Commentary

OPEN ACCESS

molecular systems biology

Community standards to facilitate development and address challenges in metabolic modeling

Maureen A Carey[1], Andreas Dräger[2,3,4], Moritz E Beber[5], Jason A Papin[1,6,*] & James T Yurkovich[7,**]

Standardization of data and models facilitates effective communication, especially in computational systems biology. However, both the development and consistent use of standards and resources remain challenging. As a result, the amount, quality, and format of the information contained within systems biology models are not consistent and therefore present challenges for widespread use and communication. Here, we focused on these standards, resources, and challenges in the field of constraint-based metabolic modeling by conducting a community-wide survey. We used this feedback to (i) outline the major challenges that our field faces and to propose solutions and (ii) identify a set of features that defines what a "gold standard" metabolic network reconstruction looks like concerning content, annotation, and simulation capabilities. We anticipate that this community-driven outline will help the long-term development of community-inspired resources as well as produce high-quality, accessible models within our field. More broadly, we hope that these efforts can serve as blueprints for other computational modeling communities to ensure the continued development of both practical, usable standards and reproducible, knowledge-rich models.

Mol Syst Biol. (2020) 16: e9235

Introduction

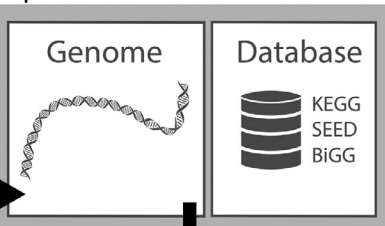
Systems biology uses holistic approaches to understand the networks that comprise biological systems. Computational models that attempt to represent these systems are inherently complex with many interacting components, requiring the mathematical formalization of biological phenomenon. Standardizing how these phenomena are represented is thus required to make these formalizations interpretable and accessible. Many resources—including databases, algorithms, file formats, software, and compiled "best practices"—exist to facilitate standardization (e.g., Le Novère *et al*, 2005; Waltemath *et al*, 2011; Dräger & Palsson, 2014; Ravikrishnan & Raman, 2015; Stanford *et al*, 2015; Keating *et al*, 2020), but the consistent use and application of these standards can pose a significant challenge (Ebrahim *et al*, 2015).

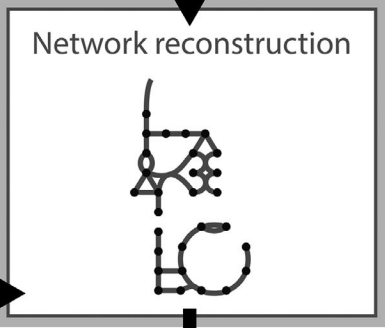
Here, we discuss existing standards in computational modeling in biology and when and why they are not met, building on previous efforts to assess standardization in computational systems biology (Stanford *et al*, 2015). The modeling process has two phases: model construction and simulation. Decisions about technical approaches and biological content to include in the model are made throughout both the construction and simulation processes, influencing the downstream use of the model. These implicit and explicit decisions affect the reusability; if the design decisions made during the model building process do not match well with a particular application, the quality of the simulation results will suffer. Such design decisions are influenced by a scientist's perspective, a motivating biological question, and data availability, as well as a scientist's familiarity with and access to existing resources. Manual steps of this process are particularly vulnerable to potential biases and thus are inherently irreproducible, emphasizing the role of diligent tracking of references and design decisions. Field-defined best practices and standards can help control for or evaluate quality and facilitate iteratively cycling between construction and simulation to improve the process.

In this Commentary, we use metabolic network modeling as a case study in which to

1 Division of Infectious Diseases and International Health, Department of Medicine, University of Virginia, Charlottesville, VA, USA
2 Computational Systems Biology of Infection and Antimicrobial-Resistant Pathogens, Institute for Biomedical Informatics (IBMI), University of Tübingen, Tübingen, Germany
3 Department of Computer Science, University of Tübingen, Tübingen, Germany
4 German Center for Infection Research (DZIF), partner site Tübingen, Tübingen, Germany
5 Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kemitorvet, Denmark
6 Department of Biomedical Engineering, University of Virginia, Charlottesville, VA, USA
7 Institute for Systems Biology, Seattle, WA, USA
*Corresponding author. E-mail: papin@virginia.edu
**Corresponding author. E-mail: jtyurkovich@gmail.com
DOI 10.15252/msb.20199235 | Mol Syst Biol. (2020) 16: e9235

Box 1: Pipeline for genome-scale metabolic network modeling, including existing and proposed standards. Reconstructions represent powerful tools for the interrogation and understanding of metabolism. Here we outline the modeling pipeline and the decisions made throughout this process, as well as existing and proposed standards.

