ReMatch: a web-based tool to construct, store and share stoichiometric metabolic models with carbon maps for metabolic flux analysis

Esa Pitkänen[1,4], Arto Åkerlund[1], Ari Rantanen[2], Paula Jouhten[1,3] and Esko Ukkonen[1]

[1]Department of Computer Science, University of Helsinki, Finland

[2]Institute of Molecular Systems Biology, ETH Zurich, Switzerland

[3]VTT Technical Research Centre of Finland

Summary

ReMatch is a web-based, user-friendly tool that constructs stoichiometric network models for metabolic flux analysis, integrating user-developed models into a database collected from several comprehensive metabolic data resources, including KEGG, MetaCyc and CheBI. Particularly, ReMatch augments the metabolic reactions of the model with carbon mappings to facilitate ^{13}C metabolic flux analysis.

The construction of a network model consisting of biochemical reactions is the first step in most metabolic modelling tasks. This model construction can be a tedious task as the required information is usually scattered to many separate databases whose interoperability is suboptimal, due to the heterogeneous naming conventions of metabolites in different databases. Another, particularly severe data integration problem is faced in ^{13}C metabolic flux analysis, where the mappings of carbon atoms from substrates into products in the model are required.

ReMatch has been developed to solve the above data integration problems. First, ReMatch matches the imported user-developed model against the internal ReMatch database while considering a comprehensive metabolite name thesaurus. This, together with wild card support, allows the user to specify the model quickly without having to look the names up manually. Second, ReMatch is able to augment reactions of the model with carbon mappings, obtained either from the internal database or given by the user with an easy-to-use tool.

The constructed models can be exported into 13C-FLUX and SBML file formats. Further, a stoichiometric matrix and visualizations of the network model can be generated. The constructed models of metabolic networks can be optionally made available to the other users of ReMatch. Thus, ReMatch provides a common repository for metabolic network models with carbon mappings for the needs of metabolic flux analysis community.

ReMatch is freely available for academic use at `http://www.cs.helsinki.fi/group/sysfys/software/rematch/`.

1 Introduction

Cellular metabolism has two vital main functions to carry: it produces energy for different cellular operations and provides the cell with precursor molecules for synthesis of more complex

[4]To whom correspondence should be addressed. E-mail: `epitkane@cs.helsinki.fi`

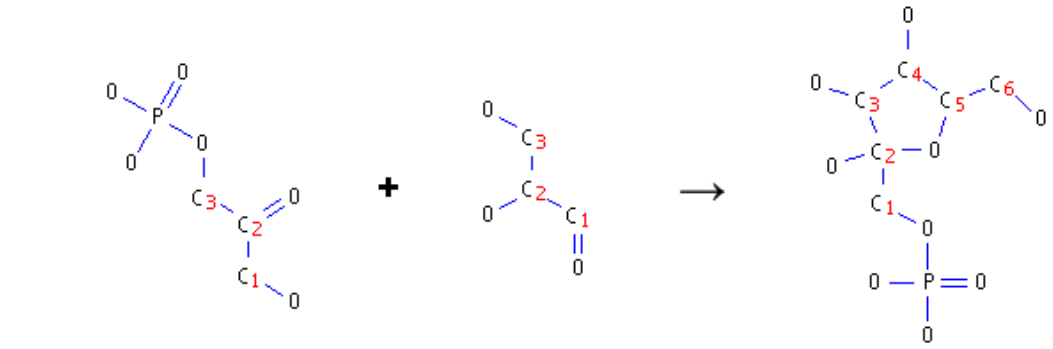

Figure 1: An example of a reaction (fructose 1-phosphate aldolase) with two substrates and a product showing the molecular structures of reactants as well as the reaction formula. EC number (here 4.1.2.13) refers to the Enzyme Commission numbering, which is a functional classification for enzymes [13].

molecules, such as proteins. To carry out these functions efficiently in a changing environment, the cell regulates the velocities of its metabolic reactions, or *metabolic fluxes*, to match its activity to current living conditions. Thus, information about metabolic fluxes provides important knowledge about cellular regulation [1] that can be utilized in many applications, such as in metabolic engineering of microorganisms for better yields of valuable target compounds [2], in improving the plant fitness [3] or in mammalian drug development [4].

The first step in the systematic analysis of metabolism, as in any systems biology research, is to obtain a network model of the system under study. In metabolic modelling, the simplest framework available are the *stoichiometric* models, where the metabolic network is described as a collection of biochemical reactions [5] (see Figure 1 for an example of one such reaction). In particular, each reaction consists of substrate and product metabolites, together with the molar ratios in which these reactants take part in the reaction. As different reactions share common reactants, the reactions span an intertwined network of metabolic pathways. Despite their simplicity, stoichiometric models allow the analysis of global structure and robustness of metabolism [6, 7, 8, 9]. Most importantly, stoichiometric models facilitate different types of analyses related to the metabolic fluxes [10, 11, 12].

