

Improving Synthetic Biology Communication: Recommended Practices for Visual Depiction and Digital Submission of Genetic Designs

Nathan J. Hillson,^{*,†,‡,§,||} Hector A. Plahar,^{†,‡,||} Jacob Beal,^{*,‡,⊥} and Ranjini Prithvira^{‡,#}

[†]Fuels Synthesis and Technology Divisions, DOE Joint BioEnergy Institute (JBEI), Emeryville, California 94608, United States

[‡]Biological Systems and Engineering Division, Lawrence Berkeley National Lab, Berkeley, California 94720, United States

[§]DOE Joint Genome Institute, Walnut Creek, California 94598, United States

^{||}Synthetic Biology Engineering Research Center, Emeryville, California 94608, United States

[⊥]Raytheon BBN Technologies, Cambridge, Massachusetts 02138, United States

[#]ACS Synthetic Biology, American Chemical Society, Washington, D.C. 20036, United States

ABSTRACT: Research is communicated more effectively and reproducibly when articles depict genetic designs consistently and fully disclose the complete sequences of all reported constructs. *ACS Synthetic Biology* is now providing authors with updated guidance and piloting a new tool and publication workflow that facilitate compliance with these recommended practices and standards for visual representation and data exchange.



In every scientific and engineering discipline, communication about research and development is made more effective by the adoption of community standards.¹ For synthetic biology, it has previously been suggested that research articles should consistently depict genetic designs^{2,3} and fully disclose the complete sequences of all the constructs they describe.^{4–6} To this end, *ACS Synthetic Biology* editorial policy⁷ has, from its beginning, stated that authors should deposit their reported organisms and sequence data into established public repositories (where possible) or in Supporting Information. *ACS Synthetic Biology* is now providing authors with updated guidance along with new tools that make it easier for authors to follow these recommended practices, using community-developed standards for visual representation^{2,3} and data exchange,^{8–10} along with digital and physical repositories.^{11–15}

■ VISUAL REPRESENTATION OF GENETIC DESIGNS

People engineering biological organisms often need to communicate in diagrams, both about the structure of nucleic acid sequences and about the functional relationships of nucleic acid sequences, their components, and other biological substances. Standardized languages for such diagrams can improve communication by decreasing ambiguity and increasing the amount of information that can be conveyed without accompanying explanation. At the same time, languages are more widely applicable when they allow as much flexibility and freedom as possible in how diagrams are organized, presented, and styled, either by hand or with the aid of various software programs.

For these reasons, *ACS Synthetic Biology* now recommends that authors visually depict genetic designs in their manuscript figures using the Synthetic Biology Open Language (SBOL)

Visual standard^{2,3} whenever applicable. SBOL Visual codifies and refines what are already fairly common practices in the representation of genetic sequence designs, in which glyphs representing sequence features are attached to lines representing nucleic acid backbone (Figure 1). Use of SBOL Visual will not

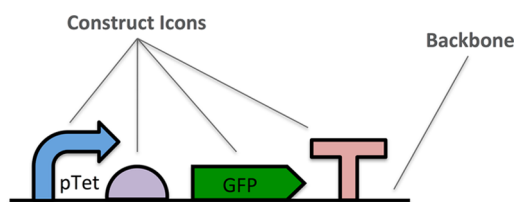


Figure 1. Example of SBOL visual representation of a genetic design.

inhibit the stylistic choices of authors, as it imposes no constraints on color, text, scaling, etc. The main effect of adopting SBOL Visual is that certain glyphs are now given formal definitions and must only be used to represent sequence features with those types, e.g., authors may not arbitrarily choose that in their diagrams the “bent arrow” commonly used to represent a promoter will instead mean a terminator. For any type of sequence feature or glyph not specifically defined by SBOL Visual, however, authors are as always encouraged to create the clearest imagery they can to represent it, and to submit their new glyphs to be considered for adoption into SBOL Visual. Full information about SBOL Visual, including a table of glyphs and

Special Issue: IWBD 2015

Received: May 20, 2016

Published: June 7, 2016

simple introductory slides, can be found at <http://sbolstandard.org/visual>.