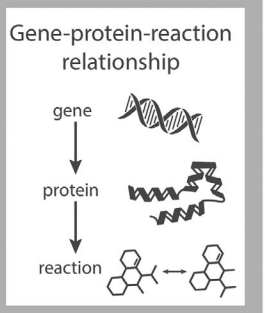




De novo reconstruction

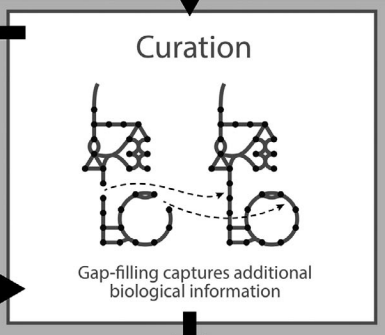
An organism's metabolic network is the collection of biochemical reactions that interconvert chemical compounds to maintain its physiological state (Feist et al., 2009). Defining this network—the 'reconstruction' process—is a time-consuming process that involves enumerating and annotating all components within a cell (Blais et al., 2013; Reed et al., 2006; Yurkovich and Palsson, 2016).

A de novo metabolic network reconstruction starts with the annotated genome sequence for an organism, mapping genes to biochemical functions (reactions) using biochemical databases to identify gene-protein-reaction relationships. An existing well-defined protocol outlines the reconstruction process (Thiele and Palsson, 2010) and has helped produce many metabolic network reconstructions.



Existing standards: (1) Evaluate with MEMOTE, (2) save as SBML, (3) follow MIRIAM guidelines

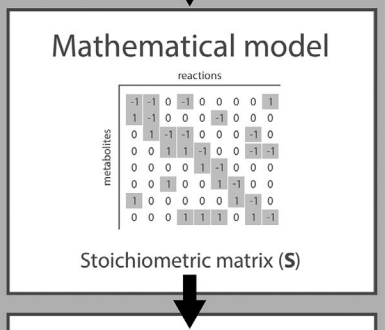
Proproposed standards: Existing standards plus (1) document genome, annotation pipeline, database versions, (2) store with README file

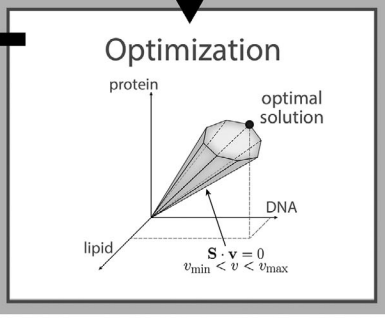


Curation

Reconstructions must be curated to ensure proper representation of the organism's biology. This augments analyses beyond mere genetic comparisons by inferring function to connect gaps in metabolism (Pan and Reed, 2018) using automated tools (Benedict et al., 2014; Thiele et al., 2014) or manual curation (Carey et al., 2017). Gap-filled additions can be used to infer biological function or generate hypotheses (Medlock and Papin, 2018; Pan and Reed, 2018). Additionally, integrating experimental data through manual curation (Tang et al., 2019; Thiele and Palsson, 2010) can help overcome challenges presented by the similarity of new reconstructions to model organisms like E. coli (Monk et al., 2014).

Proposed standards: (1) Document all automated and manual curation, (2) document all modifications in the reference field with proper citations providing evidence for model decisions





Simulation

A reconstruction can be parameterized with condition-specific data to produce a genome-scale metabolic model that can simulate metabolic physiology of a particular state (O'Brien et al., 2015, Palsson, 2015). Simulations of biological phenotypes can be generated using linear optimization techniques, a model, and an appropriate set of constraints. These constraints represent a variety of biological phenomena (Reed, 2012), from thermodynamics (Henry et al., 2007) to integrating -omics data (Bordbar et al., 2017).

Simulations can be used to interpret complex or sparse datasets (Tian and Reed, 2018), reconcile experimental data (Blazier and Papin, 2019), make predictions about the functional consequences of environmental changes (Medlock et al., 2018), or infer mechanisms behind phenotypic changes (Long et al., 2018). In other words, the process of reconstruction results in the construction of a 'knowledgebase' for a given organism because it involves the structured collation of -omics measurements, phenotypic and biochemical data, and can enable the generation of mechanistic hypotheses about biochemical function through simulation.

Existing standards: (1) MIASE and (2) COBRA software

Proposed standards: Existing standards plus (1) document software, solver, and reconstruction versions, (2) full documentation of constraints on bounds or fluxes, (3) store information in README or in COMBINE