Currently, the most accurate experimental method for estimating metabolic fluxes is the ^{13}C metabolic flux analysis [14]. In ^{13}C metabolic flux analysis, a cell is fed with ^{13}C labelled substrates. The ^{13}C labels are distributed in the metabolic network by biochemical reactions. As two metabolic pathways that converge to a common metabolite may manipulate the carbon chains of metabolites differently, the relative abundances of the different ^{13}C labelling patterns may differ in the molecules produced by different pathways. Thus, by measuring the abundances of different ^{13}C labelling patterns from the metabolites, information about the metabolic fluxes can be inferred [14, 15, 16]. For the quantitative estimations of metabolic fluxes from ^{13}C labelling measurements, basic stoichiometric models of metabolism have to be augmented with mappings that describe how carbon atoms are transferred from substrates to products in each reaction (see Figure 2).

Recent work has provided us with new sources of metabolite, metabolic reaction and network data [17, 18]. Unfortunately, integration of these data sources to construct complete models of

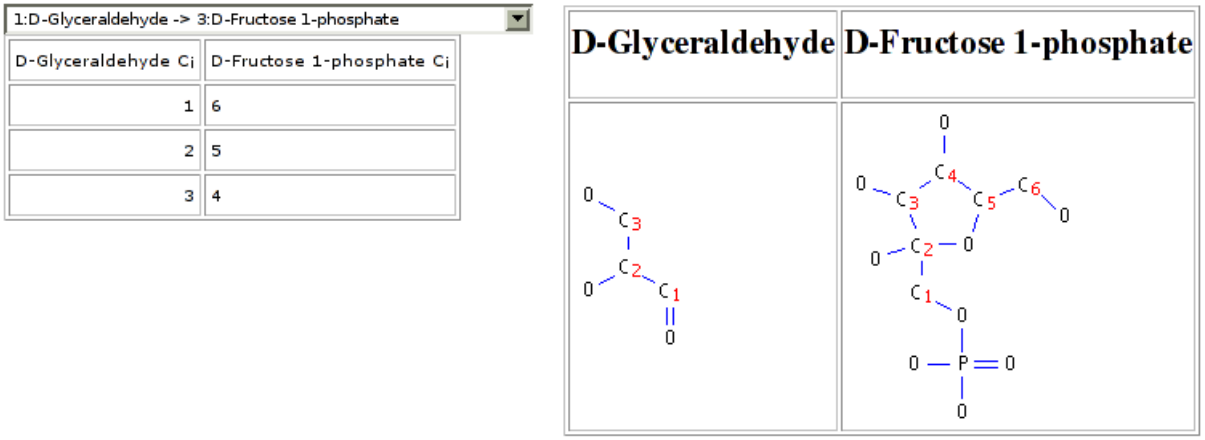

Figure 2: The mapping of carbon atoms between D-Glyceraldehyde and D-Fructose 1-phosphate in the fructose 1-phosphate aldolase reaction (EC 4.1.2.13) as seen in ReMatch. Here the carbons (1, 2, 3) in D-Glyceraldehyde are mapped to carbons (6, 5, 4) in D-Fructose 1-phosphate. ReMatch provides carbon mappings for many reactions through the ARM database, and also easy-to-use tools for adding new user-defined carbon maps.

metabolism is not always seamless. A crucial bottleneck slowing the interoperability is caused by heterogeneous naming conventions for metabolites and genes that encode the enzymes catalyzing metabolic reactions. Recently, cross-linking between different metabolite databases has been improving [19, 20, 21] but this alone does not solve all the problems hindering the easy construction of metabolic models. Additionally, easy-to-use tools for mapping the metabolite nomenclature between two existing user-developed models and between a user-developed model and databases are required. For instance, the metabolite *D-xylulose 5-phosphate* might be called "XU5P-D" in the first model, while another model may use the systematic name. If both names are carelessly used in the same model constructed by combining the models, the stoichiometry of the combined model fails to represent the underlying metabolism as reactions are incorrectly linked to each other through metabolites. Analogously, the mapping of the reactions in the existing user-developed model to databases such as KEGG LIGAND [21], is a tedious task if metabolite name conflicts have to be resolved manually.

The construction of new models for ^{13}C metabolic flux analysis is even more time-consuming as the current metabolic model repositories do not contain curated, good-quality carbon mappings in a systematic fashion [22, 23]. Adding carbon mappings manually to basic stoichiometric models is an extremely error prone task as even small models of central carbon metabolism contain hundreds, even thousands of carbon mappings. On the other hand, recent advancements in the computational analysis techniques [24, 16, 25] and in the techniques for conducting isotopomer measurements [26, 27] would facilitate the use of more detailed models in ^{13}C metabolic flux analysis. These detailed models are not readily available, but have to be constructed before ^{13}C metabolic flux analysis can take place. Thus, new tools are required to rapidly build and share metabolic network models with carbon mapping information.

In this work we concentrate on solving the above data integration problems by introducing *ReMatch*, a web-based tool for constructing, storing and sharing metabolic network models, especially for the needs of metabolic flux analysis.

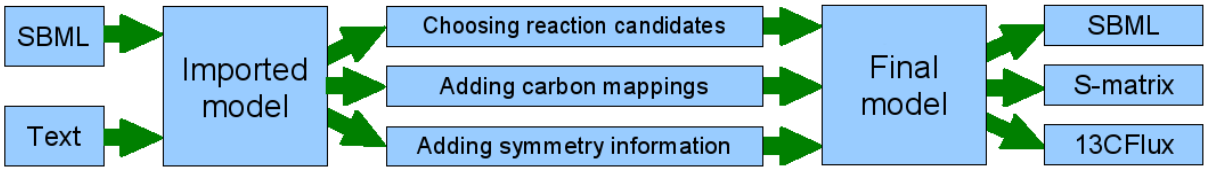

Figure 3: Workflow through the main functionality of ReMatch.

