

An ontology-based approach to streamline the reconstruction of genome-scale metabolic models

Nahim Alves de Souza¹, Renata Wassermann¹

¹Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão, 1010, Cidade Universitária, São Paulo, SP, Brasil

Abstract

The reconstruction of genome-scale metabolic models (GEMs) is a complex and laborious process that depends significantly on expert manual curation. It involves integrating data from diverse sources, such as biochemical databases and scientific literature, which often contain inconsistencies due to the lack of standardized representations for metabolites and reactions. Since current solutions cannot fully resolve these discrepancies, domain experts have to manually identify and correct them, which is a time-consuming task. This paper proposes an ontology-based approach to streamline the reconstruction of GEMs. This work proposes developing a GEM ontology to formally represent both GEM structures and the expert knowledge used during reconstruction. In the future, this ontology can be integrated into the reconstruction workflow through an application that takes draft models as input and produces enhanced models by incorporating additional information from biochemical datasets. This approach is expected to reduce manual curation effort and consequently simplify the overall reconstruction process.

Keywords

Genome-scale metabolic models, Ontology, Semantic interoperability, Data integration

1. Introduction

Metabolism is defined as a series of chemical reactions that occur continuously within living organisms to sustain life, particularly those associated with energy production and growth [1]. Nevertheless, mapping the entire metabolism of an organism is a highly complex task due to the vast number of compounds and reactions involved. The process of building a computational representation of an organism's metabolism is known as **reconstruction**, while the resulting model is often referred to as a **genome-scale metabolic model (GEM)** – since it is based on organism's genome.

Over the years, several computational resources have been developed to support the reconstruction of metabolic models, including databases, tools, and ontologies. However, these resources present significant limitations, especially in unifying data from multiple sources [2, 3, 4]. Moreover, none of the available ontologies (e.g., ChEBI [5], GO [6], SBO [7]) comprehensively represent GEMs or are actively applied in GEM reconstruction, as they contain little information about the complex relationships between reactions, metabolites, and genes – which are essential for integrating data during GEM reconstruction. Consequently, experts must rely on their domain knowledge to manually resolve data inconsistencies, suggesting that the necessary semantic information exists but is not explicitly represented in current databases and ontologies.

Based on these premises, this work proposes the development of a new ontology for representing GEMs with two main objectives: (1) to enable the integration and reconciliation of models across different datasets, and (2) to facilitate the quality assessment of GEMs through the use of logical inferences to identify (and potentially repair) inconsistencies in models. The following sections present the GEM reconstruction process and its main challenges, along with the proposed approach to address them.

Proceedings of the 18th Seminar on Ontology Research in Brazil (ONTOBRAS 2025) and 9th Doctoral and Masters Consortium on Ontologies (WIDO 2025), São José dos Campos (SP), Brazil, September 29 – October 02, 2025.

✉ nahim@usp.br (N. A. d. Souza); renata@ime.usp.br (R. Wassermann)

ORCID 0009-0007-8051-558X (N. A. d. Souza); 0000-0001-8065-1433 (R. Wassermann)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

2. GEM reconstruction

GEM reconstruction is a complex process that involves integrating data from diverse sources, conducting thorough literature reviews, performing manual curation, running mathematical simulations, and validating results through biological experiments. Each of these activities need to be carefully conducted in order to ensure the accuracy and quality of the resulting model. Thiele and Palsson [8] proposed a detailed five-stage protocol for the construction high-quality GEMs, summarized in Figure 1. The reconstruction process begins with the creation of a **draft model**, derived from the organism's genome and biochemical datasets. This initial model is then refined by experts, supported by computational tools that assess model quality and simulate the organism's metabolic behavior. Finally, the curated model, along with documentation of the reconstruction process, is compiled and published.

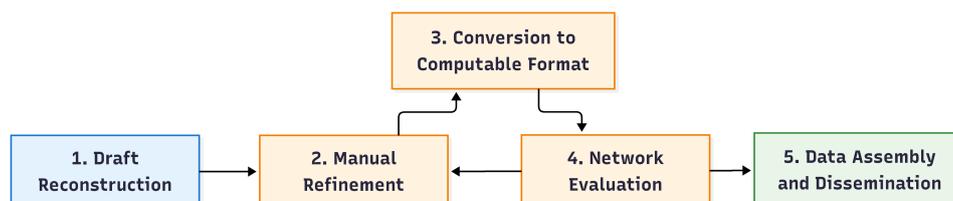


Figure 1: GEM Reconstruction stages, based Thiele and Palsson's protocol [8].

