MetExplore: collaborative edition and exploration of metabolic networks

Ludovic Cottret[1,*], Clément Frainay[2], Maxime Chazalviel[2,3], Floréal Cabanettes[1], Yoann Gloaguen[4,5,6], Etienne Camenen[2], Benjamin Merlet[2], Stéphanie Heux[7,8,9], Jean-Charles Portais[7,8,9], Nathalie Poupin[2], Florence Vinson[2] and Fabien Jourdan[2,*]

[1]LIPM, Université de Toulouse, INRA, CNRS, Castanet-Tolosan, France, [2]INRA, UMR1331, Toxalim, F-31000 Toulouse, France, [3]MedDay Pharmaceuticals, Paris, France, [4]Berlin Institute of Health Metabolomics Platform, 10178 Berlin, Germany, [5]Core Unit Bioinformatics, Berlin Institute of Health, 10178 Berlin, Germany, [6]Max Delbrück Center for Molecular Medicine, 13125 Berlin, Germany, [7]Université de Toulouse; INSA, UPS, INP; LISBP, 135 Avenue de Rangueil, F-31077 Toulouse, France, [8]INRA, UMR792, Ingénierie des Systèmes Biologiques et des Procédés, F-31400 Toulouse, France and [9]CNRS, UMR5504, F-31400 Toulouse, France

ABSTRACT

Metabolism of an organism is composed of hundreds to thousands of interconnected biochemical reactions responding to environmental or genetic constraints. This metabolic network provides a rich knowledge to contextualize omics data and to elaborate hypotheses on metabolic modulations. Nevertheless, performing this kind of integrative analysis is challenging for end users with not sufficiently advanced computer skills since it requires the use of various tools and web servers. MetExplore offers an all-in-one online solution composed of interactive tools for metabolic network curation, network exploration and omics data analysis. In particular, it is possible to curate and annotate metabolic networks in a collaborative environment. The network exploration is also facilitated in MetExplore by a system of interactive tables connected to a powerful network visualization module. Finally, the contextualization of metabolic elements in the network and the calculation of over-representation statistics make it possible to interpret any kind of omics data. MetExplore is a sustainable project maintained since 2010 freely available at https://metexplore.toulouse.inra.fr/metexplore2/.

INTRODUCTION

Metabolism modulation is a key systemic answer to biotic and abiotic stresses. To decipher these metabolic adaptations, global and non-targeted approaches such as metabolomics, proteomics and transcriptomics are cornerstone technologies. Nevertheless, the interpretation of the generated omics data requires taking into account the complex interplay between metabolites and reactions. This relational information is encoded in genome scale metabolic network reconstructions, which aim at gathering and organizing synthetic knowledge regarding the metabolic capacities of an organism, especially the properties that can be used for modeling (e.g. graph (1) or flux balance analyses (2)). Such networks include all the reactions an organism can perform, metabolites which are substrates and products of these reactions and the genes coding for the enzymes catalyzing these reactions. MetExplore has been specifically designed as an 'all-in-one' easy-to-use solution to help biologists in exploiting this valuable information in order to interpret multi-omics data in a more systemic and holistic perspective. Genome scale metabolic networks are usually built in a bottom-up manner, starting from the annotated genome of an organism and retrieving enzymatic reactions it can achieve. This process can be performed in a semi-automatic manner using dedicated tools (3–5). However, a tedious and time-consuming manual curation step is often necessary to complete and refine this reconstructed network before it can be used for data interpretation (6). Besides, a growing number of dedicated programs offer different approaches to interpret omics data in the context of metabolic pathways and networks. Visualization tools (4,7–9), enabling the location of biological entities of interest in a network, over-representation tools (10), aiming to identify metabolic pathways with a higher representation in a dataset, and flux balance modelling tools (11,12), allowing hypothesis generation and testing of system responses to genetic or environmental changes, present a large range of features. However, most of the available tools are dedicated solely to one specific analysis and a single type of

data. One of the main advantages of MetExplore is that it provides an end-to-end solution for such analyses. MetExplore thus constitutes a flexible solution to build tailored metabolic networks, either from scratch or by curating existing networks. It greatly facilitates this curation process by providing collaborative tools to curate and manage networks. MetExplore also constitutes a comprehensive solution for the interpretation of omics data into the context of metabolism, by contextualizing biological entities of interest in the metabolic network and by switching between a global overview (entire network) and detailed (sub-networks) analyses. To facilitate this data mining loop, MetExplore provides a modular environment enabling simultaneous network representations in the form of grids (spreadsheet-like views of reactions, genes, metabolites . . .) and interactive network visualization, seamlessly propagating user interactions from one view to another, and allowing focusing on parts of interest using data-based filters.

MetExplore project was first released in 2010 (13) and has undergone regular improvements in its architecture, database and functionalities. It has been used for a large range of research projects, among which network reconstruction projects (14,15), analyses of metabolomics data in toxicological studies (16), and analyses of gene expression data in plants (17). Two use cases were described in this article to illustrate MetExplore main features.