In short, the main workflow of ReMatch consists of the following five steps:

1. Import a model to ReMatch or create a new model. For each metabolic reaction in the model, the user needs to specify the substrate and product metabolites and their molar ratios, or stoichiometric coefficients. The naming of metabolites does not need to be consistent throughout the input file. Thus, the set of reactions can be compiled from multiple sources.
2. Match the user-given reactions to the reactions stored in the database of ReMatch automatically or semi-automatically. Matching resolves conflicts between the nomenclature used in the user-developed model and in the database, and augments reactions with carbon maps, if available.
3. Visualise the metabolic network with the `BMVis` viewer.
4. Export the model in SBML [28], stoichiometric matrix [5] or 13C-FLUX [29] formats to further analyse it or use as a part of a systems biology modelling workflow.
5. (Optionally) Share the model with others by exporting it into a suitable file format or declaring it as a public model in ReMatch.

This workflow is illustrated in Figure 3 and it is discussed, together will the main functionality of ReMatch, in depth in the following subsections.

2 ReMatch

2.1 Importing or creating a new model

As a starting point for model construction the user may either import a model from an SBML formatted file or a text file, copy one of the public models stored in ReMatch, or create a new model from scratch. The imported or created model is stored in ReMatch until the user decides to delete the model. It is particularly easy to switch between the models via the model selection view: the models are persistent, needing no explicit effort from the user to save the model to ensure that the current version stays stored. This enhances the usability of the software, and allows editing multiple models at the same time.

ReMatch supports the SBML level 2 version 3, being able to both import and export SBML formatted models. SBML is a popular XML file format designed for representing biochemical reaction models that is supported by over 100 software tools [30].

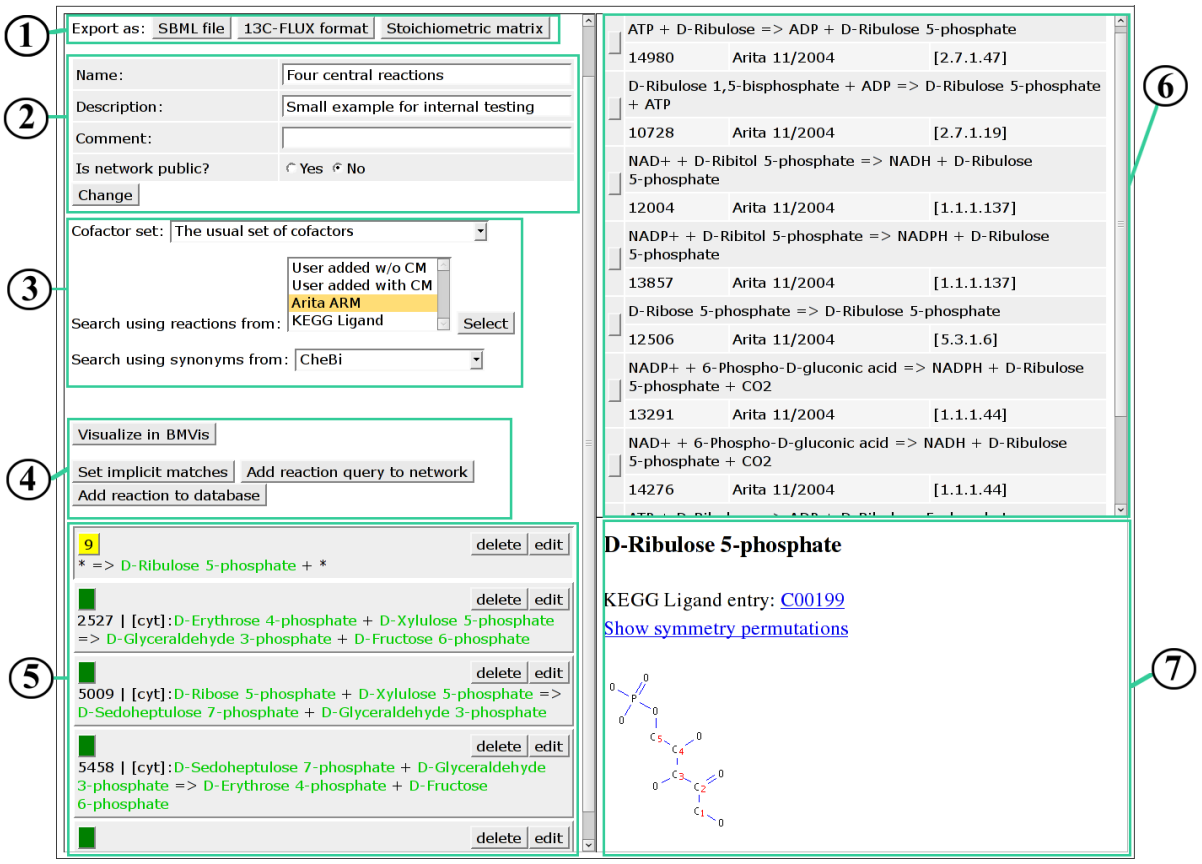

Figure 4: ReMatch reaction list view, showing model export options (1), model description (2), data source options (3), access buttons for visualising the model and adding reactions to the model (4), model reactions (5), candidate reactions for the selected reaction equation (6), and finally the molecular structure of a particular reactant, namely D-Ribulose 5-phosphate (7).