2.1. Key challenges

Integrating information on metabolites and reactions is fundamental to GEM reconstruction. Although various efforts have been made over the years to establish representation standards for biochemical data, a universal consensus has not yet been reached [4]. For instance, while the chemical structures of metabolites can be represented as SMILES strings [9], this format does not establish a unique representation for each molecule¹ [2] – water, for example, can be represented as [OH2], [H]O[H], or simply O². The InChI system [10] addresses some of these ambiguities by providing a more detailed representation of molecular structures. However, certain molecules representable in SMILES may lack proper encoding in InChI, and the optional nature of some fields in the InChI format can result in incomplete representations in certain contexts [2]. Beyond variations in metabolite naming, the representation of reactions adds another layer of complexity to the reconciliation of biochemical data. The multiple ways of expressing reactions – ranging from chemical equations [3] to more complex representations such as reaction graphs, hypergraphs, or stoichiometric matrices [11, 12, 13, 14] – hinder data comparison across different datasets.

In GEMs, each reaction is associated with genes and enzymes through GPRs (gene-protein-reaction) rules, along with other attributes essential for simulating organism behavior. SBML remains the dominant format for representing GEMs, as its design facilitates model exchange, reuse, and supports simulation-related properties [15]. While its flexible design allows core capabilities to be extended, it lacks a formal specification for mandatory fields, which results in inconsistently populated fields, containing incomplete or inappropriate information [16]. Furthermore, even when the same data sources are used, GEM reconstructions can differ due to variations in methods, algorithms, and expert decisions throughout the process [17, 18].

Some approaches, such as MNXref [2] and MetRxn [3], were designed to provide a unified, reliable data source by implementing iterative reconciliation processes that utilize identifiers, names and cross-references to resolve ambiguities. Nevertheless, these methods may overlook crucial factors such as reaction directionality, compartmentalization, mass- and charge-balance, potentially leading to

¹While Isomeric SMILES addresses some ambiguities, multiple representations remain possible [2].

²Although this last representation may seem unusual, it is adopted by ModelSEED (<https://modelseed.org/biochem/compounds/cpd00001>) and PubChem (<https://pubchem.ncbi.nlm.nih.gov/compound/Water#section=SMILES>)

Table 1

Competency questions were designed to reflect typical activities in GEM reconstruction. This table presents the most relevant ones.

Competency Questions
CQ01: Are metabolites <i>M1</i> and <i>M2</i> the same?
CQ02: Are reactions <i>R1</i> and <i>R2</i> the same?
CQ03: Does metabolite <i>M1</i> have the correct chemical formula?
CQ04: Is reaction <i>R1</i> mass- and charge-balanced?
CQ05: Is reaction <i>R</i> associated with a single gene or multiple genes?
CQ06: Which reactions are catalyzed by enzyme <i>E</i> ?
CQ07: Is reaction <i>R</i> blocked?
CQ08: Which reactions in the network produce metabolite <i>M</i> ?
CQ09: Which metabolites are dead ends in the network?
CQ10: Which genes can be knocked out to prevent the production of metabolite <i>M</i> ?
CQ11: Which metabolites are produced but not consumed by any reaction in the network?
CQ12: Which metabolites are required but not produced in the network?
CQ13: Does the model generate biomass?

inaccurate results. Additionally, internal and external inconsistencies in names and identifiers, often found in biochemical databases [19], make automated data unification particularly challenging.

3. An ontology for GEMs

Ontologies have proven to be powerful tools for addressing challenges related to data integration and system interoperability [20, 21]. In the fields of chemistry and biology, numerous ontologies have been developed to establish unified vocabularies and facilitate integration across heterogeneous systems. The Basic Formal Ontology (BFO) [22], for example, is a widely adopted upper ontology that provides foundational concepts for constructing domain-specific ontologies. Notable examples include ChEBI, which models chemical entities and their relationships [5], and the Gene Ontology (GO), which represents genes along with their functions and associated products [6].