COLLABORATIVE CURATION AND ANNOTATION OF METABOLIC NETWORKS

Automatic reconstruction tools from genomes (5,18,19) and semi-automatic gap-filling methods (20,21) produce draft reconstructions which need intensive manual curation by several experts (6). MetExplore offers curation and annotation functions which complement these automatic methods. Figure 1 summarizes these MetExplore functions from network import to the export of a curated and better quality metabolic reconstruction. Metabolic network curation being a collaborative process, MetExplore also provides an online environment to easily share this manual duty.

Importing a metabolic network

The exchange format commonly used for metabolic networks is the Systems Biology Markup Language (SBML) (22). Many network repositories like BIGG (23) or BioModels (24) offer to download metabolic networks in that format. However, the generated SBML files can be formatted in significantly different ways depending on the SBML version used. Furthermore, metabolic elements such as pathways or gene-reaction links may be described in non-standard data fields which vary from one SMBL provider to another one. MetExplore thus provides a very versatile SBML import web interface which allows importing the latest SBML versions with the Flux Balance Constraints (FBC) package as well as SBML files with specific annotation. Patterns corresponding to the format of the information stored in the SBML can thus be defined by the user (see Supplementary Figure S1).

MetExplore also provides a tool to convert KEGG pathway maps to MetExplore networks. A selection of metabolic

networks (called BioSources in MetExplore) is publicly available. When a user imports a network, it is by default added in his (her) private account.

Additionally, MetExplore offers the opportunity to create a personal copy of public metabolic networks stored in MetExplore, which can then be modified and tailored by the users. MetExplore therefore provides a way to manage working copies of metabolic networks which can finally be exported in SBML files. Easy creation of SBML files after network curation will contribute to facilitate the submission in public repositories.

Sharing metabolic networks in a collaborative project

Metabolic network reconstruction projects often require the intervention of several collaborators who can be spread in various laboratories and countries. In MetExplore, to bridge this gap, private metabolic networks can be shared using the collaborative 'Project' functionality. To ensure proper management of collaborative reconstruction projects, collaborators can be granted different curation rights from the read-only role to the ability of adding or editing network elements (see Supplementary Figure S2). Collaborative curation of metabolic networks by several annotators is further facilitated by several project management tools, such as the capacity to assign specific tasks to annotators, with deadlines and priority levels (Supplementary Figure S3).

Finally, following the guidelines of the framework of reproducibility in science, MetExplore stores and makes available curation history for each reconstruction project (Supplementary Figure S4).

Sharing annotations on network elements: votes and comments

Manual curation of a metabolic network often begins by reviewing each element starting from the higher-level ones (pathways, compartments) to the constitutive elements of metabolism (metabolites, reactions, genes). In MetExplore, before editing network elements, annotators can provide feedback regarding the accuracy or veracity of an entry, through an Approval/Disapproval voting system (Supplementary Figure S5). For each type of network element, a summary of all votes is accessible in the corresponding grid (e.g. reactions), allowing sorting and filtering to focus on problematic cases (Supplementary Figure S6). Votes can then be confronted to reach a consensus about the annotation of a network element. It is also possible to add comments to give additional pieces of information about a reaction.

Users have the possibility to attach files (e.g. experimental data) to complete comments. Bibliographic references added from PubMed identifiers (PMID) make it easy to find the source which confirms the presence of specific reactions.

Editing a metabolic network

The edition functions of MetExplore allow metabolic elements to be modified, removed or added in private networks. A collection of network elements can be removed
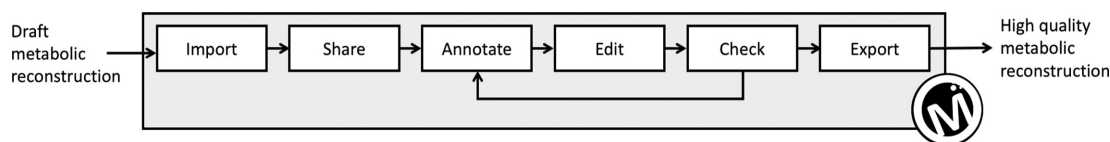
Figure 1. Metabolic network curation steps into MetExplore.

all at once using multi-selection in grids. Adding or connecting new network elements is performed by simple online forms. Specifically, the creation of a metabolite is facilitated by auto filling data fields like InChI identifier (25), formula, or molecular weight, when a KEGG or ChEBI identifier (26) is provided (Supplementary Figure S7).

Metabolic reconstruction validation

A metabolic network can contain thousands of elements tightly interconnected. It is therefore important to be able to visually check the coherence of modified parts of the network. To this end, MetExplore offers a visualization system allowing sub-networks from a selection of network elements (e.g. a pathway, a set of genes) to be represented. Gene-Protein-Reaction (GPR) links can also be visually reviewed using the dedicated graphical representation provided in the reaction information cards (Supplementary Figure S8).

