

**5th Conference on Constraint-Based
Reconstruction and Analysis**

**COBRA
2018**

October 14 – 16, 2018
Sheraton | Seattle, Washington

5th Conference on Constraint-Based Reconstruction and Analysis

BRONZE





TABLE OF CONTENTS

Welcome Address	2-3
Conference Organizers	4
Panel Discussions	5
Technical Program.....	6-9
Poster Titles.....	10-20
Invited Speaker Biographies.....	21-24
Oral Abstracts.....	25-43
Poster Abstracts.....	44-87
Code of Conduct	88
Notes.....	89

Abstracts appear as submitted by their authors. Neither the American Institute of Chemical Engineers (AIChE) and its entities, nor the employers affiliated with the authors or presenting speakers, are responsible for the content of the abstracts.

WELCOME ADDRESS

Greetings!

We want to welcome you to Seattle, WA for the 5th Conference on Constraint-Based Reconstruction and Analysis (COBRA 2018), brought to you by the International Metabolic Engineering Society (IMES), an AIChE Technical Community, in partnership with the Institute for Systems Biology (ISB).

Metabolic processes and their regulation in response to environmental and genetic perturbations are fundamental to the growth and maintenance of all cells. Linking the genome to its functioning metabolism is therefore of substantial interest in fields ranging from industrial biotechnology to human health. Constraint-based Reconstruction and Analysis (COBRA) has over recent years rapidly expanded and been established as the method of choice for studying cellular metabolism in a global and unbiased fashion. The COBRA 2018 Conference aims to bring together the worldwide COBRA community to discuss latest developments, novel tools and emerging applications in diverse biological research areas, such as metabolic engineering and medicine. The meeting will offer a unique platform for the exchange of ideas and catalyze new, multi-disciplinary collaborations.

We hope you are ready for a tremendous conference. The conference is headlined by a veritable pantheon of notable speakers including:

- Evrim Acar, University of Copenhagen
- Vassily Hatzimanikatis, Ecole Polytechnique Federale de Lausanne
- Matthias Heinemann, University of Groningen
- Joo Sang Lee, Cancer Data Science Lab, NCI/NIH
- Nathan Lewis, University of California San Diego
- Christian Lieven, Technical University of Denmark
- Costas Maranas, The Pennsylvania State University
- Jens Nielsen, Chalmers University of Technology
- Bernhard Palsson, University of California San Diego; Novo Nordisk Center for Biosustainability, Denmark
- Jennifer Reed, University of Wisconsin-Madison
- Uwe Sauer, Eidgenössische Technische Hochschule Zürich

WELCOME ADDRESS

COBRA 2018 truly encourages the growth and continued support of the larger COBRA Community. Attendees will have opportunities to participate during The Future of COBRA Discussion Panel and the Community Standards and Resources Discussion Panel. Similar to past years, this conference will also feature a poster reception where authors can share and discuss their work with their colleagues. Further, lunches and a welcome dinner will ensure that you have the opportunity to hear every speaker while also having time to interact with your peers.

For the first time, COBRA comes to the United States and offers time to explore the beautiful city of Seattle. Take time to network and enjoy the food at Pike Place Market during the two-hour lunch break on Tuesday, October 15.

A lot of work has gone into making this conference a success. We would not have been able to do so without the contributions of our Organizing Committee, who were instrumental in selecting our speakers and shaping the program. We extend additional thanks to all of our corporate sponsors, academic supporters, and media partner without whom the conference could not happen. Moreover, the tremendous support of the IMES staff has played an invaluable role.

Finally, we would like to thank you for attending the conference. We hope these three days will be pleasant, educational, and inspiring.

Sincerely,



Nathan Price
Associate Director and Professor
Institute for Systems Biology (ISB)
nprice@systemsbiology.org

CONFERENCE ORGANIZERS

Conference Co-Chairs

Nathan Price, *Institute for Systems Biology*

Organizing Committee

Markus Herrgård, *Novo Nordisk Foundation Center for Biosustainability*

Jason Papin, *University of Virginia*

Kiran Patil, *European Molecular Biology Laboratory*

Tomer Shlomi, *Technion - Israel Institute of Technology*

Vangelis Simeonidis, *Institute for Systems Biology*

Ines Thiele, *Luxembourg Centre for Systems Biomedicine*



SAVE THE DATE

Omni La Costa Resort & Spa Carlsbad • CA • February 17-19, 2019

The International Conference on Accelerating Biopharmaceutical Development (AccBio) is dedicated to strategies, technologies, and capabilities that advance biopharmaceutical development. Be part of the discussion at AccBio 2019.

Themes include:

- New Modalities
- Advancements in Data Technologies
- Manufacturing Technologies
- Patient-Centric Process Development

The 2019 conference program— along with registration and abstract information— is currently in development and will be available soon. For the latest updates, please visit www.aiche.org/accbio.

ORGANIZED BY THE SOCIETY FOR BIOLOGICAL ENGINEERING

© 2018 AIChE 3146_18 • 09.18

The Future of the COBRA Field

Nathan Price, *Institute for Systems Biology (Moderator)*

Vassily Hatzimanikatis, *Ecole Polytechnique Federale de Lausanne*

Bernhard Palsson, *University of California San Diego; Novo Nordisk Center for Biosustainability, Denmark*

Jennifer Reed, *University of Wisconsin-Madison*

Uwe Sauer, *Eidgenössische Technische Hochschule Zürich*

Community Standards and Resources

Jason Papin, *University of Virginia (Moderator)*

Maureen Carey, *University of Virginia*

Chris Henry, *Argonne National Laboratory*

Christian Lieven, *Technical University of Denmark*

TECHNICAL PROGRAM

Sunday October 14th, Day 1	
12:00 PM to 1:00 PM	REGISTRATION
1:00 PM to 1:15 PM	WELCOMING REMARKS
1:15 PM to 3:00 PM	SESSION 1: NEW FRONTIERS I
1:15 PM to 1:45 PM	INVITED SPEAKER: Developing a seamless pipeline from ^{13}C labeling data to kinetic models with a genome-wide coverage Costas Maranas , <i>The Pennsylvania State University</i>
1:45 PM to 2:15 PM	INVITED SPEAKER: Introducing thermodynamics into ME-models and crowders into kinetic models Vassily Hatzimanikatis , <i>Ecole Polytechnique Federale de Lausanne (EPFL)</i>
2:15 PM to 2:30 PM	Machine Learning for Uncertainty Reduction in Biochemical Kinetic Models Ljubisa Miskovic , <i>Ecole Polytechnique Federale de Lausanne (EPFL)</i>
2:30 PM to 2:45 PM	Expanding Metabolic Models to Three Dimensions Elizabeth Brunk , <i>University of California San Diego</i>
2:45 PM to 3:00 PM	Key Decisions for the Development of the Next-Generation of Context-Specific Models Anne Richelle , <i>University of California San Diego</i>
3:00 PM to 3:30 PM	BREAK
3:30 PM to 5:15 PM	SESSION 1: NEW FRONTIERS II
3:30 PM to 4:00 PM	INVITED SPEAKER: Metabolic Coordination through Metabolite-Protein Interactions Uwe Sauer , <i>Eidgenössische Technische Hochschule (ETH) Zürich</i>
4:00 PM to 4:30 PM	INVITED SPEAKER: An Upper Limit in Gibbs Energy Dissipation Governs Cellular Metabolism Matthias Heinemann , <i>University of Groningen</i>
4:30 PM to 4:45 PM	Compass: FBA-Powered Modeling of Single-Cell RNA-Seq Characterizes Cell-to-Cell Metabolic Heterogeneity and Reveals Novel Therapeutic Targets in Autoimmunity Allon Wagner , <i>University of California Berkeley</i>
4:45 PM to 5:00 PM	Metabolic Model-Based Evaluation of Microbiome-Metabolome Association Studies Cecilia Noecker , <i>University of Washington</i>
5:00 PM to 5:15 PM	Yeast-GEM: Reviving the Consensus Genome-Scale Model of <i>S. Cerevisiae</i> As a Standard in the Community Benjamín J. Sánchez , <i>Chalmers University of Technology</i>
5:15 PM to 6:00 PM	BREAK/POSTER SET-UP
6:00 PM to 7:00 PM	POSTER RECEPTION
7:00 PM to 8:00 PM	DINNER

TECHNICAL PROGRAM

Monday October 15th, Day 2	
8:00 AM to 11:00 AM	REGISTRATION
8:45 AM to 9:00 AM	OPENING REMARKS
9:00 AM to 10:30 AM	SESSION 2: APPLICATIONS IN MEDICINE I
9:00 AM to 9:30 AM	INVITED SPEAKER: Urea cycle dysregulation, emerging pyrimidines mutation bias and enhanced response to immunotherapy in cancer Joo Sang Lee , <i>Cancer Data Science Lab, NCI/NIH</i>
9:30 AM to 9:45 AM	Modeling Physiological Responses to Stress: Metabolism of Fasting and Acetaminophen-Induced Liver Damage in the Laboratory Rat Anders Wallqvist , <i>U.S. Army Medical Research and Materiel Command</i>
9:45 AM to 10:00 AM	Differential Metabolic Functionality in Antibiotic-Resistant <i>Pseudomonas aeruginosa</i> Revealed with an Integrated Computational and Experimental Approach Laura J. Dunphy , <i>University of Virginia</i>
10:00 AM to 10:15 AM	Model-Based Prediction of Functional SNPs Suggests Factors for Metabolic Diversity and Drug Resistance across Human-Associated <i>Mycobacterium Tuberculosis</i> Ove Øyås , <i>Eidgenössische Technische Hochschule (ETH) Zürich</i>
10:15 AM to 10:30 AM	Synthetic Lethality in Cancer Research Via Genetic Minimal Cut Sets - Francisco J. Planes , <i>TECNUN, University of Navarra</i>
10:30 AM to 11:00 AM	BREAK
11:00 AM to 11:30 AM	INVITED SPEAKER: COBRA Distinguished Young Investigator Award Winner: Pathogen Metabolism and Antibiotic Resistance Sriram Chandrasekaran , <i>University of Michigan</i>
11:30 AM to 12:15 PM	SESSION 2: APPLICATIONS IN MEDICINE II
11:30 AM to 11:45 AM	Identifying and Targeting Key Cellular Mechanisms for Proliferation in Malaria Parasites: A Combined Experimental and Computational Strategy Anush Chiappino-Pepe , <i>Swiss Federal Institute of Technology (EPFL)</i>
11:45 AM to 12:00 PM	Inferring Metabolic Mechanisms of Interaction within a Defined Gut Microbiota Gregory L. Medlock , <i>University of Virginia</i>
12:00 PM to 12:15 PM	Understanding the Role of Bile Acids in Alzheimer's Disease Priyanka Baloni , <i>Institute for Systems Biology</i>
12:15 PM to 2:30 PM	LUNCH on your own at Pike Place Market
2:30 PM to 3:45 PM	SESSION 3: MULTI-SCALE MODELING I
2:30 PM to 3:00 PM	INVITED SPEAKER: Towards a more complete view of a cell's functions through genome scale models of metabolism, protein synthesis, post-translational modification, and secretion Nathan Lewis , <i>University of California San Diego</i>
3:00 PM to 3:15 PM	Deciphering the Mechanisms Underlying Nutrient Exchange in Insect-Microbe Symbioses Nana Ankrah , <i>Cornell University</i>

TECHNICAL PROGRAM

3:15 PM to 3:30 PM	Incorporation of Dynamic pH into an Integrative Model of a Nitrification Microcosm Co-Culture of <i>Nitrosomonas Europaea</i> and <i>Nitrobacter Winogradskyi</i> Frank Chaplen , <i>Oregon State University</i>
3:30 PM to 3:45 PM	Dynamic Side of the Warburg Effect: Glycolytic Intermediate Storage in Tumor Cells to Buffer Fluctuating Glucose and O ₂ Supply Johannes H.G.M. van Beek , <i>Vrije University Medical Center, Amsterdam</i>
3:45 PM to 4:15 PM	BREAK
4:15 PM to 5:15 PM	SESSION 3: MULTI-SCALE MODELING II
4:15 PM to 4:45 PM	INVITED SPEAKER: New Directions for Cobra Bernhard Palsson , <i>University of California San Diego; Novo Nordisk Center for Biosustainability, Denmark</i>
4:45 PM to 5:00 PM	Toward a Proteome-Complete Model of the Human Red Blood Cell James T. Yurkovich , <i>University of California, San Diego</i>
5:00 PM to 5:15 PM	Koptic: A Novel Approach for in silico Prediction of Enzyme Kinetics and Regulation Wheaton Schroeder , <i>University of Nebraska - Lincoln</i>
5:15 PM to 6:15 PM	The Future of COBRA Discussion Panel

Tuesday October 16th, Day 3

8:00 AM to 11:00 AM	REGISTRATION
8:45 AM to 9:00 AM	OPENING REMARKS
9:00 AM to 10:00 AM	SESSION 4: APPLICATIONS IN METABOLIC ENGINEERING I
9:00 AM to 9:30 AM	INVITED SPEAKER: Systems Biology of Yeast Metabolism Jens Nielsen , <i>Chalmers University of Technology</i>
9:30 AM to 9:45 AM	Experimental Design for Parameter Estimation in Kinetic Models of Metabolism Christian Euler , <i>University of Toronto</i>
9:45 AM to 10:00 AM	Essential Metabolism for a Minimal Cell Marian Breuer , <i>University of Illinois at Urbana-Champaign</i>
10:00 AM to 10:30 AM	BREAK
10:30 AM to 11:15 AM	SESSION 4: APPLICATIONS IN METABOLIC ENGINEERING II
10:30 AM to 10:45 AM	Characterization and Design of Phase Spaces and Yield Spaces in Genome-Scale Metabolic Models Jürgen Zanghellini , <i>Austrian Centre of Industrial Biotechnology, University of Natural Resources and Life Sciences</i>
10:45 AM to 11:00 AM	DD-Decaf: Data-Driven Design of Cell Factories and Communities Nikolaus Sonnenschein , <i>Technical University of Denmark</i>
11:00 AM to 11:15 AM	Dynamic Enzyme-Cost Flux Balance Analysis (deFBA) Modelling for an Industrially Relevant Methanotroph <i>Methylobacterium Buryatense</i> Kobe De Becker , <i>KU Leuven</i>
11:15 AM to 12:15 PM	Community standards and resources Discussion Panel

TECHNICAL PROGRAM

12:15 PM to 1:30 PM	LUNCH
1:30 PM to 3:30 PM	SESSION 5: METHODS AND SOFTWARE I
1:30 PM to 2:00 PM	INVITED SPEAKER: Data Fusion Based on Coupled Matrix and Tensor Factorizations Evrim Acar , <i>University of Copenhagen</i>
2:00 PM to 2:30 PM	INVITED SPEAKER: COBRA and Machine Learning Approaches for Engineering Microbial Biocatalysts Jennifer Reed , <i>University of Wisconsin-Madison</i>
2:30 PM to 2:45 PM	How Accurate Is Automated Gap Filling of Metabolic Models? Peter Karp , <i>SRI International</i>
2:45 PM to 3:00 PM	Bridging the Gap between Structural Systems Biology and Constraint-Based Modeling: Tools for Defining the Functional Structural Proteome and Extracting Genome-Scale Structural Information for Metabolic and Macromolecular Expression Models Edward Catoiu , <i>University of California, San Diego</i>
3:00 PM to 3:15 PM	Scalable Tools for Analyzing Steady-State Microbial Communities Using Standardized Genome-Scale Metabolic Models Siu Hung Joshua Chan , <i>Colorado State University</i>
3:15 PM to 3:30 PM	Genome-Scale Metabolic Model Embedding for Fed-Batch Optimal Feed Policy Determination Raul Conejeros , <i>Pontificia Universidad Catolica de Valparaiso</i>
3:30 PM to 4:00 PM	BREAK
4:00 PM to 5:45 PM	SESSION 5: METHODS AND SOFTWARE II
4:00 PM to 4:30 PM	INVITED SPEAKER: Memote: A Community Driven Effort Towards a Standardized Genome-Scale Metabolic Model Test Suite Christian Lieven , <i>Technical University of Denmark</i>
4:30 PM to 4:45 PM	An Ecology of Modeling Methods, Data, and Tools to Decipher Functional Delegation and Species Interactions within a Microbiome Christopher S. Henry , <i>Argonne National Laboratory</i>
4:45 PM to 5:00 PM	Thermodynamic Topology of Metabolite Networks Elucidates Regulatory Responses to Genetic and Environmental Perturbations Tolutola Oyetunde , <i>Washington University in St. Louis</i>
5:00 PM to 5:15 PM	Dynamic FBA with Global Constraints on Cellular Protein Fraction Eivind Almaas , <i>Norwegian University of Science and Technology</i>
5:15 PM to 5:30 PM	Incorporating Flux Sampling into a Minimal Assumption Dynamic Flux Balance Analysis Algorithm St. Elmo Wilken , <i>University of California, Santa Barbara</i>
5:30 PM to 5:45 PM	Creation and Analysis of Biochemical Constraint-Based Models: The Cobra Toolbox v3.0 Ronan M.T. Fleming , <i>Leiden University</i>
5:45 PM to 6:00 PM	COBRA AWARDS
6:00 PM to 6:15 PM	CLOSING REMARKS