The ontology most directly associated with GEM reconstruction is the Systems Biology Ontology (SBO) [7], which defines terms related to Systems Biology, including physical entities (e.g., metabolites, genes, biomass) and processes (e.g., biochemical reactions). SBO terms can be used to annotate tags in SBML files, facilitating the integration of models originated from different sources [23]. However, despite these features, SBO does not adequately capture the complex relationships between entities, limiting its effectiveness in representing the full structure and semantics of GEMs. Furthermore, to the best of our knowledge, no existing ontology addresses this level of detail, highlighting an opportunity to overcome these challenges through the development of an ontology for GEMs, capable of enabling concept disambiguation and data integration.

3.1. Ontology development

Initially, we conducted a literature review to understand the GEM reconstruction process, existing solutions, and available databases and tools. Subsequently, in collaboration with systems biology experts, the project scope and objectives were established, and competency questions (Table 1) were formulated based on the activities typically performed during GEM reconstruction (e.g., comparing metabolites and reactions, evaluating the metabolic network). Based on these questions we identified the core concepts and relationships of the ontology and designed an initial domain model (Figure 2).

Although this model still requires refinement, it already provides a strong foundation for understanding the domain. The central component of the model is the **MetabolicModel** class, which serves as an information aggregator, encompassing properties of the organism and a list of metabolites, reactions, genes, and compartments. A **BiologicalEntity**, corresponds to entities that physically exist in the

real world, which contains a unique ID, name, and a list of cross references including external data and metadata (**CrossReference** class). A **Metabolite** correspond to any molecule participating in metabolic reactions, either as a reactant, product, cofactor, or intermediate. In GEMs, each metabolite is located in a specific **Compartment** (a region within the cell, e.g., glucose in the cytosol vs. extracellular glucose are distinct). A **Reaction** is a biochemical transformation that converts a set of reactants into products. Reactions can be enzymatic (catalyzed by a single enzyme or by an enzymatic complex) or non-enzymatic (occurring spontaneously). An **Enzyme** is a special type of protein involved in the catalysis of biochemical reactions. A single enzyme may catalyze multiple reactions, and a single reaction may be catalyzed by different enzymes (the same applies to enzymatic complexes). Additionally, each enzyme can be encoded by one or more genes. A **Gene** is a DNA sequence that encodes a protein (in this context, an enzyme). Genes and reactions are linked through GPR associations, which are boolean formulas represented in the diagram through the relationships among the **Gene**, **Enzyme**, **EnzymeAssociation**, and **Reaction** classes.

In order to fully support concept disambiguation, and promote explainability and systems interoperability, the current model still requires further enhancements. For instance, several terms mentioned in the competency questions lack clear and explicit definitions, such as “same” in CQ02, “blocked reaction” in CQ07, and “required” in CQ12. Therefore, the next step, which we are currently working on, involves using the initial model as a foundation to create a more explicit model with the necessary semantic information. Recent studies have demonstrated that the application of **ontological unpacking** techniques, combined with the modeling language OntoUML, yields promising results by enriching concept definitions and making implicit knowledge explicit [24, 25, 20]. Applying these techniques to the current model can lead to a more comprehensive and semantically rich representation of GEMs.

3.2. Evaluation

The evaluation of the proposed approach must be twofold: (1) assessing the consistency and accuracy of the GEM ontology in representing domain knowledge and (2) verifying the quality of the GEMs generated using this ontology-based approach. For the first part, the representation can be assessed by encoding the ontology in OWL and using automated reasoners to detect logical inconsistencies in the knowledge base – for example, defining a reaction as occurring in one compartment while its metabolites are located in another, or incorrectly declaring two entities equivalent when they have different property values. In addition, competency questions can be translated into SPARQL queries and used to evaluate whether the ontology can answer the questions proposed by domain experts, thereby assessing the correctness and completeness of the representation [26].

The second part of the evaluation can be carried out by comparing the model generated by an application based on the GEM ontology with manually curated models from well-studied organisms, such as *Escherichia coli* (an approach commonly adopted in the literature [27, 28, 29, 21]). In this scenario, the evaluation pipeline would consist of: (1) generating a draft model from the genome of the selected organism (e.g., *E. Coli*³); (2) using the application to load and enhance the model; and (3) comparing the resulting model with a manually curated reference model (e.g., model iJO1366⁴). To achieve a more comprehensive evaluation of the application’s effectiveness, the comparison should include models of varying quality (highly curated and poorly curated) from organisms of different types (eukaryotes and prokaryotes, well-studied and less-studied). The criteria for comparing models may include, for instance, the number and presence or absence of metabolites, genes, and reactions [30, 27], the ability to directly perform FBA (Flux Balance Analysis) [28, 21], and reports from tools such as FROG [31] and MEMOTE [32], which provide additional insights into the metabolic network (e.g., mass and charge balancing, stoichiometric consistency, FVA, and gene/reaction deletion fluxes).