Further than visual inspection, it is necessary for flux modelling to check that the created network can carry fluxes and achieve essential metabolic functions. MetExplore flux computation features make it possible to check the flux-consistency of the metabolic model under curation. The identification of dead-end reactions and orphan metabolites allows the detection of incomplete parts of the network. Flux Balance Analysis (27), Flux Variability Analysis (28) and reaction or gene knockout analyses enable the comparison between simulations and experimental evidences. This step may lead in the identification of inconsistent parts of the metabolic network which can be further edited within MetExplore. Hence, MetExplore users can perform this iterative curation loop in the same environment hence saving time and avoiding file transfers between tools.

Exporting a metabolic network

Each network can be exported as SBML or Excel files. Generated SBML files can then be uploaded in metabolic model repositories (e.g. BioModels) or used in analysis tool boxes (e.g., COBRA Toolbox (11)). Excel generated files correspond exactly to the grids of the network elements as displayed in the MetExplore interface at the time of the export. For instance, if the user has performed a data mapping, as described in the next section, all the corresponding columns will be available in the Excel file. These files are easier to read and can be used as tabulated files for other kinds of computations.

DYNAMIC AND INTERACTIVE EXPLORATION

Data mapping

Metabolic networks contain knowledge on genes, proteins, reactions and small molecules (metabolites) related to metabolism. Hence it provides a relevant biochemical context to interpret omics data. MetExplore offers a simple way to map data on a given metabolic network: the user can easily copy and paste data from common spreadsheet editors or simple tabulated text files. Mapping can be achieved for all network elements based on names or network identifiers. Additionally, for metabolites, mapping can be made based on monoisotopic mass, InChI or InChIKey.

The mapping results are reported in all grids: for mapped elements, a column containing the corresponding input values is added for each condition. Moreover, over-representation analyses are performed, using hypergeometric tests (corrected with Bonferroni and Benjamini Hochberg methods), and reported in the pathway grid to facilitate future interpretation of the mapping results.

Smart filters

MetExplore strongly facilitates the exploration of metabolic network contents by providing spreadsheets (grids) for all constitutive elements of the network (from compartments to genes). These grids offer search, sort and selection functions. Users often explore network content by establishing links between elements, for instance one may want to know in which reactions and pathways a metabolite is involved. Hence, MetExplore implements a filtering function allowing all grid contents to be dynamically changed based on a selection from a particular grid.

These smart filters are of particular interest when analyzing data mapping results. In fact, since mapping provides pathway enrichment, users can then filter grid contents based on the selection of significantly enriched metabolic pathways. This filtering step allows focusing on a subnetwork of interest which can then be further mined using visualization (see Use Case Section).

Visualization and subnetwork extraction

Network visualization is one of the main features in MetExplore. It is possible to create a graphical representation of a network based on any selection of reactions. For instance, one can visualize all network reactions at once or select the reactions involved in a combination of pathways and visualize how they interconnect. Representation is achieved using MetExploreViz web module (29) which allows data mapping on reactions and metabolites, side compound duplications (e.g. water, ATP), compartments and pathways highlighting, edition, graphical exports (e.g. SVG, png, jpeg), graph format exports (e.g. gml).

Visual data analysis in the context of large networks is a challenging task when networks contain hundreds of elements (reactions and metabolites). Methods have been developed to extract subnetworks using network topology

(graph based methods) (30). Among these methods, lightest paths (31) allow to automatically find metabolic paths between pairs of metabolites of interest. For a list of metabolites, a subnetwork can be extracted by putting together lightest paths between each pair of elements in the list. This algorithm is implemented in MetExploreViz and allows reducing visual complexity of metabolic networks to facilitate analysis.

MetExplore being an integrated solution, it is also possible to select elements in the network representations and see where they stand in the grids.

USE CASES

MetExplore as shared platform to accelerate and facilitate Jamborees

The metabolic network management capabilities of MetExplore are very well adapted to the organization of a reconstruction annotation jamboree. The goal of a jamboree is to reconcile and refine knowledge about the metabolism of an organism (32). In order to refine the metabolic reconstruction of the bacterium *Agrobacterium fabrum*, a Jamboree in two sessions of two days has been organized. The original metabolic network was built using the MAGE platform (33) and publicly available in MetExplore. A total of 11 participants were involved in the two jamboree sessions. One participant, with more advanced bioinformatics skills, was designated as the project manager. He created the project, imported the metabolic network and managed the rights of the other participants.

Depending on their skills, different roles were assigned to the other participants. For example, those who had the broadest knowledge of the bacterium were responsible for voting for the presence/ accuracy of each metabolic pathway and reaction and for commenting on certain parts of the metabolism. Other participants were in charge of the addition of metabolic pathways containing reactions absent from the original network.

The statistics after these four days show the benefit of using MetExplore to organize Jamborees and improve their outcomes. Each of the 391 metabolic pathways received at least one vote. More than 1600 votes and 250 comments had been added to pathways or reactions. Four compartments, 27 metabolic pathways, 149 reactions, 76 metabolites and 56 enzymes have been integrated into the metabolic network. Annotators highly appreciated the network and Gene-Protein-Reaction visualizations to check the links between the elements they just added.