POSTER TITLES

1. **Towards a Metabolic and Expression Model of the Metabolically Versatile Bacterium *Rhodopseudomonas Palustris*.**
Adil Alsiyabi¹, Cheryl Immethun², and Rajib Saha³
(1)Chemical and Biomolecular Engineering, University of Nebraska - Lincoln, Lincoln, NE, (2)Chemical & Biomolecular Engineering, University of Nebraska, Lincoln, NE, (3)Department of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE
2. **Genetic Optimization Algorithm for Metabolic Engineering Revisited.**
Tobias B. Alter, Lars M. Blank, and Birgitta E. Ebert
iAMB - Institute of Applied Microbiology, ABBt - Aachen Biology and Biotechnology, RWTH Aachen University, Aachen, Germany
3. **New Methods to Constrain Genome Scale Models with Stable Isotope Labeling Data.**
Tyler Backman¹, David E. Ando², Jay D. Keasling³, and Hector Garcia Martin²
(1)Joint BioEnergy Institute, Emeryville, CA, (2)Joint BioEnergy Institute (JBEI), Emeryville, CA, (3)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby, Denmark
4. **Mgpipe: A Toolbox for Investigating Metagenomics Data through Personalized Microbiota Metabolic Models.**
Federico Baldini¹, Almut Heinken¹, Laurent Heirendt¹, Ronan M.T. Fleming², and Ines Thiele¹
(1)Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg, (2)Leiden Academic Centre for Drug Research, Leiden University, Leiden, Netherlands
5. **Systems Approaches to Identify Metabolite Signatures in Placental Development.**
Priyanka Baloni¹, Alison Paquette¹, Heather Brockway², Yoel Sadovsky³, Louis Muglia², and Nathan D. Price¹
(1)Hood-Price Lab, Institute for Systems Biology, Seattle, WA, (2)Cincinnati Children's Hospital Medical Center, Cincinnati, OH, (3)Magee-Womens Research Institute, Pittsburgh, PA
6. **Cobrapy: A Solid Foundation for a Diverse Python Metabolic Modeling Community.**
Moritz E. Beber¹, João G.R. Cardoso², Christian Diener³, Mátyás Fodor⁴, Ali Kaafarani¹, Zachary A. King⁵, Matthias König⁶, Svetlana Kutuzova¹, Christian Lieven¹, Alba Lopez¹, Michael Mundy⁷, Vivek Rai⁸, Henning Redestig⁹, Nikolaus Sonnenschein¹, Peter C. St. John¹⁰, and Sascha Steinbüß¹¹
(1)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark, (2)Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark, (3)National Institute for Genomic Medicine, Mexico City, Mexico, (4)Genialis, Inc, Ljubljana, Slovenia, (5)Bioengineering, University of California, San Diego, La Jolla, CA, (6)Institute of Theoretical Biology, Humboldt University Berlin, Berlin, Germany, (7)Center for Individualized Medicine, Mayo Clinic, Rochester, MN, (8)Department of Computational Medicine & Bioinformatics, University of Michigan, Ann Arbor, MI, (9)Genencor International, DuPont, Leiden, Netherlands, (10)Biosciences Center, National Renewable Energy Laboratory, Golden, CO, (11)Debian Project, New York, NY
7. **Probabilistic Metabolic Network Reconstruction: A Modular Approach to Integrate Omics Data and Represent Uncertainty.**
David Bernstein^{1,2}, Shany Ofaim^{2,3}, Luca Zoccarato⁴, Daniel Sher⁵, and Daniel Segrè^{1,2,3,6,7}
(1)Department of Biomedical Engineering, Boston University, Boston, MA, (2)Biological Design Center, Boston University, Boston, MA, (3)Bioinformatics Program, Boston University, Boston, MA, (4)Department of Experimental Limnology, Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Neuglobsow, Germany, (5)Department of Marine Biology, University of Haifa, Haifa, Israel, (6)Department of Biology, Boston University, Boston, MA, (7)Department of Physics, Boston University, Boston, MA
8. **Essential Metabolism for a Minimal Cell.**
Marian Breuer
Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL

9. **A Distance Measure for Heterogeneity Using Genome Scale Metabolic Networks.**
Andrea Cabbia¹, Peter A. J. Hilbers¹, and Natal A. W. van Riel^{1,2}
 (1)Biomedical Engineering, Eindhoven University of Technology, Eindhoven, Netherlands, (2) Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands
10. **Thermodynamics As an Optimization Goal for Metabolism: Prediction of Metabolite Levels, Rate Constants and Post-Translational Regulation.**
William R. Cannon¹, Jeremy D. Zucker¹, Douglas J. Baxter², Neeraj Kumar¹, Jennifer M. Hurley³, Scott E. Baker⁴, and Jay C. Dunlap⁵
 (1)Computational Biology, Pacific Northwest National Laboratory, Richland, WA, (2)Research Computing, Pacific Northwest National Laboratory, Richland, WA, (3)Department of Biological Sciences, Rensselaer Polytechnic Institute, Troy, NY, (4)Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, (5)Department of Molecular and Systems Biology, Geisel School of Medicine at Dartmouth, Hanover, VT
11. **Comparative Genomics and Network Modeling of Parasites.**
Maureen A. Carey¹, Michal Stolarczyk², Ana Untariou³, Gregory L. Medlock⁴, Jennifer L. Guler², and Jason A. Papin⁴
 (1)Microbiology, Immunology, and Cancer Biology, University of Virginia, Charlottesville, VA, (2) Biology, University of Virginia, Charlottesville, VA, (3)Biomedical Engineering, University of Virginia, Charlottesville, VA, (4)Department of Biomedical Engineering, University of Virginia, Charlottesville, VA
12. **Models Constrained with Transcriptomics or Proteomics Data Generate Discordant Predictions.**
Maureen A. Carey¹, Ana Untariou², and Jason A. Papin³
 (1)Microbiology, Immunology, and Cancer Biology, University of Virginia, Charlottesville, VA, (2) Biomedical Engineering, University of Virginia, Charlottesville, VA, (3)Department of Biomedical Engineering, University of Virginia, Charlottesville, VA
13. **Toward a Whole-Cell Model of Human Embryonic Stem Cells: A Composable Model of Metabolism Built on a Scalable Framework for Integrating Multiple Pathway Submodels.**
Yin Hoon Chew, Andrew Sundstrom, and Jonathan R. Karr
 Department of Genetics & Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York City, NY
14. **Metabolic Analysis of Nitrogen Fixation in the Soybean Micro-Symbiont Sinorhizobium Fredii Using Constraint-Based Modeling.**
Carolina A. Contador¹, Qianwen Wang¹, Siu Hung Joshua Chan², and Hon-Ming Lam¹
 (1)School of Life Sciences and Center for Soybean Research of the Partner State Key Laboratory of Agrobiotechnology, Chinese University of Hong Kong, Hong Kong, China, (2)Department of Chemical and Biological Engineering, Colorado State University, Fort Collins, CO
15. **Metabolic Network Reconstruction for the Pan-Genome: A Scalable Method to Get High-Quality Metabolic Models across the Tree of Life.**
Kevin Correia and Radhakrishnan Mahadevan
 Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON, Canada
16. **Reconciling Barseq Datasets for *Z. Mobilis* Using a Genome-Scale Metabolic Model.**
Dylan Courtney¹, Wai Kit Ong¹, Shu Pan¹, and Jennifer Reed²
 (1)Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI, (2)Chemical and Biological Engineering, University of Wisconsin Madison, Madison, WI
17. **Developing and Evaluating Integrated Metabolic Regulatory Models for Microbial Life.**
Molly Creagar¹, Thomas Kamp², Ruby Fore³, Derek Friend⁴, Craig Disselkoen⁵, Christopher S. Henry⁶, Matthew DeJongh⁷, Aaron Best⁸, William Lindsey², and Nathan Tintle⁹
 (1)University of San Francisco, San Francisco, CA, (2) Dordt College, Sioux Center, IA, (3)Brown University,

POSTER TITLES

- Providence, RI, (4)University of Nevada, Reno, Reno, NV, (5)University of California, San Diego, San Diego, CA, (6)Argonne National Laboratory, Argonne, IL, (7)Computer Science, Hope College, Holland, MI, (8)Biology, Hope College, Holland, MI, (9)Mathematics, Dordt College, Sioux Center, IA
18. **Constructing a High Quality Genome-Scale Metabolic Network for Streptomyces Lividans TK24.**
Wouter Daniels¹, **Kjerstin De Winter**¹, Justien Ghesquière¹, Kenneth Simoens¹, Tobias Busche², Christian Rückert², Bart Lievens³, Tomer Shlomi⁴, Jörn Kalinowski², and Kristel Bernaerts¹
(1)Bio- and Chemical Systems Technology, Reactor Engineering, and Safety (CREaS), Department of Chemical Engineering, KU Leuven, University of Leuven, Leuven, Belgium, (2)Center for Biotechnology (CeBiTec), University of Bielefeld, Bielefeld, Germany, (3)Laboratory for Process Microbial Ecology and Bioinspirational Management, Department of Microbial and Molecular Systems, KU Leuven, University of Leuven, Sint-Katelijne Waver, Belgium, (4)Computer Science, Technion IIT, Haifa, Israel
19. **Construction of a Dynamic Flux Balance Analysis Model for Streptococcus Gordonii.**
Kjerstin De Winter¹, Wim Teughels², Steffen Waldherr³, and Kristel Bernaerts⁴
(1)Department of Chemical Engineering, KU Leuven, University of Leuven, Leuven, Belgium, (2)Periodontology and Oral Microbiology, Department of oral health science, KU Leuven, University of Leuven, Leuven, Belgium, (3)Bio- and Chemical Systems Technology, Reactor Engineering and Safety (CREaS), Department of Chemical Engineering, KU Leuven, University of Leuven, Leuven, Belgium, (4)Bio- and Chemical Systems Technology, Reactor Engineering, and Safety (CREaS), Department of Chemical Engineering, KU Leuven, University of Leuven, Leuven, Belgium
20. **Identification and Analysis of Bacterial Genomic Metabolic Signatures in the Department of Energy's Systems Biology Knowledgebase.**
Matthew DeJongh¹, Megan Grimes¹, Lauren Satkiewicz¹, Nathan Tintle², and Aaron Best³
(1)Computer Science, Hope College, Holland, MI, (2)Mathematics, Dordt College, Sioux Center, IA, (3)Biology, Hope College, Holland, MI
21. **Untangling Dynamic Protein Behaviour with Rule-Based Modelling and Sensitivity Analysis.**
Robin Denhardt-Eriksson
EPFL - LCSB, Lausanne, Switzerland
22. **Towards a Genome-Wide Transport Systems Encoding Genes Tracker.**
Lagoa D. (1), Liu F. (2), Ferreira E.C. (1), Faria J. (2), Henry C. (2), Dias O. (1)
(1) CEB - Centre of Biological Engineering, University of Minho
(2) Mathematics and Computer Science Division, Argonne National Laboratory
23. **Predicting Shifts in Cardiomyocyte Metabolism during Heart Failure.**
Bonnie Dougherty¹, Kristopher Rawls¹, Glynis Kolling¹, Vinnakota Kalyan², Venkat Pannala², Anders Wallqvist², and Jason A. Papin¹
(1)Department of Biomedical Engineering, University of Virginia, Charlottesville, VA, (2)Department of Defense Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick, MD
24. **Metabolic Network Reconstruction of Treponema Pallidum Spp. Pallidum.**
Silvia Morini^{1,2}, Isabella Casini^{3,4}, Thomas M. Hamm^{1,5}, Kay Nieselt^{5,6}, and **Andreas Dräger**^{1,5}
(1)Computational Systems Biology of Infection and Antimicrobial-Resistant Pathogens, Center for Bioinformatics Tübingen (ZBIT), Tübingen, Germany, (2)Department of biotechnologies and biosciences, Università degli Studi di Milano - Bicocca, Milano, Italy, (3)Biological and Environmental Engineering, Max Planck Institute for Developmental Biology, Tübingen, Germany, (4)Department of Geosciences, University of Tübingen, Tübingen, Germany, (5)Department of Computer Science, University of Tübingen, Tübingen, Germany, (6)Integrative Transcriptomics, Center for Bioinformatics Tübingen (ZBIT), Tübingen, Germany

25. Modeling of Potentially Virulence-Associated Metabolic Pathways in *Pseudomonas Aeruginosa* PA14 Including Experimental Verification.

Alina Renz^{1,2}, Erwin Bohn³, Monika Schütz³, and Andreas Dräger^{1,2}

(1)Computational Systems Biology of Infection and Antimicrobial-Resistant Pathogens, Center for Bioinformatics Tübingen (ZBIT), Tübingen, Germany, (2)Department of Computer Science, University of Tübingen, Tübingen, Germany, (3)Institute for Medical Microbiology and Hygiene, University Hospital Tübingen, Tübingen, Germany

26. Representation of ME-Models in Sbml.

Marc Alexander Voigt^{1,2}, Colton J. Lloyd³, Laurence Yang³, Zachary A. King³, Oliver Kohlbacher^{2,4,5,6}, Kay Nieselt^{2,7}, and Andreas Dräger^{1,2}

(1)Computational Systems Biology of Infection and Antimicrobial-Resistant Pathogens, Center for Bioinformatics Tübingen (ZBIT), Tübingen, Germany, (2)Department of Computer Science, University of Tübingen, Tübingen, Germany, (3)Bioengineering, University of California, San Diego, La Jolla, CA, (4)Max Planck Institute for Developmental Biology, Tübingen, Germany, (5)Center for Personalized Medicine (ZPM), University of Tübingen, Tübingen, Germany, (6)Quantitative Biology Center (QBiC), University of Tübingen, Tübingen, Germany, (7)Integrative Transcriptomics, Center for Bioinformatics Tübingen (ZBIT), Tübingen, Germany

27. Automated Reconstruction, Evaluation and Comparison of Diverse Genome Scale Fungal Models.

Janaka N Edirisinghe^{1,2}, José P. Faria¹, Sara Calhoun³, Filipe Liu¹, Igor Grigoriev³, and Christopher S. Henry^{1,2}

(1)Argonne National Laboratory, Argonne, IL, (2)Computation Institute, University of Chicago, Chicago, IL, (3)Lawrence Berkeley National Laboratory, Berkeley, CA

28. Metabolic Supply and Demand Constrain Optimal Regulatory Strategies: Structural Evidence from the *E coli* Small Molecule Regulatory Network.

Christian Euler¹ and Radhakrishna Mahadevan²

(1)Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON, Canada, (2)Department of Chemical Engineering and Applied

Chemistry, University of Toronto, Toronto, ON, Canada

29. Improving High Throughput Genome-Scale Metabolic Model Reconstruction and Validation with Tnseq Data Using Modelseed 2.

José P. Faria¹, Janaka N Edirisinghe¹, Filipe Liu², Samuel M.D. Seaver¹, James G. Jeffryes¹, Qizh Zhang¹, Pamela Weisenhorn¹, Boris Sadkhin¹, Nidhi Gupta¹, Tian Gu¹, and Christopher S. Henry¹

(1)Argonne National Laboratory, Argonne, IL, (2)Centre of Biological Engineering, University of Minho, Braga, Portugal

30. A Deep Dive into False Positive Gene Essentiality Predictions with Modeling and Multi-Omics Analysis.

Adam M. Feist

Bioengineering, University of California, San Diego, La Jolla, CA

31. A Multiobjective Strain Design Platform for Modular Cell Engineering.

Sergio Garcia¹ and Cong T. Trinh²

(1)Chemical and Biomolecular Engineering, University of Tennessee, Knoxville, TN, (2)Department of Chemical and Biomolecular Engineering, University of Tennessee, Knoxville, TN

32. Development of an Accelerated Workflow for Parameterizing Kinetic Models of Metabolism.

Satyakam Dash, Saratram Gopalakrishnan, and Costas D. Maranas

Department of Chemical Engineering, The Pennsylvania State University, University Park, PA

33. Mlproscope: Machine Learning (ML) Based Method for Engineering Enzymes Faster By Modeling Protein Fitness Landscape (ProScape).

Sanjan T.P. Gupta^{1,2}, Evan Glasgow^{2,3}, Brian G. Fox^{2,3}, Parmeswaran Ramanathan⁴, and Jennifer Reed^{1,2}

(1)Chemical and Biological Engineering, University of Wisconsin Madison, Madison, WI, (2)Great Lakes Bioenergy Research Center, Madison, WI, (3)Biochemistry, University of Wisconsin Madison, Madison, WI, (4)Electrical and Computer Engineering, University of Wisconsin Madison, Madison, WI

POSTER TITLES

34. Constraint-Based Modeling Reveals the Distinct Metabolic Potential in Gut Microbiomes of Inflammatory Bowel Disease Patients with Dysbiosis.

Almut Heinken¹, Laurent Heirendt¹, Federico Baldini¹, Ronan M.T. Fleming², and Ines Thiele¹
(1)Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg, (2)Leiden Academic Centre for Drug Research, Leiden University, Leiden, Netherlands

35. COBRA.jl - Gearing up for the huge scale.

Laurent Heirendt¹, Sylvain Arreckx¹, Venkata P. Satagopam¹, Reinhard Schneider¹, Ines Thiele¹, and Ronan M.T. Fleming²
(1)Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg, (2)Leiden Academic Centre for Drug Research, Leiden University, Leiden, Netherlands

36. Genome-Scale Flux Elucidation in the Fast-Growing Cyanobacterium *Synechococcus Elongatus* UTEX 2973.

John I Hendry¹, Saratram Gopalakrishnan¹, Himadri B. Pakrasi², Yinjie J. Tang³, and Costas D. Maranas¹
(1)Department of Chemical Engineering, The Pennsylvania State University, University Park, PA, (2)Department of Biology, Washington University, St. Louis, MO, (3)Department of Energy, Environmental and Chemical Engineering, Washington University, St. Louis, MO

37. Miren: An Optimization Tool for Data-Driven Discovery of Global Regulatory Phenomena Used to Elucidate the Heat Stress Response Mechanism in Rice Seed.