DIGITAL REPRESENTATION OF SEQUENCE INFORMATION

ACS *Synthetic Biology* has long recommended that authors should make their reported organisms and sequence data fully available. On the basis of recent advancements in both representation format and available repositories, ACS *Synthetic Biology* now recommends that all authors digitally deposit their genetic designs using the Synthetic Biology Open Language (SBOL) data exchange format.^{8,9} Full information about SBOL can be found at <http://sbolstandard.org>. In particular, it is recommended that authors use SBOL version 2.0 or above, as this allows representation of not only sequence and sequence features, but also the expected functional interactions between design elements.

Authors may benefit from using a variety of community software tools¹⁶ when preparing genetic design figures and digital files. One tool in particular for authors to consider when preparing their genetic design files is the SBOL Validator/Converter tool,¹⁷ that will validate that a given file complies with the SBOL 2 data exchange standard as well as convert other file formats (e.g., Genbank, FASTA, SBOL 1) to SBOL 2.

WORKFLOW FOR GENETIC DESIGN DEPOSIT AND REVIEW

In order to simplify sharing of genetic design information for both authors, editors, and reviewers, ACS *Synthetic Biology* has instantiated a digital registry of genetic designs and microbial strains at <https://acs-registry.jbei.org>. Authors are encouraged to use this registry to deposit their sequence and design information, as well as organism information, during the manuscript submission process.

Authors using the ACS *Synthetic Biology* Registry will follow the following steps:

1. Create a user account in the ACS *Synthetic Biology* Registry, and log in.
2. Create a personal private collection (named following the convention “FirstAuthorLastName *et al.* Year”, e.g., “Hillson *et al.* 2012”) for the manuscript to be submitted.
3. Create private entries for each reported strain and DNA sequence (preferably deposited in SBOL format, automatically validated during submission to the Registry, but alternatively in Genbank or FASTA formats, which will be automatically converted to SBOL). If physical samples of the reported strains and DNA sequences are available from public repositories (e.g., from AddGene¹⁵), the author can provide the corresponding repository link in the samples section of a given entry.
4. Place all of these entries in the manuscript’s collection.
5. In the Supporting Information, provide the URL to the manuscript’s private collection (e.g., <https://acs-registry.jbei.org/folders/1>), which during the review process only ACS *Synthetic Biology* editorial staff will have access to. Further details can be found in the journal’s Information to Authors.⁷

A series of video tutorials (for the ICE repository platform,¹⁴ the open source foundation of the ACS *Synthetic Biology* Registry), which are available *via* a YouTube playlist,¹⁸ may be helpful for authors new to the ACS *Synthetic Biology* Registry. For entries associated with DNA sequences, the ACS *Synthetic*

Biology Registry also provides an SBOL Visual rendering of the genetic design (*via* Pigeon¹⁹) that may be useful during the figure preparation process.

Once an author using the ACS *Synthetic Biology* Registry has submitted a manuscript for publication (it is crucial that a link to the collection of entries in the ACS *Synthetic Biology* Registry be included in the Supporting Information), ACS *Synthetic Biology* editorial staff will generate a zip file that contains a validated SBOL 2 (as well as a Genbank) format file for each reported DNA sequence, in addition to a comma-separated value (CSV) spreadsheet file that lists additional information associated with each of the reported strains and DNA sequences. This zip file will then be uploaded as additional Supporting Information for the manuscript, which reviewers will have access to during the review process. Once the manuscript is published, the ACS *Synthetic Biology* editorial staff will set the manuscript’s collection to public, so that anyone with an ACS *Synthetic Biology* Registry account (or with an account on another registry connected *via* Web of Registries functionality, Plahar *et al.*, unpublished) will have informatic access to all of the reported strains and DNA sequences. Should the manuscript not be published, however, the manuscript’s collection and associated strain and DNA sequence entries will remain private to the author.

FUTURE DIRECTIONS

Community standards and practices are not static, but may be expected to continue evolving over time, as the needs and knowledge of the community evolves. The adoption of these standards by ACS *Synthetic Biology* and the use of the registry workflow we describe is at the moment in its pilot phase, and compliance with SBOL Visual and SBOL 2 data standards will (for the moment at least) remain a voluntary opt-in recommended practice, and not an editorial mandate. Manuscripts that do follow these standards, however, will be marked with symbols indicating that they follow community standards for diagrams, data representation, and sharing of data in an approved public repository (such as the ACS *Synthetic Biology* Registry).

