protein, Carbohydrate or Lipid etc.) to view a pop-up window containing all metabolites within that category. Within this window, the user can edit the stoichiometric values of the metabolites individually. Similarly, product stoichiometries can be edited directly within the panels situated below the reactant categories.

3.2 User input

3.2.1 Transfer mode

Input files

The transfer window provides two panels – one for loading the template model (left panel) and the other for the target model (right panel).

Required input: modifying an existing biomass reaction

Biomass reaction ID: The user can supply the biomass reaction for each model by typing the IDs directly into the associated text control boxes (one for each model) or they can select from the associated list of all reaction IDs contained within each model.

Species ID type: The ID type can be selected from a drop-down list (one for each model). Note that if 'Other' is selected as the ID type, the user must provide an ID conversion dictionary as a flat file – refer to section 3.1 for formatting requirements.

Required input: creating a new biomass reaction

Species ID type: As above.

New biomass reaction ID: The user must type the desired new biomass reaction ID directly into the associated text control box.

Compartment: The user must select the reaction compartment from the drop-down list provided.

Upper bound: The user must type the upper bound value (numbers only) into the associated text control box.

Lower bound: The user must type the lower bound value (numbers only) into the associated text control box.

Flux value: The user must type the flux value (numbers only) into the associated text control box.

Flux units: The user must select the flux units from the drop-down list provided.

Objective coefficient: If the user ticks this box, the objective coefficient of their new reaction will be set to 1 and all other reactions will have their

objective coefficients set to 0. Alternatively, if the user chooses to leave this box un-ticked, the objective coefficient of their new reaction will be set to 0 and the objective coefficient of all other reactions will be left as they were within the original model.

3.2.2 Edit mode

Input file

In the transfer mode, the user must load one model: a target model.

Required input

Biomass reaction ID: This is the same process as required for the transfer mode (refer to section 3.2.1).

Species ID type: This is the same process as required for the transfer mode (refer to section 3.2.1).

3.3 Output

Transient output

The biomass composition diagram is updated whenever a new reaction is selected or changes are made to an existing reaction.

Saved output

Both modes of function allow the user to save the altered model to a new XML file, which can be viewed within a text editor. The output file follows standard SBML format and contains all of the changes made to the original model. It can also be loaded into BROOMM for further changes at a later date or, if the biomass equation is complete, it can be used for constraint-based analyses such as flux balance analysis (FBA) [13].

Chapter 4

Example

4.1 Setting up

We have transferred information from a curated template model of *E. coli* (*i*JO1366) using non-standard IDs and a semi-automatically generated *E. coli* target model downloaded from the BioModels database (BMID000000140222) that uses KEGG IDs ([4], [14], [15]). The flat-file for converting the IDs in *i*JO1366 was created using the conversion file provided with the model and conversion dictionaries obtained from BioDB ([12], [14]). A copy of this file has been provided as an example conversion flat-file under the name of `other_id_type_Template_example.txt`. Both models are also provided with the software as an example to users.

BMID000000140222 contains only one biomass equation with *'biomass_reaction'* as its ID. *i*JO1366 has two biomass equations – we utilised the reaction with the following ID: *'R_Ec_biomass_iJO1366_WT_53p95M'*. Here we present the results from transferring information from *i*JO1366 to BMID000000140222.

The biomass reaction within the target model is typical of semi-automatically generated metabolic models created using the SuBliMinaL Toolbox [4]. These models contain a generic biomass reaction containing a standard list of reactants and products, all of which have a stoichiometry of 1 (refer to table 4.1 for all reactants and table 4.2 for all products). These biomass reactions require curation in order to generate accurate results from quantitative analyses. The process of curation involves providing accurate stoichiometric values as well as adding or removing metabolites to reflect the requirements of the individual cell being modelled.

Table 4.1: *The reactants within the target biomass reaction before the transfer of information, showing KEGG IDs, metabolite names and ChEBI categories.*

Metabolite ID	**Metabolite name**	**Category**
C00025	L-glutamate	Protein
C00037	Aminoacetic acid	Protein
C00148	L-proline	Protein
C00123	L-leucine	Protein
C00082	L-tyrosine	Protein
C00152	L-asparagine	Protein
C00073	L-methionine	Protein
C00135	L-histidine	Protein
C00062	L-arginine	Protein
C00065	L-serine	Protein
C00064	L-glutamine	Protein
C00041	L-alanine	Protein
C00079	Phenylalanine	Protein
C00407	L-isoleucine	Protein
C00188	L-threonine	Protein
C00078	L-tryptophan	Protein
C00047	L-lysine	Protein
C00049	L-aspartate	Protein
C00183	L-valine	Protein
C00097	L-cysteine	Protein
C00002	ATP	GAM/NGAM
C00459	dTTP	Nucleic acids
C00131	dATP	Nucleic acids
C00458	dCTP	Nucleic acids
C00286	dGTP	Nucleic acids
_a8dc8dc1_b8b0_4206_b692_6f114244bbb8		Other (Unidentified lipid)