Mohammad Mazharul Islam¹, Jaspreet Sandhu², Harkamal Walia², and Rajib Saha³
(1)Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE, (2)Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE, (3)Department of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE

38. Leveraging a *Clostridium Difficile* Genome and Metagenomics to Identify Candidate Probiotic Bacterial Strains.

Matthew Jenior¹ and Jason Papin²
(1)Biomedical Engineering, University of Virginia, Charlottesville, VA, (2)Department of Biomedical Engineering, University of Virginia, Charlottesville, VA

39. Metabolic Chaos during Microbial Stress Responses.

Paul A. Jensen
Bioengineering Department and Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL

40. Fine Tuning Thresholds to Facilitate Integration of Transcriptomics Data.

Chintan Joshi, Song-Min Schinn, Austin Chiang, Anne Richelle, and Nathan Lewis
University of California, San Diego, La Jolla, CA

41. A Multi-Omics Investigation Unveiling the Tiered Regulation of Breast Cancer Cell Metabolism.

Rotem Katzir
University of Maryland, College Park, MD

42. Constraint-based modeling of allelic variation mechanistically links genome-wide associations

Erol Kavvas
Bioengineering, UCSD, La Jolla, CA

43. Laboratory Evolution Reveals a Two-Dimensional Rate-Yield Tradeoff in Microbial Metabolism.

Zachary A. King¹ and Chuankai Cheng²
(1)Bioengineering, University of California, San Diego, La Jolla, CA, (2)University of California, San Diego, La Jolla, CA

44. A Metabolic Reconstruction of *Lactobacillus Reuteri* and Analysis of Its Potential As a Cell Factory.

Thordis Kristjansdottir¹, Elleke F. Bosma², Filipe Branco dos Santos³, Emre Özdemir⁴, Alex Toftgaard Nielsen⁴, and Steinn Gudmundsson¹
(1)Centre For Systems Biology, University Of Iceland, Reykjavik, Iceland, (2)DTU Biosustain, (3)University of

- Amsterdam, Amsterdam, Netherlands, (4)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark
45. **Metabolic Modeling of a Model Diatom.**
Manish Kumar¹, Shawn W. Polson², Lisa Zeigler Allen³, and Karsten Zengler⁴
 (1)Department of Pediatrics, University of San Diego, San Diego, CA, (2)Department of Computer and Information Sciences, University of Delaware, Newark, DE, (3)Department of Microbial and Environmental Genomics, J. Craig Venter Institute, La Jolla, CA, (4)Department of Pediatrics, University of California, San Diego, CA
46. **Computational Modeling of *Ostreococcus Tauri* Central Metabolism Based on Statistical Thermodynamics.**
Neeraj Kumar¹, William R. Cannon¹, James E. Evans², and Jeremy D. Zucker³
 (1)Computational Biology, Pacific Northwest National Laboratory, Richland, WA, (2)Environmental Molecular Sciences Laboratory, Pacific Northwest National Lab, Richland, WA, (3)Pacific Northwest National Laboratory, Richland, WA
47. **Bofdat: Generating Biomass Objective Function for Constraint-Based Metabolic Models from Experimental Data.**
Jean-Christophe Lachance
 Biology, Université de Sherbrooke, Sherbrooke, QC, Canada
48. **Elucidating the Metabolic Rewiring of Pluripotent Stem Cells during Differentiation into Adult Progenitors Using an Updated and Enzyme-Constrained Human Genome-Scale Model.**
Meiyappan Lakshmanan¹, Dong-Yup Lee^{2,3}, and Lokanand Koduru⁴
 (1)Bioinformatics Group, Bioprocessing Technology Institute, Singapore, Singapore, (2)Bioprocessing Technology institute, A*STAR, Singapore, Singapore, (3)School of Chemical Engineering, Sungkyunkwan University, Gyeonggi-do, Korea, Republic of (South), (4)National university of Singapore, Singapore, Singapore
49. **Implementation and Application of a New Combined Stoichiometric and Thermodynamic Flux Balance Analysis.**
Simeon Leupold, Bastian Niebel, and Matthias Heinemann
 Molecular Systems Biology, University of Groningen, Groningen, Netherlands
50. **Implementing and Evaluating a Gaussian Mixture Framework for Identifying Gene Function from Tnseq Data.**
Kevin Li¹, Rachel Chen², William Lindsey³, Aaron Best⁴, Matthew DeJongh⁵, Christopher S. Henry⁶, and Nathan Tintle⁷
 (1)Columbia University, New York, NY, (2)North Carolina State University, Raleigh, NC, (3)Dordt College, Sioux Center, IA, (4)Biology, Hope College, Holland, MI, (5)Computer Science, Hope College, Holland, MI, (6)Argonne National Laboratory, Argonne, IL, (7)Mathematics, Dordt College, Sioux Center, IA
51. **Sbmltools: A Sbml Integration Tool for Kbase.**
Filipe Liu¹, José P. Faria¹, Qizhi Zhang¹, Miguel Rocha², Isabel Rocha², and Christopher S. Henry¹
 (1)Argonne National Laboratory, Argonne, IL, (2)Centre of Biological Engineering, Department of Biological Engineering, University of Minho, Braga, Portugal
52. **The genetic basis for adaptation of model-designed syntrophic co-cultures**
Colton J. Lloyd¹, Ali Ebrahim¹, Laurence Yang¹, Zachary A. King¹, Edward Catoiu¹, Edward J. O'Brien², Joanne Liu³, and Bernhard O. Palsson¹
 (1)Bioengineering, University of California, San Diego, La Jolla, CA, (2)Bioinformatics, University of California, San Diego, La Jolla, CA, (3)University of California, San Diego, La Jolla, CA
53. **Exploring the Metabolic Role of Multidrug-Resistance Efflux Pumps in *Salmonella Typhimurium*.**
Sean G. Mack¹, Xuan Wang-Kan², Laura J. V. Piddock², and Daniel J. Dwyer^{1,3,4,5}
 (1)Dept. of Chemical and Biomolecular Engineering, University of Maryland, College Park, MD, (2)

POSTER TITLES

- Inst. of Microbiology and Infection, University of Birmingham, Birmingham, United Kingdom, (3)Dept. of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD, (4)Inst. for Physical Science and Technology, University of Maryland, College Park, MD, (5)Dept. of Bioengineering, University of Maryland, College Park, MD*
54. **Thermodynamic Metabolic Flux Analysis with Consistent Handling of Errors in the Estimates from the Component Contribution Method.**
Vishnuvardhan Mahamkali¹, Kaspar Valgepea¹, Tim McCubbin¹, Esteban Marcellin¹, and Lars K. Nielsen^{1,2}
(1)AIBN, The University of Queensland, Brisbane, Australia, (2)DTU Biosustain, Technical University of Denmark, Kongens Lyngby, Denmark
55. **Developement and Application of Constraint-Based Modeling Methods to Study Vulnerabilities Associated to Lipid Metabolism in Prostate Cancer.**
Igor Marín de Mas¹, Esther Aguilar², Carmen Bedia³, Timothy M. Thomson⁴, Romà Tauler³, Balazs Papp⁵, Marta Cascante², and Lars K. Nielsen¹
(1)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark, (2)Departamento de Bioquímica I Biología Molecular, University of Barcelona, Barcelona, Spain, (3)IDAEA, CSIC, Barcelona, Spain, (4)Department of Environmental Chemistry, Molecular Biology Institute of Barcelona, National Research Council (IBMB-CSIC), (5)Department of Biochemistry and Molecular Biology, Biological Research Centre of the Hungarian Academy of Sciences, Szeged, Hungary
56. **Modelling of Cell Metabolism to Reduce the Lactate Generation in CHO Cell Cultures.**
Iván Martínez-Monge^{1,2}, Svetlana Volkova¹, Igor Marín de Mas¹, Hooman Hefzi³, Pere Comas², Nathan Lewis³, Martí Lecina^{2,4}, Jordi Joan Cairó², and Lars Keld Nielsen¹
(1)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby, Denmark, (2)Department of Chemical, Biological and Environmental Engineering, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Spain, (3)Department of Bioengineering, University of California, San Diego, CA, (4)Bioengineering Department, IQS-Universitat Ramon Llull, Barcelona, Spain
57. **Systematic Reduction of Genome-Scale Models for the Study of Metabolic Phenotypes of Human Cells.**
Maria Masid¹, Meriç Ataman¹, and Vassily Hatzimanikatis²
(1)Laboratory of Computational Systems Biotechnology (LCSB), EPFL, Lausanne, Switzerland, (2)Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland
58. **Elementary Mass Action Stoichiometric Simulation Models Predict Non-Negligible Concentrations of Enzyme-Bound Metabolites.**
Marta R. A. Matos¹, Daniel C. Zielinski², Lars K. Nielsen³, and Nikolaus Sonnenschein⁴
(1)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs Lyngby, Denmark, (2)Bioengineering, University of California, San Diego, CA, (3)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark, (4)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark
59. **Medusa: A Software Package for Construction and Analysis of Ensembles of Metabolic Networks.**
Gregory L. Medlock and Jason A. Papin
Department of Biomedical Engineering, University of Virginia, Charlottesville, VA
60. **A Systematic Assessment of Current Genome-Scale Metabolic Reconstruction Tools.**
Sebastián N. Mendoza, Brett G. Olivier, Douwe Molenaar, and Bas Teusink
Systems Bioinformatics, Amsterdam Institute for Molecules, Medicines and Systems, Vrije Universiteit Amsterdam, Amsterdam, Netherlands
61. **Genome-Scale Metabolic Model of Chromochloris, an Emerging Model Organism for Sustainable Fuel Production.**
Alexander Metcalf and Nanette R. Boyle
Chemical and Biological Engineering, Colorado School of Mines, Golden, CO

62. Exploring Biosynthetic Pathways for the Production of Five Methyl Ethyl Ketone Precursors.

*Milenko Tokic¹, Noushin Hadadi², Meric Ataman³, Dário Neves⁴, Birgitta E. Ebert⁴, Lars M. Blank⁴, **Ljubisa Miskovic³**, and Vassily Hatzimanikatis⁵*

(1)Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), 1015 Lausanne, Switzerland, (2)Swiss Institute of Bioinformatics (SIB), 1015 Lausanne, Switzerland, (3)Laboratory of Computational Systems Biotechnology (LCSB), EPFL, Lausanne, Switzerland, (4)iAMB - Institute of Applied Microbiology, ABBt - Aachen Biology and Biotechnology, RWTH Aachen University, Aachen, Germany, (5)Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland

63. Predicted Responses of a Large-Scale Pseudomonas Putida KT2440 Kinetic Metabolic Model to Several Single-Gene Knockouts Are Consistent with Experimental Observations.

*Milenko Tokic¹, **Ljubisa Miskovic²**, and Vassily Hatzimanikatis³*

(1)Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), 1015 Lausanne, Switzerland, (2) Laboratory of Computational Systems Biotechnology (LCSB), EPFL, Lausanne, Switzerland, (3)Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland

64. Genome-Scale Metabolic Modeling of Actinomycetes for Secondary Metabolites Production.

***Omkar Mohite¹**, Tilmann Weber¹, Hyun Uk Kim², and Sang Yup Lee^{1,2}*

(1)Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark, (2)Department of Chemical and Biomolecular Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, Republic of (South)

65. Genome Scale Metabolic Analysis of the Lactobacillus Genus with a Focus on Probiotic Strains.

***Thomas J. Moutinho Jr.**, Benjamin C. Neubert, and Jason A. Papin*

Department of Biomedical Engineering, University of Virginia, Charlottesville, VA

66. Multiple Reconstructions of P. Putida for Metabolic Engineering Applications.

Joshua Mueller

Chemical and Biomolecular Engineering, University of Nebraska - Lincoln, Lincoln, NE; Biological Engineering, University of California, San Diego, San Diego, CA

67. System-Level Examination of Metabolism in Rhizosphere Grown Burkholderiales.

***Ali Navid¹**, Jennifer Pett-Ridge², Evan Star³, Jillian F. Banfield⁴, Mary Firestone⁵, and Erin Nuccio⁶*

(1)Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA, (2) Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, (3)2Department of Plant & Microbial Biology, University of California, Berkeley, CA, (4)University of California, Berkeley, CA, (5)Department of Environmental Science, Policy & Management, University of California, Berkeley, CA, (6)Lawrence Livermore National Laboratory, Livermore, CA

68. The Strain Optimization API: A Flexible Strain Design Formalism for an Automated High-Throughput Industrial Pipeline.

Chiam Yu Ng

Bioinformatics and Software Engineering, Amyris, Inc, Emeryville, CA

69. Genome Scale Metabolic Modeling and Analysis of Clostridium Difficile.

***Charles J. Norsigian**, Bernhard Ø. Palsson, and Jonathan M. Monk*

Bioengineering, University of California, San Diego, La Jolla, CA

POSTER TITLES

70. **An Integrated COBRA-PBPK Model to Study Interactions between the Gut Microbiome and the Brain in Autism.**
Meghana Palukuri¹, Shruti Shivakumar², Swagatika Sahoo³, and Raghunathan Rengaswamy³
(1)Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX, (2)College of Computing, Georgia Institute of Technology, Atlanta, GA, (3)Indian Institute of Technology Madras, Chennai, India
71. **Genome-Scale Constraint-Based Metabolic Modeling and Analysis of *Cryptococcus Curvatus*.**
Nhung Pham, Maarten Reijnders, Maria Suarez-Diez, Bart Nijse, Peter J Schaap, and Vitor A. P. Martins Dos Santos
Laboratory of Systems and Synthetic Biology, Wageningen University and Research, Wageningen, Netherlands
72. **Identifying Orphan Enzymes in *Pseudomonas Putida* Using Cobra-Based Methods.**
Brandon Phan¹, Shu Pan², Christine Vera Colon³, Nathaniel Bennett³, and Jennifer L. Reed³
(1)Cellular and Molecular Biology, University of Wisconsin-Madison, Madison, WI, (2)Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI, (3)Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI
73. **Enzyme Multicollinearity in Genome-Scale Metabolic Models Revealed By an Efficient Coupling Algorithm**
Dikshant Pradhan and Paul A. Jensen
University of Illinois at Urbana-Champaign Bioengineering Department
74. **Model-Guided Engineering of Cyanobacteria for Improved Biofuel Production.**
Hugh M. Purdy and Jennifer L. Reed
Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI
75. **Kinetic Model of *Clostridium Beijerinckii* Based on Phenotypic States Superposition.**
Marcelo Rivas-Astroza, Ivan Paredes, German Aroca, and Raul Conejeros
School of Biochemical Engineering, Pontificia Universidad Catolica de Valparaiso, Valparaiso, Chile
76. **EtfI: A Formulation of ME-Models Accounting for Expression and Thermodynamics in Constraint-Based Models.**
Pierre Salvy¹ and Vassily Hatzimanikatis²
(1)Laboratory of Computational Systems Biotechnology (LCSB), EPFL, Lausanne, Switzerland, (2)Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland
77. **Effective Visualization for Investigating Elementary Flux Modes in Genome-Scale Metabolic Models.**
Chaitra Sarathy¹, Michael Lenz¹, Martina Kutmon^{1,2}, Chris T. Evelo^{1,2}, and Ilja C.W. Arts^{1,3}
University of California, San Diego, La Jolla, CA
78. **Deconvoluting Independent Regulatory Signals in the *Escherichia coli* Transcriptome.**
Anand Sastry¹, Ye Gao², Richard Szubin¹, Ying Hefner¹, Sibe Xu¹, Donghyuk Kim^{1,3}, Kumari Sonal Choudhary¹, Laurence Yang¹, Zachary A. King¹, and Bernhard O. Palsson¹
(1)Bioengineering, University of California, San Diego, La Jolla, CA, (2)Biological Sciences, University of California, San Diego, San Diego, CA, (3)Department of Genetic Engineering and Graduate School of Biotechnology, Kyung Hee University, Yongin, Korea, Republic of (South)
79. **Characterizing and Ranking Computed Metabolic Engineering Strategies.**
Philipp Schneider and Steffen Klamt
Analysis and Redesign of Biological Networks, Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany

80. **Optfill: A Novel Optimization-Based Tool to Automate the Gapfilling of Genome-Scale Metabolic Models.**

Wheaton Schroeder¹ and Rajib Saha²

(1)Chemical and Biomolecular Engineering, University of Nebraska - Lincoln, Lincoln, NE, (2)Department of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE

81. **Sammi: An Interactive, Semi-Automated Tool for the Illustration and Visualization of Metabolic Networks.**

Andre Schultz and Rehan Akbani

Bioinformatics and Computational Biology, University of Texas MD Anderson Cancer Center, Houston, TX

82. **Genome-Scale Metabolic Reconstructions for Phylogeny?**

Christian Schulz and Eivind Almaas

Dept. of Biotechnology, NTNU - Norwegian University of Science and Technology, Trondheim, Norway

83. **A Computational Knowledge-Base Elucidates the Response of Staphylococcus Aureus to Different Media Types.**

Bernhard O. Palsson¹, **Yara Seif²**, Jon Monk³, Nathan Mih⁴, Saugat Poudel², Jared T. Broddrick¹, Cristal Zuniga P⁵, Hannah Tsunemoto⁶, and Karsten Zengler⁷

(1)Bioengineering, University of California, San Diego, La Jolla, CA, (2)UCSD, San Diego, CA, (3)UCSD, La Jolla, CA, (4)UCSD, (5)Pediatrics, University of California, San Diego, CA, (6)University of California San Diego, San Diego, (7)Department of Pediatrics, University of California, San Diego, CA

84. **Describing the Glucose-Lactate Consumption Rate during Expansion and Osteogenic Differentiation of Human Bone Marrow Derived Mesenchymal Stem Cells: A Premise for Building a Systems Biology Model of Osteogenesis Using Metabolomics Analysis.**

Póra B. Sigmarsdóttir¹, Sarah McGarrity², Óttar Rolfsson³, and Olafur Sigurjonsson⁴

(1)Reykjavík University, Reykjavík, Iceland, (2)Center for Systems Biology, University of Iceland, 101 Reykjavik, Iceland, (3)Centre For Systems Biology, University Of Iceland, Reykjavik, Iceland, (4)Reykjavik University

85. **Mighty Morphing Metabolic Models: Leveraging Manual Curations for Automatic Metabolic Reconstruction of Clades.**

Evangelos Simeonidis, Brendan King, Matthew A. Richards, and Nathan D. Price

Hood-Price Lab, Institute for Systems Biology, Seattle, WA

86. **Utilizing RNA-Seq Data in Bayesian Estimation of Gene Activity States.**

Eldon Sorensen¹, Jaden Boehme², Angela Gasdaska³, Matthew DeJongh⁴, Aaron Best⁵, William Lindsey⁶, and Nathan Tintle⁷

(1)Pacific Lutheran University, Tacoma, WA, (2)Oregon State University, Corvallis, OR, (3)Emory University, Atlanta, GA, (4)Computer Science, Hope College, Holland, MI, (5)Biology, Hope College, Holland, MI, (6)Dordt College, Sioux Center, IA, (7)Mathematics, Dordt College, Sioux Center, IA

87. **Bayesian Inference of Metabolic Kinetics from Genome-Scale Multiomics Data.**

Peter C. St. John¹, Jonathan Strutz², and Yannick J. Bomble¹

(1)Biosciences Center, National Renewable Energy Laboratory, Golden, CO, (2)Chemical and Biological Engineering, Northwestern University, Evanston, IL

88. **Dynamic FBA with Time-Course Transcriptomics.**

Snorre Sulheim^{1,2}, Alexander Wentzel¹, and Eivind Almaas³

(1)Dept. of Biotechnology and Nanomedicine, SINTEF Industry, Trondheim, Norway, (2)Dept. of Biotechnology and Food Science, NTNU - Norwegian University of Science and Technology, Trondheim, Norway, (3)Dept. of Biotechnology, NTNU - Norwegian University of Science and Technology, Trondheim, Norway

89. **Determination of CHO Biomass Composition.**

Diana Szeliova^{1,2}, David Ruckerbauer^{2,3}, Sarah Galleguillos^{1,2}, Stephan Hann², Michael Hanscho^{1,2}, and Nicole Borth^{1,2}

(1)The Austrian Centre of Industrial Biotechnology, Vienna, Austria, (2)University of Natural Resources and Life Sciences, Vienna, Austria, (3)Austrian Centre of Industrial Biotechnology, Vienna, Austria

POSTER TITLES

90. Grid-Based Computational Methods for the Design of Constraint-Based Parsimonious Chemical Reaction Networks to Simulate Metabolite Production: Gridprod.