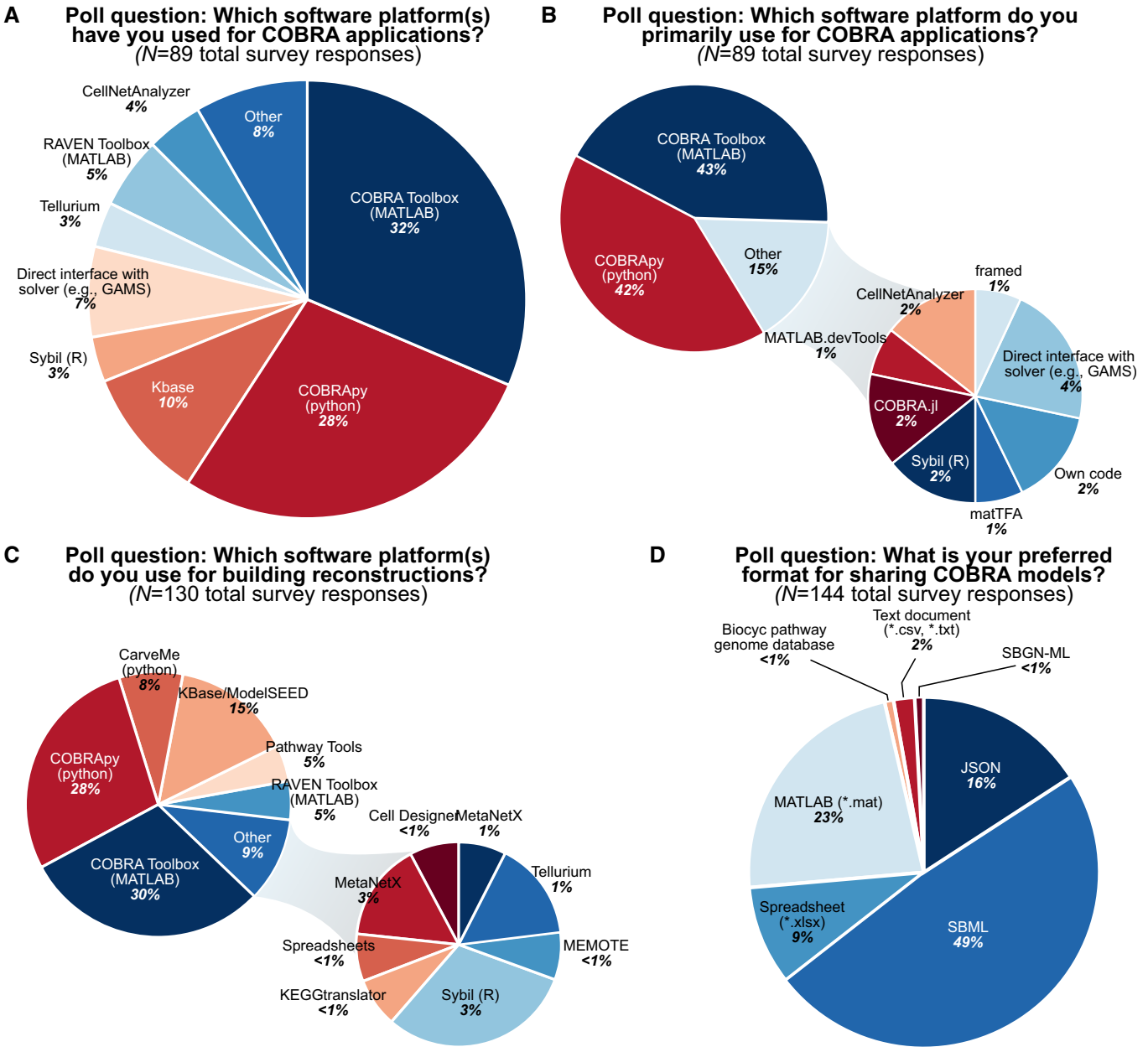


Figure 1. Poll results from the COBRA community survey.

The survey was initially compiled and released at the 5th Annual Conference on Constraint-based Reconstruction and Analysis (COBRA, October 14–16, 2018); feedback from the conference was used to refine the survey, with an updated version later shared via social media (results are shown here; raw data provided in Dataset EV1). The survey included 16 multiple-choice and three open-ended questions to summarize the field's use and awareness of existing standards, as well as collect community-identified challenges. A total of 89 researchers completed the survey, representing different levels of expertise in the field; some questions permitted multiple responses (panels C and D).

discuss the challenges to accept and implement standards. We first discuss how metabolic models are built, reviewing existing standards and their application to metabolic modeling. Next, we highlight challenges that the metabolic modeling community faces in effectively utilizing these resources, identified from a community survey. Finally, we propose an integrated set of standards which we hope will serve as a checklist to improve accessibility, interpretability, and consistency of metabolic network reconstructions. We hope that our proposed checklists will help lower the activation energy required for both experts and newcomers alike to build new reconstructions or use existing reconstructions, as well as provide a model for sustainable standardization for other systems biology fields.

Standardization in metabolic modeling: a case study

The metabolic modeling community frequently utilizes COnstraint-Based Reconstruction and Analysis (COBRA) methods to build and compute computational models that represent an organism's metabolic phenotype. The construction of genome-scale

Table 1. Resources for using community standards and software tools.

Resource	Description	Link/references
MIRIAM*	Minimum Information Required In the Annotation of biochemical Models	(Le Novère *et al*, 2005)
MIASE*	Minimum Information About a Simulation Experiment	(Waltemath *et al*, 2011)
MEMOTE	MEtabolic MOdel TEsts	https://memote.io/
COBRA-related Google groups	Help for users of COBRApy, the python implementation of COBRA software	https://groups.google.com/forum/#!forum/cobra-pie
	Help for users of The COBRA Toolbox, the MATLAB implementation of COBRA software	https://groups.google.com/forum/#!forum/cobra-toolbox
	Discussion forum of the systems modeling community of the International Society for Computational Biology (ISCB)	https://groups.google.com/forum/#!forum/sysmod
COBRA GitHub	Repository for COBRA software, includes issue and help pages	https://github.com/opencobra/
COMBINE*	Community for coordinating standards for modeling in biology (umbrella organization for SBML, MIRIAM, MIASE, and more)	http://co.mbine.org
Kbase Help Board	Issue-tracking system to aid users to utilize tools and datasets	https://kbase.us/help-board/
SBML Validator*	Tests the syntax and internal consistency of an SBML file	http://sbml.org/Facilities/Validator/
SBML Test Suite*	Conformance testing system to test the degree and correctness of the SBML support provided in a software package	http://sbml.org/Software/SBML_Test_Suite/
BiGG Models	Freely accessible database of GEMs	http://bigg.ucsd.edu
BioModels*	Repository of mathematical models of biological and biomedical systems	https://www.ebi.ac.uk/biomodels/
MetaNetX	Platform for accessing, analyzing and manipulating GEMs	https://www.metanetx.org
SBO terms*	Systems Biology Ontology terms are a nested classification scheme to group model components	http://www.ebi.ac.uk/sbo/

