Combining standards for tomorrow's models: Report of the 2015 Whole-Cell Modeling Summer School

Dagmar Waltemath[1], Falk Schreiber[2,3], Jonathan R. Karr[4], Michael Hucka[5], Chris J. Myers[6], Frank T. Bergmann[5,7,8], Vijayalakshmi Chelliah[9], Marcus Krantz[10], Wolfram Liebermeister[11], Pedro Mendes[12–15], Pinar Pir[16] and 2015 Whole-Cell Modeling Summer School Consortium[17]

[1]Department of Systems Biology and Bioinformatics, University of Rostock
[2]Clayton School of Information Technology, Monash University
[3]Institute of Computer Science, Martin Luther University Halle-Wittenberg
[4]Department of Genetics & Genomic Sciences, Icahn School of Medicine at Mount Sinai
[5]Department of Computing and Mathematical Sciences, California Institute of Technology
[6]Department of Electrical and Computer Engineering, University of Utah
[7]BioQuant, University of Heidelberg
[8]Centre for Organismal Studies, University of Heidelberg
[9]European Molecular Biology Laboratory, European Bioinformatics Institute
[10]Department of Biology, Humboldt University of Berlin
[11]Institute of Biochemistry, Charité Medical University of Berlin
[12]Manchester Institute of Biotechnology, University of Manchester
[13]School of Computer Science, University of Manchester
[14]Center for Quantitative Medicine, University of Connecticut Health Center
[15]Department of Cell Biology, University of Connecticut Health Center
[16]Babraham Institute
[17]Membership of the 2015 Whole-Cell Modeling Summer School Consortium is listed in Table SI

Computational modeling is an increasingly important tool for biological research, bioengineering, and medicine. Recently, researchers developed the first whole-cell model which represents every individual molecular species and gene function. However, significant work remains to develop fully complete and accurate models. We organized the 2015 Whole-Cell Modeling Summer School to teach students the latest cell modeling methods, as well as to assess the need for whole-cell modeling standards. We found that standards would accelerate whole-cell modeling and that our current standards SBML, SED-ML, and SGBN must be expanded to support whole-cell modeling.

Index Terms—**Whole-cell modeling, Systems biology, Computational biology, Mathematical modeling, Simulation, Standards, Education**

I. INTRODUCTION

OVER the past twenty years, computational modeling has become an essential and powerful tool for biological research, bioengineering, and medicine to analyze high-throughput molecular measurements and understand the molecular details of complex biological systems. Computational modeling has already been used to identify new metabolic genes [1], add metabolic pathways to bacteria [2], and identify potential new antimicrobial drug targets [3]. Computational models also have the potential to enable bioengineers to design new microorganisms for industrial applications such as chemical synthesis, biofuel production, and waste decontamination, as well as to enable clinicians to design personalized medical therapies tailored to each individual patient's unique genome. Realizing this potential requires more comprehensive and accurate computational models which are capable of predicting cellular behavior from genotype, as well

as standardized methods for exchanging models, simulation experiments, and model visualizations [4], [5], [6].

Recently, researchers at Stanford University developed the first whole-cell model of the gram-positive bacterium *Mycoplasma genitalium* [7]. The model represents the life cycle of a single Mycoplasma cell including the copy number dynamics of each metabolite, RNA, and protein species and accounts for every known gene function. The model is comprised of 28 sub-models, each of which was implemented using different mathematical representations including ordinary differential equations (ODEs), flux balance analysis (FBA), and Boolean rules (BRs), and trained using different experimental data.

The *M. genitalium* whole-cell model was implemented in MATLAB, is available open-source under the MIT license, and is extensively documented. This has enabled other researchers to use the model for their own research and enabled educators to use the model to teach systems biology.

Although MATLAB is commonly used by academic researchers, it is proprietary and expensive. In addition, because many of the biological details of the model are intertwined

with the MATLAB code, significant domain expertise is currently required to understand, use, or modify the Mycoplasma model or construct new whole-cell models. Transparent whole-cell modeling standards and simulation tools are needed to enable more researchers to develop and use whole-cell models for their own research. Such standards would accelerate the whole-cell modeling field. They would enable researchers to develop models more quickly, more deeply explore model predictions, and more rigorously evaluate models.