Takeyuki Tamura

Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji, Kyoto, Japan

91. Identification of the Enzymes Responsible for Diverse Phenotypic States of Yeast Lipid Metabolism Using Comprehensive Mechanistic Models.

Sophia Tsouka¹ and Vassily Hatzimanikatis²

(1)Laboratory of Computational Systems Biotechnology (LCSB), EPFL, Lausanne, Switzerland,

(2)Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland

92. Selective Metabolic Vulnerabilities in Multiple Myeloma.

Luis Vitores Valcárcel^{1,2}, Raquel Ordoñez², Cem Meydan³, Iñigo Apaolaza¹, Ari Melnick³, Xabier Agirre², Felipe Prosper², and **Francisco J. Planes**¹
(1)TECNUN, University of Navarra, San Sebastian, Spain, (2)CIMA, University of Navarra, Pamplona, Spain, (3)Cornell University, New York, NY

93. Computational Pathway Design for Funneling Lignin Intermediates to Aromatic Products.

Lin Wang¹, Gregg T. Beckham², and Costas D. Maranas¹

(1)Department of Chemical Engineering, The Pennsylvania State University, University Park, PA, (2)National Bioenergy Center, National Renewable Energy Laboratory, Golden, CO

94. Template of Metabolic Reprogramming in Cancer and Healthy Cells for Inferring Oncogenes.

Wu-Hsiung Wu¹, Fan-Yu Li¹, Yi-Chen Shu¹, Jin-Mei Lai², Peter Mu-Hsin Chang³, Chi-Ying F. Huang⁵, and **Feng-Sheng Wang**¹

(1)Department of Chemical Engineering, National Chung Cheng University, Chiayi, Taiwan, (2) Department of Life Science, Fu-Jen Catholic University, New Taipei City, Taiwan, (3)Faculty of Medicine, National Yang Ming University, Taipei, Taiwan, (4)Institute of Biopharmaceutical Sciences, National Yang Ming University, Taipei, Taiwan

95. Gapseq: A Novel Approach for in silico Prediction and Analysis of Bacterial Metabolic Pathways and Genome-Scale Networks.

Johannes Zimmermann, Christoph Kaleta, and **Silvio Waschina**

Institute for experimental medicine, Christian-Albrechts-University Kiel (UKSH Campus), Kiel, Germany

96. Erythrodb: A Multi-Omic Visual Knowledge Base for the Human Red Blood Cell.

James T. Yurkovich and Bernhard O. Palsson

Bioengineering, University of California, San Diego, La Jolla, CA

97. Multiomics Integration for Prediction of Complex Metabolic Phenotypes.

Aleksej Zelezniak

Chalmers University of Technology, Gothenburg, Sweden

98. Genome Scale Metabolic Model Assisted Strain Designs for Itaconic Acid Production in Yeast.

Eric M. Young¹, **Zheng Zhao**², Bianca Gielesen², Liang Wu², Ben Gordon³, Johannes A. Roubos², and Christopher A. Voigt⁴

(1)Synthetic Biology Center, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, (2)DSM Biotechnology Center, Delft, Netherlands, (3)The Foundry, MIT / Broad Institute, Cambridge, MA, (4)Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA

99. A Machine Learning-Empowered Kinetic Reconstruction of E. coli Metabolism.

Daniel C. Zielinski¹ and Bernhard O. Palsson²

(1)Bioengineering, University of California, San Diego, CA, (2)Bioengineering, University of California - San Diego, La Jolla, CA

INVITED SPEAKER BIOGRAPHIES

Evrin Acar

Simula Metropolitan Center for Digital Engineering

Evrin Acar is a Senior Research Scientist at Simula Metropolitan Center for Digital Engineering. Her research focuses on data mining, in particular, tensor factorizations, data fusion using coupled factorizations of higher-order tensors and matrices, and their applications in diverse disciplines. Prior to joining Simula, Evrin was with the Chemometrics and Analytical Technology group at the University of Copenhagen, and was awarded the Danish Council for Independent Research Sapere Aude Elite Young Researcher Award in 2012. Evrin received her MS and PhD in Computer Science from Rensselaer Polytechnic Institute (Troy, NY) in December 2006 and May 2008, respectively. She got her BS in Computer Engineering from Bogazici University (Istanbul, Turkey) in July 2003.

Vassily Hatzimanikatis

Ecole Polytechnique Federale de Lausanne (EPFL)

Dr. Vassily Hatzimanikatis is currently Associate Professor of Chemical Engineering and Bioengineering at Ecole Polytechnique Federale de Lausanne (EPFL), in Lausanne, Switzerland. He received a PhD and an MS in Chemical Engineering from the California Institute of Technology, and his Diploma in Chemical Engineering from the University of Patras, in Greece. After the completion of his doctoral studies, he held a research group leader position at the Swiss Federal Institute of Technology in Zurich (ETHZ), Switzerland. Prior to joining EPFL, Dr. Hatzimanikatis was Assistant Professor at Northwestern University, at Illinois, USA, and he worked for three years at DuPont and Cargill.

Dr. Hatzimanikatis is Editor-in-Chief of Metabolic Engineering Communication, and associate editor of the journals Metabolic Engineering, Biotechnology & Bioengineering and Integrative Biology, and Senior Editor of Biotechnology Journal. He has published over 70 articles and he is co-inventor in three patents and patent applications.

Dr. Hatzimanikatis is a fellow of the American Institute for Medical and Biological Engineering (2010) and he is in the founding Board of Director of the International Metabolic Engineering Society. He was a DuPont Young Professor (2001-2004), and he received the Jay Bailey Young Investigator Award in Metabolic Engineering (2000), and the ACS Elmar Gaden Award (2011).

Matthias Heinemann

University of Groningen, Netherlands

Matthias Heinemann is full professor for molecular systems biology at the University of Groningen in the Netherlands. Matthias earned his PhD in engineering from the RWTH Aachen University in Germany and mutated into a biologist during his postdoc time at the Institute of Molecular Systems Biology (ETH Zurich). In his research, he aims to understand how primary carbon and energy metabolism functions and how it controls other cellular processes. To achieve this goal, the members of his lab use wet and dry lab approaches of systems biology, in combination with the more classical approaches to biological research, with a particular emphasis on zooming into metabolism of individual cells.

In recent years, his lab has found that cells can measure intracellular flux (i.e. the rate of metabolic activity) and use this information for regulation of other metabolic fluxes, opening up a new view on metabolism (Mol Sys Biol 2010, PNAS 2012). Further, the lab showed that such flux-sensing can lead to bistability in metabolism (Mol Sys Biol 2014), to antibiotic tolerant persisters (Mol Sys Biol 2016), and has relevance for aging in yeast (eLife 2015). Recently, the lab discovered that an upper limit in Gibbs energy dissipation rate governs cellular metabolism (Nat Metabolism 2018) and that metabolism of yeast is an autonomous oscillator (Mol Cell 2017).

Nathan E. Lewis

University of California, San Diego

Dr. Lewis is an Assistant Professor of Pediatrics and Bioengineering at the University of California, San Diego. He received his BS in biochemistry at Brigham Young University, and his PhD at UC San Diego, where he focused on proteomics and developing novel approaches for analyzing biological big data using genome-scale systems biology modeling techniques. Dr. Lewis completed his postdoctoral training at the Wyss Institute at Harvard Medical School, where he worked on genome editing and the use of systems biology for the interpretation of genetic screens. Dr. Lewis' lab integrates all of his previous work by focusing heavily on the use of systems biology and genome editing techniques to map out and engineer the cell pathways controlling mammalian cell growth, protein synthesis, and protein glycosylation.

Research

Living systems have classically been studied in-depth one gene at a time. However, each gene and protein exists

INVITED SPEAKER BIOGRAPHIES

and functions within complex matrices consisting of a vast array of biomolecules, ranging from small metabolites to large macromolecules. Thus, in vivo, the thousands of other unique molecules and their interactions will influence the function and evolution of each individual protein. Systems-engineering principles are now being applied to elucidate how the cellular context influences each protein and how each protein influences cellular phenotype. This can be accomplished by treating each enzyme as a component in a vast network of proteins, and then modeling their interactions, as if the system were a chemical plant or electrical circuit. We develop novel algorithms to integrate genome-scale data with these models to gain insight into how the network context influences how each protein contributes to phenotypes, such as disease. We also leverage this knowledge to guide cell engineering efforts for biotherapeutic development.

Christian Lieven

Novo Nordisk Foundation Center for Biosustainability

Christian Lieven is a Post-Doc at the Novo Nordisk Foundation Center for Biosustainability.

He received his B.Sc. in molecular biotechnology at the Ruprecht-Karls University of Heidelberg in 2012. For his M.Sc. thesis in 2015 at the Institute for Applied Microbiology in Aachen, he reconstructed a genome-scale metabolic model of *Ustilago maydis*, a basidiomycete fungus relevant in the production of organic acids and lipids from crude glycerol.

During his PhD in the Design group of the iLoop Core, a translational research unit focused on microbial cell factory engineering at the Center for Biosustainability, Christian has been working on the EFPro2 project: in collaboration with other researchers from the University of Southern Denmark, Aarhus University, Vestjyllands Andel and Unibio A/S he reconstructed a genome-scale metabolic model for the methanotroph *Methylococcus capsulatus*. Coordinating with Unibio's specialists, Christian has used this model to consolidate data and explore fermentation strategies. Furthermore, by engaging the COBRA community, Christian and colleagues developed 'memote', a software for quality control of genome-scale metabolic models inspired by common software development practises (<https://memote.io/upload>).

Christian is interested in emerging technologies to harness the potential of C1 feedstocks, high-quality software development for the scientific community, and the standardization and maintenance of genome-scale

metabolic reconstructions. Recently, he has joined the DD-DeCaF project which is focused on building services for data-driven design of cell factories and communities (<http://dd-decaf.eu/>).

Costas Maranas

The Pennsylvania State University

Dr. Costas D. Maranas (b. 1967) is the Donald B. Broughton Professor in the Department of Chemical Engineering at The Pennsylvania State University. He received his Diploma in Chemical Engineering from the Aristotle University, Greece in 1990 and a Ph.D. in Chemical Engineering from Princeton University in 1995. He has been in the faculty of the department of Chemical Engineering at Penn State since 1995. He is the recipient of the Allan P. Colburn Award for Excellence in Publications by a Young Member of AIChE (2002), the Outstanding Young Investigator Award of the Computing and Systems Technology AIChE Division (2006), the S.V. Sotirchos Lectureship at 6th Panhellenic Chemical Engineering Conference (2007), the Penn State Engineering Alumni Society (PSEAS) Premier Research Award (2016) and Outstanding Research Award in (2012). He is a member of a number of journal Editorial Boards including PLOS Computational Biology, BMC Systems Biology, Biotechnology Journal and Metabolic Engineering. He is a Fellow of the American Institute of Medical and Biological Engineering (AIMBE). He is a member of advisory/steering committees for PNNL/EMSL and EcoCyc and the "Use Inspired Research" Lead in the Center for Bioenergy Innovation (CBI) DOE center.

The C. Maranas group develops and deploys computational framework informed by systems engineering and mathematical optimization to understand, analyze and redesign metabolism and proteins. Research interests include: Computational protein design; enzyme and antibody engineering; reconstruction, curation and analysis of metabolic networks; computational strain design and synthetic biology; metabolism of photosynthetic organisms; metabolism of obligatory anaerobes; modeling of microbial communities; optimization theory and algorithms. He has co-authored over 160 refereed journal publications including a textbook on "Optimization Methods in Metabolic Networks" (2016). He has supervised 29 PhD theses with many group alumni occupying leading positions in industry and academia.

INVITED SPEAKER BIOGRAPHIES

Jens Nielsen

Chalmers University of Technology, Sweden

Jens Nielsen has an MSc degree in Chemical Engineering and a PhD degree (1989) in Biochemical Engineering from the Danish Technical University (DTU), and after that established his independent research group and was appointed full Professor there in 1998. He was Fulbright visiting professor at MIT in 1995-1996. At DTU he founded and directed Center for Microbial Biotechnology. In 2008 he was recruited as Professor and Director to Chalmers University of Technology, Sweden, where he is currently directing a research group of more than 50 people. At Chalmers he established the Area of Advance Life Science Engineering, a cross departmental strategic research initiative and was founding Head of the Department of Biology and Biological Engineering, which now encompass more than 170 people.

Jens Nielsen has published so far more than 600 papers that have been cited more than 47,000 times (current H-factor 106), co-authored more than 40 books and he is inventor of more than 50 patents. He was identified by Thompson Reuter as a highly cited researcher in 2015 and 2016.

Jens Nielsen founded Fluxome A/S that raised more than M20EUR in venture capital. This company metabolically engineered yeast for production of resveratrol and used this yeast for commercial production of this compound. This process was acquired by the company Evolva. Jens Nielsen has founded several other biotech companies, including Metabogen AB and Biopetrolia AB, and he has served in the scientific advisory board of a range of different biotech companies in the USA and Europe.

Jens Nielsen has received numerous Danish and international awards including the Villum Kann Rasmussen's Årslegat, Merck Award for Metabolic Engineering, Amgen Award for Biochemical Engineering, Nature Mentor Award, the Gaden Award, the Norblad-Ekstrand gold medal, the Novozymes Prize, the ENI Award and the Eric and Sheila Samson Prize. He is member of several academies, including the National Academy of Engineering in USA, the Royal Swedish Academy of Science, the Royal Danish Academy of Science and Letters, the Royal Swedish Academy of Engineering Sciences and the American Academy of Microbiology. He is a founding president of the International Metabolic Engineering Society.

Bernhard Palsson

Novo Nordisk Foundation Center for Biosustainability

Bernhard Palsson is the Galletti Professor of Bioengineering, the Principal Investigator of the Systems Biology Research Group in the Department of Bioengineering, and Professor of Pediatrics at the University of California, San Diego. He is also the CEO of the Novo Nordisk Center for Biosustainability in Denmark, working in this capacity since 2011. Dr. Palsson has co-authored more than 490 peer-reviewed research articles and has authored four textbooks. His research includes the development of methods to analyze metabolic dynamics (flux-balance analysis, and modal analysis), and the formulation of complete models of selected cells (the red blood cell, E. coli, CHO cells, and several human pathogens). He sits on the editorial board of several leading peer-reviewed microbiology, bioengineering, and biotechnology journals. He previously held a faculty position at the University of Michigan for 11 years and was named the G.G. Brown Associate Professor at Michigan in 1989, a Fulbright fellow in 1995, and an Ib Henriksen Fellow in 1996. He is the author of over 40 U.S. patents, the co-founder of several biotechnology companies, and holds several major biotechnology awards. He received his PhD in Chemical Engineering from the University of Wisconsin, Madison. Dr. Palsson is a member of the National Academy of Engineering and is a Fellow of the AIMBE, AAAS, and the AAM.

Jennifer Reed

University of Wisconsin-Madison

Jennifer Reed is a Professor at the University of Wisconsin-Madison. Jennifer obtained her B.S., M.S., and Ph.D. in bioengineering from the University of California, San Diego. Her research interests involve building, analyzing, and utilizing metabolic and regulatory models of organisms involved in bioremediation, biofuels, and pharmaceutical applications. Once developed these models can be used to evaluate the capabilities of different organisms from a network-based perspective and to identify ways in which genetic manipulations could enhance their productivity. She is also interested in developing computational methods for designing strains or cell lines with enhanced production yields of desired products. Models can also be used to identify potential metabolic or regulatory roadblocks that might be limiting production in developed strains.

INVITED SPEAKER BIOGRAPHIES

Joo Sang Lee

Cancer Data Science Lab, NCI/NIH

Joo's research is focused on developing and harnessing data science approaches for the integration of multi-omics data to better understand the pathogenesis of cancer, its evolution and treatment. We collaborate with many experimental labs, aiming to develop and utilize computational approaches to jointly gain an integrative view of the systems we study. Together with our collaborators, we aim to predict and test novel drug targets and biomarkers to treat cancer more effectively. Before joining the NCI, Joo has been a postdoc at Center for Bioinformatics and Computational Biology at the University of Maryland after finishing his PhD at Northwestern University.

Uwe Sauer

ETH Zurich

Uwe Sauer is Professor of Systems Biology in the Department of Biology at the ETH Zurich. He earned his MS and PhD in microbiology from the University of Göttingen. During his postdoc work on metabolic

engineering in the chemical engineering lab of Jay Bailey in Zurich, he became interested in computational modelling of cellular behavior and quantitative analysis of intracellular fluxes in particular. In the late nineties, his research began to focus on quantitative understanding of interactions and regulation within complex microbial metabolic networks by using and developing cutting-edge experimental and computational methods.

For this purpose, research in the Sauer lab takes an interdisciplinary approach that combines quantitative experimentation and modelling to solve fundamental questions of how bacteria and yeasts coordinate their metabolism and how they interact with each other and various hosts. In particular, his lab has pioneered quantitative mass spectrometry-based methods for ¹³C-flux analysis and high-throughput metabolomics, enabling rapid hypothesis generation on the molecular and functional operation of complex metabolic networks.

Uwe has over 200 publications in peer-reviewed journals, and is a member of various international editorial boards, scientific steering and advisory committees of organizations and companies in systems biology and biotechnology.