Resources developed for broad applications in computational systems biology are denoted with an asterisk; unmarked resources are specific to the COBRA field.

metabolic network reconstructions and models is a multi-step process that involves the reconstruction of a metabolic network, manual curation to incorporate known physiology, computation of metabolic phenotypes, and the distribution of the models and results (Box 1). The COBRA field has been led by community-driven, open-source software efforts (Ebrahim *et al*, 2013; Heirendt *et al*, 2019) developed to enable these kinds of analyses, building on existing systems biological standards and principles.

Model structure

SBML is the *de facto* standard file format for storing and sharing biological data and systems biology models (Keating *et al*, 2020). SBML files encode biological models in a machine-readable format and are the most common format for editing and sharing metabolic reconstructions (Fig 1). SBML files contain lists of system components with corresponding parameters linking these components (e.g., metabolites in a reaction) and constraints (e.g., compartmentalization, reaction bounds). Saving a reconstruction as an SBML file thus inherently reinforces a set of standards. Further, the SBML field also offers several model validators and a test suite to identify non-standard formatting in COBRA models (Table 1).

Ultimately, SBML is just a serialization of a particular data model and other formats for sharing models exist. The format of the serialization itself is not crucial; what matters is the format's ability to represent the necessary data structures and whether information can be unambiguously encoded and made freely accessible. These standards must be widely accepted to be easily used in multiple software tools. This pervasiveness is essential—especially for network reconstructions—where the same knowledgebase could prove useful in various applications, requiring multiple tools in a complex analysis pipeline.

Model testing

There are different types of model evaluation processes. An important first step is to ensure a model is saved as a syntactically valid and machine-readable SBML file with a SBML validator (Table 1); however, valid syntax does not imply biological meaning or computational correctness. Thus, a model must also be evaluated for biological sense. A recent effort to improve standardization in the COBRA community resulted in MEMOTE, a set of MEtabolic MOdel TEsts (Lieven *et al*, 2020) to increase reproducibility and model quality through model evaluation. With this tool, users can generate a report to evaluate a reconstruction, including (i) namespace of components, (ii) biochemical consistency, (iii) network topology, and (iv) versioning. MEMOTE focuses on both the technical correctness (i.e., syntax) of a model while also providing metrics that can help users to evaluate the biological correctness of the model.

Namespaces are evaluated for metabolites, genes, and reactions to check annotations for coverage, consistency, and redundancy. To check for coverage, we might ask how many metabolites have an InChI key. To ensure consistency, we evaluate if the metabolites have the correct InChI keys. Namespaces can be evaluated for redundancy by identifying how many components have additional identifiers to more thoroughly document the component. Biochemical consistency is evaluated to verify the preservation of mass and charge across both individual reactions and the entire network. MEMOTE also reports the state of the software and environment versions used by the reconstruction and during the process of

Box 2: Proposed minimum standardized content for a metabolic network reconstruction. We propose that modelers use this list as a guide to help standardize accessibility, content, and quality; however, more comprehensive documentation and more interpretable and accessible information can only improve the usability and biological relevance of the shared reconstruction. See https://github.com/maureencarey/community_standards_supplemental for tutorials demonstrating the implementation of these requirements.

Model

- Recognized naming convention
  - historical approach: i + *authors initials* + *number of genes in model*, *e.g.*, iJE660 for the *E. coli* model constructed by Jeremy Edwards with 660 genes
  - recommended approach: i + *species indicator* + *iteration identifier*, *e.g.*, iPfal17 for *P. falciparum* published in 2017
- Machine-readable reference to organism embedded via MIRIAM annotation
  - full species name, including relevant identifiers if available (*e.g.*, NCBI reference genome)
  - taxonomy ID
  - strain ID, if necessary
  - tissue type, if necessary
  - URL to obtain genome
- Reference information
  - DOI
  - Author(s) names and contact information embedded
- Consistent namespace for all model identifiers

Metabolite

- Human-readable, descriptive name (e.g., D-Glucose)
- Charge (e.g., 0)
- Chemical formula (e.g., C6H12O6)
- Structural identifiers
  - InChI strings (if pH is known, pH-relevant InChI ID for each metabolite)
  - SMILES (optional)
- At least one database identifier from a reliable resource, such as
  - MetaNetX
  - BiGG
  - KEGG Compound
  - ChEBI
  - ModelSEED
  - HMDb
  - MetaCyc