Importing the model as a text file offers a light-weight alternative to SBML import. The model is specified by listing each reaction of the model in one line of the file. Each line specifies the substrates and products of the reaction. An example of such a reaction description is `GLU + ATP => G6P + ADP`. Compartments are supported by prepending reactions or metabolites with tags, such as `cyt` for cytosol (see Figure 4, panel 5). Alternatively, a reaction can be specified with its EC number [13]. ReMatch will then give the user a list of all reactions matching the EC number in the database to choose the correct reaction or reactions from.

Finally, it is also possible to start editing a copy of any of the public models stored in ReMatch. Currently, ReMatch contains the KEGG models for *Saccharomyces cerevisiae*, *Escherichia coli* and *Bacillus subtilis*, as well as a published model containing the central carbon metabolism and amino acid biosynthesis pathways for *S. cerevisiae* [31]. These models, as well as other public models from the users, are available for basis of model construction. A description on how one of the current public models of *S. cerevisiae* was imported and processed in ReMatch is given in Section 3.

2.2 Mapping reactions to database

After a model has been imported into ReMatch, the user needs to resolve any naming ambiguities, if any, remaining after the import step. In this phase, the user-given reactions are mapped

to corresponding reactions in the database. This is a non-trivial task, as the model may have been specified, partly or fully, using a vocabulary not recognized by the public databases integrated into ReMatch. However, ReMatch assists in the matching phase by showing for each user-given reaction all matching reactions in the database. A user-given reaction r matches to a database reaction r' when each reactant in r can be matched to some synonym of reactants in r'. Additionally, we require that no two reactants in r are mapped to the same reactant of r'.

The collection of synonyms, or thesaurus, used by ReMatch was collected from ChEBI [19], KEGG [21] and MetaCyc [20]. At first, we searched for additional synonyms to the KEGG metabolites from ChEBI and MetaCyc, adding a synonym whenever two metabolites in KEGG and either in ChEBI or MetaCyc already shared a name but ChEBI or MetaCyc also had another name for the metabolite. After this step, we further expanded the thesaurus by considering metabolites with similar molecular structures in KEGG and MetaCyc. If two metabolites were found to share a similar structure but no common name, all the names given for the two metabolites in both databases were assigned to the KEGG metabolite in the thesaurus. This way we were able to further extend the thesaurus even in absence of common metabolite names. Particularly, we needed to consider not only exact structural matches because the molecule structures are described in KEGG and MetaCyc at different levels of abstraction, for example with respect to chirality. In such situation, identical molecules in the two databases cannot be exactly matched, but the match can be found when one allows inexact matches as well. The structural comparison was done using the `nauty` library for isomorphism detection [32].

ReMatch will show the number of matching reactions in the database for each user-given reaction (see Figure 4, panel 5). Clicking the reaction will open a list of matching candidate reactions in the upper right panel (Figure 4, panel 6). If the database contains exactly one match, there is no ambiguity to be resolved. If two or more matching reactions were found, the user needs to choose which of these candidate reactions he or she wants to include in the model.

It is also possible that no matching reactions are found for a particular reaction formula. In this case, the user needs to modify the reaction formula to find candidate reactions. Wildcard support provided by ReMatch can be useful in this situation in finding partial matches to user-given reactions. Partial matching helps to resolve metabolite naming ambiguities that remain in the model after vocabulary matching: the correct metabolite can be found from the reaction context. For instance, replacing the `ATP` and `ADP` in `GLU + ATP => G6P + ADP` with question marks `?`, will show all reactions in the database with two substrates, two products, and having Glucose as a substrate and Glucose-6-phosphate as a product. Again, the user is able to choose the correct reaction from the list provided by the software. Asterisk `*` can be given instead of question mark `?`, if the user wants to match the wildcard with one or more reactants instead of exactly one (see Figure 4, panel 5, for an example).

For the purposes of ^{13}C metabolic flux analysis, it is often feasible to limit the search for matching reactions to reactions with carbon maps. ReMatch allows this restriction by letting the user to choose the individual databases where the reaction matches are searched from. In this particular case, we would want to restrict the search to the ARM database [33, 34] that contains the carbon mapped reactions (see Figure 4, panel 3), although it is possible to use also carbon maps from the KEGG database in the future versions of ReMatch.