SUNDAY, OCTOBER 14

SESSION 1: New Frontiers

INVITED SPEAKER

Developing a seamless pipeline from 13C labeling data to kinetic models with a genome-wide coverage

Costas Maranas

Department of Chemical Engineering, The Pennsylvania State University of State, University Park, PA 16802; <http://www.maranasgroup.com/>

Kinetic models of metabolic networks offer the potential of truly predictive models of metabolism. The mechanistic characterization of enzyme catalyzed reactions allows for accurate prediction of perturbations in metabolite concentrations and reaction fluxes in response to genetic and environmental perturbation that are beyond the scope of stoichiometric models. Despite their potential, the application of kinetic models to microbial strain design and metabolic engineering has been limited by (i) the paucity of fluxomic datasets that span the entire metabolism needed for parametrization and (ii) the computational expense associated with kinetic model parameterization.

In this talk, we will highlight a two-step pipeline that promises to alleviate some of the inherent computational challenges. It requires as an input a curated genome-scale metabolic (GSM) model and 13C-labeling distributions for intracellular metabolites under multiple genetic and environmental perturbations. Additional data such as biomass yield, metabolite concentrations, enzyme kinetics and/or fermentation data can be integrated whenever available. The key concept here is to use the same GSM metabolic map for performing both metabolic flux elucidation and kinetic model parameterization. This eliminates any errors due to ad hoc model aggregation/simplification. The first step involves the application of improved algorithms for performing both stationary and instationary MFA using atom mapping model with genome-wide coverage. The second step uses as input the elucidated metabolic flux ranges to estimate kinetic parameters that agree with all fluxomic datasets using the newly developed K-FIT parameterization algorithm. The K-FIT algorithm relies on a combination of nested decomposition schemes and iterative solution techniques to evaluate steady-state fluxes in response to genetic perturbations. It achieves orders of magnitude CPU improvements compared to a genetic algorithm

(GA)-based procedure. The approach has been tested for a core model of *E. coli* containing 109 reactions and 61 metabolites, a large-scale *E. coli* model with 319 reactions and 276 metabolites as well as models for clostridia. Using this workflow, 13C data can be seamlessly translated into parameterized kinetic models.

INVITED SPEAKER

Introducing thermodynamics into ME-models and crowders into kinetic models

Vassily Hatzimanikatis

Laboratory of Computational Systems Biotechnology (LCSB)

Ecole polytechnique fédérale de Lausanne (EPFL)

CH-1015 Lausanne, Switzerland.

vassily.hatzimanikatis@epfl.ch

The field of COBRA has impressively evolved since the first meeting in Iceland 7 years ago. The recent trends are moving towards the introduction of constraints that capture mRNA and protein expression (ME-models), thermodynamics and metabolomics (TFA), and most recently kinetics. While these additional constraints expand the capabilities and the scope of the models, they also introduce computational and modeling challenges. We will discuss here these challenges and we will present some new and computationally efficient problem formulations that allow the introduction of thermodynamic constraints into ME-models, and they also consider the effect of intracellular crowding conditions.

Machine Learning for Uncertainty Reduction in Biochemical Kinetic Models.

Ljubisa Miskovic¹, Jonas Beal², Michael Moret², and Vassily Hatzimanikatis³

(1)Laboratory of Computational Systems Biotechnology (LCSB), EPFL, Lausanne, Switzerland, (2)EPFL, Lausanne, Switzerland, (3)Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland

The primary goal of kinetic models is to capture the systemic properties of the metabolic networks, and we need large-scale kinetic models for reliable *in silico* analyses of the complex dynamic behavior of metabolism. However, parameter uncertainty hinders the development of kinetic models and uncertainty levels increase with the model size. Current methods for building kinetic models within constraint-based modeling frameworks address uncertainty indirectly by integrating data from different biological origins. We recently proposed iSCHRUNK, a

computational approach which combines Monte Carlo sampling methods and machine learning techniques to characterize the uncertainties and to reveal complex relationships between the kinetic parameters and the responses of the metabolic networks. Monte Carlo sampling methods allow us to exploit synergies between different data sources and generate a population of kinetic models that are consistent with the available data and physicochemical laws. The machine learning allows us to data-mine the *a priori* generated kinetic parameters together with the integrated datasets and derive values of kinetic parameters consistent with the observed physiology. In this work, we modified iSCHRUNK to address a design question: can we identify the kinetic parameters and their values that give rise to a desired metabolic behavior? Such information is important for a wide variety of studies ranging from biotechnology to medicine. As an illustration, we used the proposed methodology to find parameters that ensure a rate improvement of the xylose uptake (XTR) in a glucose-xylose co-utilizing *S. cerevisiae* strain. Our results indicate that only three kinetic parameters need to be accurately characterized to reduce the uncertainty, and ultimately increase confidence in the design and control the metabolism desired responses. This framework paves the way for a new generation of methods that will systematically integrate the wealth of omics data and extract the information necessary for metabolic engineering and synthetic biology decisions.

Expanding Metabolic Models to Three Dimensions.

Elizabeth Brunk

Bioengineering, University of California, San Diego, La Jolla, CA

With the increasing coverage of genomic, proteomic and metabolomic data, we are witnessing the birth of a new era in biology- one which will change the face of medical research and clinical practice in unprecedented ways. A significant challenge that impedes further advancement of genomic medicine is understanding how to effectively integrate and translate big biological data into transformative health practices that impact patients lives. Recently, we have integrated protein structural information with the human metabolic network which provides a compelling new angle and scale for studying disease (Brunk et al. Nature Biotechnology 2018). Expanding reconstructions in this way provides new avenues for understanding how biochemical processes relate to mechanisms at the atomic scale. In this talk, I will discuss several recent applications that exemplify a new scientific frontier which has been termed “structural systems biology”. For the first time, relationships between

human metabolic genes, their encoded proteins, and reactions they catalyze can be described in the context of specific 3D configurations, interactions, and properties. Such capabilities open the door to exploring the spatial relationships of cancer mutations, the influence of SNPs on drug or metabolic responses, as well as the mechanistic underpinnings of basic biological processes like protein translation.

Key Decisions for the Development of the Next-Generation of Context-Specific Models.

Anne Richelle, Austin Chiang, Jahir Gutierrez, Chintan Joshi, Benjamin Kellman, Shangzong Li, Joanne Liu, and Nathan Lewis

University of California, San Diego, La Jolla, CA

Genome-scale metabolic models provide a valuable context for analyzing data from diverse high-throughput experimental techniques. Since some enzymes are only active in specific environments, several algorithms have been developed to build context-specific models. However, the content and associated predictive capacity of resulting models is considerably impacted by the different assumptions used at each step of the extraction process: from the decisions on how to overlay data onto networks to the parameter choices for extraction algorithms^{1,2}. These choices lead to a poor consensus in generated models, which may limit the use of context-specific methods for data-driven hypothesis.

We present a comparative analysis of existing extraction algorithms and an assessment of the key decisions influencing the data contextualization methods. From this work, we propose an approach to obtain better consensus across existing extraction algorithms and more accurate models. This approach is enabled with a framework we built for inferring metabolic functions that should be active in a specific context directly from transcriptomic data. These functions can be used to protect associated biochemical reactions during the implementation of extraction methods. The promising results obtained using this protectionist approach underline the potential interest of describing genome-scale metabolic reconstructions as more than a network of reactions but rather as an interconnected map of cellular functionalities.

1. Opdam, S., Richelle, A., Kellman, B., Li, S., Zielinski, D. C., & Lewis, N. E. (2017). A Systematic Evaluation of Methods for Tailoring Genome-Scale Metabolic Models. *Cell Systems*, 4(3). <http://doi.org/10.1016/j.cels.2017.01.010>

2. Richelle, A., Joshi, C., & Lewis, N. E. (2018). Assessing key decisions for transcriptomic data integration in biochemical networks. *BioRxiv*, 301945. <http://doi.org/10.1101/301945>

INVITED SPEAKER

Metabolic Coordination through Metabolite-Protein Interactions.

Uwe Sauer

Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland

How do bacteria know what goes on in their environment and how to they make appropriate decisions? While some bona fide extracellular sensors are known, there are far more environmental conditions and cellular responses than could possibly be dealt with through dedicated sensors. Instead, most microbial responses are based on intracellular changes to environmental changes. One of the first affected networks to just about any extracellular change is metabolism that passively responds to nutritional or chemical/physical challenges. Since fluxes and intracellular metabolite levels respond within seconds, allosteric binding of metabolites to regulatory proteins and enzymes is a highly effective and rapid sensing mechanism. Different from well-establish methods to assess physical interaction between proteins and between proteins and nucleic acids, however, methods to assess metabolite-protein interactions are still in its infancy. In this talk I will focus first on reconstructing the regulatory network of metabolite-enzyme interactions from databases, which currently consists of 1500 unique regulatory interactions (1). I will then describe out efforts to experimentally map this network out further. The current results indicate that the known interactions are only the tip of the iceberg (2). Beyond knowing the interaction topology, I will conclude this talk on the even more challenging and conceptual problem: understanding which of the many regulation mechanisms actually matter for a given adaptation to elicit an appropriate physiological system response.

1. Reznik, Christodoulou, Goldford, Briars, Sauer, Segre & Noor. *Cell Reports* 20: 2666-2677 (2017).

2. Piazza, Kochanowski, Cappelletti, Fuhrer, Noor, Sauer & Picotti. *Cell* 72:358-372 (2018).

INVITED SPEAKER

An Upper Limit in Gibbs Energy Dissipation Governs Cellular Metabolism.

Matthias Heinemann, Bastian Niebel, and Simeon Leupold

Molecular Systems Biology, University of Groningen, Groningen, Netherlands

The principles governing cellular metabolic operation are still poorly understood. Because very diverse organisms show relatively comparable physiologies, we hypothesized that a fundamental thermodynamic constraint might govern cellular metabolism. To investigate this, we developed a novel constraint-based model for *Saccharomyces cerevisiae* with a comprehensive description of the biochemical thermodynamics and including a Gibbs energy balance. Nonlinear regression analyses of metabolome and physiology data revealed the existence of an upper rate limit for the cellular Gibbs energy dissipation. Applying this limit in flux balance analyses using growth maximization as objective, the model correctly predicted physiologies, intracellular metabolic fluxes, as well as even some metabolite levels, for different glucose uptake rates. Our work indicates that cells arrange their metabolic fluxes such that with increasing substrate uptake rates, an optimal growth rate is accomplished, but the critical rate limit in Gibbs energy dissipation is never exceeded. Once all possibilities for further intracellular flux redistribution towards 'saving' Gibbs energy dissipation are exhausted, cells reach their maximal growth rate. We found that this principle also holds for *Escherichia coli* and on different carbon sources. Our work suggests that metabolic reaction stoichiometry, a limit in the cellular Gibbs energy dissipation rate, and the growth maximization objective might shape metabolism across different organisms and conditions.

Compass: FBA-Powered Modeling of Single-Cell RNA-Seq Characterizes Cell-to-Cell Metabolic Heterogeneity and Reveals Novel Therapeutic Targets in Autoimmunity.

Allon Wagner^{1,2}, David Detomoso², Chao Wang³, Johannes Fessler³, Arman Koul³, Aviv Regev^{4,5}, Vijay K. Kuchroo^{3,4}, and Nir Yosef^{1,2}

(1)Electrical Engineering and Computer Science, University of California, Berkeley, Berkeley, CA, (2)Center for Computational Biology, University of California, Berkeley, Berkeley, CA, (3)Evergrande center for immunologic diseases, Harvard Medical School and Brigham and Women's Hospital, Boston, MA, (4)Broad Institute of MIT and Harvard, Boston, MA, (5)Howard Hughes Medical Institute and the Department of Biology, Massachusetts Institute of Technology, Boston, MA

The rapid advance in single-cell RNA-Sequencing (scRNA-Seq) is one of the most exciting recent developments in biomedical research. By comprehensively quantifying the transcriptome of individual cells, rather than the average over many cells, scRNA-Seq allows, for example, to discover rare cell sub-types and to retrace intermediate steps in lineage development. However, scRNA-Seq requires tailor-made computational methods that take

advantage of its novel properties while accounting for its unique technical limitations (e.g., dropouts), and its immense data magnitude, which is fast approaching millions of cells per experiment (Wagner et al., Nature Biotechnology 2016). Here, we answer these challenges in the realm of cellular metabolism. We present COMPASS, an FBA algorithm for comprehensive characterization of single-cell metabolic states. COMPASS capitalizes on the magnitude of scRNA-Seq data by modeling cells as points in a high-dimensional metabolic space, while mitigating scRNA-Seq data sparsity by smoothing kNN neighborhoods. The resulting metabolic space is directly interpretable in mechanistic terms. It allows data-driven characterization of metabolic heterogeneity among cells and associating metabolic programs with phenotypes of interest. To demonstrate COMPASS, we study T helper 17 (Th17) cells, which are remarkably plastic and mediate both autoimmunity and immune tolerance. A COMPASS characterization of Th17 metabolic heterogeneity recovered known associations of Th17 effector states with metabolic programs and predicted novel metabolic intervention targets. We show that the glycolytic shift in pathogenic Th17 is associated with extensive remodeling of anabolic pathways, which depends on pyruvate dehydrogenase (PDH). We then implicate the polyamine pathway, which has been scarcely studied in autoimmunity, in Th17 pathogenicity. *In vivo* deletion of either PDH or polyamine synthesis led to considerably better clinical outcome in a murine model of multiple sclerosis (MS). The clinical significance of our findings is supported by differential abundance of polyamines in the blood of healthy human donors compared with MS patients.

Metabolic Model-Based Evaluation of Microbiome-Metabolome Association Studies.

Cecilia Noecker¹, Hsuan-Chao Chiu^{1,2}, Colin McNally¹, and Elhanan Borenstein^{1,3,4}

(1)Department of Genome Sciences, University of Washington, Seattle, WA, (2)Office of Chief Technology Officer, MediaTek, Hsinchu City, Taiwan, (3)Department of Computer Science and Engineering, University of Washington, Seattle, WA, (4)Santa Fe Institute, Santa Fe, NM

Identifying specific microbial drivers of variation in metabolic phenotypes is a major goal in the study of host-associated and environmental microbiomes. Correlation-based analysis of paired microbiome-metabolome datasets is a widespread approach to this objective. To date, however, the efficacy and limitations of this approach have not been evaluated. To address this challenge, we developed a mathematical definition of the contribution of each taxon to metabolite variation

based on its uptake and secretion fluxes. We applied a multi-species dynamic genome-scale metabolic modeling pipeline to simulate simplified gut communities, generating idealized microbiome-metabolome datasets with precisely known microbial metabolic fluxes. Comparing the observed taxon-metabolite correlations in this simulated setting with the true taxonomic contributors, we found that correlation-based analysis poorly predicts key contributors, with low accuracy and a high false discovery rate. Importantly, however, the predictive value was strongly influenced by both metabolite and taxon properties, as well as exogenous environmental variation. These findings have practical implications for the analysis and interpretation of microbiome-metabolome studies. This study illustrates the utility of multi-scale metabolic modeling as a simulation tool to inform study design and methods development in microbiome research.

Yeast-GEM: Reviving the Consensus Genome-Scale Model of *S. Cerevisiae* As a Standard in the Community.

Benjamín J. Sánchez, Hongzhong Lu, Feiran Li, Iván Domenzain, Eduard J Kerkhoven, and Jens Nielsen
Department of Biology and Biological Engineering,
Chalmers University of Technology, Gothenburg, Sweden

Genome-scale models (GEMs) are essential tools for understanding and computing metabolism. However, as they are so large (thousands of metabolites and reactions), it becomes challenging to keep track of changes between two versions of the same model. Furthermore, a framework is needed for researchers to work on a single model at the same time, without creating conflicts.

Here, we have brought back to life the GEM for *Saccharomyces cerevisiae* as yeast-GEM: a version controlled model publicly available for open collaboration. Changes are tracked with Git, can be easily visualized on Github, and are linked to complementary data and scripts, so anyone can know what was changed, who did it and why. By hosting the model in Github we also allow model developers to work on parallel, we periodically release new versions of the model, and we publicly display current and future work, organized in different projects. Additionally, we take advantage of memote, the metabolic model test suite, to display the quality of our model as we work on it.

We will present numerous model improvements that we have performed using this framework, such as an extensive curation to metabolites, reactions and genes based on new genome annotation; the introduction of SLIME reactions, which Split Lipids Into Measurable Entities for properly representing limitations on both the

lipid classes and at the same time the acyl chains; and a growth validation study under different carbon/nitrogen/sulphur/phosphate sources.

Finally, we will show how we connect this model to GECKO, a toolbox for adding enzyme constraints to a GEM, and achieve a continuous synchrony of the metabolic model (yeast-GEM) and the enzyme-constrained model (ecYeast-GEM), to offer to the community a set of ready to use constrained-based models of yeast. See more of yeast-GEM, and contribute if you wish, at <https://github.com/SysBioChalmers/yeast-GEM>

MONDAY, OCTOBER 15

SESSION 2: Applications in Medicine

INVITED SPEAKER

Urea cycle dysregulation, emerging pyrimidines mutation bias and enhanced response to immunotherapy in cancer

Eytan Rupp, Joo Sang Lee

Cancer Data Science Lab, NCI, NIH

The urea cycle (UC) is the main metabolic pathway by which mammals dispose waste nitrogen. Here we show that the expression of most UC enzymes is altered in many tumors, leading to a general metabolic hallmark that we term UC dysregulation (UCD). UCD elicits nitrogen diversion towards carbamoyl-phosphate synthetase2, aspartate transcarbamylase and dihydroorotase (CAD) activation and enhances pyrimidine synthesis, resulting in detectable changes in nitrogen metabolites in both patient tumors and bio-fluids. The accompanying excess of pyrimidine vs purine nucleotides results in a novel genomic signature consisting of transversion mutations at the DNA, RNA and protein levels. This mutational bias is associated with more hydrophobic tumor antigens and with a better response to immune checkpoint inhibitors, independent of mutational load. Taken together, our findings demonstrate that UCD is a common feature of tumors, which profoundly impacts carcinogenesis, mutagenesis and immunotherapy response.

[Joint work with the Erez lab at the Weizmann Institute, Israel]

Modeling Physiological Responses to Stress: Metabolism of Fasting and Acetaminophen-Induced Liver Damage in the Laboratory Rat.