Biochemical reaction

- Human-readable, descriptive name (e.g., phosphofructokinase)
- Reaction formula (e.g., ATP + L-glutamate + ammonium $\rightleftharpoons$ ADP + L-glutamine + H^+ + phosphate)
- At least one database identifier from a reliable resource, such as
  - MetaNetX
  - Rhea
  - BiGG
  - KEGG Reaction
  - ModelSEED
  - MetaCyc
- EC Number
- Associated genes (gene-protein-reaction rule, or GPR)
- Recommended: systems biology ontology (SBO) terms

Gene

- Name or gene symbol
- DNA and/or Protein sequence ID (i.e., a mechanism for mapping this information to a sequence)
  - Entrez
  - Ensembl
  - UniProt
  - Other field-specific database identifier
- Position (including chromosome, if applicable)

testing the reconstruction. The recent development of such a community-defined testing suite should improve the rigor of the field, particularly by tailoring general systems biology resources to our specific use cases. We encourage the community to make the use of MEMOTE an expected standard for newly published models, when applicable.

Models should also be evaluated for biological accuracy, whenever possible. Tests evaluating network topology can be used to evaluate the more subjective features of the model, by using connectedness as a proxy for inferring the scope of manual curation or the quality of a reconstruction. However, this requires such topological measures (and/or machine learning; Medlock & Papin, 2020) to be combined with biological knowledge of the system. Condition-specific tests (often referred to as "metabolic tasks" in the COBRA field) are developed to evaluate the biological meaning of the network and attempt to represent specific biochemical experiments. Examples include production or consumption of particular metabolites given a set of constraints. Metabolic tasks can be generated for each model (including tissue-specific models) and to evaluate iterative rounds of curation.

Model quality and content

Many of the standards used in the COBRA field were developed by interdisciplinary teams of modelers and software developers for broad use in the computational biological modeling field (Table 1); we can use these existing resources or adapt them for use in our field as done with MEMOTE (Box 1). Minimum and recommended quality standards have been formulated and presented as a set of expectations for biological models and simulations through (i) the Minimum Information Required In the Annotation of biochemical Models (MIRIAM; Le Novère *et al*, 2005) and (ii) Minimum Information About a Simulation Experiment (MIASE; Waltemath *et al*, 2011), respectively. However, engagement in the COBRA field in particular has been modest, likely due to community members' lack of familiarity with these resources and the challenges associated with updating these recommendations with new data types and applications. In the following sections, we discuss potential challenges facing the widespread adoption of these standards in the COBRA field and possible solutions.

Box 3: Proposed checklist for reviewers. We propose that reviewers of manuscripts that include a novel metabolic network reconstruction use this list as a guide to help standardize accessibility, content, and quality

Reconstruction

Availability:
- Is the reconstruction publicly available?
  - On BioModels?
  - Elsewhere? (optional)
- Is the reconstruction shared on an accessible database?

Formatting:
- Is the model saved in a language-independent format (i.e., SBML)?
- Optional: additional formats (e.g., XLS, JSON)

Nomenclature:
- Does the reconstruction's name indicate a version?
- Does the reconstruction's name indicate an organism?
- Are identifiers (genes, metabolites, reactions) consistently from one namespace?

Optimization:
- Have MEMOTE (Lieven *et al*, 2020) tests been run?
- Is the objective reaction indicated?
- Is evidence (i.e., references) provided for use of the objective function(s)?
- Are exchange, sink, and demand reactions and all necessary constraints included as defaults or in code?
- Can the reconstruction be instantiated without error with COBRA software?

Simulations:
- Are simulation parameters (objective reaction, constraints, etc.) provided in any of the following formats (include at least one):
  - README.md file?
  - COMBINE repository?
  - full analytic code? *e.g.*, iPython notebook or equivalent
- Is the COBRA software version documented?
- Are the solvers documented?

Manuscript:
- Are COBRA software efforts appropriately credited?
- Are previous iterations and/or other versions of the reconstruction appropriately credited?
- Is the model clearly referenced (i.e., with a resolvable link and identifier) in Materials & Method or in the Data Availability Statement?

Challenges preventing the use of standards

Despite these efforts, many genome-scale metabolic network models fail to meet minimum standards and quality metrics. Ravikrishnan and Raman found that almost 60% of models had no standardized (i.e., interpretable) metabolite identifiers, 36% could not be evaluated for mass imbalances due to unstandardized formatting, and 35% did not contain gene-protein-reaction associations in the SBML file (Ravikrishnan & Raman, 2015). This is a broad challenge throughout systems biology fields (Stanford *et al*, 2015). As a community, we must therefore ask why standards are not used more broadly if they enable the sharing, reuse, and evaluation of biological models and associated simulations. At the 5th Annual Conference on Constraint-based Reconstruction and Analysis (COBRA, October 14–16, 2018), we surveyed the COBRA community regarding the use of community standards. This survey identified two major causes for the lack of standardization in the COBRA field (full anonymized survey results provided in Dataset EV1).

First, the responses identified several biological phenomena that are not captured by current standards. For example, modelers of intracellular pathogen metabolism struggle to comply with nomenclature and mass balance when adding both pathogen and host biochemistry (Box 4; Carey *et al*, 2017). Similarly, it is challenging to use the correct and sufficiently detailed nomenclature for biologically relevant tautomers and polymers. While such issues will likely only be relevant in specific biological applications, it is vital that community-adopted standards can and do evolve to address these increasingly-common edge cases.