Several systems biology standards have already been developed by the COmputational Modeling in BIology NEtwork (COMBINE) [8] including the Systems Biology Markup Language (SBML) [9], the Cell Markup Language (CellML) [10], the Simulation Experiment Description Markup Language (SED-ML) [11], and the Systems Biology Graphical Notation (SBGN) [12]. SBML and CellML are standard languages for describing mathematical models including ODE, logical, and FBA models. Both have been used to build hundreds of models of various intracellular pathways. However, neither language supports several of the features needed for whole-cell modeling including large, sparsely occupied state spaces and multi-algorithm simulations. SED-ML is a standard language for describing computational experiments. SED-ML enables computational scientists to reproduce simulations by tracking the parameter values used in each simulation. SBGN is a standard language for describing visual representations of models. SBGN also has restrictions regarding the visualization of whole-cell models. Taken together, further work is needed to expand the capabilities of our current systems biology standard languages SBML, CellML, SED-ML, and SBGN to support whole-cell modeling.

We organized the 2015 Whole-cell Modeling Summer School to train students in whole-cell modeling and assess the limitations of the current systems biology standards for whole-cell modeling. The first and foremost goal of the school was to train young researchers how to build whole-cell models, including how to develop sub-models of individual cellular pathways using different mathematical formalisms; how to assemble sub-models into a single model; and how to encode models using SBML, simulate models using SED-ML, and visualize models using SBGN. The second goal of the school was to identify how the SBML, SED-ML, and SBGN languages must be expanded to support whole-cell modeling. The third goal of the school was to reimplement the *M. genitalium* whole-cell model using SBML.

The summer school arose from discussions at the 2012 COMBINE tutorial [13] about how to best engage and inform modelers about COMBINE standards. The discussions concluded that a summer school focused on a specific well-known model would be an excellent forum to educate young researchers on how to build models and encode them using COMBINE standards. Dagmar Waltemath and Falk Schreiber organized the school over the following three years, obtained funding for the school from the Volkswagen Foundation, recruited nine tutors with broad experience in computational systems biology and the COMBINE standards to help teach the school, and selected 45 students to attend the school.

Here, we report the educational and scientific outcomes

of the summer school. First, we summarize the format and content of the summer school. Second, we describe the educational outcomes of the school and our progress toward encoding the *M. genitalium* whole-cell model using SBML. Third, we describe the limitations we learned for using SBML to encode whole-cell models. We also describe several lessons that we learned about how to successfully organize a summer school. Lastly, we outline a path toward expanding the SBML, SED-ML, and SBGN standards to support whole-cell modeling and encoding the *M. genitalium* whole-cell model using SBML.

II. THE 2015 WHOLE-CELL MODELING SUMMER SCHOOL

The goal of the summer school was to teach students how to build models and encode models using COMBINE standards through encoding the *M. genitalium* model using only standard representation formats and open-source simulation software. To achieve this goal, we devoted most of the five-day school to hands-on active learning sessions in which ten groups of 4-5 students, each led by one tutor, were tasked with encoding parts of the *M. genitalium* model using COMBINE standards. Prior to the school, we assigned students, tutors, and sub-models to groups based on a survey of the students' backgrounds and interests. The groups also began meeting online using Google Hangouts and encoding sub-models prior to the school. Jonathan Karr also worked with all of the groups to teach students about cell modeling. Table SI lists the ten groups and all of the students and tutors. In addition to the hands-on active learning sessions, the school featured two lectures on whole-cell modeling and modeling standards from Jonathan Karr and Michael Hucka, two group discussions on model integration, daily oral progress reports from each of the working groups, a poster session, and three networking activities.

A. Introductory lectures

We began the school with two lectures to introduce the students to whole-cell modeling and the existing systems biology standards. Jonathan Karr from the Icahn School of Medicine at Mount Sinai, USA presented an overview of whole-cell modeling, the latest methods for constructing whole-cell models, the *M. genitalium* whole-cell model and his own research on developing whole-cell models for bio-engineering and precision medicine. Michael Hucka from the California Institute of Technology, USA presented an overview of the main systems biology standards SBML, SED-ML, and SBGN; the most popular open-source software tools which support these standards; and the COMBINE initiative.