Table 4.2: *The products within the target biomass reaction before the transfer of information, showing KEGG IDs and metabolite names.*

Metabolite ID	**Metabolite name**
C00080	H^+
C00008	ADP
C00009	Orthophosphate
biomass_b	

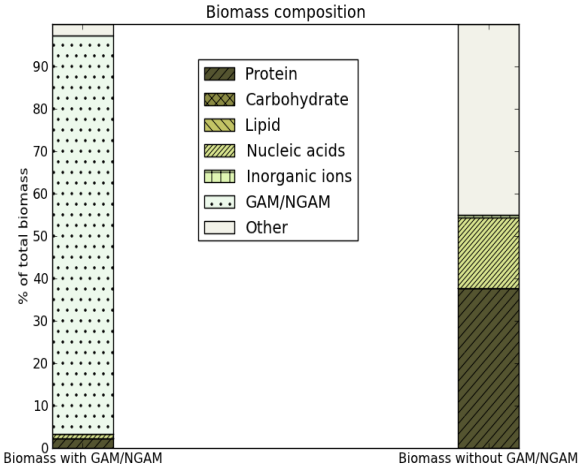
4.2 Transfer mode

The software was launched via the command line by navigating to the directory containing the software ('src' directory within the expanded BROOMM directory) and typing the following command: `python BROOMM.py`

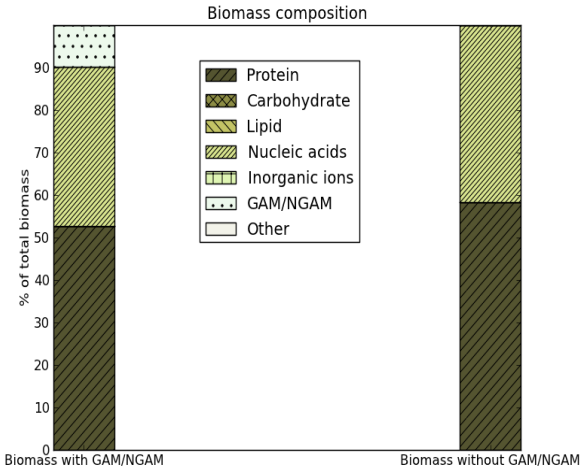
Of the 26 metabolites within the 'Proteins' category, only 2 of the template metabolites were not transferred because their corresponding IDs within the target model could not be determined, one of which is within the target model – refer to table 4.3 for details of the changes made. All 13 metabolites within the 'Nucleic acids' category were transferred successfully. Of the 8 'Inorganic ions', 6 were not transferred to the target model because it does not contain these ions and they can therefore be neglected. All four metabolites within the 'GAM/NGAM' category were transferred successfully. Of the 51 metabolites within the 'Other' category, 32 were not transferred because the ChEBI IDs could not be determined when generating the flat-file and another 8 were not contained within the target model and could therefore be neglected. There were no metabolites to transfer within the 'Carbohydrate' and 'Lipids' reactant categories. All four products were transferred.

Table 4.3: *The metabolites and stoichiometries observed within the Proteins category of the biomass reaction of the target model after the transfer of information from the template model. Note the failed transfer of L-tyrosine, which is discussed in section 4.3 below.*

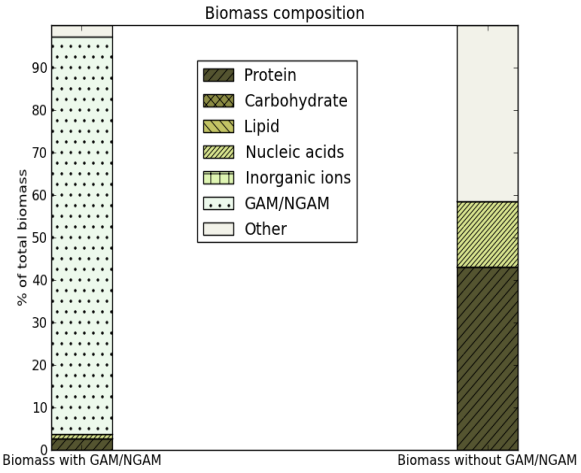
Metabolite ID	**Metabolite name**	**Stoichiometry**
C00025	L-glutamate	0.255712
C00037	Aminoacetic acid	0.595297
C00148	L-proline	0.214798
C00123	L-leucine	0.437778
C00082	L-tyrosine	1.0
C00152	L-asparagine	0.234232
C00073	L-methionine	0.149336
C00135	L-histidine	0.092056
C00062	L-arginine	0.28742
C00065	L-serine	0.209684
C00064	L-glutamine	0.255712
C00041	L-alanine	0.499149
C00079	Phenylalanine	0.180021
C00407	L-isoleucine	0.282306
C00188	L-threonine	0.246506
C00078	L-tryptophan	0.055234
C00047	L-lysine	0.333448
C00049	L-aspartate	0.234232
C00183	L-valine	0.411184
C00097	L-cysteine	0.088988