Finally, it is possible to easily add new reactions into the ReMatch database. The added reactions are available for subsequent searches in a separate database that may be included or

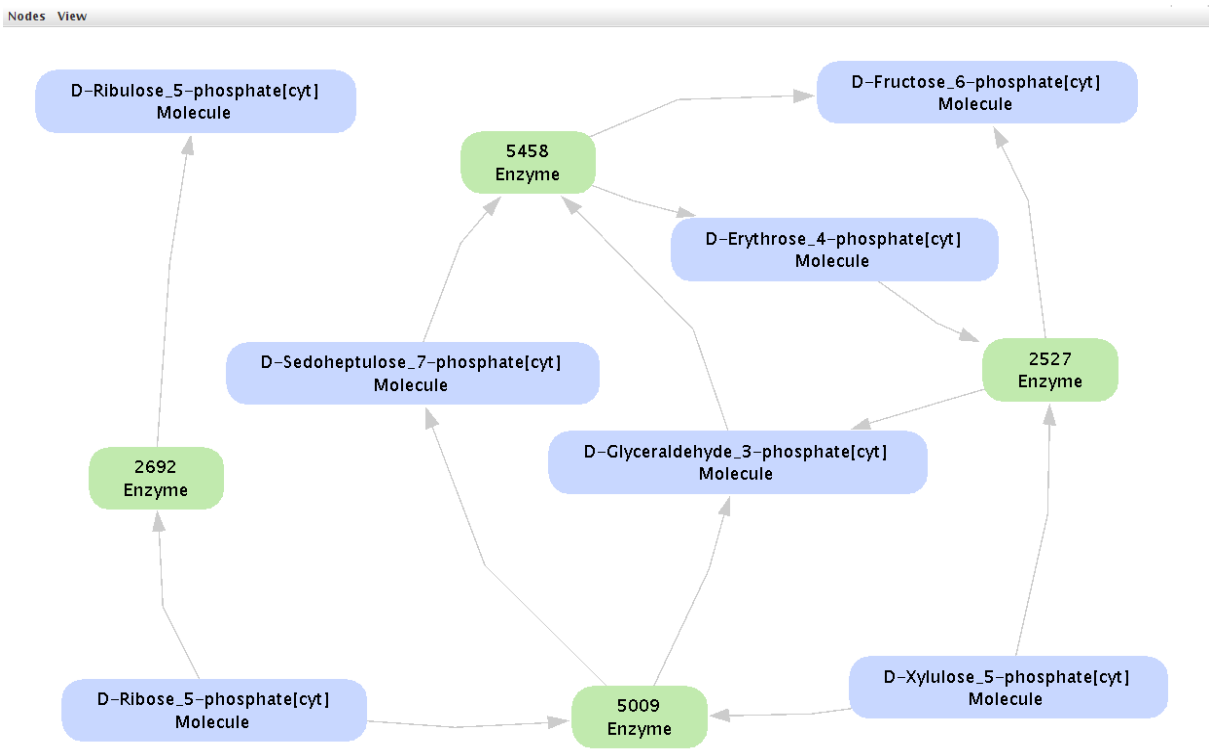

Figure 5: Visualisation of a small example model with four enzymatic reactions (green nodes) and seven metabolites (blue nodes) in the `BMVis` viewer.

excluded with the option described above. ReMatch also provides an easy-to-use tool for manually adding carbon mappings to reactions that currently do not have them (see Figure 2). In ^{13}C metabolic flux analysis, besides carbon mappings, one also needs information on molecule symmetry. The user is able to add new symmetry information to the molecules, if needed.

In summary, the underlying database contains 14013 reactions and 11030 metabolites gathered from the above-mentioned three databases. The metabolite thesaurus holds 43186 synonyms in total, an average of 3.92 synonyms per metabolite. Finally, 5178 reactions have an associated carbon mapping.

2.3 Exploring the model

ReMatch incorporates the `BMVis` graph visualisation tool[1] for quick visual exploration of the metabolic model. The model visualisation allows the user to get a topological overview of the network as well as obtain illustrations by taking screenshots. Figure 5 shows a `BMVis` visualisation of a small model with four reactions.

`BMVis` is an efficient yet easy-to-use Java applet that is launched from ReMatch to a separate tab or window in the browser. The applet allows interactive exploration of the metabolic model by letting the user to drag the metabolites and reactions around while the layout is still in progress. Thus, it is possible by quick inspection to check the model consistency, for example.

Metabolic models are visualised as reaction-metabolite graphs where both reactions and

[1]BMVis, BioMine project, Department of Computer Science, University of Helsinki, `http://www.cs.helsinki.fi/group/biomine/`.

metabolites are represented as nodes. Consumption and production relationships between reactions and metabolites are given by edges between the nodes. Moreover, metabolites in different cellular compartments are drawn as separate nodes, allowing exploration of compartmentalized models as well.

2.4 Sharing and subsequent use of ReMatch models

ReMatch provides several ways to subsequently use the processed models. These features enable the software to both act as a public repository for carbon-mapped metabolic models for ^{13}C metabolic flux analysis and as a part of a metabolic modelling workflow. In particular, ReMatch can be used as the first phase of the workflow where a large model is being put together from smaller models given inconsistent reaction and metabolite name vocabulary.

Firstly, it is possible to share a model with other users of ReMatch by declaring it public in the reaction list view. A public model remains editable only by its creator. However, any user is able to make a copy of a public model and edit the copy. By default all new models are private. A public model may be set private again at any time by its owner. As discussed in Section 2.1, we have provided ReMatch users a set of models to serve as the basis of new models.