B. Group discussions

In addition to the introductory lectures, we organized three group discussions on three common SBML encoding problems. The first discussion focused on how to sample from random distributions within the context of SBML. The second discussion focused on the interfaces that sub-models must expose to facilitate integration into a single model. The third

discussion focused on how to model large, sparsely occupied state spaces such as that which represents the locations of DNA-bound proteins.

C. Active learning sessions

The majority of days 2-4 of the school were dedicated to hands-on active learning sessions in which students learned about whole-cell modeling and the COMBINE standards by encoding parts of the *M. genitalium* whole-cell model using SBML. Specifically, eight of the ten groups were responsible for encoding one or more sub-models. The ninth group focused on developing a standards-compliant scheme for integrating SBML-encoded sub-models into a single model. This group was responsible for defining the sub-model interfaces. The tenth group focused on helping the other groups document and visualize their sub-models. Each group was led by an experienced tutor.

The groups pursued various strategies to encode the sub-models using COMBINE-compliant representations. Several of the groups encoded sub-models by first reading the sub-model documentation, second drawing pathway maps using software tools such as Cell Designer [14] and VANTED [15], and then writing scripts to automatically generate SBML representations from their maps. Other groups used modeling software tools such as COPASI [16] and BioUML [17] to encode sub-models based on their documentation. A third set of groups encoded sub-models by converting the original MATLAB code to SBML. These groups then automatically generated SBGN maps from their SBML to better understand their sub-models.

D. Progress reports

We concluded each day with brief oral progress presentations from one member of each of the ten groups. These progress reports facilitated discussion on common standards encoding challenges, provided groups opportunities to get feedback from the rest of the school, and helped the integration group plan the assembly of the integrated model. On the first day, the groups presented overviews of the biology of their sub-models and their progress to date. On the last day, the groups summarized their progress from the school and their plan to complete their sub-model encoding. Throughout the school, each student presented at least once.

E. Poster session

To provide students opportunities to share their own research, we organized a poster session on the second night of the school. 27 students presented posters. Many of the students presented projects on whole-cell modeling or other computational systems biology projects.

F. Networking activities

We organized several social activities to facilitate student networking. On the first night, we organized a dinner reception. On the third night, we organized a city tour led by a local guide. On the last evening, we organized a group dinner and dance party.

III. RESULTS

A. Educating young systems biologists

The primary goal of the summer school was to educate young scientists on how to build and apply whole-cell models. Because the students came from a variety of backgrounds ranging from computer science to biology, the tutors provided students one-on-one instruction on molecular cell biology; common dynamical modeling methods including ODEs, FBA, and BRs; model integration; and the SBML, SED-ML, and SBGN systems biology standards.

Following the summer school, we surveyed students to evaluate the course. Almost all of students reported that they gained a deep knowledge of whole-cell modeling. Most of the students also reported increased appreciation for the importance of reproducibility and increased understanding of the SBML, SED-ML, and SBGN standards and how they facilitate reproducible simulations. In addition, many students reported that they learned a lot about open-source modeling software tools which are applicable to their graduate research.

The students also reported that the course was a valuable networking event. Students reported that they had many opportunities to connect with other young computational systems biology researchers from around the world. In particular, students found that the poster session provided them an opportunity to share their own work and get valuable feedback from each other, the tutors, and other scientists from the University of Rostock. Several of the students also reported that the course introduced them to other opportunities to learn more about whole-cell modeling including open postdoctoral positions and the 2016 Whole-Cell Modeling Summer School which Jonathan Karr, Luis Serrano, Maria Lluch-Senar, and Javier Carrera are organizing in Barcelona, Spain.

B. Progress toward an SBML-encoded whole-cell model

In addition to training young systems biology researchers, the course produced preliminary SBML and SBGN-ML versions of each of the *M. genitalium* whole-cell model sub-models. Since the course, the project groups have continued to meet online to complete the encoding and testing of their sub-models. We plan to complete the encoding and integration of the sub-models into a single model within the next few months.