(a)



(b)



(c)

Figure 4.1: Figure 4.1a shows the composition of the biomass reaction selected for use within the template model [14]. Figure 4.1b shows the composition of the biomass reaction within the target model before transferring information from the template model [15]. Figure 4.1c shows the composition of the biomass reaction within the target model after transferring the information from the template model – as expected, the new composition of the biomass reaction is far more similar to that of the template model.

Figure 4.1 shows that the composition of the biomass reaction within the target model is similar to that of the template model after transferring as much information as possible. The results of the transfer also indicate that a high-quality flat-file for converting other IDs to ChEBI IDs is vital in order to get the best quality transfer. Here, the ChEBI IDs that could not be provided in the flat-file led to a higher proportion of metabolites being left uncategorised and not transferred (refers to 32/51 metabolites within the 'Other' category). The flat-file used here was generated using the metabolite information provided with the template model as well as conversion dictionaries to obtain ChEBI IDs. Inevitably, not all of the non-standard IDs could be converted to ChEBI IDs in this case but it is expected that users with their own models will be able to generate higher quality flat-files than the one generated here.

4.3 Edit mode

To complete the biomass reaction, the two remaining unaltered reactant stoichiometries were changed. These metabolites were tyrosine (one of the metabolites that was not transferred from the protein category, KEGG ID = C00082) and the other was the metabolite used to represent lipid within the biomass of the target model (ID = _a8dc8dc1_b8b0_4206_b692_6f114244bbb8). This lipid is a special metabolite found within the model that has a non-standard ID and it is formed from the contribution of various lipids. This generic lipid is then used as a reactant within the biomass reaction to take into account the contribution of lipids. Refer to table 4.4 for the lipid components identified.

For tyrosine, the stoichiometry within the template biomass reaction was copied directly. In contrast, the stoichiometry of the lipid component was calculated by summing the stoichiometries of the lipids observed within the template model (which were under the 'Other' category).

Table 4.4: *Lipids identified within the biomass reaction of the template model and their stoichiometries. All of these lipids had been characterised within the 'Other' ChEBI category due to the absence of corresponding ChEBI IDs for the conversion flat-file.*

Template lipid ID	**Stoichiometry**
M_pe161_c	0.009618
M_pg161_c	0.004439
M_pe160_p	0.031798
M_pe181_c	0.004957
M_pg160_p	0.004892
M_pe161_p	0.024732
M_pg161_p	0.003805
M_pe160_c	0.012366
M_pg160_c	0.005707
M_pe181_p	0.012747
M_clpn181_p	0.00118
M_pg181_c	0.002288
M_pg181_p	0.001961
M_clpn160_p	0.002944
M_clpn161_p	0.00229

4.4 Gene essentiality assessment and flux balance analysis

Flux balance analysis (FBA) can be used to assess gene essentiality within models – the more closely a given model represents the true biological system, the more closely it's essential genes will match with those observed within wet-lab experiments [13]. In order to assess the improvement in the accuracy of the target model, the newly-curated target model and the original 'raw' model were assessed using FBA and their gene essentiality results were compared to each other and the expected results. The COBRApy package was used to cary out these analyses [16].

4.4.1 Defining the growth medium

The glucose minimal growth medium composition was determined by the Model SEED database of growth media and is shown in table 4.5. Out of the growth medium components, nine were not contained within the target model and could therefore be neglected.

Out of all the other nine components, only two (Fe^{2+} and Fe^{3+}) did not already have import reactions within the model. Therefore, these two import reactions were added. The upper and lower bounds were set to zero for all import reactions for metabolites other than those within the growth medium. In contrast, the upper bounds for all elements of the growth medium were set to 100 (the highest upper bounds within the model).

Table 4.5: *Composition of the glucose minimal medium, showing whether the metabolites and import reactions are present in the target model.*

Metabolite name	**KEGG ID**	**Contained in target model**	**Import reaction in target model**
D-glucose	C00031/C00267/C00293	Yes	Yes
Ca^{2+}	C000076	No	No
Fe^{3+}	C14819	Yes	No
H^{+}	C00080	Yes	Yes
H_2O	C00001/C01328	Yes	Yes
K^{+}	C00238	No	No
Mg	C00305	No	No
Na^{+}	C01330	No	No
NH_3	C01342/C00014	Yes	Yes
Phosphate	C00009	Yes	Yes
Sulfate	C00059	Yes	Yes
O_2	C00007	Yes	Yes
Cl^{-}	C00698/C01327	No	No
Cu^{2+}	C00070	No	No
Co^{2+}	C00175	No	No
Mn^{2+}	C00034	No	No
Zn	C00038	No	No
Fe^{2+}	C14818	Yes	No