“As the complexity of synthetic biology projects grow, it will be critical to standardize the exchange of designs and data”, says Christopher Voigt, Editor-in-Chief, ACS *Synthetic Biology*.

We expect that both the above-described workflow and the standards that support it will be continually improved over time. The standards are open community efforts to which all practitioners are encouraged to contribute, in order to ensure they best serve the needs of the greatest number of synthetic biologists. For the workflow, it is anticipated that it will in the future expand to also incorporate other community repositories that provide the same functionality prerequisites (*i.e.*, stable public access for the community, SBOL 2.0+ data-exchange, granular authentication/authorization access controls, URL addressable collections of entries, Web of Registries integration, and editor tools for collection generation and publication). Future workflows may also integrate reviewing more directly with repositories, so that reviewers can benefit from repository tools for exploration and visualization of design collections.

It is hoped that these new recommendations for authors and the availability of the new ACS *Synthetic Biology* Registry will add value to the manuscript publications process for authors, reviewers, editors, and the broader scientific community alike. Feedback from the community will contribute to the continual improvement of the synthetic biology publishing process.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: njhillson@lbl.gov.

*E-mail: jakebeal@bnn.com.

Author Contributions

NJH and RP designed the genetic design digital submission workflow. HAP developed the software user interface (available under a freely open-source BSD license from <https://github.com/JBEI/ice>) supporting the genetic design digital submission workflow (specifically, the new Editor Tools functionality). NJH and JB wrote the initial draft of the manuscript, and all authors edited the manuscript.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors would like to thank Oge Nnadi for an initial prototype of a software user interface supporting the genetic design digital submission workflow; Chris Myers, Michal Galdzicki, Douglas Densmore, and Anil Wipat for their contributions to the conception of the genetic design digital submission workflow; and the entirety of the Synthetic Biology Open Language Developers Group (<http://sbolstandard.org/development/developers/>) for their helpful suggestions and feedback on this manuscript including genetic design visual depiction, data representation, and the digital submission workflow. The authors would also like to thank Chris Myers and Herbert Sauro for leading the development of the SBOL software infrastructure that has made this work possible (supported by National Science Foundation Grants DBI-1356041 and DBI-1355909). This work has been supported in part by the DOE Joint BioEnergy Institute (<http://www.jbei.org>) and the DOE Joint Genome Institute (<http://jgi.doe.gov>) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy. This work was also supported by the Synthetic Biology Engineering Research Center (SynBERC) through National Science Foundation grant NSF EEC 0540879. Any opinions, findings, conclusions, or recommendations expressed herein are those of the authors and do not necessarily reflect the views of these funding agencies. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