The sub-models and integrated model will be COMBINE-compliant, and simulatable with open-source software tools such as COPASI, BioUML, and iBioSim [18]. The sub-models and integrated model will be available open-source at http://github.com/whole-cell-tutors/wholecell. We believe the sub-models and integrated model will be a valuable community resource which will enable other scientists to develop, expand, and apply whole-cell models to their own research.

IV. LESSONS LEARNED

A. Limitations of SBML, SED-ML, and SBGN for whole-cell modeling

Prior to the summer school, SBML, SED-ML, and SBGN had never been used to represent similarly large or complex models. Consequently, the summer school revealed several

limitations of SBML, SBGN, and the existing COMBINE-compliant software for large models. First, no COMBINE-compliant software supports multi-algorithm models such as the *M. genitalium* whole-cell model. Second, SBML does not support sparse representations of large, combinatorial state spaces. Consequently, every state space must be enumerated. For the case of the *M. genitalium* model, this means that the chromosome state space must be enumerated, resulting in millions of state variables and unmanageably large XML files. Third, the core SBML language does not support arrays. Consequently, matrix operations that were efficiently implemented and executed in MATLAB, must be enumerated, resulting in verbose, opaque SBML files and inefficient simulations. Fourth, SBML does not support random number generation which is required to encode many of the sub-models. Consequently, many of the sub-models must be recast as Gillespie algorithm simulations to be supported by SBML.

The summer school also revealed several limitations of SBGN for whole-cell models. First, whole-cell modeling requires SBGN maps which contain a mixture of Process Description, Entity Relationship, and Activity Flow nodes. Currently, SBGN maps can only contain one of these three types of nodes. Second, whole-cell modeling demands improved support for automatic layout of large maps.

Most importantly, the summer school facilitated discussions among modelers, software tools developers, and model curators about how to best overcome these limitations.

B. Organizing a summer school

We also learned several valuable lessons about how to organize a productive, educational summer school from analyzing student surveys of the school. First, we found that summer schools must have clear learning objectives. Second, we found that students greatly enjoy hands-on active learning. Furthermore, we found that a high teacher to student ratio and a flexible schedule are necessary to successfully run open-ended project-based courses. Third, we found that students should be divided into multidisciplinary teams. This enables students to learn from other scientists with alternative perspectives from other fields and builds strong teams that are well-equipped to solve difficult problems. In the case of this summer school, it also enabled us to quickly assess potential expansions to the SBML language to support whole-cell models from multiple perspectives, ensuring that the expanded standards will meet the needs of modelers, software developers, and model curators. Fourth, we found that social activities greatly enhance summer schools by encouraging informal networking interactions among students. Lastly, we found that summer schools must have generous financial support. This enables organizers to accept students irrespective of their funding status. In turn, this enables the course to include a broader range of students than would otherwise be possible.

V. FUTURE STEPS

The summer school made significant strides toward expanding the whole-cell modeling field by educating 45 young scientists in multi-algorithm modeling. The students had two introductory lectures which outlined the whole-cell modeling field and the existing systems biology standards; three discussions on model integration and how to efficiently represent large state spaces; and three full days of hands-on active learning exercises to recode sub-models using SBML. Jonathan Karr, Luis Serrano, Maria Lluch-Senar, and Javier Carrera are organizing a second whole-cell modeling summer course in Barcelona in April 2016. Chris Myers at the University of Utah, USA and Vincent Danos at the National Centre for Scientific Research, France have also begun to offer graduate courses in whole-cell modeling.

Additional summer schools and courses are needed to provide students deeper theoretical training in the mathematical foundation of whole-cell modeling, as well as practical training in the construction of large dynamical models. These courses should provide students opportunities to learn about techniques for integrating multiple sub-models and resolving conflicts including state locking, adaptive time-stepping, controllers and semaphores; how to identify the parameters of large models including numerical optimization, distributed optimization, and model reduction; and how to test models using unit testing and continuous integration.

The summer school made also significant strides toward reimplementing the *M. genitalium* whole-cell model using SBML and COMBINE-compliant software. However, significant work remains to expand the SBML language to support whole-cell models, recode the *M. genitalium* model using SMBL, and develop COMBINE-compliant software programs which are capable of efficiently simulating multi-algorithm models. Currently, we are working to finish recoding the sub-models using SBML and mapping each of the sub-models using SBGN. Since the Rostock meeting, the working groups have continued to meet virtually through Google Hangouts and share materials using GitHub and Google Drive. Several of the students and tutors also plan to participate in a second working meeting in conjunction with the COMBINE conference in Salt Lake City, USA in October 2015. We plan to publish these sub-models through the BioModels database [19] along with tests implemented using SED-ML.