4.4.2 Cull metabolites blocking flux

In order to gain accurate predictions of essential genes under specific conditions (e.g. simulating growth in glucose minimal medium), the cell must be able to pass flux through the biomass reaction, otherwise all genes are predicted as essential. Once the target model's biomass reaction has been curated/altered and the growth medium set, it is compulsory to determine which metabolites (if any) within the biomass reaction block cell growth (i.e.

flux through the biomass reaction). This is because it is possible that the target model ma not be able to fully support the production of all the biomass reactants at a high enough level to support biomass production (e.g. some automatically generated metabolic models contain holes in less well characterised areas of metabolism). In order to make sure that the target metabolic model can support the production of its biomass reactants, each one must be tested separately. Here, this was done by selecting each biomass reactant in turn and setting all other biomass reactant stoichiometries to 0. In each of these analyses, if the cell could grow, the selected metabolite would remain in the biomass reactant; conversely, if the model had a growth rate of 0, the selected metabolite would be removed from the biomass reaction. The metabolites that were removed are listed in table 4.6.

Table 4.6: *Metabolites removed from the biomass reaction due to flux blockage (i.e. preventing growth). Most of these metabolites were newly added to the target biomass reaction due to their presence in the template reaction.*

Metabolite name	**KEGG ID**	**Newly added to target reaction**
L-leucine	C00123	No
FAD	C000016	Yes
Thiamin diphosphate	C00068	Yes
Pyridoxal phosphate	C00018	Yes
Ubiquinol	C00390	Yes
Riboflavin	C00255	Yes
Biotin	C00120	Yes
Menaquinol	C05819	Yes

4.5 Results

The predicted essential genes from the raw and altered target model were compared to experimental results (obtained from Orth *et al.*, 2011). The raw model predicted 30 essential genes with a sensitivity of 0.11, specificity of 0.99 and F-measure of 0.20; the altered model predicted 69 essential genes with a sensitivity of 0.24, specificity of 0.98 and an F-measure of 0.37.

4.6 Conclusion

The results show an improvement in sensitivity with very little reduction in specificity in identifying the essential genes in the altered target model compared to its raw counterpart. These results show the successful functionality of the BROOMM software and prove its worth to the wider community of metabolic modellers for aiding biomass reaction curation within metabolic models.

Bibliography

[1] Kanehisa, M. *et al.* (2012) KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Research*, **40**, D109–D114.

[2] Henry, C.S. *et al.* (2010) High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nature Biotechnology*, **28**, 977–982.

[3] Caspi, R. *et al.* 2012, The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Research* **40**, D742 – D753.

[4] Swainston, N. *et al.* The SuBliMinaL Toolbox: automating steps in the reconstruction of metabolic networks. *Journal of Integrative Bioinformatics* **8** (2).

[5] Thiele, I. & Palsson, B. O. 2010, A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocols* **5** (1), 93 – 121.

[6] van Rossum, G. and Drake F.L. (eds), Python Reference Manual, PythonLabs, Virginia, USA, 2001. Available at http://www.python.org.

[7] Bornstein, B. *et al.* (2008) LibSBML: An API Library for SBML. *Bioinformatics* **24**, 880 – 881.

[8] wxPython, 2012. *wxPython.* [online] Available at: http://www.wxpython.org [Accessed 12 August 2013]

[9] NumPy, 2013. *NumPy.* [online] Available at: http://www.numpy.org [Accessed 12 August 2013]

[10] Hunter, J.D. (2007) Matplotlib: A 2D graphics environment. *Computing In Science & Engineering*, **9** (3), 90–95.

[11] Hastings, J. *et al.* (2013) The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. *Nucleic acids research*, **41**(Database issue), D45663. doi:10.1093/nar/gks1146.

[12] Imanishi, T., & Nakaoka, H. (2009). Hyperlink Management System and ID Converter System: enabling maintenance-free hyperlinks among major biological databases. *Nucleic acids research*, **37**(Web Server issue), W1722. doi:10.1093/nar/gkp355.

[13] Orth, J. D., Thiele, I. & Palsson, B. O. 2010, What is flux balance analysis? *Nature Biotechnology* **28** (3), 245 – 248.

[14] Orth, J.D. *et al.* (2011) A comprehensive genome-scale reconstruction of Escherichia coli metabolism–2011. *Molecular Systems Biology*, **7**, 535.

[15] Li, C. *et al.* (2010) BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Systems Biology*, **4**, 92.

[16] Schellenberger, J. *et al.* 2011, Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nature Protocols* **6** (9), 1290 – 1307.