Venkat Pannala¹, Martha Wall², Shanea Estes³, Trenary Irina², O'Brien Tracy³, Printz Richard³, Vinnakota Kalyan¹, Jaques Reifman¹, Masakazu Shiota³, Jamey Young², and Anders Wallqvist¹

(1)Department of Defense Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick, MD, (2)Department of Chemical and Biomolecular Engineering, Vanderbilt University School of Engineering, Nashville, TN, (3)Department of Molecular Physiology and Biophysics, Vanderbilt University School of Medicine, Nashville, TN

Metabolism is an integral component of our response to physiological stress. We need to convert nutrients to energy, make the requisite macromolecular components, and use these components, to function normally and handle varying physiological stress. Capturing these physiological responses using metabolic network modeling is instrumental for obtaining scientific insights from transcriptomic and metabolomic data. Here, we studied coupled liver and kidney metabolism in the laboratory rat under mild stress conditions, i.e., during short-term fasting and during the initial period of acute acetaminophen toxicity before the onset of cellular damage. We used a genome-scale network reconstruction of rat liver and kidney metabolism, together with a modeling approach constrained by stress-induced gene expression data, central carbon fluxes derived from isotope tracer techniques, and physiological flux bounds. We gauged the performance of the approach by comparing metabolite levels secreted in plasma and urine predicted by the model with those obtained from non-targeted metabolic profiling. In the case of short-term fasting, we correctly accounted for glucose/glycogen metabolism in the liver, using a modeling approach based on the metabolic-network structure coupled only with measured central carbon fluxes. For acetaminophen—a known hepatotoxicant—we further incorporated changes in liver and kidney gene expression into multiple modeling frameworks to study the resultant changes in endogenous metabolism. A comparison of the modeled results with the global metabolic profiling data revealed that, our approach satisfactorily predicted altered plasma metabolite levels as early as 5 h after exposure to 2 g/kg of acetaminophen, and that, after 10 h of exposure the predictions significantly improved. We achieved a 70% correspondence between predicted and measured changes of metabolite levels in plasma and

urine. The coupled multi-tissue modeling framework of *in vivo* metabolism provides both mechanistic insights and a capability to identify plasma and urine metabolites as early markers of toxicant-induced organ damage.

Differential Metabolic Functionality in Antibiotic-Resistant *Pseudomonas aeruginosa* Revealed with an Integrated Computational and Experimental Approach.

Laura J. Dunphy¹, Phillip Yen², and Jason Papin³

(1)Biomedical Engineering, University of Virginia, Charlottesville, VA, (2)University of Virginia, (3)Department of Biomedical Engineering, University of Virginia, Charlottesville, VA

Changes in bacterial metabolism accompanying the development of antibiotic resistance remain poorly understood. In this study, we performed a single-carbon source utilization screen on lab-evolved antibiotic-resistant *Pseudomonas aeruginosa* to investigate these changes. The metabolic capabilities of piperacillin-resistant, tobramycin-resistant, and ciprofloxacin-resistant *P. aeruginosa* as well as paired ancestral and media-evolved control lineages were evaluated by measuring growth curves on 190 unique carbon sources. Our resulting 950 growth curves revealed that resistant lineages exhibit mainly decreased catabolic function with occasional gains of function compared to antibiotic-sensitive *P. aeruginosa*. In resistant lineages we also observed changes in growth dynamics, including growth rate and time to reach mid-exponential phase. A genome-scale metabolic network reconstruction of *P. aeruginosa* strain UCBPP-PA14, iPau1129, was used to contextualize whole-genome sequencing data of the resistant lineages. The model was used to predict the impact of resistance mutations on loss of catabolic function. For example, five genes deleted in the piperacillin-resistant lineage were predicted to drive loss of the ability to utilize L-leucine. Model predictions were experimentally validated with a transposon mutant library. Our results show that metabolism is altered through the evolution of antibiotic resistance. Instances where our model failed to correctly predict genotype-phenotype relationships highlight gaps in our current understanding of *P. aeruginosa* metabolism. Our combined computational and experimental framework can be applied to identify metabolic limitations in other antibiotic-resistant pathogens. Drug-driven metabolic limitations have the potential to be targeted to select against antibiotic-resistant populations.

Model-Based Prediction of Functional SNPs Suggests Factors for Metabolic Diversity and Drug Resistance across Human-Associated *Mycobacterium Tuberculosis*.

Ove Øyås¹, Sonia Borrell^{2,3}, Andrej Trauner^{2,3}, Michael Zimmermann⁴, Sébastien Gagneux^{2,3}, Jörg Stelling¹, Uwe Sauer⁴, and Mattia Zampieri⁴

(1)D-BSSE, ETH Zurich, Basel, Switzerland, (2)Swiss TPH, Basel, Switzerland, (3)University of Basel, Basel, Switzerland, (4)Institute of Molecular Systems Biology, ETH Zurich, Zurich, Switzerland

The *Mycobacterium tuberculosis* complex (MTBC) is a group of closely related pathogenic bacteria that can cause tuberculosis (TB) upon infection. TB is a leading cause of human mortality worldwide and threatens to remain so for the foreseeable future due to the emergence of multidrug-resistant MTBC strains. A characteristic feature of the MTBC is that it harbors comparatively little genetic variation, but a large share of its single-nucleotide polymorphisms (SNPs) are non-synonymous and experimental evidence is challenging the long-held assumption that phenotypic diversity in the MTBC is negligible. Bridging the genotype-phenotype gap in the MTBC could therefore provide valuable insights into the evolution of phenotypes such as drug resistance.

To address this problem, we built a single constraint-based model that integrates the exometabolomes and genomes of 18 MTBC strains from six lineages native to different parts of the world with strain-specific genome-scale metabolic models. We used the model to predict the metabolic effects of non-synonymous SNPs in enzyme-encoding genes via a three-step optimization approach, aiming to explain as much of the observed metabolic variation as possible by as few SNPs as possible while obtaining consistent flux distributions. Using our predicted SNP effects, we classified 88 SNPs (15%) as functional. These functional SNPs affect 67 unique enzymes across most metabolic pathways and include a SNP in pyruvate kinase previously shown to be functional in *Mycobacterium bovis*. In addition, we predicted three functional SNPs in enzymes involved in folate metabolism and we suggest a possible explanation for differential sensitivity to *para*-aminosalicylic acid, one of the antibiotics currently used to treat multidrug-resistant TB. Concluding, our method is capable of predicting the metabolic effects of genetic variation in microbes and allowed us to connect genetic and metabolic diversity in the MTBC.

Synthetic Lethality in Cancer Research Via Genetic Minimal Cut Sets.

Iñigo Apaolaza¹, Edurne San Jose², Luis Vitores Valcárcel^{1,2}, Xabier Agirre², Felipe Prosper², and **Francisco J. Planes¹**

(1)TECNUN, University of Navarra, San Sebastian, Spain, (2)CIMA, University of Navarra, Pamplona, Spain

Synthetic lethality is a promising approach in precision medicine and cancer as it largely expands the number of possible drug targets and creates an opportunity for selectivity. The increasing evidence of metabolic reprogramming of cancer cells makes it ideal to exploit the concept of synthetic lethality. A number of in-silico tools have been developed to target cancer metabolism using a synthetic lethality approach. In particular, constraint-based modeling (CBM) for genome-scale metabolic networks has received much attention. Here, we present a novel CBM approach to synthetic lethality that is based on the concept of genetic Minimal Cut Sets (gMCSs). With respect to existing methods, our approach avoids the step of network contextualization and integrates -omics data in a more natural and objective manner. In addition, it enables not only the detection metabolic targets but also response biomarkers for them. To illustrate our approach, we first show the results of an experimental proof-of-concept in multiple myeloma (MM), where we validated the therapeutic potential of RRM1 inhibition in different MM cell lines. We also predicted a metabolic signature based on gene expression data that explained the response to RRM1 inhibition in different cancer cell lines. Second, we show preliminary results of a study where response biomarkers for the effectiveness of Methotrexate in different cancer types is predicted following our methodology. This new algorithm, freely available in the COBRA Toolbox, undoubtedly opens new avenues to develop precision medicine strategies in complex and unaddressed clinical questions involving heterogeneous molecular data.

INVITED SPEAKER

**COBRA Young Investigator Award Winner:
Pathogen Metabolism and Antibiotic Resistance**

Sriram Chandrasekaran
University of Michigan

A major focus of my lab is to apply systems approaches to tackle antibiotic resistance by understanding pathogen metabolism. Antibiotics need to be effective in diverse environments in vivo. However, the pathogen

microenvironment can have a significant impact on antibiotic potency. To exhaustively explore the impact of diverse metabolic environments, we developed a computational framework - Metabolism And GENomics-based Tailoring of Antibiotics (MAGENTA). We uncovered antibiotic regimens with robust efficacy across distinct environments against both *E. coli* and *A. baumannii*, which we confirmed experimentally. Designing effective antibiotics requires understanding of both the pathogen and the host immune response. Hence, we recently used genome-scale metabolic modeling to uncover the unique multi-pronged mechanism-of-action of an anti-microbial immune protease produced by T-cells on *E. coli*, *M. tuberculosis* and *Listeria monocytogens*; this strategy could be used as a template for designing new therapies.

Identifying and Targeting Key Cellular Mechanisms for Proliferation in Malaria Parasites: A Combined Experimental and Computational Strategy.

Anush Chiappino-Pepe¹, Vikash Pandey¹, Ellen Bushell², Rebecca Limenitakis³, Julian C Rayner², Volker Heussler³, Oliver Billker², and Vassily Hatzimanikatis¹

(1)Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, (2)Wellcome Trust Sanger Institute, Cambridge, United Kingdom, (3)Institute of Cell Biology, University of Bern, Bern, Switzerland

Recent advances in cell genome editing techniques enable the generation of high-throughput gene knockout data in the malaria parasites *in vivo*. Integrative analysis of this data can lead to the identification of biological mechanisms that explain the observed phenotypes and that provide testable hypotheses for further discoveries. Metabolic modeling can cope with the tangled metabolism of the malaria parasites, and hence is a compelling approach for understanding the parasites physiology.

In this study, we present a combined experimental and computational approach that suggests cellular mechanisms for targeting the malaria parasites. We predict *in silico* and test *in vivo* lethal knockouts and synthetic lethal pairs in the blood and liver stages of the malaria infection. We perform computational analyses on a newly developed genome-scale model of the malaria parasite *Plasmodium berghei* (iPbe), and we use high-throughput gene knockout data generated in the PlasmogEM project. The comparison between data and gene essentiality predictions allow the understanding of the parasite's physiology in the blood and liver stages. We identify the thermodynamic bottlenecks, genetic interactions, and the accessibility to nutrients behind the

phenotypes. When we simulate in iPbe the hypothesized physiology, we achieve 80% consistency between the prediction of essential genes and the experimental data. This result indicates that our model iPbe is a valuable framework for the generation of testable hypothesis on antimalarial targets. Overall, the knowledge generated in this framework will serve to tackle more efficiently the malaria parasites' metabolism during infection.

Inferring Metabolic Mechanisms of Interaction within a Defined Gut Microbiota.

Gregory L. Medlock¹, Maureen A. Carey², Dennis McDuffie¹, Michael Mundy³, Natasa Giallourou⁴, Jonathan Swann⁴, Glynis Kolling¹, and Jason A. Papin¹

(1)Department of Biomedical Engineering, University of Virginia, Charlottesville, VA, (2)Microbiology, Immunology, and Cancer Biology, University of Virginia, Charlottesville, VA, (3)Center for Individualized Medicine, Mayo Clinic, Rochester, MN, (4)Division of Computational and Systems Medicine, Imperial College London, London, United Kingdom

Within the gastrointestinal tract of mammals, the diversity and number of species present enable a multitude of metabolic interactions. Methods for predicting these interactions are attractive because they may enable engineering of microbiome function. However, identifying the mechanism and consequences of metabolic interactions between even two species is incredibly challenging. In this work, we developed, applied, and experimentally tested a framework for identifying potential metabolic mechanisms associated with interspecies interactions.

We performed pairwise co-culture growth experiments using bacterial species a model mouse microbiota. After 72 hours of growth, we quantified the abundance of each strain and determined metabolite consumption and production using untargeted supernatant metabolomics. We applied our framework, which we call the Constant Yield Expectation (ConYE) model, to dissect emergent metabolic behaviors that occur in co-culture.

The ConYE model assumes the yield of a metabolite is constant in monoculture and co-culture. Using ConYE, we identified widespread indications of increased efficiency of biomass production in co-culture. We interrogated an amino acid crossfeeding interaction that is likely to confer a growth benefit to one ASF strain (*Clostridium* sp. ASF356) in coculture with another strain (*Parabacteroides goldsteinii* ASF519). We attempted *in silico* identification of growth-enhancing metabolites by constructing genome-scale metabolic network reconstructions for the ASF strains. Using the simulations together with the ConYE results, we designed media supplementation

experiments and verified that the proposed interaction leads to a growth benefit for this strain.

Our results reveal the types and extent of emergent metabolic behavior in microbial communities and demonstrate how metabolomic data can be used to identify potential metabolic. Although we focus on growth-modulating interactions, the framework we develop can be applied to generate specific hypotheses about mechanisms of interspecies interaction involved in any phenotype of interest within a microbial community.

Understanding the Role of Bile Acids in Alzheimer's Disease.

Priyanka Baloni¹, Cory Funk¹, Evangelos Simeonidis¹, Jingwen Yan², Kwangsik Nho², Matthias Arnold³, Gabi Kastenmueller⁴, Gregory Louie⁵, Alexandra Kueider-Paisley⁵, Andrew J. Saykin², Rima Kaddurah-Daouk⁵, and Nathan D. Price¹

(1)Hood-Price Lab, Institute for Systems Biology, Seattle, WA, (2)Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN, (3)Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany, (4)Helmholtz Zentrum München, German Research Center for Environmental Health, Institute of Bioinformatics and Systems Biology, Neuherberg, Germany, (5)Department of Psychiatry and Behavioral Medicine, Duke Institute for Brain Sciences, Durham, NC

Alzheimer's disease (AD) is the leading cause of dementia. AD-related neurodegeneration has been correlated with metabolic dysfunction. An integrative analysis of transcriptomic, proteomic and metabolomic data from the brains of post-mortem AD patients and controls has provided evidence of genes and metabolites involved in altered metabolic pathways in AD. We have built context-specific draft metabolic networks for multiple brain regions, using high throughput transcriptome data of post-mortem brain samples from AD patients and controls, and integrated them with Recon3D, the most comprehensive and recent human metabolic reconstruction. We previously identified a role for circulating bile acids in AD, as well as upstream changes in cholesterol metabolism. Primary bile acids are synthesized in the liver, whereas secondary bile acids are typically produced by bacteria in the gut. Increased levels of secondary bile acids and ratios to their primary bile acid educts have been linked to AD and cognitive decline. Expression analysis shows that genes involved in alternative bile acid synthesis pathways are expressed in the brain, while genes in the classical pathway are not, thus providing insights into bile acid

metabolism in the brain. By integrating transcriptomics data and metabolomics measurements of bile acids from brain tissue of AD patients and controls within these metabolic networks, our models can capture *in silico* changes in these pathways, highlighting the role of brain bile acid metabolism in AD pathophysiology. Our brain-tissue metabolic models can be exploited to capture *in silico* changes and possibly identify metabolic biomarkers prior to disease manifestation, making them useful for identifying at-risk individuals. Using the information of brain-specific transcriptional regulatory networks, we aim to identify key genes involved in regulating metabolic changes. Our approach combines multi-omics information and provides insights into interactions of primary and secondary bile acids and their possible connection to cognitive decline and AD.

SESSION 3: Multi-Scale Modeling

INVITED SPEAKER

Towards a more complete view of a cell's functions through genome scale models of metabolism, protein synthesis, post-translational modification, and secretion

Nathan Lewis

University of California San Diego

In mammalian cells, metabolism is a core process driving homeostasis, but variations in other cell processes largely define cell type identity and cell-type specific functions. The profile of secreted and membrane proteins show substantial cell-type specificity and drive many tissue specific functions. These proteins, encoded by up to 1/3 of mammalian protein-coding genes, include hormones, membrane proteins, and enzymes that modulate the extracellular space, and these are synthesized and trafficked through the secretory pathway. The pathway complexity, however, obscures its impact on the secretion of different proteins. Unraveling its impact on diverse proteins is particularly important since the pathway is implicated in many diseases and harnessed for biopharmaceutical production. We have delineated the core secretory pathway functions and integrated them with genome-scale metabolic models of human, mouse, and Chinese hamster ovary cells. Here I will discuss our efforts to use this conceptual expansion of constraint-based metabolic modeling to obtain insights into bioenergetic demands imposed by protein synthesis and secretion, and how these models can be deployed in efforts to engineer mammalian cells for enhanced secretion of high-value biologic drugs. Thus, this

work represents a knowledge-base of the mammalian secretory pathway that serves as a novel tool for systems biotechnology.

Deciphering the Mechanisms Underlying Nutrient Exchange in Insect-Microbe Symbioses.

Nana Ankrah, Cindy Wu, and Angela Douglas
Cornell University, Ithaca, NY

Many insects of medical and agricultural importance depend on symbiotic associations with nutrient producing bacteria. Very commonly, net nutrient production by the bacteria is dictated by metabolic competition and cooperation among multiple bacterial taxa and strongly influenced by the metabolites derived from the insect host, but the mechanistic details of the multi-way metabolic interactions are poorly understood. To investigate the processes shaping insect-microbe metabolic interactions and to derive estimates of the composition and amount of nutrients exchanged between insect and microbe, we reconstructed genome and transcriptome-informed metabolic models of sap-feeding insects that harbor intracellular bacterial partners and the bacterial community in the *Drosophila* gut. Our simulations reveal that among intracellular bacterial partners the insect host controls bacterial production of nutrients, by precise controls over the concentrations of substrate metabolites in essential nutrient biosynthetic pathways. The net result is the structuring of the metabolic networks of the bacteria, to promote cooperative cross-feeding of metabolites and minimize competition for host-derived substrates. Among gut microbes competitive interactions dominate and the quality and quantity of nutrients available to the host is influenced by the composition of gut microbiota which synthesize and consume host and microbe derived nutrients. Our studies elucidate the metabolic basis of insect-microbe interactions and provide the methodology to identify specific gene targets for novel pest control strategies and to select the optimal microbial consortia for personalized microbial therapies in treating animal diseases associated with microbial dysbiosis.

Incorporation of Dynamic pH into an Integrative Model of a Nitrification Microcosm Co-Culture of Nitrosomonas Europaea and Nitrobacter Winogradskyi.