Ultimately, we hope to integrate the individual sub-models into a single SBML-encoded model using the SBML-comp package. To achieve this, we must expand the SBML language to support multi-algorithm models and efficiently represent large, combinatorial state spaces. We plan to present proposals to expand SBML to support these features at the 2015 COMBINE conference in Salt Lake City, USA.

VI. CONCLUSION

We believe that whole-cell modeling has the potential to be an important tool for biological discovery, bioengineering, and medicine. Achieving this potential requires improved standards for transparently communicating whole-cell models, as well as COMBINE-compliant software for reproducibly simulating whole-cell models. In turn, this requires expanding the whole-cell modeling field including training young researchers in whole-cell modeling.

We organized the 2015 Whole-Cell Modeling Summer School to train young researchers how to building whole-

cell models using standardized languages through hands-on experience recoding the *M. genitalium* whole-cell model using SBML. The school brought together 45 students including modelers, modeling software developers, and model curators. The school successfully trained 45 students, made significant progress toward recoding the *M. genitalium* whole-cell model using SBML, and initiated a conversation about how to expand the SBML language to support multi-algorithm whole-cell models. We plan to publish the recoded sub-models through the BioModels database in the Fall and hope to eventually combine the SBML-encoded sub-models into a single model. We also plan to propose extensions to the SBML language to support whole-cell models at the 2015 COMBINE meeting in Salt Lake City, USA and to disseminate the proposed extensions through EraNet SysBio [20], CaSYM [21] and ISBE [22]. Finally, we hope the summer school will be the first of many whole-cell modeling meetings and courses.

REFERENCES

[1] J. L. Reed, T. R. Patel, K. H. Chen, A. R. Joyce, M. K. Applebee, C. D. Herring, O. T. Bui, E. M. Knight, S. S. Fong, and B. O. Palsson, "Systems approach to refining genome annotation," *Proc Natl Acad Sci U S A*, vol. 103, no. 46, pp. 17 480–17 484, 2006.

[2] D. S. Lee, H. Burd, J. Liu, E. Almaas, O. Wiest, A. L. Barabási, Z. N. Oltvai, and V. Kapatral, "Comparative genome-scale metabolic reconstruction and flux balance analysis of multiple *Staphylococcus aureus* genomes identify novel antimicrobial drug targets," *J Bacteriol*, vol. 191, no. 12, pp. 4015–4024, 2009.

[3] J. W. Lee, D. Na, J. M. Park, J. Lee, S. Choi, and S. Y. Lee, "Systems metabolic engineering of microorganisms for natural and non-natural chemicals," *Nat Chem Biol*, vol. 8, no. 6, pp. 536–546, 2012.

[4] D. N. Macklin, N. A. Ruggero, and M. W. Covert, "The future of whole-cell modeling," *Curr Opin Biotechnol*, vol. 28, pp. 111–115, 2014.

[5] J. R. Karr, K. Takahashi, and A. Funahashi, "The principles of whole-cell modeling," *Curr Opin Microbiol*, vol. (in press), 2015.

[6] E. Klipp, W. Liebermeister, A. Helbig, A. Kowald, and J. Schaber, "Systems biology standards - the community speaks," *Nat Biotech*, vol. 25, no. 4, pp. 390–391, 2007.

[7] J. R. Karr, J. C. Sanghvi, D. N. Macklin, M. V. Gutschow, J. M. Jacobs, B. Bolival, Jr, N. Assad-Garcia, J. I. Glass, and M. W. Covert, "A whole-cell computational model predicts phenotype from genotype," *Cell*, vol. 150, no. 2, pp. 389–401, 2012.

[8] N. Le Novère, M. Hucka, N. Anwar, G. D. Bader, E. Demir, S. Moodie, and A. Sorokin, "Meeting report from the first meetings of the Computational Modeling in Biology Network (COMBINE)," *Standards in Genomic Sciences*, vol. 5, no. 2, p. 230, 2011.