³*E. Coli* genome is available at NCBI website: https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_000005845.2

⁴This is the most recent *E. Coli* model available in BiGG: <http://bigg.ucsd.edu/models/iJO1366>

information to assist domain experts in refining the model, either manually or automatically, depending on the nature of the issues detected. Finally, it should also allow the experts to edit, refine, and subsequently export the model in multiple formats.

The technical details of such an application should be defined during the implementation phase. This includes selecting the most appropriate biochemical databases and ontologies for data enrichment, defining mechanisms for data access (e.g., APIs, SPARQL endpoints), and establishing strategies for data management and storage. In addition, the deployment and documentation of the application should be carefully planned to ensure long-term maintainability, reproducibility, and ease of use for both users and developers.

4. Conclusion

This work proposed an ontology-based approach to streamline the reconstruction of genome-scale metabolic models (GEMs), comprising both a GEM ontology and an application to integrate it into the reconstruction workflow. The ontology – currently under development – is expected to facilitate data reconciliation across biochemical databases and enable automated reasoning. The application is intended to centralize information retrieval from multiple datasets, thereby reducing manual effort and improving efficiency. Developing both the ontology and the application is complex and time-consuming; therefore, the work can be divided into two phases. The first phase should focus on ontology development and evaluation, while the second should focus on building an application on top of the ontology to generate enhanced models.

In addition, the comparison of GEMs remains an open challenge, as it requires the unambiguous identification of model components such as metabolites and reactions. Defining comparison criteria for metabolic models is expected to provide a novel standard for evaluating the quality of GEMs generated by different tools in future studies.

Acknowledgments

This work was supported by a doctoral scholarship from CAPES/PROEX and CNPq (140323/2025-2), and by funding from Inria ERABLE – University of Lyon, which enabled a research visit to the LBBE (Laboratoire de Biométrie et Biologie Évolutive) in Lyon, France, to foster collaboration between our research teams.

Special thanks to Gabriela Torres Montanaro (ICB-USP), Ariel Mariano Silber (ICB-USP), Lucas Gentil Azevedo (Cidacs/FIOCRUZ, BA), Mariana Galvão Ferrarini (Max Planck Institute for Chemical Ecology in Jena, Germany) and Marie-France Sagot (University of Lyon, France) for their valuable contributions to this work.

Declaration on Generative AI

During the preparation of this work, the author(s) used ChatGPT (GPT-4o) and Claude Sonnet 4 in order to: Grammar and spelling check, paraphrase and reword. After using these tool(s)/service(s), the author(s) reviewed and edited the content as needed and take(s) full responsibility for the publication's content.

References

- [1] A. Judge, M. S. Dodd, *Metabolism, Essays in Biochemistry* 64 (2020) 607–647. URL: <https://doi.org/10.1042/EBC20190041>. doi:10.1042/EBC20190041.