Two Jamboree sessions were not enough to finalize the metabolic reconstruction. However, collaborators are now autonomous to proceed further with the reconstruction in their respective workplaces. Moreover, pruning or correcting the metabolic elements, a task devoted here to the project manager, will be highly facilitated by the votes and comments. The logs (history) of edition actions performed on the metabolic network will also help the project manager to detect annotation errors. At last, centralizing all annotations in MetExplore will facilitate the resumption of reconstruction work in future sessions.

Network exploration to decipher Yeast metabolic response to Cadmium exposure

MetExplore is largely surveyed as a key server for metabolomics data analysis (10,34–36), in particular to biochemically interpret metabolic fingerprints (list of significant metabolites for a given experimental condition). To highlight this feature, we will use a dataset extracted from a paper by Madalinski *et al.* (37). The aim of the experiment described in this article was to decipher yeast metabolic modulations induced by cadmium (toxicological metal) exposure. The list of metabolites, obtained using direct infusion high resolution mass spectrometry, contains 10 metabolites for which concentrations have significantly increased or decreased between control and exposed cells. Following this metabolomics study, MetExplore has been used to make sense of this list in terms of metabolic modulations.

Several networks in MetExplore are available for *Saccharomyces cerevisiae*, to exemplify the exploration pipeline we use KEGG network. Metabolites belonging to the fingerprint can be listed in an Excel spreadsheet (see supplementary material 2) and copy pasted in the MetExplore mapping tool. Once the mapping is achieved, pathway enrichment is computed to associate a corrected p-value to each pathway (see Figure 2.1). By applying MetExplore smart filters, user can select the significant pathways and filter all grids based on this selection. Hence, the filtered reaction grid will then only contain reactions involved in at least one of the selected pathways (see Figure 2.2). These reactions can then be put together to create a network corresponding to the union of all selected pathways. The network can be displayed in MetExplore and fingerprint metabolites can be highlighted using various colour schemes (see Figure 2.3). This network still contains few hundreds elements (reactions and metabolites) making visual interpretation challenging. MetExplore graph based sub-network extraction method can be applied to get the core part connecting metabolites in the fingerprint belonging to these pathways.

This pathway-based method provides a first overview of the metabolic modulation. However, it does not include all metabolites from the fingerprint (8 out of 10). In fact, two metabolites of interest are isolated in other pathways which are not found to be significantly enriched. It is the case for glutamyl-cysteine for which the concentration fold change is larger than 200. To include all metabolites into the interpretation, user can apply the sub-network extraction approach to the entire network by selecting all reactions. It results in a sub-network (see Supplementary Figure S9) containing all metabolites from the fingerprint.

Once the analysis is achieved, it is possible to export all data in an Excel spreadsheet. If filters have been applied, only selected elements will be exported in the spreadsheet. Visualisation output can also be exported in various formats for publication purposes.

In the original article, Madalinski and co-authors suggest a biochemical interpretation of metabolomics fingerprint which does not involve changes in methionine concentration. Using MetExplore, it was possible to explore the metabolic modulations in a holistic manner and to propose
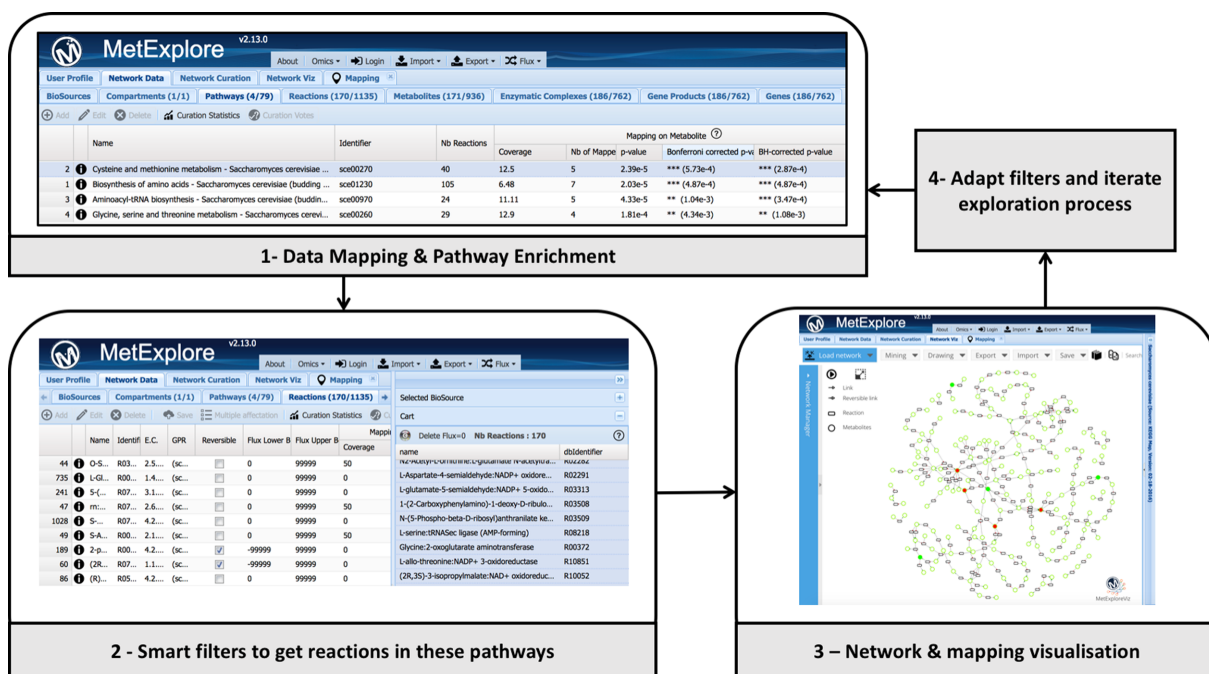
Figure 2. Analysis cycle combining omics data mapping and sub-network extraction from grids and network visualisation.