[9] M. Hucka, A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle, H. Kitano, A. P. Arkin, B. J. Bornstein, D. Bray, A. Cornish-Bowden, A. A. Cuellar, S. Dronov, E.-D. Gilles, M. Ginkel, V. Gor, I. I. Goryanin, W. J. Hedley, T. C. Hodgman, J.-H. Hofmeyr, P. J. Hunter, N. S. Juty, J. L. Kasberger, A. Kremling, U. Kummer, N. L. Novère, L. M. Loew, D. Lucio, P. Mendes, E. D. Mjolsness, Y. Nakayama, M. R. Nelson, P. F. Nielsen, T. Sakurada, J. C. Schaff, B. E. Shapiro, T. S. Shimizu, H. D. Spence, J. Stelling, K. Takahashi, M. Tomita, J. Wagner, and J. Wang, "The Systems Biology Markup Language (SBML): A medium for representation and exchange of biochemical network models," *Bioinformatics*, vol. 19, no. 4, pp. 524–531, 2003.

[10] W. J. Hedley, M. R. Nelson, D. P. Bullivant, and P. F. Nielson, "A short introduction to CellML," *Philosophical Transactions of the Royal Society of London A*, vol. 359, pp. 1073–1089, 2001.

[11] D. Waltemath, R. Adams, F. Bergmann, M. Hucka, F. Kolpakov, A. Miller, I. Moraru, D. Nickerson, S. Sahle, J. Snoep, and N. Le Novère, "Reproducible computational biology experiments with SED-ML—the Simulation Experiment Description Markup Language," *BMC Systems Biology*, vol. 5, no. 1, p. 198, 2011.

[12] M. P. van Iersel, A. Villéger, T. Czauderna, S. E. Boyd, F. T. Bergmann, A. Luna, E. Demir, A. A. Sorokin, U. Dogrusöz, Y. Matsuoka, A. Funahashi, M. I. Aladjem, H. Mi, S. L. Moodie, H. Kitano, N. Le Novère, and F. Schreiber, "Software support for SBGN maps: SBGN-ML and LibSBGN," *Bioinformatics*, vol. 28, no. 15, pp. 2016–2021, 2012.

[13] COMBINE. COMBINE 2012. [Online]. Available: http://co.mbine.org/events/COMBINE_2012

[14] A. Funahashi, Y. Matsuoka, A. Jouraku, M. Morohashi, N. Kikuchi, and H. Kitano, "Celldesigner 3.5: a versatile modeling tool for biochemical networks," *Proceedings of the IEEE*, vol. 96, no. 8, pp. 1254–1265, 2008.

[15] H. Rohn, A. Junker, A. Hartmann, E. Grafahrend-Belau, H. Treutler, M. Klapperstück, T. Czauderna, C. Klukas, and F. Schreiber, "VANTED v2: a framework for systems biology applications," *BMC Syst Biol*, vol. 6, p. 139, 2012.

[16] P. Mendes, S. Hoops, S. Sahle, R. Gauges, J. Dada, and U. Kummer, "Computational modeling of biochemical networks using COPASI," *Methods Mol Biol*, vol. 500, pp. 17–59, 2009.

[17] F. Kolpakov, "BioUML: visual modeling, automated code generation and simulation of biological systems," *Proc. BGRS*, vol. 3, pp. 281–285, 2006.

[18] J. T. Stevens and C. J. Myers, "Dynamic modeling of cellular populations within iBioSim," *ACS Synth Biol*, vol. 2, no. 5, pp. 223–229, 2013.

[19] C. Li, M. Donizelli, N. Rodriguez, H. Dharuri, L. Endler, V. Chelliah, L. Li, E. He, A. Henry, M. Stefan, J. Snoep, M. Hucka, N. Le Novere, and C. Laibe, "BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models," *BMC Syst Biol*, vol. 4, no. 1, p. 92, 2010.