- [2] T. Bernard, A. Bridge, A. Morgat, S. Moretti, I. Xenarios, M. Pagni, Reconciliation of metabolites and biochemical reactions for metabolic networks, *Briefings in Bioinformatics* 15 (2012) 123–135. URL: <https://doi.org/10.1093/bib/bbs058>. doi:10.1093/bib/bbs058.
- [3] A. Kumar, P. F. Suthers, C. D. Maranas, Metrxn: a knowledgebase of metabolites and reactions spanning metabolic models and databases, *BMC Bioinformatics* 13 (2012) 6. URL: <https://doi.org/10.1186/1471-2105-13-6>. doi:10.1186/1471-2105-13-6.
- [4] A. Hari, A. Zarrabi, D. Lobo, mergem: merging, comparing, and translating genome-scale metabolic models using universal identifiers, *NAR Genomics and Bioinformatics* 6 (2024) 1–21. URL: <https://doi.org/10.1093/nargab/lqae010>. doi:10.1093/nargab/lqae010.
- [5] K. Degtyarenko, P. de Matos, M. Ennis, J. Hastings, M. Zbinden, A. McNaught, R. Alcántara, M. Darsow, M. Guedj, M. Ashburner, ChEBI: a database and ontology for chemical entities of biological interest, *Nucleic Acids Research* 36 (2007) D344–D350. URL: <https://doi.org/10.1093/nar/gkm791>. doi:10.1093/nar/gkm791.
- [6] M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, G. Sherlock, Gene ontology: tool for the unification of biology, *Nature Genetics* 25 (2000) 25–29. URL: <https://doi.org/10.1038/75556>. doi:10.1038/75556.
- [7] M. Courtot, N. Juty, C. Knüpfer, D. Waltemath, A. Zhukova, A. Dräger, M. Dumontier, A. Finney, M. Golebiewski, J. Hastings, S. Hoops, S. Keating, D. B. Kell, S. Kerrien, J. Lawson, A. Lister, J. Lu, R. Machne, P. Mendes, M. Pocock, N. Rodriguez, A. Villegier, D. J. Wilkinson, S. Wimalaratne, C. Laibe, M. Hucka, N. Le Novère, Controlled vocabularies and semantics in systems biology, *Molecular Systems Biology* 7 (2011) 543. URL: <https://www.embopress.org/doi/abs/10.1038/msb.2011.77>. doi:<https://doi.org/10.1038/msb.2011.77>.
- [8] I. Thiele, B. Ø. Palsson, A protocol for generating a high-quality genome-scale metabolic reconstruction, *Nature Protocols* 5 (2010) 93–121. URL: <https://doi.org/10.1038/nprot.2009.203>. doi:10.1038/nprot.2009.203.
- [9] D. Weininger, Smiles, a chemical language and information system. 1. introduction to methodology and encoding rules, *Journal of Chemical Information and Computer Sciences* 28 (1988) 31–36. URL: <https://doi.org/10.1021/ci00057a005>. doi:10.1021/ci00057a005.
- [10] A. McNaught, The iupac international chemical identifier: InChI—a new standard for molecular informatics, *Chemistry international* 28 (2006) 12–14. URL: <https://api.semanticscholar.org/CorpusID:89057853>.
- [11] L. Cottret, P. Vieira Milreu, V. Acuña, A. Marchetti-Spaccamela, F. Viduani Martinez, M.-F. Sagot, L. Stougie, Enumerating precursor sets of target metabolites in a metabolic network, in: K. A. Crandall, J. Lagergren (Eds.), *Algorithms in Bioinformatics*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008, pp. 233–244.
- [12] A. Amara, C. Frainay, F. Jourdan, T. Naake, S. Neumann, E. M. Novoa-del Toro, R. M. Salek, L. Salzer, S. Scharfenberg, M. Witting, Networks and graphs discovery in metabolomics data analysis and interpretation, *Frontiers in Molecular Biosciences* 9 (2022). URL: <https://www.frontiersin.org/journals/molecular-biosciences/articles/10.3389/fmolb.2022.841373>. doi:10.3389/fmolb.2022.841373.
- [13] V. Lacroix, L. Cottret, P. Thébault, M.-F. Sagot, An introduction to metabolic networks and their structural analysis, *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 5 (2008) 594–617. doi:10.1109/TCBB.2008.79.
- [14] L. Wang, S. Dash, C. Y. Ng, C. D. Maranas, A review of computational tools for design and reconstruction of metabolic pathways, *Synthetic and Systems Biotechnology* 2 (2017) 243–252. URL: <https://www.sciencedirect.com/science/article/pii/S2405805X17300820>. doi:<https://doi.org/10.1016/j.synbio.2017.11.002>.
- [15] L. Strömbäck, P. Lambrix, Representations of molecular pathways: an evaluation of sbml, psi mi and biopax, *Bioinformatics* 21 (2005) 4401–4407. URL: <https://doi.org/10.1093/bioinformatics/bti718>. doi:10.1093/bioinformatics/bti718.
- [16] T. Pfau, M. P. Pacheco, T. Sauter, Towards improved genome-scale metabolic network recon-