Frank Chaplen¹ and Brett Mellbye²

(1)Biological & Ecological Engineering, Oregon State University, Corvallis, OR, (2)Oregon State University, Oregon State University, Corvallis, OR

Understanding coupled biogeochemical systems is important for elucidating anthropogenic environmental impacts. Nitrifying bacteria are fundamental elements of the nitrogen (N) Cycle. N gases production by soils through microbial action can be a significant source of atmospheric NO_x and N₂O. Multi-level genome-scale modeling of *N. europaea* and *N. winogradskyi* in a microcosm suggests a key role for NO in the abiotic and biotic chemistry of nitrifying systems[1]. Further refinement of the integrative model to incorporate dynamic pH profiling resulted in additional insights into the physiology of *N. europaea*. Matching the dynamic pH profile of the microcosm required the use of the Euclidean norm for flux distribution minimization during optimization of the genome-scale model component of the integrative model. The requirement for the Euclidean norm suggested that *N. europaea* in the microcosm operates in a manner that optimizes metabolite channeling[2]. Modeling with the enhanced integrative model also suggested that *N. europaea* facilitates conversion of ammonium ion to ammonia. A possible mechanism for ammonium conversion is uptake by ammonium transporter (*amtB*; NE0448) or active membrane-potential-driven mechanisms into a more basic cytoplasm[3].

[1] Mellbye, B.L., Giguere, A., Murthy, G.S., Bottomley, P.J. Sayavedra-Soto, L.A. and Chaplen F. (2018) Genome-scale, constraints-based modeling of nitrogen oxide fluxes during co-culture of *Nitrosomonas europaea* and *Nitrobacter winogradskyi*. *mSystems* 3(3): e00170-17. (Editor's Pick) Retrieved from <http://dx.doi.org/10.1128/mSystems.00170-17>

[2] Wu, H.-Q., Cheng, M.-L., Lai, J.-M., Wu, H.-H., Chen, M.-C., Liu, W.-H., Wu, W.-H., Chang, P., M.-H., Huang, C.-Y. H., Tsou, A.-P., Shao, M.-S., and Wang, F.-S. (2017) Flux balance analysis predicts Warburg-like effects of mouse hepatocyte deficient in miR-122a. *PLOS Comput. Biol.* 13(7):e1005618. <https://doi.org/10.1371/journal.pcbi.1005618>.

[3] Cherif-Zahar, B., Durand, A., Schmidt, I., Hamdaoui, N., Matic, I., Merrick, M., and Matassi, G. (2007) Evolution and functional characterization of the RH50 gene from ammonia-oxidizing bacterium *Nitrosomonas europaea*. *J. Bacteriol.* 189(24):9090-2100. Doi:10.1128/JB.01089-07

Dynamic Side of the Warburg Effect: Glycolytic Intermediate Storage in Tumor Cells to Buffer Fluctuating Glucose and O₂ Supply.

Johannes H.G.M. van Beek

Clinical Genetics, VU University medical center, Amsterdam, Netherlands; Experimental Vascular Medicine, Amsterdam University Medical Centers, location AMC, Amsterdam, Netherlands

Tumor cells show the Warburg effect: high glucose uptake and lactate production despite sufficient oxygen supply. A multi-scale computational model of tumor cell

metabolism, intracellular metabolite storage and transport in tissue by blood flow and diffusion is developed here. This model is used to investigate the dynamic behavior of the metabolic system underlying the Warburg glycolytic phenotype. Remarkably, experiments have shown that the ascites tumor cells studied by Otto Warburg can transiently take up glucose an order of magnitude faster than the already high glycolytic rate that he measured for hours. This transiently very high glucose uptake is investigated here with a computational model that reproduces kinetic experiments on these ascites tumor cells. Model analysis of the data shows that the head section of the glycolytic chain in the tumor cells is partially inhibited in about a minute when substantial amounts of glucose have been taken up intracellularly. This head section of the glycolytic chain is subsequently disinhibited slowly when concentrations of glycolytic intermediates fall to low levels again. Based on these dynamic characteristics, simulations of tissue with fluctuating O₂ and glucose supply predict that tumor cells greedily take up glucose when this periodically becomes available at low concentrations, leaving very little for other cells. The glucose is stored as fructose 1,6-bisphosphate and other glycolytic intermediates, which are used for ATP production during O₂ and glucose shortages that last several minutes. This dynamic energy buffer capacity may have selective advantages for tumor cells experiencing cycling hypoxia and nutrient shortages, which are often found in cancer tissue. The hypothesis is put forward here that dynamic regulation of the powerful glycolytic enzyme system in tumor cells is used to buffer oxygen and nutrient fluctuations in tissue. The stored glycolytic intermediates can support the ATP requirements of the tumor cells for 2-10 minutes.

INVITED SPEAKER

New Directions for Cobra.

Bernhard O. Palsson

Bioengineering, University of California, San Diego, La Jolla, CA

Calls for effective multi-omic data integration are growing, and with it, the need for explanatory-AI (or xAI) has arisen. COBRA methods provide a structured and fundamental basis for data analysis and interpretation and should offer opportunities for xAI broadly for the life sciences. Three new directions for COBRA will be discussed: 1. BD2K; 2. Structural proteomics as an additional omics data type; and 3. The pangenome goldmine.

Toward a Proteome-Complete Model of the Human Red Blood Cell.

James T. Yurkovich, Laurence Yang, and Bernhard O. Palsson

Bioengineering, University of California, San Diego, La Jolla, CA

The human red blood cell has served as a starting point for the application and development of systems biology approaches due to its simplicity, intrinsic experimental accessibility, and importance in human health applications. Here, we present a multi-scale computational model of the human red blood cell that accounts for metabolism and macromolecules. Proteomics data are used to place quantitative constraints on individual protein complexes that catalyze metabolic reactions, as well as a total proteome capacity constraint. We explicitly describe molecular mechanisms—such as hemoglobin binding and the formation and detoxification of reactive oxygen species—and account for sequence variations between individuals, allowing for personalized physiological predictions. This model allows for direct computation of the oxyhemoglobin curve as a function of model species and a more accurate computation of the flux state of the metabolic network. More broadly, this work represents another step toward building increasingly more complete systems biology whole-cell models.

Koptic: A Novel Approach for in silico Prediction of Enzyme Kinetics and Regulation.

Wheaton Schroeder¹ and Rajib Saha²

(1)Chemical and Biomolecular Engineering, University of Nebraska - Lincoln, Lincoln, NE, (2)Department of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE

Computational modeling of metabolism enables the design of engineering interventions directed to the overproduction of a specific bioproduct or improvement of plant performance. Flux Balance Analysis (FBA) is the primary tool used for this purpose, but has significant limitations due to the lack of reaction kinetics, chemical species concentration, and metabolic regulation. In contrast, kinetic models of metabolism (kMMs) provide not only a more accurate method for designing novel biological systems but also for the characterization of reaction kinetics, metabolite concentration, and metabolic regulation in these systems; however, the multi-omics data required for their construction is prohibitive to their development and widespread use.

Here, we introduce **K**inetic **OPT**imization using Integer Conditions (**KOPTIC**), which can circumvent the omics

data requirement and semi-automate kMM construction by using reaction rates and concentration data derived from a metabolic network model to return plausible kinetic mechanisms through an optimization-based approach. *Arabidopsis thaliana*'s (hereafter *Arabidopsis*'s) prominent role in 'omics' and plant science research makes it an ideal organism for the verification of KOPTIC-predicted kinetic mechanisms. While, several metabolic network models for *A. thaliana* already exist, the core carbon metabolism of this organism was chosen as the test system. A four-tissues (leaf, root, seed, and stem) metabolic network model, was reconstructed for *Arabidopsis* (1015 reactions, 901 metabolites, 508 genes). FBA was performed at 71 time-points to simulate the *Arabidopsis* lifecycle, and KOPTIC was applied to the FBA data. In total KOPTIC predicted 3577 regulatory interactions with a median fit error of 13.44%. More than 30 verified by existing literature. This research showcases how an optimization-based approach can be used to create meaningful hypotheses of reaction kinetics and increase mechanistic understanding of metabolism.

TUESDAY, OCTOBER 16

SESSION 4: Applications in Metabolic Engineering

INVITED SPEAKER

Systems Biology of Yeast Metabolism

Jens Nielsen^{1,2}

¹Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg, Sweden

²Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, DK-2800 Lyngby, Denmark

Metabolic Engineering relies on the Design-Build-Test cycle. This cycle includes technologies like mathematical modeling of metabolism, genome editing and advanced tools for phenotypic characterization. In recent years there have been advances in several of these technologies, which has enabled faster development of metabolically engineered strains that can be used for production of fuels and chemicals, but it is still challenging to perform efficient design. There is therefore in particular a need for advancing our ability to model metabolism, and this can be achieved through integration of systems biology tools. The yeast *Saccharomyces cerevisiae* is widely used for production of fuels, chemicals, pharmaceuticals

and materials. Through metabolic engineering of this yeast a number of novel industrial processes have been developed over the last 10 years. Besides its wide industrial use, *S. cerevisiae* also serves as an eukaryal model organism, and many systems biology tools have therefore been developed for this organism. These tools can be used for detailed phenotypic characterization as well as for metabolic design. In this lecture, advances in metabolic modeling of yeast will be presented. It will be shown how integration of kinetic information into genome-scale metabolic models can significantly improve their predictive strength. Furthermore, through expanding these models to describe protein synthesis and translocation, it is possible to get new insight into what is constraining flux through different metabolic pathways. Finally, it will be shown how advances in quantitative omics allows us to provide a global map of flux control in the metabolic network.

Experimental Design for Parameter Estimation in Kinetic Models of Metabolism.

Shyam Srinivasan¹, William Cluett¹, Radhakrishnan Mahadevan^{1,2}, and **Christian Euler**^{1,2}

(1)Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON, Canada,

(2)Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON, Canada

In kinetic models of metabolism, the parameter values determine the dynamic behaviour predicted by these models. Estimating parameters from in vivo experimental data requires the parameters to be structurally identifiable, and the data to be informative enough to estimate these parameters. Existing methods to determine the structural identifiability of parameters in kinetic models of metabolism can only be applied to models of small metabolic networks due to their computational complexity. Additionally, a priori experimental design, a necessity to obtain informative data for parameter estimation, also does not account for using steady state data to estimate parameters in kinetic models. We present a scalable methodology to structurally identify parameters for each flux in a kinetic model of metabolism based on the availability of steady state data. We determine the number and nature of experiments for generating steady state data to estimate the enzyme kinetic parameters in a kinetic model of a small metabolic network. We show that most parameters in fluxes expressed by mechanistic enzyme kinetic rate laws (Michaelis-Menten, Monod-Wyman-Changeux and Hill kinetics) can be identified using steady state data. We also show that the requisite steady state data can be obtained from experiments

involving both substrate and enzyme level perturbations. While substrate perturbation experiments are necessary for estimating parameters of uptake fluxes, enzyme level perturbation experiments may be necessary to estimate parameters of intracellular fluxes. The proposed methodology can be used in combination with other identifiability and experimental design algorithms that use dynamic data in order to determine the most informative experiments requiring the least resources to perform.

Essential Metabolism for a Minimal Cell.

Marian Breuer

Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL

The question of the core requirements of cellular life led to the construction of the “minimal cell” JCVI-syn3A: A cell where practically all genes were removed that were not essential for robust growth in a stress-free laboratory environment. With only 493 genes in a 543 kbp genome, JCVI-syn3A has a genome smaller than that of any independently-replicating cell found in nature. It provides a versatile platform to study the basics of cellular life and is small enough that a complete description of all cellular functions can be pursued.

Here, we present an extensively curated metabolic reconstruction and flux balance analysis (FBA) model of this minimal cell, using the vast amount of experimental information available on its natural precursor, *Mycoplasma mycoides capri*. The model, featuring 337 reactions involving 304 metabolites, is near-complete with 98 % of enzymatic reactions justified through gene assignments and/or experimental evidence, and agrees well with gene essentiality data from transposon mutagenesis experiments. The 155 genes included in the reconstruction have a high *in vivo* essentiality or quasi-essentiality of 92 %, compared to 79 % *in silico* essentiality—underscoring the minimality of the network. The reconstruction itself and the comparison of *in vivo* and *in silico* essentialities lead to new hypotheses on particular metabolic functions, suggesting specific experiments. Thus, the model provides a solid foundation for further experimental and computational studies on the minimal cell.

Characterization and Design of Phase Spaces and Yield Spaces in Genome-Scale Metabolic Models.

Steffen Klamt¹, Stefan Müller², Georg Regensburger³, and **Jürgen Zanghellini**^{4,5}

(1)Analysis and Redesign of Biological Networks, Max Planck Institute for Dynamics of Complex Technical

Systems, Magdeburg, Germany, (2)Faculty of Mathematics, University of Vienna, Vienna, Austria, (3)Institute for Algebra, Johannes Kepler University Linz, Linz, Austria, (4)Austrian Centre of Industrial Biotechnology, Vienna, Austria, (5)University of Natural Resources and Life Sciences, Vienna, Austria

Production rates and yields are key parameters of biochemical transformation processes. While optimal rates and phase spaces are readily studied with flux-balance analysis (FBA) approaches, optimal yields and yield spaces are rarely systematically analyzed. Often elementary flux modes (EFMs) are used to characterize yields and yield-optimal pathways. However, EFMs characterize the unbounded flux cone and are incompatible with non-zero flux bounds and allocation constraints often used in FBA.

By resorting to the concept of elementary flux vectors (EFVs), it is possible to generalize the idea of unique metabolic pathways to also account for inhomogeneous linear constraints. We show that any rate-optimal FBA solution sits in an optimal polyhedron spanned by (certain) EFVs. This holds true not only for rate-optimal but for yield-optimal solutions too, which cannot be found by standard FBA approaches.

Next, we demonstrate that (optimal) yield spaces can be readily calculated even in genome-scale metabolic models by linear-fractional programming without explicitly enumerating EFVs. Although phase spaces and yield spaces often are of similar shapes (and therefore sometimes confused), they carry very different information. In a realistic analysis based on *E. coli*, we show how these complementary pieces of information can be used to understand and optimally shape the metabolic capabilities of cell factories with any desired yield and/or rate requirements.

We conclude that EFVs provide an unifying framework for the theoretical description and analysis of any constraint-based model under arbitrary linear constraints. More specifically, EFVs close the gap between biased FBA approaches and unbiased EFM approaches and allow one to fully characterize and shape metabolic phase spaces and yield spaces. This reinforces the fundamental importance of EFVs (or EFMs) as the “coordinates of metabolism”. However, an explicit enumeration of EFVs is not required as phase spaces and yield spaces can be efficiently computed even in genome-scale metabolic networks.

DD-Decaf: Data-Driven Design of Cell Factories and Communities.

Svetlana Kutuzova¹, Moritz E. Beber¹, Ali Kaafarani¹, Alba Lopez¹, Mátyás Fodor², Christian Lieven¹, Danny Dannaher¹, Henning Regestig¹, Markus J. Herrgård¹, and **Nikolaus Sonnenschein¹**

(1)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark, (2)Genialis, Inc, Ljubljana, Slovenia

With ultra-precise genome editing tools at our disposal, we will see a transformation of the life sciences. We will shift away from experiments manipulating one factor at a time towards simultaneous changes at multiple loci. The success of those increasingly complex experiments will crucially depend on our ability to accurately predict system's level behavior and will often require non-intuitive designs.

In principle, integrating omics data with systems biology models provides the means for optimizing knowledge gain through rational target selection and experimental design. This combination of methods is not leveraged effectively, however, due to a lack of readily available tools that allow us to rapidly access and analyze public and private data.

With this project, we aim to make a broad spectrum of omics data useful for biotechnology and life science research by integrating systems biology with design in a one-stop source.

A group of five renowned academic partners (DTU, Chalmers, EMBL, EPFL and UMinho) is driving research on integrative model-based omics data analysis to enable:

Metagenomics-enabled design of novel enzymes and biochemical pathways.

Omics data-driven design of cell factories for the production of chemicals and proteins.

Analysis and design of microbial communities relevant to human health, industrial biotechnology and agriculture.

All research efforts are integrated in an interactive, web-based platform available at <http://dd-decaf.eu>.

Dynamic Enzyme-Cost Flux Balance Analysis (deFBA) Modelling for an Industrially Relevant Methanotroph *Methylobaculum Buryatense*.

Kobe De Becker¹ and Steffen Waldherr²

(1)Bio- & Chemical Systems Technology, Reactor Engineering and Safety (CREaS), Department of Chemical Engineering, KU Leuven, Leuven, Belgium, (2)Bio- and Chemical Systems Technology, Reactor Engineering and

Safety (CREaS), Department of Chemical Engineering, KU Leuven, University of Leuven, Leuven, Belgium

Constraint-based modelling has made significant progress in the last few years. Starting from metabolic network reconstructions a whole range of different modelling frameworks has been developed. The static flux balance analysis (FBA) method allows calculation of intracellular and exchange fluxes. Mahadevan et al. (2002) extended this approach to predict dynamic concentration profiles of extracellular metabolites, the so-called dynamic flux balance analysis method (DFBA). Recently, Waldherr et al. (2015) developed the dynamic enzyme-cost flux balance analysis (deFBA) framework. This approach includes enzyme production and takes into account that reaction fluxes are limited by the amount of catalyzing enzyme present inside the cells.

In this contribution, the construction of a deFBA model for the methanotroph *Methylobaculum buryatense* is discussed. *M. buryatense* consumes methane and produces small organic acids. The deFBA model is built starting from a metabolic network reconstruction from de la Torre et al. (2015). Enzyme composition and catalytic information are obtained from online databases such as BRENDA, Kegg, GenBank and Uniprot. A deFBA model is obtained containing 546 reactions from which 149 are enzyme production reactions. The model simulations predict specific substrate uptake and growth rates comparable to experimentally obtained values by Gilman et al. (2015). An exponential growth profile for biomass and its components is predicted. Production fluxes for acetic and formic acid are non-zero. However, lactic acid is not produced which contradicts experimental measurements.

Mahadevan et al., Dynamic flux balance analysis of diauxic growth in *Escherichia coli*, Biophys J., 2002;83(3):1331-1340.