possible biochemical connections between methionine decrease and glutathione biosynthesis increase in response to cadmium exposure. MetExplore all-in-one solution allows an iterative process between data filtering and visualization which leads to enhanced interpretation fully exemplifying the benefit of using genome scale networks in that context.

This example focuses on metabolomics data analysis, for which MetExplore is often used. However, the new version of MetExplore makes it possible to carry out the same type of analysis, whatever the type of input omics data (transcriptomics, proteomics, fluxomics …). For instance, Jardinaud *et al.* used MetExplore to predict which metabolic pathways were regulated during the nodulation in legumes from transcriptomics data (17).

DISCUSSION

MetExplore is an integrative web platform that brings together several features, which had never been put together in the same framework up to now. Its initial and still widely used functionality is the interpretation of metabolomics data through the analysis of a metabolic network at the genome level. However, since the first release, new features have greatly expanded the use of MetExplore. Figure 3 summarizes the features in MetExplore and compares them with those found in other web servers dedicated to metabolism analysis (4,5,8,9,23,38–41). We focused our comparison on web servers providing several similar functionalities as the ones available in MetExplore. For the pathway enrichment topic specifically, a recent review compared several tools including MetExplore (10). These tools are very efficient and informative for specific tasks but none fulfill simultaneously the pipeline from network curation to data analysis.

The organization of metabolic data into grids makes the use of MetExplore intuitive for biologists who generally explore their data with Excel. Moreover, thanks to its dynamic filters between the different grids, absent in the previous version, MetExplore greatly facilitates the exploration of the metabolic network content. The subnetwork corresponding to the filters applied to the grids can be directly visualized in the form of a bipartite graph linking reactions to metabolites. Since version 2.0, metabolism-specific drawing options can reduce visualization complexity and add information about metabolic pathways or compartments. Different types of omics data (metabolomics, proteomics, transcriptomics) can now be mapped on the same network, and the mapping results appear both on the network drawing and in the grids where enrichment values are displayed. As for mapping, since version 2.0, flux analyses can be launched on the whole network and results displayed only on a sub-network, which allows the user to focus on specific metabolic processes. Filtered grids with additional columns corresponding to analysis results can be exported in tabulated files and then easily shared or published. MetExplore is designed to allow users with no coding skills to perform omics data analysis. By providing an export in tabulated format, MetExplore offers the opportunity to perform more advanced data analysis methods using other platforms like R.

The mapping functions of MetExplore depend on identifiers provided by the metabolic databases. The lack of standardization of identifiers makes this task sometimes difficult. For metabolite mapping, to overcome the lack of a unified identifier system shared by all databases, the mapping can be performed in MetExplore using InChI strings or InChI keys, which provide a unique identifier for each metabolite, based on chemical structures. This is also true

		MetExplore 2018	MetExplore 2010	MetaNetX	Reactome	Kbase	BioCyc	VMH	BIGG	MetaboAnalyst	SMPDB	KEGG
SCOPE	Multi-organisms	●	●	●		●	●		●	●		●
	Pathway-focused	●	●		●		●			●	●	●
	Network-focused	●	●	●		●	●	●	●			
EXPLORATION	Network visualisation	●			●		●				●	●
	Dynamic network layout	●										
	Smart filtering	●					●	●				
	Omics Data Mapping	●	●		●		●			●	●	
	Pathway enrichment	●			●		●			●	●	
	Flux Analysis	●		●		●	◐					
CURATION & Input Output	Network import	●		●								
	Network export	●	●	●	●		●	●	●			
	Automatic reconstruction			●		●	◐					●
	Manual curation	●				●	◐					
	Collaborative curation	●					◐					
	Programmatic access	●			●		●		●			●

Figure 3. Comparison between MetExplore and other web servers integrating several metabolic network functionalities. Light gray cells in the BioCyc column correspond to the Pathway Tools feature which is not a web server but was added to the list since it is strongly related to the BioCyc server.

for genes, which can be identified by identifiers from different databases such as KEGG or NCBI. Globally, the standardization of identifiers for metabolic network identifiers is certainly a great challenge and we think that an effort of the whole community should be made to tackle this issue, which is often a limiting step in omics data interpretation. The collaborative curation and annotation functionalities of MetExplore, introduced in version 2.0, will help metabolic reconstruction to be organised between several collaborators. Until now, a thousand private networks, about 60 projects and more than 500 user accounts have been created in MetExplore. MetExplore makes available 285 metabolic networks coming from different metabolic databases: Micro-Cyc (33), BioCyc (41), BIGG (23), KEGG (4), BioModels (24) or publications.