[20] ERASysBio consortium. ERA-NET for systems biology. [Online]. Available: http://www.erasysbio.net

[21] CaSYM consortium. Coordinating Action Systems Medicine: Implementation of Systems Medicine Across Europe. [Online]. Available: https://www.casym.eu

[22] O. Wolkenhauer, D. Fell, P. De Meyts, N. Blüthgen, H. Herzel, N. Le Novère, T. Höfer, K. Schürrle, and I. van Leeuwen, "SysBioMed report: advancing systems biology for medical applications," *IET Syst Biol*, vol. 3, no. 3, pp. 131–136, 2009.

Dagmar Waltemath received her Ph.D. in Computer Science from the University of Rostock, Germany in 2011. Dr. Waltemath is a Junior Research Group Leader in the Department of Systems Biology and Bioinformatics at the University of Rostock. Her group focuses on developing methods to reproducibly simulate computational biology models including managing, annotating, searching and retrieving models. Dr. Waltemath is also a COMBINE Coordinator and an SBML Editor. Contact: dagmar.waltemath@uni-rostock.de.

Falk Schreiber received his Ph.D. (2001) and habilitation (2006) in Computer Science from the University of Passau, Germany. He is a Professor in the School of Information Technology at Monash University and an Adjunct Professor in the Institute of Computer Science at Martin Luther University Halle-Wittenberg. His research is focused on visual computing and visual analysis of biological data, analysis of the structure and dynamics of biological networks, integration of multimodal data, and standards for systems biology, as well as modeling and analysis of metabolism. He is a COMBINE Coordinator and has been an SBGN Editor. Contact: falk.schreiber@monash.edu.

Jonathan R. Karr received his Ph.D. in Biophysics and M.S. in Medicine from Stanford University, USA in 2014 and his S.B.s in Physics and Brain & Cognitive Sciences from the Massachusetts Institute of Technology, USA in 2006. Dr. Karr is a Fellow at the Icahn School of Medicine at Mount Sinai, USA. His research focuses on the development of comprehensive whole-cell computational models and their applications to bioengineering and medicine. Contact: karr@mssm.edu.

Michael Hucka received his Ph.D. in Electrical Engineering and Computer Science from the University of Michigan in 1998 and his B.S.s in Computer Science and Electrical Engineering in 1986. Dr. Hucka is a Member of the Professional Staff of the California Institute of Technology. Contact: mhucka@caltech.edu.

Chris J. Myers received his Ph.D. in Electrical Engineering from Stanford University in 1995 and his B.S.s in Electrical Engineering and Chinese History from the California Institute of Technology in 1991. Dr. Myers is a Professor of Electrical and Computer Engineering at the University of Utah. Contact: myers@ece.utah.edu.

Frank T. Bergmann received his Ph.D. in Computational and Systems Biology from Claremont Graduate University in 2010 and his M.S. in Computer Science from Goethe University Frankfurt in 2004. Dr. Bergmann is a Software Developer at the California Institute of Technology and a Member of the Scientific Staff at the University of Heidelberg. Contact: fbergman@caltech.edu.

Vijayalakshmi Chelliah received her Ph.D. in Structural Bioinformatics & Computational Biology from the University of Cambridge in 2005 and her B.Sc. in Mathematics, Physics and Computer Science from Manonmaniam Sundaranar University in 1995. Dr. Chelliah is a Senior Scientific Curator at the European Bioinformatics Institute. Contact: viji@ebi.ac.uk.

Marcus Krantz received his Ph.D. in Microbiology from the University of Gothenburg in 2005. Dr. Krantz is a Junior Research Group Leader at the Humboldt University of Berlin. Contact: marcus.krantz@biologie.hu-berlin.de.

Wolfram Liebermeister received his Ph.D. in Theoretical Biophysics from the Humboldt University of Berlin and his B.S. in Physics from the University of Tübingen. Dr. Liebermeister is a Postdoctoral Fellow at the Charité Medical University of Berlin. Contact: wolfram.liebermeister@charite.de.

Pedro Mendes received his Ph.D. in Biological Sciences from the University of Wales Aberystwyth in 1994 and his B.Sc in Biochemistry from the University of Lisbon in 1988. Dr. Mendes is a Professor of Computer Science at the University of Manchester and a Professor of Cell Biology at the University of Connecticut Health Center. Contact: pedro.mendes@manchester.ac.uk.