Second, users identified a set of novel analyses that current standards do not sufficiently support. Existing standards are inherently insufficient for novel techniques. Extensive community networks—such as modeling multiple members of the microbiota—and modeling macromolecular expression mechanisms represent current areas in metabolic modeling where some standards are currently lagging. Although standards evolve as the field progresses, they inherently cannot capture the latest cutting-edge developments. This "lag" in standardization is not field-specific and such cutting-edge examples will likely only be identified in novel methods development. Both of these user-identified limitations require community-driven efforts to update standards as the field expands into new application areas and with novel analytic approaches.

We hypothesize that two additional factors play a role in these standardization challenges. First, biologists, modelers, and software developers are sometimes "siloed" into separate communities and with distinct motivating factors (e.g., research interests, funding mechanisms). As a result, biologists and modelers are often not aware of relevant resources generated by software developers. Our survey identified that fewer than 25% of researchers in the COBRA field were familiar with MIASE and only 56% were aware of MIRIAM; these best practices cannot be used if they are not known. In turn, biological limitations—like those discussed above—might not be relayed to software developers focusing on a standard formulation. Thus, even community-driven efforts do not necessarily move laterally across subdisciplines. Second, as users, the lack of standardization often makes it easier to generate a novel reconstruction or analytic tool than to improve upon an existing version, further diversifying the set of existing approaches and amplifying the challenge of developing unifying standards.

Community-driven solutions

To remove these barriers, we suggest the field shifts to incentivize standardization by promoting model reuse and markers of quality; ultimately, this practice will improve

Box 4: In Box 2, we identified a proposed minimum standardized content for a metabolic network reconstruction. Here, we list the associated components in the example model, iPfal19, and discuss some of the challenges in implementing these standards. This is the third iteration of the *Plasmodium falciparum* 3D7 genome-scale metabolic network reconstruction. The original reconstruction was generated using a custom pipeline and multiple rounds of curation were conducted (Carey *et al*, 2017). iPfal19 fails to compile with several of the recommended guidelines, see notes for explanations and the MEMOTE report for other issues. Additionally, the README file associated with this model is sparser than ideal due to the lack of documentation associated with the original curation efforts.

Model

- Recognized naming convention
  - iPfal19: i + *species indicator* + *iteration identifier*
- model metadata (organism name, curation history, genome, authors, etc.)
- ORFs were called manually using proteomics and RNASeq data and compiled on the malaria parasite database; thus, no NCBI/refseq IDs (etc.) would accurately represent the genome used

Metabolite

- 100% of metabolites have a human readable, descriptive name
- 94% of metabolites have a charge and chemical formula[a]
- 63% of metabolites have InChI strings, although because the pH for each subcellular compartment is known, these strings might not represent the appropriate species (i.e., protonation status)[b]
- 100% of metabolites have an ID from BiGG (100% have BiGG-like IDs)[c]
- 91, 64.4, 68.7, 75, 47.5, and 68.4% of metabolites also have an ID from MetaNetX, KEGG Compound, ChEBI, ModelSEED, HMDb, or MetaCyc, respectively

Biochemical reaction

- 100% of reactions have a human readable, descriptive name
- 100% of reactions have a reaction formula
- 60.6% of reactions have an ID from BiGG (100% have BiGG-like IDs)[c]
- 58.6, 20.7, 0, and 23.3% of metabolic reactions have an ID from MetaNetX, KEGG Reaction, ModelSEED, and MetaCyc, respectively
- 27.6% of reactions have an EC Number[b,d]

Gene

- Gene IDs use PlasmoDB gene nomenclature, consistent with malaria field. These IDs map to a genomic location, DNA sequence, and protein sequence on PlasmoDB.org
- 71.1% of reactions have GPRs

[a]Some metabolites do not have a charge or formula associated with them including metabolites representing host or parasite proteins. If many different proteins can participate in a reaction, the reaction contains a generic reactant to represent all of the possible protein reactants. Of the 6% problematic metabolites, nearly all are proteins or aggregate metabolites.
[b]Not all BiGG IDs are mapped to InChI strings, EC numbers, or other useful identifiers (e.g., https://www.metanetx.org/chem_info/MNXM4217). This interferes with some MEMOTE functionality, such as identifying duplicate reactions.
[c]BiGG-like IDs are proposed new BiGG IDs consistent with the general naming approach in BiGG. For example, new IDs (pheme_fv, pheme_ap) have been created for protoheme corresponding to protoheme located in parasite-specific compartments, the food vacuole and apicoplast, respectively, consistent with existing BiGG IDs for protoheme (http://bigg.ucsd.edu/universal/metabolites/pheme). New reactions are created when the existing BiGG reaction occurs in only one compartment but should be present elsewhere in iPfal19. For example, "PLIPA2A120pp" is a periplasmic reaction in BiGG but occurs in the cytoplasm of P. falciparum; "pp" is the suffix used to denote the periplasm so the new cytoplasmic version is named "PLIPA2A120." Additionally, new aggregate reactions (i.e., relevant pseudoreactions) were created and named intuitively (e.g., lipid1, lipid2).
[d]Transporters, exchange reactions, and aggregate reactions should not have an EC number and these make up 35% of all reactions.

communication among biologists, modelers, and software developers. One solution to increase the standardization in the field is to mandate compliance during the manuscript publication process—more than 85% of community-survey responders think this should occur. However, sharing a noncompliant reconstruction or even for failing to make a reconstruction publicly available has unfortunately little consequences for authors. Incentivizing standardization through funding models is inherently challenging: Funding for science is evaluated in the short term, whereas the benefits of software or resource development are observed on a longer time scale. We can integrate funding for infrastructure (e.g., software and standards development) into applied projects (i.e., research projects). This funding approach will reflect the paired nature of these two kinds of COBRA projects.