However, the objective of MetExplore is not to become a model repository, such as BioModels or BIGG. Once the metabolic curation is performed, metabolic networks can be exported in SBML format and then, uploaded in such model repositories. Automatic reconstruction from genomic data is also not in the scope of MetExplore. Indeed, there are already many efficient automatic tools to automat-

ically reconstruct metabolic networks (5,18,42) but it was clearly lacking an effective tool for curation steps which are essential for the quality of a reconstruction and the results that will be based on its analysis (6).

At last, programmatic access (web service), introduced in version 2.0, allows MetExplore functionalities to be integrated in other websites, like it was done in PiMP (43) or Metabolights (44).

MetExplore is a sustainable project maintained by a group of permanent scientists. New features, methods and improvements will be included. Complex queries, as proposed in BioCyc, could be incorporated to filter metabolic data in an 'expert' mode. Graph analyses could complete flux analyses by computing topological indices, as proposed by the web server NeAT (45). Over-representation analysis could be also refined by taking into account the values associated to the mapped elements rather than just their presence.

Besides, current cutting-edge sequencing techniques now allow the reconstruction (at least semi-automatically) of hundreds of metabolic networks (46). The annotation and curation tools will have to adapt to this change of scale. Easy

propagation of annotations to several related metabolic networks could speed up the manual curation of such a large number of networks. Comparison and network merging tools would also greatly help the analysis of metabolic interaction in multi-organism communities.

CONCLUSION

Thanks to the integration of several original refinement and metabolic network exploration tools, MetExplore offers a powerful and easy-to-use solution to many biologists without advanced bioinformatic skills who would like to study and build their own metabolic networks. While initially intended for metabolomics data analysis, MetExplore's usage has now broadened to other omics data analyses (e.g. transcriptomics, fluxomics ...). To better accompany this large community interested by MetExplore's functionalities and its ease of use, the MetExplore team has organised numerous trainings for both students and principal investigators. Thus, since 2012, 401 users have been trained in 25 sessions in Europe, South America or Africa. MetExplore is in fact a very versatile and easy to use solution to grasp the concepts of metabolic networks and how they can greatly enhance multi-omics data analysis.

METEXPLORE TECHNICAL REQUIREMENTS

We recommend using MetExplore on recent versions of Chrome or Firefox.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

REFERENCES

1. Lacroix,V., Cottret,L., Thébault,P. and Sagot,M.-F. (2008) An introduction to metabolic networks and their structural analysis. *IEEE/ACM Trans. Comput. Biol. Bioinform.*, **5**, 594–617.