REFERENCES

- (1) Slattery, W. J. (1971) *An Index of US Voluntary Engineering Standards; Covering Those Standards, Specifications, Test Methods, and*
- (2) Quinn, J. B., Bhatia, S., Cai, P., Chen, J., Hill, N., Clancy, K., Maheshwari, A., Umesh, P., Pocock, M., Rodriguez, C., Stan, G.-B., and Endy, D. (2013) *Synthetic Biology Open Language Visual (SBOL Visual)*, version 1.0.0, BioBricks Foundation Request for Comments (BBF RFC).
- (3) Quinn, J. Y., Cox, R. S., 3rd, Adler, A., Beal, J., Bhatia, S., Cai, Y., Chen, J., Clancy, K., Galdzicki, M., Hillson, N. J., Le Novere, N., Maheshwari, A. J., McLaughlin, J. A., Myers, C. J., P, U., Pocock, M., Rodriguez, C., Soldatova, L., Stan, G. B., Swainston, N., Wipat, A., and Sauro, H. M. (2015) SBOL Visual: A Graphical Language for Genetic Designs. *PLoS Biol.* 13, e1002310.
- (4) Galdzicki, M., Chandran, D., Gennari, J. H., and Sauro, H. M. (2011) Data Model Standardization for Synthetic Biomolecular Circuits and Systems, In *Design and Analysis of Biomolecular Circuits: Engineering Approaches to Systems and Synthetic Biology* (Koepl, H., Setti, G., di Bernardo, M., and Densmore, D., Eds.), pp 281–293, Springer: New York.
- (5) Peccoud, J., Anderson, J. C., Chandran, D., Densmore, D., Galdzicki, M., Lux, M. W., Rodriguez, C. A., Stan, G. B., and Sauro, H. M. (2011) Essential information for synthetic DNA sequences. *Nat. Biotechnol.* 29, 22.
- (6) Blaxter, M., Danchin, A., Savakis, B., Fukami-Kobayashi, K., Kurokawa, K., Sugano, S., Roberts, R. J., Salzberg, S. L., and Wu, C.-I. (2016) Reminder to deposit DNA sequences. *Science* 352, 780–780.
- (7) ACS Synthetic Biology Information for Authors. http://pubs.acs.org/paragonplus/submission/asbcd6/asbcd6_authguide.pdf.
- (8) Bartley, B., Beal, J., Clancy, K., Misirli, G., Roehner, N., Oberortner, E., Pocock, M., Bissell, M., Madsen, C., Nguyen, T., Zhang, Z., Gennari, J. H., Myers, C., Wipat, A., and Sauro, H. (2015) Synthetic Biology Open Language (SBOL) Version 2.0.0. *J. Integr. Bioinform.* 12, 272.
- (9) Roehner, N., Beal, J., Clancy, K., Bartley, B., Misirli, G., Grunberg, R., Oberortner, E., Pocock, M., Bissell, M., Madsen, C., Nguyen, T., Zhang, M., Zhang, Z., Zundel, Z., Densmore, D., Gennari, J. H., Wipat, A., Sauro, H. M., and Myers, C. J. (2016) Sharing Structure and Function in Biological Design with SBOL 2.0. *ACS Synth. Biol.*, DOI: 10.1021/acssynbio.5b00215.
- (10) Galdzicki, M., Clancy, K. P., Oberortner, E., Pocock, M., Quinn, J. Y., Rodriguez, C. A., Roehner, N., Wilson, M. L., Adam, L., Anderson, J. C., Bartley, B. A., Beal, J., Chandran, D., Chen, J., Densmore, D., Endy, D., Grunberg, R., Hallinan, J., Hillson, N. J., Johnson, J. D., Kuchinsky, A., Lux, M., Misirli, G., Peccoud, J., Plahar, H. A., Sirin, E., Stan, G. B., Villalobos, A., Wipat, A., Gennari, J. H., Myers, C. J., and Sauro, H. M. (2014) The Synthetic Biology Open Language (SBOL) provides a community standard for communicating designs in synthetic biology. *Nat. Biotechnol.* 32, 545–550.
- (11) SynBioHub: A parts repository for synthetic biology. <http://sbolhub.org/>.
- (12) Registry of Standard Biological Parts. <http://parts.igem.org>.
- (13) Cooling, M. T., Rouilly, V., Misirli, G., Lawson, J., Yu, T., Hallinan, J., and Wipat, A. (2010) Standard virtual biological parts: a repository of modular modeling components for synthetic biology. *Bioinformatics* 26, 925–931.
- (14) Ham, T. S., Dmytriv, Z., Plahar, H., Chen, J., Hillson, N. J., and Keasling, J. D. (2012) Design, implementation and practice of JBEI-ICE: an open source biological part registry platform and tools. *Nucleic Acids Res.* 40, e141.
- (15) Kamens, J. (2015) The Addgene repository: an international nonprofit plasmid and data resource. *Nucleic Acids Res.* 43, D1152–1157.
- (16) SBOL Software Tools. <http://sbolstandard.org/software/tools/>.
- (17) SBOL Validator/Converter tool. <http://www.async.ece.utah.edu/sbol-validator/>.
- (18) ICE Platform Tutorial Video YouTube Playlist. <https://www.youtube.com/playlist?list=PLAkVdehjfmlLYSdU2Jq6XZIWNiz-Fev>.
- (19) Bhatia, S., and Densmore, D. (2013) Pigeon: a design visualizer for synthetic biology. *ACS Synth. Biol.* 2, 348–350.