SBML facilitates the addition of information that is specific to one particular tool or use case; once this function becomes more widely used and necessary, it can be turned into an extension package for SBML. Thus, SBML supports these edge-case to be encoded in a standardized fashion. Hence, standards should provide such features to give enough freedom to developers of models and tools. However, this process relies on the use of such tools (e.g., SBML) and communication between the communities that design standards and the communities that use them. Such interactions could be stimulated through scientific meetings: Each conference could have a dedicated keynote presentation by a representative from the other community, followed by a panel discussion led by the presenting representative. By maintaining clear contributing instructions for the COBRA software suites, associated analysis packages, and infrastructure (e.g., SBML model format and associated API libraries), the community can update and extend standards to address edge cases.

Recommendations for standards

In response to some of the issues and challenges outlined above, we propose a set of guidelines to help improve the accessibility, content, and quality of metabolic network reconstructions—both for those creating reconstructions/models (Box 1 and 2) and those peer-reviewing reconstructions/models (Box 3). The suggestion of these standards was informed by panel discussions at the COBRA 2018 conference and from the community poll results (Dataset EV1), as well as previous community efforts (Stanford *et al*, 2015). Our recommendations here represent field-specific implementation of the FAIR Data Principles (Wilkinson *et al*, 2016), a set of guidelines intended to improve reproducible research (Sansone *et al*, 2019).

First, focusing on the reconstruction process, we propose that a reconstruction metadata file is shared and includes model building information, such as the genome, database, and software versions [example README.md is provided at https://github.

com/maureencarey/community_standards_
supplemental, see also discussion in Box 4, or COMBINE archive in Additional file 2 of (Bergmann *et al*, 2014)]. Although this information is likely in the original manuscript, this format would link the reconstruction to the reconstruction file. The COMBINE archive also facilitates including details on gene and protein sequences, rather than mere IDs. Models should be shared on at least one publicly available repository (e.g., BioModels); because visibility increases usability, authors may want to share a model via other means as well (e.g., laboratory website, BiGG).

Second, we encourage the use of version control and specific effort to document automated and manual curation. Version control can be implemented in multiple ways, mainly through a publicly available repository that includes all iterations or by making all versions publicly available and identifiable through clear naming conventions. Further, we propose that all curation efforts be documented in the reconstruction and explicitly include a literature reference and notes in the annotations field of a reaction.

Third, we emphasize the need for MIASE requirements (Waltemath *et al*, 2011) when sharing simulation results. These data about experimental data, constraints, and versioning can be stored in a COMBINE repository or the analytic code, if publicly available. Ultimately, a standardized format (like COMBINE) could enable minor advances in COBRA software to facilitate the re-implementation of a simulation.

Looking ahead

Here, we have summarized existing standards in the COBRA field and identified challenges associated with both the development and compliance of software and model standards. We have proposed "checklists" for use during both the reconstruction and peer review processes that will help improve the accessibility, content, and quality of metabolic network reconstructions. Additional community-inspired challenges and results from the COBRA community survey conducted in early 2019 are documented in Dataset EV1; we hope these examples will inspire new discussions and novel solutions.

There exist several open challenges for the field regarding the adoption of and development of new standards. We must embrace flexible standardization to facilitate their adoption and to build upon existing work. For example, although resources like MetaNetX (Moretti *et al*, 2016) and the BiGG Models database (Norsigian *et al*, 2020) facilitate the mapping of genes, reactions, and metabolites across the different namespaces, nomenclature discrepancies remain a challenge and sometimes result in redundant nonstandardized efforts. Another challenge is for community standard development to be derived from the community instead of in a top-down manner. While this organizational structure is currently in effect for the SBML community, it only functions if there is community participation—we need those who use the standards (i.e., modelers) to raise their hands and participate in the decision making process.

Ultimately, community adherence to standards will improve modeling reproducibility and better document the reconstruction process. We hope that the community embraces existing standards and our community-driven suggestions moving forward—both during the preparation of manuscripts and during the peer review process—and anticipate that compliance will increase the rigor of the field while simultaneously making it easier for scientists from other disciplines to build and use metabolic models.