2. Orth,J.D., Thiele,I. and Palsson,B.Ø. (2010) What is flux balance analysis? *Nat. Biotechnol.*, **28**, 245–248.
3. DeJongh,M., Formsma,K., Boillot,P., Gould,J, Rycenga,M. and Best,A. (2007) Toward the automated generation of genome-scale metabolic networks in the SEED. *BMC Bioinformatics*, **8**, 139.
4. Kanehisa,M., Furumichi,M., Tanabe,M., Sato,Y. and Morishima,K. (2017) KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.*, **45**, D353–D361.
5. Karp,P.D., Latendresse,M., Paley,S.M., Krummenacker,M., Ong,Q.D., Billington,R., Kothari,A., Weaver,D., Lee,T., Subhraveti,P. *et al.* (2016) Pathway tools version 19.0 update: software for pathway/genome informatics and systems biology. *Brief. Bioinform.*, **17**, 877–890.
6. Thiele,I. and Palsson,B.Ø. (2010) A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.*, **5**, 93–121.
7. Paley,S.M. and Karp,P.D. (2006) The Pathway Tools cellular overview diagram and Omics Viewer. *Nucleic Acids Res.*, **34**, 3771–3778.
8. Noronha,A., Daníelsdóttir,A.D., Gawron,P., Jóhannsson,F., Jónsdóttir,S., Jarlsson,S., Gunnarsson,J.P., Brynjólfsson,S., Schneider,R., Thiele,I. *et al.* (2017) ReconMap: an interactive visualization of human metabolism. *Bioinformatics*, **33**, 605–607.
9. Fabregat,A., Jupe,S., Matthews,L., Sidiropoulos,K., Gillespie,M., Garapati,P., Haw,R., Jassal,B., Korninger,F., May,B. *et al.* (2018) The reactome pathway knowledgebase. *Nucleic Acids Res.*, **46**, D649–D655.
10. Marco-Ramell,A., Palau-Rodriguez,M., Alay,A., Tulipani,S., Urpi-Sarda,M., Sanchez-Pla,A. and Andres-Lacueva,C. (2018) Evaluation and comparison of bioinformatic tools for the enrichment analysis of metabolomics data. *BMC Bioinformatics*, **19**, 1.
11. Schellenberger,J., Que,R., Fleming,R.M.T., Thiele,I., Orth,J.D., Feist,A.M., Zielinski,D.C., Bordbar,A., Lewis,N.E., Rahmanian,S. *et al.* (2011) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nat. Protoc.*, **6**, 1290–1307.
12. Marmiesse,L., Peyraud,R. and Cottret,L. (2015) FlexFlux: combining metabolic flux and regulatory network analyses. *BMC Syst. Biol.*, **9**, 93.
13. Cottret,L., Wildridge,D., Vinson,F., Barrett,M.P., Charles,H., Sagot,M.-F. and Jourdan,F. (2010) MetExplore: a web server to link metabolomic experiments and genome-scale metabolic networks. *Nucleic Acids Res.*, **38**, W132–W137.
14. Rao,Q., Rollat-Farnier,P.-A., Zhu,D.-T., Santos-Garcia,D., Silva,F.J., Moya,A., Latorre,A., Klein,C.C., Vavre,F., Sagot,M.-F. *et al.* (2015) Genome reduction and potential metabolic complementation of the dual endosymbionts in the whitefly Bemisia tabaci. *BMC Genomics*, **16**, 226.
15. Shameer,S., Logan-Klumpler,F.J., Vinson,F., Cottret,L., Merlet,B., Achcar,F., Boshart,M., Berriman,M., Breitling,R., Bringaud,F. *et al.* (2015) TrypanoCyc: a community-led biochemical pathways database for Trypanosoma brucei. *Nucleic Acids Res.*, **43**, D637–D644.
16. Zalko,D., Soto,A.M., Canlet,C., Tremblay-Franco,M., Jourdan,F. and Cabaton,N.J. (2016) Bisphenol a exposure disrupts neurotransmitters through modulation of transaminase activity in the brain of rodents. *Endocrinology*, **157**, 1736–1739.
17. Jardinaud,M.-F., Boivin,S., Rodde,N., Catrice,O., Kisiala,A., Lepage,A., Moreau,S., Roux,B., Cottret,L., Sallet,E. *et al.* (2016) A laser dissection-RNAseq analysis highlights the activation of cytokinin pathways by nod factors in the medicago truncatula root epidermis. *Plant Physiol.*, **171**, 2256–2276.
18. Moriya,Y., Itoh,M., Okuda,S., Yoshizawa,A.C. and Kanehisa,M. (2007) KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.*, **35**, W182–W185.
19. Devoid,S., Overbeek,R., DeJongh,M., Vonstein,V., Best,A.A. and Henry,C. (2013) Automated genome annotation and metabolic model reconstruction in the SEED and Model SEED. *Methods Mol. Biol.*, **985**, 17–45.
20. Latendresse,M. (2014) Efficiently gap-filling reaction networks. *BMC Bioinformatics*, **15**, 225.
21. Prigent,S., Frioux,C., Dittami,S.M., Thiele,S., Larhlimi,A., Collet,G., Gutknecht,F., Got,J., Eveillard,D., Bourdon,J. *et al.* (2017) Meneco, a Topology-Based Gap-Filling tool applicable to degraded Genome-Wide metabolic networks. *PLOS Comput. Biol.*, **13**, e1005276.

22. Smith,L.P., Hucka,M., Hoops,S., Finney,A., Ginkel,M., Myers,C.J., Moraru,I. and Liebermeister,W. (2015) SBML Level 3 package: Hierarchical model composition, version 1 release 3. *J. Integr. Bioinform.*, **12**, 268.

23. King,Z.A., Lu,J., Dräger,A., Miller,P., Federowicz,S., Lerman,J.A., Ebrahim,A., Palsson,B.O. and Lewis,N.E. (2016) BiGG Models: a platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res.*, **44**, D515–D522.

24. Glont,M., Nguyen,T.V.N., Graesslin,M., Hälke,R., Ali,R., Schramm,J., Wimalaratne,S.M., Kothamachu,V.B., Rodriguez,N., Swat,M.J. *et al.* (2018) BioModels: expanding horizons to include more modelling approaches and formats. *Nucleic Acids Res.*, **46**, D1248–D1253.

25. Heller,S.R., McNaught,A., Pletnev,I., Stein,S. and Tchekhovskoi,D. (2015) InChI, the IUPAC International Chemical Identifier. *J. Cheminform.*, **7**, 23.

26. Hastings,J., Owen,G., Dekker,A., Ennis,M., Kale,N., Muthukrishnan,V., Turner,S., Swainston,N., Mendes,P. and Steinbeck,C. (2016) ChEBI in 2016: Improved services and an expanding collection of metabolites. *Nucleic Acids Res.*, **44**, D1214–D1219.