Secondly, a model can be exported into SBML format [28]. As currently it is not possible to include carbon mapping information into SBML data, we have used SBML annotation fields to store the mappings augmented to the user-developed model. The encoding used is general and flexible, and able to describe reactions with an arbitrary number of reactants and associated mappings, in contrast to the 13C-FLUX format, where the reaction formula is allowed to contain a maximum of two substrates and two products. The carbon mapping annotation, in addition to the representation of symmetry and molecular structure information, is described in detail on the ReMatch web page. Annotations are used because the SBML level 2 does not support carbon mappings, and this way the carbon mapping information can be added to the SBML file without breaking SBML compliance.

Thirdly, models can be exported into 13C-FLUX compatible text files [29]. 13C-FLUX is a popular computational tool for ^{13}C metabolic flux analysis. However, the amount of manual work involved in writing compatible inputs for 13C-FLUX can be significant due to complexity of carbon maps for large molecules. This phase is automated by the 13C-FLUX export in ReMatch, removing this hindrance from the ^{13}C metabolic flux analysis modelling workflow. ReMatch will also support the file format of the upcoming version of 13C-FLUX [25], as well FluxDirect [16], an upcoming software for ^{13}C metabolic flux analysis developed in the same research group as ReMatch.

Lastly, ReMatch is able to write the stoichiometric matrix [5] corresponding to the model in text format. The matrix can be then read into any general computing package, such as R or MATLAB, for further analysis.

2.5 Software architecture and supported browsers

ReMatch is a web application implemented in Java, which runs on the Apache Tomcat 6.0[2] platform. Java was a logical choice as the implementation language, as Tomcat implements

[2] Apache Tomcat, `http://tomcat.apache.org`

Java Servlets and Java Server Pages. Metabolite name thesaurus was built by using a graph isomorphism library `nauty` [32] to compare the structures of metabolites in KEGG and MetaCyc databases. A relational database management system PostgreSQL[3] is used to store all the data used by ReMatch, including reactions, metabolites, synonyms and models. SBML support is implemented with the libSBML library [35].

ReMatch was developed using Mozilla Firefox 2.0 on Linux as a client platform. Firefox 2.0+ is also the only Linux web browser currently supported. On Microsoft Windows the supported browsers are Mozilla Firefox 2.0+ and Microsoft Internet Explorer 7.0. Future work may be done to support also some other major browsers, such as Opera on Windows and Safari on Mac OSX. ReMatch requires JavaScript to be enabled in the browser.

3 Case study

To demonstrate the practical usefulness of ReMatch, we next describe how we used the software to construct a model of the central carbon metabolism for *S. cerevisiae* [31]. We constructed the model to estimate metabolic fluxes based on isotopomer measurement data from metabolites and biomass amino acids and to generate artificial ^{13}C isotopomer data for testing the framework for ^{13}C metabolic flux analysis under development, FluxDirect [16]. Therefore each reaction of the model had to be associated with a carbon map and the model had to contain the necessary reactions for amino acid synthesis.

We first imported a stoichiometric model of yeast central carbon metabolism into ReMatch as a text file containing a list of 34 reactions [31]. Majority of metabolite names in the imported model were automatically mapped to the corresponding KEGG metabolites by the thesaurus of ReMatch. We added information on carbon position symmetry to symmetrically reacting metabolites such as succinate and fumarate. Since comprehensive carbon mapping of model reactions was required, we limited the search for matching our user-given reactions to the ARM database. For many reactions in the imported model, a corresponding reaction with carbon mappings was readily available.

Often in metabolic modelling, sets of subsequent reactions in a real metabolic pathway are fused together into single reactions, because of the complexity of metabolic networks. In particular, the fused reactions do not correspond to any real biochemical reactions, and hence they are not found from the public databases. Therefore, we manually added the fused reactions found in the imported model and the corresponding carbon mappings to the ReMatch database from where they can be utilized in future models. Also transport reactions with carbon mappings were added in this fashion. We then expanded the model to include the amino acids by inserting fused amino acid synthesis reactions with carbon mappings into the model and the ReMatch database. The manual carbon mapping was facilitated considerably by the side-by-side visualizations of metabolite structures and default carbon mappings provided by ReMatch. We took into account the compartmentalization of yeast metabolism as in the imported model by assigning the isolated metabolite pools with compartment tags.

The final consistent carbon mapped metabolic model included 111 metabolites and 105 reactions, of which 40 were transport reactions. In total, 1165 mapped carbon atoms were associated with the reactions.

[3] PostgreSQL, `http://www.postgresql.org`

The final model stored into ReMatch was utilized for various tasks. For generation of artificial ^{13}C distributions the model was exported into the 13C-Flux file format. Also the stoichiometric matrix of the model, directly exported from ReMatch, was used in the data generation. More importantly, the constructed model was applied to analysis of isotopomer measurement data with FluxDirect [16]. The constructed model was further used as a starting point in construction of metabolic models of *B. subtilis* and *E. coli* by first taking a copy of the model of *S. cerevisiae* and then modifying it. All the above models were made publicly available for users of ReMatch and can also be used as starting points for a quick model construction.

The case study was carried out by an experimentalist not involved in the software development. Based on the study we conclude that ReMatch considerably facilitates the construction of metabolic models for ^{13}C metabolic flux analysis and is user-friendly for experimentalists with basic knowledge of metabolic modeling.