- structions: unification, transcript specificity and beyond, *Briefings in Bioinformatics* 17 (2015) 1060–1069. URL: <https://doi.org/10.1093/bib/bbv100>. doi:10.1093/bib/bbv100.
- [17] M. A. Carey, A. Dräger, M. E. Beber, J. A. Papin, J. T. Yurkovich, Community standards to facilitate development and address challenges in metabolic modeling, *Molecular Systems Biology* 16 (2020) e9235. URL: <https://www.embopress.org/doi/abs/10.15252/msb.20199235>. doi:<https://doi.org/10.15252/msb.20199235>.
- [18] D. B. Bernstein, S. Sulheim, E. Almaas, D. Segrè, Addressing uncertainty in genome-scale metabolic model reconstruction and analysis, *Genome Biology* 22 (2021) 64. URL: <https://doi.org/10.1186/s13059-021-02289-z>. doi:10.1186/s13059-021-02289-z.
- [19] N. Pham, R. G. A. van Heck, J. C. J. van Dam, P. J. Schaap, E. Saccenti, M. Suarez-Diez, Consistency, Inconsistency, and Ambiguity of Metabolite Names in Biochemical Databases Used for Genome-Scale Metabolic Modelling, *Metabolites* 9 (2019). URL: <https://www.mdpi.com/2218-1989/9/2/28>. doi:10.3390/metabo9020028.
- [20] G. Guizzardi, N. Guarino, Explanation, semantics, and ontology, *Data and Knowledge Engineering* 153 (2024) 102325. URL: <https://www.sciencedirect.com/science/article/pii/S0169023X24000491>. doi:<https://doi.org/10.1016/j.datak.2024.102325>.
- [21] S. M. D. Seaver, F. Liu, Q. Zhang, J. Jeffryes, J. P. Faria, J. N. Edirisinghe, M. Mundy, N. Chia, E. Noor, M. E. Beber, A. A. Best, M. DeJongh, J. A. Kimbrel, P. D’haeseleer, S. R. McCorkle, J. R. Bolton, E. Pearson, S. Canon, E. M. Wood-Charlson, R. W. Cottingham, A. P. Arkin, C. S. Henry, The ModelSEED Biochemistry Database for the integration of metabolic annotations and the reconstruction, comparison and analysis of metabolic models for plants, fungi and microbes, *Nucleic Acids Research* 49 (2020) D575–D588. URL: <https://doi.org/10.1093/nar/gkaa746>. doi:10.1093/nar/gkaa746.
- [22] J. N. Otte, J. Beverley, A. Ruttenberg, Bfo: Basic formal ontology, *Applied ontology* 17 (2022) 17–43. doi:10.3233/ao-220262.
- [23] M. Hucka, F. T. Bergmann, C. Chaouiya, A. Dräger, S. Hoops, S. M. Keating, M. König, N. L. Novère, C. J. Myers, B. G. Olivier, S. Sahle, J. C. Schaff, R. Sheriff, L. P. Smith, D. Waltemath, D. J. Wilkinson, F. Zhang, The systems biology markup language (sbml): Language specification for level 3 version 2 core release 2, *Journal of Integrative Bioinformatics* 16 (2019) 20190021. URL: <https://doi.org/10.1515/jib-2019-0021>. doi:doi:10.1515/jib-2019-0021.
- [24] A. García S., A. Bernasconi, G. Guizzardi, O. Pastor, V. C. Storey, I. Panach, Assessing the value of ontologically unpacking a conceptual model for human genomics, *Information Systems* 118 (2023) 102242. URL: <https://www.sciencedirect.com/science/article/pii/S0306437923000789>. doi:<https://doi.org/10.1016/j.is.2023.102242>.
- [25] A. Bernasconi, G. Guizzardi, O. Pastor, V. C. Storey, Semantic interoperability: ontological unpacking of a viral conceptual model, *BMC Bioinformatics* 23 (2022) 491. URL: <https://doi.org/10.1186/s12859-022-05022-0>. doi:10.1186/s12859-022-05022-0.
- [26] G. K. Q. Monfardini, J. S. Salamon, M. P. Barcellos, Use of competency questions in ontology engineering: A survey, in: J. P. A. Almeida, J. Borbinha, G. Guizzardi, S. Link, J. Zdravkovic (Eds.), *Conceptual Modeling*, Springer Nature Switzerland, Cham, 2023, pp. 45–64.
- [27] S. N. Mendoza, B. G. Olivier, D. Molenaar, B. Teusink, A systematic assessment of current genome-scale metabolic reconstruction tools, *Genome Biology* 20 (2019) 158. URL: <https://doi.org/10.1186/s13059-019-1769-1>. doi:10.1186/s13059-019-1769-1.
- [28] J. Zimmermann, C. Kaleta, S. Waschina, gapseq: informed prediction of bacterial metabolic pathways and reconstruction of accurate metabolic models, *Genome Biology* 22 (2021) 81. URL: <https://doi.org/10.1186/s13059-021-02295-1>. doi:10.1186/s13059-021-02295-1.
- [29] M. Di Filippo, C. Damiani, D. Pescini, Gpruler: Metabolic gene-protein-reaction rules automatic reconstruction, *PLOS Computational Biology* 17 (2021) 1–25. URL: <https://doi.org/10.1371/journal.pcbi.1009550>. doi:10.1371/journal.pcbi.1009550.
- [30] Y. E. Hsieh, K. Tandon, H. Verbruggen, Z. Nikoloski, Comparative analysis of metabolic models of microbial communities reconstructed from automated tools and consensus approaches, *npj Systems Biology and Applications* 10 (2024) 54. URL: <https://doi.org/10.1038/s41540-024-00384-y>.