27. Orth,J.D., Thiele,I. and Palsson,B.Ø. (2010) What is flux balance analysis? *Nat Biotechnol.*, **28**, 245–248.

28. Mahadevan,R. and Schilling,C.H. (2003) The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab. Eng.*, **5**, 264–276.

29. Chazalviel,M., Frainay,C., Poupin,N., Vinson,F., Merlet,B., Gloaguen,Y., Cottret,L. and Jourdan,F. (2017) MetExploreViz: web component for interactive metabolic network visualization. *Bioinformatics*, **34**, 312–313.

30. Frainay,C. and Jourdan,F. (2017) Computational methods to identify metabolic sub-networks based on metabolomic profiles. *Brief. Bioinform.*, **18**, 43–56.

31. Faust,K. and van Helden,J. (2012) Predicting metabolic pathways by sub-network extraction. *Methods Mol. Biol.*, **804**, 107–130.

32. Thiele,I. and Palsson,B.Ø. (2010) Reconstruction annotation jamborees: a community approach to systems biology. *Mol. Syst. Biol.*, **6**, 361.

33. Médigue,C., Calteau,A., Cruveiller,S., Gachet,M., Gautreau,G., Josso,A., Lajus,A., Langlois,J., Pereira,H., Planel,R. *et al.* (2017) MicroScope-an integrated resource for community expertise of gene functions and comparative analysis of microbial genomic and metabolic data. *Brief. Bioinform.*, doi:10.1093/bib/bbx113.

34. Alonso,A., Marsal,S. and Julià,A. (2015) Analytical methods in untargeted Metabolomics: State of the art in 2015. *Front. Bioeng. Biotechnol.*, **3**, 23.

35. Booth,S.C., Weljie,A.M. and Turner,R.J. (2013) Computational tools for the secondary analysis of metabolomics experiments. *Comput. Struct. Biotechnol. J.*, **4**, e201301003.

36. Chokkathukalam,A., Kim,D.-H., Barrett,M.P., Breitling,R. and Creek,D.J. (2014) Stable isotope-labeling studies in metabolomics: new insights into structure and dynamics of metabolic networks. *Bioanalysis*, **6**, 511–524.

37. Madalinski,G., Godat,E., Alves,S., Lesage,D., Genin,E., Levi,P., Labarre,J., Tabet,J.-C., Ezan,E. and Junot,C. (2008) Direct introduction of biological samples into a LTQ-Orbitrap hybrid mass spectrometer as a tool for fast metabolome analysis. *Anal. Chem.*, **80**, 3291–3303.

38. Moretti,S., Martin,O., Van Du Tran,T., Bridge,A., Morgat,A. and Pagni,M. (2016) MetaNetX/MNXref - Reconciliation of metabolites and biochemical reactions to bring together genome-scale metabolic networks. *Nucleic Acids Res.*, **44**, D523–D526.

39. Xia,J., Sinelnikov,I. V, Han,B. and Wishart,D.S. (2015) MetaboAnalyst 3.0–making metabolomics more meaningful. *Nucleic Acids Res.*, **43**, W251–W257.

40. Frolkis,A., Knox,C., Lim,E., Jewison,T., Law,V., Hau,D.D., Liu,P., Gautam,B., Ly,S., Guo,A.C. *et al.* (2010) SMPDB: The small molecule pathway database. *Nucleic Acids Res.*, **38**, D480–D487.

41. Caspi,R., Billington,R., Ferrer,L., Foerster,H., Fulcher,C.A., Keseler,I.M., Kothari,A., Krummenacker,M., Latendresse,M., Mueller,L.A. *et al.* (2016) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.*, **44**, D471–D480.

42. Henry,C.S., DeJongh,M., Best,A.A., Frybarger,P.M., Linsay,B. and Stevens,R.L. (2010) High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.*, **28**, 977–982.

43. Gloaguen,Y., Morton,F., Daly,R., Gurden,R., Rogers,S., Wandy,J., Wilson,D., Barrett,M. and Burgess,K. (2017) PiMP my metabolome: an integrated, web-based tool for LC-MS metabolomics data. *Bioinformatics*, **33**, 4007–4009.

44. Haug,K., Salek,R.M., Conesa,P., Hastings,J., de Matos,P., Rijnbeek,M., Mahendraker,T., Williams,M., Neumann,S., Rocca-Serra,P. *et al.* (2013) MetaboLights—an open-access general-purpose repository for metabolomics studies and associated meta-data. *Nucleic Acids Res.*, **41**, D781–D786.

45. Brohée,S., Faust,K., Lima-Mendez,G., Sand,O., Janky,R., Vanderstocken,G., Deville,Y. and van Helden,J. (2008) NeAT: a toolbox for the analysis of biological networks, clusters, classes and pathways. *Nucleic Acids Res.*, **36**, W444–W451.

46. Magnúsdóttir,S., Heinken,A., Kutt,L., Ravcheev,D.A., Bauer,E., Noronha,A., Greenhalgh,K., Jäger,C., Baginska,J., Wilmes,P. *et al.* (2016) Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nat. Biotechnol.*, **35**, 81–89.