4 Discussion

Existing data integration approaches for metabolic models have focused on developing data warehouses, which collect metabolic data from different databases under integrated nomenclature, such as SABIO-RK [36], BNDB [37], Reactome [38], and model repositories, such as JWS Online [22] and BioModels Database [23], which store user-developed models but not necessarily using a unified nomenclature. In contrast, ReMatch was developed to act as a model repository that uses a unified nomenclature, improving reusability of stored metabolic models. Importantly, ReMatch is able to both read and write models in the popular SBML format which allows the software to work as a part of a systems biology workflow.

To the authors' knowledge, ReMatch is the first web-based tool which enables integration of carbon mapping information into user-developed metabolic models for ^{13}C metabolic flux analysis purposes. Moreover, we believe that the extensive metabolite thesaurus and wildcard-assisted reaction matching will prove useful when converting any metabolic model to use standardized metabolite nomenclature. This has the immediate benefit of being able to access data on reactions and metabolites that have been stored in the matched database, such as KEGG.

We argue that the ability to share the models with the ^{13}C metabolic flux analysis community is a very important feature of ReMatch. Online repositories have been developed for SBML models [22, 23], but they do not contain carbon mapping information. Thus they are not immediately usable for ^{13}C metabolic flux analysis. From ReMatch, on the other hand, public models can be immediately exported into file formats compatible with popular ^{13}C metabolic flux analysis software. In this fashion, researchers can easily test their carbon labelling data with the published models to obtain estimations of the activities of reactions in the model, without the need of going through the laborous step of model construction.

The software is easy and fast to use: due to the intuitive user interface, the user can export a stoichiometric metabolic network model of a realistic size with carbon mappings and consistently named metabolites within minutes after importing a set of reactions to ReMatch. The process is accelerated further by the possibility to take a copy of an existing public network model as the basis of the new model. Thus, it is often not necessary to start building new models for ^{13}C metabolic flux analysis from scratch.

ReMatch is constantly being updated, reflecting the feedback and needs of users. Some of the planned features include more fine-grained access control to model sharing and better visualisation options. In the future we will also provide the carbon maps from the KEGG database to the users of ReMatch.

Acknowledgments

The authors would like to thank the Biomine project, Petteri Hintsanen, Kimmo Kulovesi and Hannu Toivonen at the Department of Computer Science, University of Helsinki, for the `BMVis` network visualisation tool and assistance in integrating it into ReMatch, and Markus Heinonen for suggesting the name for the software.

References

[1] E. Fischer and U. Sauer. Large-scale in vivo flux analysis shows rigidity and suboptimal performance of Bacillus subtilis metabolism. *Nature Genetics*, 37(6):636–640, 2005.

[2] K. Sanford, P. Soucaille, G. Whited, and G. Chotani. Genomics to fluxomics and physiomics – pathway engineering. *Current Opinion in Microbiology*, 5(3):318–322, 2002.

[3] J. Schwender, F. Goffman, J. Ohlrogge, and Y. Shachar-Hill. Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds. *Nature*, 432:779–782, 12 2004.

[4] M. Hellerstein and E. Murphy. Stable isotope-mass spectrometric measurements of molecular fluxes in vivo: emerging applications in drug development. *Current Opinion on Molecular Therapeutics*, 6(3):249–264, 2004.

[5] G. N. Stephanopoulos, A. A. Aristidou, and J. Nielsen. *Metabolic Engineering, Principles and Methodologies*. Academic Press, 1998.

[6] J. Stelling, S. Klamt, K. Bettenbrock, S. Schuster, and E. D. Gilles. Metabolic network structure determines key aspects of functionality and regulation. *Nature*, 420:190–193, 2002.

[7] A. P. Burgard, E. V. Nikolaev, C. H. Schilling, and C. D. Maranas. Flux coupling analysis of genome-scale metabolic network reconstructions. *Genome Research*, 14(2):301–312, 2004.

[8] J. Stelling, U. Sauer, Z. Szallasi, F. J. Doyle, and J. Doyle. Robustness of cellular functions. *Cell*, 118:675–685, 2004.

[9] V. Hatzimanikatis, C. Li, J. A. Ionita, C. S. Henry, M. D. Jankowski, and L. J. Broadbelt. Exploring the diversity of complex metabolic networks. *Bioinformatics*, 21(8):1603–1609, 2005.

[10] A. Varma and B. O. Palsson. Metabolic flux balancing: Basic concepts, scientific and practical use. *Bio/Technology*, 12(10):994–998, 1994.

[11] S. Schuster, T. Dandekar, and D. A. Fell. Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends in Biotechnology*, 17(2):53–60, 1999.

[12] K. J. Kauffman, P. Prakash, and J. S. Edwards. Advances in flux balance analysis. *Current Opinion in Biotechnology*, 14(5):491–496, 2003.

[13] IUBMB. *Enzyme Nomenclature*. Academic Press, 1992.

[14] W. Wiechert. An introduction to ^{13}C metabolic flux analysis. *Genetic Engineering*, 24:215–238, 2002.

[15] E. Fischer, N. Zamboni, and U. Sauer. High-throughput metabolic flux analysis based on gas chromatography-mass spectrometry derived ^{13}C constraints. *Analytical Biochemistry*, 325(2):308–316, 2004.

[16] A. Rantanen, J. Rousu, P. Jouhten, N. Zamboni, E. Maaheimo, and E. Ukkonen. An analytic and systematic framework for estimating metabolic flux ratios from ^{13}C tracer experiments. *BMC Bioinformatics*, 9(266):266, 2008.