doi:10.1038/s41540-024-00384-y.

- [31] K. Raman, M. Kratochvíl, B. G. Olivier, M. König, P. Sengupta, D. K. Kuppa Baskaran, T. V. N. Nguyen, D. Lobo, S. E. Wilken, K. K. Tiwari, A. K. Raghu, I. Palanikumar, L. Raajaraam, M. Ibrahim, S. Balakrishnan, S. Umale, F. Bergmann, T. Malpani, V. P. Satagopam, R. Schneider, M. E. Beber, S. Keating, M. Anton, A. Renz, M. Lakshmanan, D.-Y. Lee, L. Koduru, R. Mostolizadeh, O. Dias, E. Cunha, A. Oliveira, Y. Q. Lee, K. Zengler, R. Santibáñez-Palominos, M. Kumar, M. Barberis, B. L. Puniya, T. Helikar, H. V. Dinh, P. F. Suthers, C. D. Maranas, I. Casini, S. B. Loghmani, N. Veith, N. Leonidou, F. Li, Y. Chen, J. Nielsen, G. Lee, S. M. Lee, G. B. Kim, P. T. Monteiro, M. C. Teixeira, H. U. Kim, S. Y. Lee, U. W. Liebal, L. M. Blank, C. Lieven, C. Tarzi, C. Angione, M. E. Blaise, Ç. P. Aytar, M. Kulyashov, I. Akberdin, D. Kim, S. H. Yoon, Z. Xu, J. Gautam, W. T. Scott, P. J. Schaap, J. J. Koehorst, C. Zuñiga, G. Canto-Encalada, S. Benito-Vaquerizo, I. P. Olm, M. Suarez-Diez, Q. Yuan, H. Ma, M. M. Islam, J. A. Papin, F. Zorrilla, K. R. Patil, A. Basile, J. Nogales, G. S. León, F. Castillo-Alfonso, R. Olivares-Hernández, G. Canto-Encalada, G. Viguera-Ramírez, H. Hermjakob, A. Dräger, R. S. Malik-Sheriff, Frog analysis ensures the reproducibility of genome scale metabolic models, *bioRxiv* (2024). URL: <https://www.biorxiv.org/content/early/2024/09/26/2024.09.24.614797>. doi:10.1101/2024.09.24.614797.
- [32] C. Lieven, M. Beber, B. Olivier, F. Bergmann, M. Ataman, P. Babaei, J. Bartell, L. Blank, S. Chauhan, K. Correia, C. Diener, A. Dräger, B. Ebert, J. Edirisinghe, J. Faria, A. Feist, G. Fengos, R. Fleming, B. García-Jiménez, V. Hatzimanikatis, W. van Helvoirt, C. Henry, H. Hermjakob, M. Herrgård, A. Kaafarani, H. Kim, Z. King, S. Klamt, E. Klipp, J. Koehorst, M. König, M. Lakshmanan, D.-Y. Lee, S. Lee, S. Lee, N. Lewis, F. Liu, H. Ma, D. Machado, R. Mahadevan, P. Maia, A. Mardinoglu, G. Medlock, J. Monk, J. Nielsen, L. Nielsen, J. Nogales, I. Nookaew, B. Palsson, J. Papin, K. Patil, M. Poolman, N. Price, O. Resendis-Antonio, A. Richelle, I. Rocha, B. Sánchez, P. Schaap, R. Malik Sheriff, S. Shoaie, N. Sonnenschein, B. Teusink, P. Vilaça, J. Vik, J. Wodke, J. Xavier, Q. Yuan, M. Zakhartsev, C. Zhang, Memote for standardized genome-scale metabolic model testing, *Nature Biotechnology* 38 (2020). doi:10.1038/s41587-020-0446-y.