Pinar Pir received her Ph.D. in Chemical Engineering from Boğaziçi University in 2005. Dr. Pir is Postdoctoral Research Associate at the Babraham Institute. Contact: pp305@cam.ac.uk.

2015 Whole-Cell Modeling Summer School Consortium includes the 56 researchers listed in Table SI who organized and attended the 2015 summer school.

Table SI
2015 WHOLE-CELL MODELING SUMMER SCHOOL CONSORTIUM MEMBERS.

Group	Participant	Affiliation
Cytokinesis	Daniel Alejandro Priego Espinosa	National Autonomous University of Mexico
	Ilya Kiselev	Siberian Branch of the Russian Academy of Sciences Novosibirsk, Russia
	Naveen Kumar Aranganathan	University Paris-Sud
	Wolfram Liebermeister	Charité Medical University of Berlin
	Yan Zhu	Monash University, Australia
DNA repair	Arne Bittig	University of Rostock, Germany
	Vijayalakshmi Chelliah	European Bioinformatics Institute, UK
	Audald Lloret-Vilas	European Bioinformatics Institute, UK
	Mahesh Sharma	National Institute of Pharmaceutical Education and Research, India
	Namrata Tomar	Friedrich-Alexander University of Erlangen-Nürnberg, Germany
Metabolism	Kambiz Baghalian	University of Oxford, UK
	Frank T. Bergmann	California Institute of Technology, USA
	Matthias König	Charité Medical University of Berlin
	Kieran Smallbone	University of Manchester, UK
	Rafeal Sousa Costa	University of Lisbon, Portugal
	Milenko Tokic	Swiss Federal Institute of Technology in Lausanne, Switzerland
Protein	Begum Alaybeyoglu	Boğaziçi University, Turkey
	Matteo Cantarelli	OpenWorm, UK
	Yin Hoon Chew	University of Edinburgh, UK
	Fiete Haack	University of Rostock, Germany
	Marcus Krantz	Humboldt University of Berlin, Germany
	Daewon Lee	KAIST, South Korea
Replication	Vincent Knight-Schrijver	Babraham Institute, UK
	Je-Hoon Song	KAIST, South Korea
	Jannis Uhlendorf	Humboldt University of Berlin, Germany
	Dagmar Waltemath	University of Rostock, Germany
	James Yurkovich	University of California, San Diego, USA
	Anna Zhukova	University of Bordeaux, France
Replication initiation	Harold Gomez	Boston University, USA
	Jens Hahn	Humboldt University of Berlin, Germany
	Michael Hucka	California Institute of Technology, USA
	Mandrik Nikita	Siberian Branch of the Russian Academy of Sciences Novosibirsk, Russia
	Martin Scharm	University of Rostock, Germany
	Florian Wendland	University of Rostock, Germany
RNA	Tuure Hameri	Swiss Federal Institute of Technology in Lausanne, Switzerland
	Jesse Kyle Medley	University of Washington, USA
	Sucheendra Kumar Palaniappan	RIKEN, Japan
	Pinar Pir	Babraham Institute, UK
	Natalie Stanford	University of Manchester, UK
	Markus Wolfien	University of Rostock, Germany
Translation	Joseph Cursons	University of Melbourne, Australia
	Muhammad Haseeb	Mohammad Ali Jinnah University, Pakistan
	Daniel Hernandez	Swiss Federal Institute of Technology in Lausanne, Switzerland
	Pedro Mendes	University of Manchester, UK
	Denis Kazakiewicz	University of Hasselt, Belgium
	Hojjat Naderi Meshkin	Academic Center for Education, Culture and Research, Iran
Integration	Paulo Eduardo Pinto Burke	University of São Paulo, Brazil
	Tobias Czauderna	Monash University, Australia
	Bertrand Moreau	CoSMo Company, France
	Thawfeek Mohamed Varusai	University College Dublin, Ireland
	Chris J. Myers	University of Utah, USA
	Argyris Zardilis	University of Edinburgh, UK
Visualization and documentation	Christian Knüpfer	University of Jena, Germany
	Falk Schreiber	Monash University, Germany
	Tom Theile	University of Rostock, Germany
Modeling tutor	Jonathan R. Karr	Icahn School of Medicine at Mount Sinai, USA