[17] J. Forster, I. Famili, P. Fu, B. O. Palsson, and J. Nielsen. Genome-scale reconstruction of the Saccharomyces cerevisiae metabolic network. *Genome Research*, 13(2):244–253, 2003.

[18] I. M. Keseler et al. EcoCyc: A comprehensive database resource for Escherichia coli. *Nucleic Acids Research*, 33:D334–D337, 2005.

[19] K. Degtyarenko, P. de Matos, M. Ennis, J. Hastings, M. Zbinden, A. McNaught, R. Alcántara, M. Darsow, M. Guedj, and M. Ashburner. ChEBI: a database and ontology for chemical entities of biological interest. *Nucleic Acids Research*, 36:D344–D350, 2008.

[20] R. Caspi, H. Foerster, C. A. Fulcher, P. Kaipa, M. Krummenacker, M. Latendresse, S. Paley, S. Y. Rhee, A. G. Shearer, C. Tissier, T. C. Walk, P. Zhang, and P. D. Karp. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Research*, 36:D623–631, 1 2008.

[21] M. Kanehisa, M. Araki, S. Goto, M. Hattori, M. Hirakawa, M. Itoh, T. Katayama, S. Kawashima, S. Okuda, T. Tokimatsu, and Y. Yamanishi. KEGG for linking genomes to life and the environment. *Nucleic Acids Research*, 36:D480–D484, 2008.

[22] J. L. Snoep and B. G. Olivier. Java Web Simulation (JWS): a web based database of kinetic models. *Molecular Biology Reports*, 29:259–263, 2002.

[23] Nicolas Le Novere, Benjamin Bornstein, Alexander Broicher, Melanie Courtot, Marco Donizelli, Harish Dharuri, Lu Li, Herbert Sauro, Maria Schilstra, Bruce Shapiro, Jacky L. Snoep, and Michael Hucka. BioModels database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Research*, 34(suppl_1):D689–691, 2006.

[24] M. Antoniewicz, J. Kelleher, and G. Stephanopoulos. Elementary Metabolite Units (EMU): a novel framework for modeling isotopic distributions. *Metabolic Engineering*, 9(1):68–86, 2006.

[25] M. Weitzel, W. Wiechert, and K. Nöh. The topology of metabolic isotope labeling networks. *BMC Bioinformatics*, 8(315):315, 2007.

[26] K. Nöh, K. Grönke, B. Luo, R. Takors, M. Oldiges, and W. Wiechert. Metabolic flux analysis at ultra short time scale: Isotopically non-stationary ^{13}C labeling experiments. *Journal of Biotechnology*, 129(2):249–267, 2007.

[27] W. van Winden, J. van Dam, C. Ras, R. Kleijn, J. Vinke, W. Gulik, and J. Heijnen. Metabolic-flux analysis of Saccharomyces cerevisiae CEN.PK113-7D based on mass isotopomer measurements of ^{13}C-labeled primary metabolites. *FEMS Yeast Research*, 5(6–7):559–568, 2005.

[28] M. Hucka, A. Finney, H. Sauro, H. Bolouri, J. Doyle, et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19:524–531, 2003.

[29] W. Wiechert, S. Möllney, Petersen, and A. de Graaf. A universal framework for ^{13}C metabolic flux analysis. *Metabolic Engineering*, 3(3):265–283, 2001.

[30] M. Hucka, A. Finney, B. J. Bornstein, S. M. Keating, B. E. Shapiro, J. Matthews, B. L. Kovitz, M. J. Schilstra, A. Funahashi, J. C. Doyle, and H. Kitano. Evolving a lingua franca and associated software infrastructure for computational systems biology: the Systems Biology Markup Language (SBML) project. *Systems Biology*, 1(1):41–53, 2004.

[31] L. M. Blank, L. Kuepfer, and U. Sauer. Large-scale 13C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome Biology*, 6(6):R49, 2005.

[32] Brendan D. McKay. *nauty User's Guide*. Department of Computer Science, Australian National University, 2007.

[33] Automatic reconstruction of metabolism. `http://www.metabolome.jp/index.html`.

[34] M. Arita. In silico atomic tracing of substrate-product relationships in Escherichia coli intermediary metabolism. *Genome Research*, 13(11):2455–2466, 2003.

[35] B. J. Bornstein, S. M. Keating, A. Jouraku, and Hucka M. LibSBML: An API library for SBML. *Bioinformatics*, 24(6):880–881, 2008.

[36] O. Krebs, M. Golebiewski, R. Kania, S. Mir, J. Saric, A. Weidemann, U. Wittig, and I. Rojas. SABIO-RK: A data warehouse for biochemical reactions and their kinetics. *Journal of Integrative Bioinformatics*, 4(1), 2007.

[37] J. Kuentzer, C. Backes, T. Blum, A. Gerasch, M. Kaufmann, O. Kohlbacher, and H.-P. Lenhof. BNDB – the biochemical network database. *BMC Bioinformatics*, 8(367):367, 2007.

[38] I. Vastrik et al. Reactome: a knowledge base of biologic pathways and processes. *Genome Biology*, 8:R39, 2007.