Acknowledgements

The authors would like to thank those who contributed to the community survey and provided feedback in the Community Standards and Resources Panel at COBRA 2020, as well as on the survey and draft manuscript. This work was funded by the Institute for Systems Biology's Translational Research Fellows Program (J.T.Y.), the University of Virginia's Engineering-in-Medicine Seed Funding (M.C. and J.P.), the PhRMA Foundation Postdoctoral Fellowship for Translational Medicine and Therapeutics (M.C.), the National Institutes of Health (2R01GM070923-13 to A.D. and R01AT010253 to J.P.), the European Union's Horizon 2020 research and innovation program under grant agreement No 686070 (DD-DeCaF to M.E.B.), the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation to A.D. via Cluster of Excellence: Controlling Microbes to Fight Infections EXC 2124), and the German Center for Infection Research (DZIF, Deutsche Zentren der Gesundheitsforschung 8020708703 to A.D.).

Conflict of interest

The authors declare that they have no conflict of interest.

References

Bergmann FT, Adams R, Moodie S, Cooper J, Glont M, Golebiewski M, Hucka M, Laibe C, Miller AK, Nickerson DP *et al* (2014) COMBINE archive and OMEX format: one file to share all information to reproduce a modeling project. *BMC Bioinformatics* 15: 369

Carey MA, Papin JA, Guler JL (2017) Novel Plasmodium falciparum metabolic network reconstruction identifies shifts associated with clinical antimalarial resistance. *BMC Genom* 18: 543

COBRA (October 14–16, 2018) COBRA 2018 (Seattle, WA). *5th Conference on Constraint-Based Reconstruction and Analysis (COBRA 2018)*.

Dräger A, Palsson BØ (2014) Improving collaboration by standardization efforts in systems biology. *Front Bioeng Biotechnol* 2: 61

Ebrahim A, Lerman JA, Palsson BO, Hyduke DR (2013) COBRApy: COnstraints-based reconstruction and analysis for python. *BMC Syst Biol* 7: 74

Ebrahim A, Almaas E, Bauer E, Bordbar A, Burgard AP, Chang RL, Dräger A, Famili I, Feist AM, Fleming RMT *et al* (2015) Do genome-scale models need exact solvers or clearer standards? *Mol Syst Biol* 11: 831

Heirendt L, Arreckx S, Pfau T, Mendoza SN, Richelle A, Heinken A, Haraldsdóttir HS, Wachowiak J, Keating SM, Vlasov V *et al* (2019) Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat Protoc* 14: 639 – 702.

Keating SM, Waltemath D, König M, Zhang F, Dräger A, Chaouiya C, Bergmann FT, Finney A, Gillespie CS, Helikar T *et al* (2020) SBML Level 3: an extensible format for the exchange and reuse of biological models. *Mol Syst Biol* https://doi.org/10.15252/msb.20199110

Le Novère N, Finney A, Hucka M, Bhalla US, Campagne F, Collado-Vides J, Crampin EJ, Halstead M, Klipp E, Mendes P *et al* (2005) Minimum information requested in the annotation of biochemical models (MIRIAM). *Nat Biotechnol* 23: 1509 – 1515

Lieven C, Beber ME, Olivier BG, Bergmann FT, Ataman M, Babaei P, Bartell JA, Blank LM, Chauhan S, Correia K *et al* (2020) MEMOTE for standardized genome-scale metabolic model testing. *Nat Biotechnol*: 38: 272 – 276

Medlock GL, Papin JA (2020) Guiding the refinement of biochemical knowledgebases with ensembles of metabolic networks and machine learning. *Cell Syst* 10: 109 – 119.e3

Monk J, Nogales J, Palsson BO (2014) Optimizing genome-scale network reconstructions. *Nat Biotechnol* 32: 447 – 452

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

Norsigian CJ, Pusarla N, McConn JL, Yurkovich JT, Dräger A, Palsson BO, King Z (2020) BiGG models 2020: multi-strain genome-scale models and expansion across the phylogenetic tree. *Nucleic Acids Res* 48: D402 – D406

Pan S, Reed JL (2018) Advances in gap-filling genome-scale metabolic models and model-driven experiments lead to novel metabolic discoveries. *Curr Opin Biotechnol* 51: 103 – 108

Ravikrishnan A, Raman K (2015) Critical assessment of genome-scale metabolic networks: the need for a unified standard. *Brief Bioinform* 16: 1057 – 1068

Sansone S-A, McQuilton P, Rocca-Serra P, Gonzalez-Beltran A, Izzo M, Lister AL, Thurston M, FAIRsharing Community (2019) FAIRsharing as a community approach to standards, repositories and policies. *Nat Biotechnol* 37: 358 – 367

Stanford NJ, Wolstencroft K, Golebiewski M, Kania R, Juty N, Tomlinson C, Owen S, Butcher S, Hermjakob H, Le Novère N *et al* (2015) The evolution of standards and data management practices in systems biology. *Mol Syst Biol* 11: 851

Tang YA, Pichler K, Füllgrabe A, Lomax J, Malone J, Munoz-Torres MC, Vasant DV, Williams E, Haendel M (2019) Ten quick tips for biocuration. *PLoS Comput Biol* 15: e1006906

Thiele I, Palsson BØ (2010) A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat Protoc* 5: 93 – 121

Waltemath D, Adams R, Beard DA, Bergmann FT, Bhalla US, Britten R, Chelliah V, Cooling MT, Cooper J, Crampin EJ *et al* (2011) Minimum information about a simulation experiment (MIASE). *PLoS Comput Biol* 7: e1001122

Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE *et al* (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3: 160018