Waldherr et al., Dynamic optimization of metabolic networks coupled with gene expression, Journal of Theoretical Biology, 2015;365:469-485

de la Torre et al., Genome-scale metabolic reconstructions and theoretical investigation of methane conversion in *Methylobaculum buryatense* strain 5G(B1), Microbial Cell Fact., 2015;14:188

Gilman et al., Bioreactor performance parameters for an industrially-promising methanotroph *Methylobaculum buryatense* 5GB1, Microbial Cell Fact., 2015;14:182

SESSION 5: Methods and Software

INVITED SPEAKER

Data Fusion Based on Coupled Matrix and Tensor Factorizations.

Evrin Acar

Simula Metropolitan Center for Digital Engineering, Oslo, Norway

When the goal is to discover the underlying patterns in a complex system such as the human metabolism or brain, the complexity of the problem requires the collection and analysis of data from multiple sources. Therefore, data fusion, i.e., knowledge extraction by jointly analyzing complementary data sets, is a topic of interest in many fields. For instance, in metabolomics, analytical platforms such as Liquid Chromatography - Mass Spectrometry and Nuclear Magnetic Resonance Spectroscopy are used for chemical profiling of biological samples. Measurements from different platforms are capable of detecting different chemical compounds with different levels of sensitivity, and their fusion has the potential to provide a more complete picture of the metabolome related to a specific condition. Similarly, neuroimaging modalities such as functional Magnetic Resonance Imaging (fMRI) and electroencephalography (EEG) provide information about the brain function in complementary spatio-temporal resolutions, and their joint analysis is expected to provide better understanding of brain activities. However, data fusion remains a challenging task since there is a lack of data mining tools that can jointly analyze incomplete (i.e., with missing entries) heterogeneous (i.e., in the form of higher-order tensors and matrices) data sets, and capture the underlying shared/unshared patterns. We formulate data fusion as a coupled matrix and tensor factorization (CMTF) problem and discuss its extension to structure-revealing data fusion, i.e., fusion models that can identify shared and unshared factors in coupled data sets. Numerical experiments on real coupled data sets demonstrate that while traditional methods based on matrix factorizations have limitations in terms of jointly analyzing heterogeneous data sets, the structure-revealing CMTF model can successfully capture the underlying patterns by exploiting the low-rank structure of higher-order tensors. We will show the broad impact of CMTF-based fusion models with applications from metabolomics, neuroscience and recommender systems.

INVITED SPEAKER

COBRA and Machine Learning Approaches for Engineering Microbial Biocatalysts

Jennifer L. Reed^{1,2}

¹*Dept. of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI USA*

²*Great Lakes Bioenergy Research Center, Madison, WI USA*

Microbes have been engineered to produce a variety of chemicals, including biofuels, commodity chemicals, specialty chemicals, and therapeutics. Chemical production can be enhanced by connecting synthesis pathways to host metabolism, re-wiring regulatory networks, improving precursor production, and optimizing gene expression. A number of computational systems biology approaches have been developed to facilitate metabolic engineering efforts by suggesting which combination of genetic changes would improve chemical production. Network analysis methods can be used to identify paths from renewable substrates to high-value chemical products and central metabolic precursors common to a variety of chemical products. Genome-scale metabolic models can be used to predict how gene deletions, gene additions, and flux changes would impact chemical product yields, growth rates, and/or productivities. Additionally, machine learning and active learning algorithms can be used to optimize gene expression constructs to efficiently convert metabolic precursors into desired products. Case studies will be presented that show how a variety of computational tools can guide development of strains with enhanced chemical production. This work will illustrate how integrating computational and experimental efforts can lead to the rapid development of microbial biocatalysts for renewable chemical production.

How Accurate Is Automated Gap Filling of Metabolic Models?

Mario Latendresse¹ and Peter Karp²

(1)SRI International, (2)SRI International, Menlo Park, CA

Reaction gap filling is a computational technique for proposing the addition of reactions to genome-scale metabolic models to permit those models to run correctly. Gap filling completes a reaction network by adding reactions that enable biosynthesis of all required metabolic products from available nutrients. The models are incomplete because they are derived from annotated genomes in which not all enzymes have been identified.

We present two studies of gap-filling accuracy. In the first study [1] we compared the results of applying an automated likelihood-based gap filler (MetaFlux) within the Pathway Tools software with the results of manually gap filling the same metabolic model. Both gap-filling exercises were applied to the same genome-derived metabolic reconstruction for *Bifidobacterium longum*. The MetaFlux gap filler attained recall of 61.5% and precision of 66.6%, taking the manual gap-filling result as the gold standard.

In the second study [2] we generated degraded versions of the EcoCyc-20.0-GEM model by randomly removing flux-carrying reactions from a growing model. We gap-filled the degraded models using 13 variations of MetaFlux (including the use of the SCIP and CPLEX Mixed Integer Linear Programming solvers and three different gap-filling algorithms) and compared the resulting gap-filled models with the original model. The best MetaFlux variation showed a best average precision of 87% and a best average recall of 61%. Although none of the 13 algorithm variations was best in all dimensions, we found one variation that was fast, accurate, and returned more information to the user. Some gap-filling variations were inaccurate, producing solutions that were non-minimum or invalid (did not enable model growth).

[1] "How Accurate is Automated Gap Filling of Metabolic Models?" Karp PD., Latendresse M., Weaver DW., BMC Syst Biol. 2018 12(1):73.

[2] "Evaluation of reaction gap-filling accuracy by randomization," Latendresse M., Karp PD., BMC Bioinformatics 2018 19(1):53.

Bridging the Gap between Structural Systems Biology and Constraint-Based Modeling: Tools for Defining the Functional Structural Proteome and Extracting Genome-Scale Structural Information for Metabolic and Macromolecular Expression Models.

Edward Catoiu¹, Nathan Mih², Jean-Christophe Lachance³, Colton J. Lloyd¹, Erol Kavvas⁴, and Bernhard O. Palsson¹

(1)Bioengineering, University of California, San Diego, La Jolla, CA, (2)UCSD, (3)Biology, Université de Sherbrooke, Sherbrooke, QC, Canada, (4)Bioengineering, UCSD, La Jolla, CA

Genome-scale models of metabolism and macromolecular expression (ME-models) have used structural information to identify properties of the structural proteome by mapping individual genes to a subunit within a 3D protein structure. While gene-mapping is a first step in structural systems biology, enzymes that catalyze biochemical reactions can be composed of multiple protein subunits from multiple genes. Thus, the proper gene-mapping of

multi-subunit protein structures to form a “functional” structural proteome is not only crucial for the integration of structural data into genome-scale models, but also provides the basis for discovery and incorporation of novel structural properties for constraint development in ME-models. Here, we describe the enhancement of a structural systems biology software (ssbio) tools to define and to gather novel genome-scale properties of the functional structural proteome of *Escherichia coli*. Specifically, tools were developed to address the technical challenges of working with structural data, ensuring that structures incorporated into ssbio are in a format that reflects the multi-subunit architecture of the enzyme. Structure quality predictions were improved to identify and remove mis-folded protein regions found in homology models. An algorithm was developed to distinguish between multiple conformations of structures available in the Protein Databank (PDB), to recreate annotated gene-enzyme stoichiometry, to define the functional structural proteome of *E.coli*, and to predict incorrect enzyme annotations. Upon the identification of the functional proteome, novel angstrom-level structural properties including cross-sectional area of transmembrane enzymes and soluble volume of enzymes were calculated. Additionally, computational tools to map residue-specific information such as enzymatic domains, secondary structure and strain-specific sequence variation onto structural data were deployed to analyze in-house mutations of interest. Together, the development of the aforementioned structural systems biology tools allows for seamless integration of functionally relevant structural information into genome-scale models, potentially leading to a new set of constraints driving biological discovery.

Scalable Tools for Analyzing Steady-State Microbial Communities Using Standardized Genome-Scale Metabolic Models.

Siu Hung Joshua Chan¹ and **Costas Maranas²**

(1)Department of Chemical and Biological Engineering, Colorado State University, Fort Collins, CO, (2)Department of Chemical Engineering, The Pennsylvania State University, University Park, PA

Extending genome-scale metabolic modeling to microbial communities is a promising way to unravel community interactions. We present three tools that significantly improve the accuracy and feasibility of simulating community metabolism. First, little emphasis has been placed on the need to impose a time-averaged constant growth rate across all community members to ensure co-existence and stability. Without this constraint, the faster growing organism will ultimately displace other microbes. We introduced the SteadyCom framework

for predicting community metabolism consistent with the steady-state requirement. SteadyCom is scalable to a large number of organisms and compatible with flux variability analysis (FVA). In contrast to flux balance analysis, SteadyCom can predict an abundance profile with good agreement to experimental gut microbiota.

Second, during simulating community metabolism, an unstandardized biomass reaction of any organisms that produces biomass with a molecular weight different from the standard 1 g/mmol introduces a systematic error. We developed a computational procedure for checking the biomass weight so as to eliminate the error. 42 out of 64 examined models show >5% discrepancies in biomass weights. We demonstrated that biomass weight discrepancies could cause significant errors in the predicted community composition that are disproportionately larger than the biomass weight discrepancies. Microbes with underestimated biomass weights are overpredicted whereas microbes with overestimated biomass weights are underpredicted.

Third, to effectively perform FVA in the absence of thermodynamically infeasible cycles in community models, we devised a method termed localized loop-less FVA (III-FVA) with superior computational performance. We identified the fewest constraints required under the loop-less requirement and put forth the concept of localized loop-less constraints. The computational time is reduced by 10-150 times compared to the original loop-less constraints and by 4-20 times compared to the currently fastest method Fast-SNP. III-FVA offers a scalable strategy for loopless flux calculations for large multi-compartment/multi-organism models (e.g., >10⁴ reactions).

Genome-Scale Metabolic Model Embedding for Fed-Batch Optimal Feed Policy Determination.

Felipe Scott¹, **Vassilios Vassiliadis²**, and **Raul Conejeros³**

(1)Faculty of Engineering and Applied Sciences, Universidad de los Andes, Santiago, Chile, (2)Department of Chemical Engineering and Biotechnology, University of Cambridge, Cambridge, United Kingdom, (3)School of Biochemical Engineering, Pontificia Universidad Catolica de Valparaiso, Valparaiso, Chile

A new method for genome-scale metabolic model into culture kinetics, resulting in a dynamic flux balance analysis model is described and used to calculate optimal feed policies for fed-batch cultures where the dynamics of the fermentation process is captured by a set of differential equations and the metabolism is described by a genome scale flux balance analysis.

The methodology transforms the bounds of the embedded linear programming problems of flux balance analysis via a logarithmic barrier (interior point) approach. Further algebraic manipulations produce an implicit ordinary differential equations system.

The system of differential equations obtained, is embedded in an optimal control solver (using gPROMS), to optimize the feed flow rate profiles, substrate concentration in the feed and dissolved oxygen profiles in the culture. For genome-scale models, the optimal control problem has thousands of ordinary differential equations. Despite its size, a simulation can be performed in less than one second, and the full optimal control task can be completed in a fraction of a minute. This represents a significant improvement from previous works presented in literature, where problem reformulation based in collocation results in a large non-linear programming problem, and where only a network describing the central metabolism was addressed using this methodology.

Case studies include genome-scale metabolic models for *Escherichia coli* where total biomass accumulation at the end of the fermentation was maximized and a case study for ethanol accumulation from xylose and glucose using the yeast *Scheffersomyces stipitis*.

INVITED SPEAKER

Memote: A Community Driven Effort Towards a Standardized Genome-Scale Metabolic Model Test Suite.

Christian Lieven¹, Moritz E. Beber¹, Brett G. Olivier², Frank T. Bergmann³, Nikolaus Sonnenschein¹, and The COBRA Community

(1)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark, (2)Systems Bioinformatics, Amsterdam Institute for Molecules, Medicines and Systems, Vrije Universiteit Amsterdam, Amsterdam, Netherlands, (3)Modelling of Biol. Processes, Heidelberg University, Heidelberg, Germany

Several studies have shown that neither the formal representation nor the functional requirements of genome-scale metabolic models (GEMs) are precisely defined. Without a consistent standard, comparability, reproducibility, and interoperability of models across groups and software tools cannot be guaranteed.

Here, we present memote (<https://github.com/opencobra/memote>) an open-source software containing a community-maintained, standardized set of metabolic model tests. The tests cover a range of aspects from annotations to conceptual integrity and can be extended

to include experimental datasets for automatic model validation. In addition to testing a model once, memote can be configured to do so automatically, i.e., while building a GEM. A comprehensive report displays the model's performance parameters, which supports informed model development and facilitates error detection.

Memote provides a measure for model quality that is consistent across reconstruction platforms and analysis software and simplifies collaboration within the community by establishing workflows for publicly hosted and version controlled models.

An Ecology of Modeling Methods, Data, and Tools to Decipher Functional Delegation and Species Interactions within a Microbiome.

Christopher S. Henry¹, Pamela Weisenhorn¹, Daniel van der Lelie², James G. Jeffryes¹, Dylan C. Chivian³, José P. Faria¹, Janaka N Edirisinghe¹, Filipe Liu¹, Samuel M.D. Seaver¹, Harry Dudley⁴, Qizhi Zhang¹, Boris Sadkhin¹, Nidhi Gupta¹, Tian Gu¹, Ronald C. Taylor⁵, Hyun-Seob Song⁵, Hans C. Bernstein⁵, Jeremy Zucker⁶, Stephen R. Lindemann⁵, Jack Gilbert⁶, Robert Cottingham⁷, and Adam Arkin⁸

(1)Argonne National Laboratory, Argonne, IL, (2)Gusto Global, Charlotte, NC, (3)Lawrence Berkeley National Laboratory, Berkeley, CA, (4)University of Colorado Boulder, Boulder, CO, (5)Pacific Northwest National Laboratory, Richland, WA, (6)Argonne National Laboratory, Lemont, IL, (7)Oak Ridge National Laboratory, Oak Ridge, TN, (8)Bioengineering, University of California, Berkeley, Berkeley, CA

The term *metagenome* was first used in the literature 20 years ago, and in that time, we have witnessed a revolution in our capacity to gather data, model, and understand microbiome systems in natural and engineered environments. Finally we have sufficient data available to support the validation and refinement of metabolic models for microbiome systems. Here, I describe an ecology of tools, pipelines, and data that we have developed to support the modeling and simulation of microbiome systems. One of the initial challenges in modeling a microbiome system is to obtain species genomes from metagenomic data. To aid with this challenge, the KBase team recently implemented a pipeline for metagenome assembly, binning, QC, and annotation comprised of available open-source tools (e.g. metaSPAdes, MaxBin2). Another challenge is to rapidly construct predictive genome-scale models from newly assembled and binned genomes. To address this challenge, we have developed publicly accessible tools

to generate interoperable models for microbial, plant, and fungal genomes (with recent significant improvements to our microbial model reconstruction pipeline). Even with models, these community systems are typically under-determined, with species interactions difficult to identify from noise. For this, we developed tools to predict auxotrophy from genomic data, revealing exciting new insights into how microbiome systems evolve and build stable interconnections. We also developed tools and pipelines to support the mapping of models and microbiome systems in general to metabolomics data, applying network approaches to enable models to slice through noise. Finally, we recently developed a new approach that explores the balance of protein-resources that plays a significant role in the delegation of functions in microbiome systems. In my talk, I describe these tools and explore the results from their application to a variety of microbiome systems, including a photo-autotrophic microbial mat, an electrosynthetic microbiome, and a microbiome on a built surface.

Thermodynamic Framework for Mutant Phenotype Prediction.

Tolutola Oyetunde¹ and **Yinjie J. Tang²**

(1)Energy, Environmental and Chemical Engineering, Washington University in St. Louis, St. Louis, MO, (2) Department of Energy, Environmental and Chemical Engineering, Washington University, St. Louis, MO

Accurate prediction of the metabolic flux redistribution within mutant strains is vital for directing rational strain design. The discrepancies between experimentally determined and computationally predicted flux distributions imply that the predictive techniques do not accurately capture the mechanisms of cellular regulation.

Previous mutant prediction algorithms have essentially two components: (1) a metric to characterize the cell's desired metabolic state (for example flux or gene expression profiles) and (2) a metric to describe the distance from the desired state (for example, Euclidean distance). The mutant flux profile is then computed as the closest possible to the wild type state subject to the constraints of genetic or environmental perturbations. Different algorithms have reported significant gains in accuracy by tweaking the definitions of the two metrics.

To further improve the fidelity of knockout predictions and subsequent computational strain design, we developed a thermodynamics-based method, RElative MEtabolite Patterns (REMEP). REMEP hypothesizes that the optimum metabolic state is reflected in the energetic requirements to sustain flux through each metabolite

node, and thus cell fluxomes adapt to perturbations from a reference state by preserving relative pattern of metabolite energy flows. REMEP performs better than comparable algorithms across different experimental datasets for *E. coli* and *S. cerevisiae* (in terms of lower root mean square errors and higher Pearson's correlation coefficients). These improvements support the REMEP assumption that cellular mechanisms of response to genetic and environmental perturbations leaves signatures that can be inferred from thermodynamics-derived metabolite patterns. The findings provide a new paradigm for genotype to phenotype mapping and insights into microbial flux network plasticity.

Dynamic FBA with Global Constraints on Cellular Protein Fraction.

Eivind Almaas^{1,2}, **Emil Karlsen¹**, **Christian Schulz¹**, and **Nikolay Martyushenko¹**

(1)Dept. of Biotechnology, NTNU - Norwegian University of Science and Technology, Trondheim, Norway, (2) K.G. Jebsen Center for Genetic Epidemiology, NTNU - Norwegian University of Science and Technology, Trondheim, Norway

In this presentation, I will first briefly discuss recently developed software in our group that aids in the process of model reconstruction and curation: (i) *AutoKEGGRec* [1], a COBRA toolbox pipeline based on KEGG that creates first draft metabolic network reconstructions of single organisms, reconstructions based on a list of organisms, and finally a consolidated reconstruction for a list of organisms or strains. (ii) *ModelExplorer* [2], a visual software framework to assist the user in exploring and curating genome-scale metabolic models provided in the sbml format. ModelExplorer automatically visualizes a metabolic network as a graph with a distinct separation and delineation of cellular compartments. Additionally, ModelExplorer highlights reactions and species that are unable to carry flux or participate in flux generation (ie. blocked), with the use of several different consistency checking modes available.

Using high-quality models resulting from curation efforts, it is possible to apply a plethora of constraint-based methods to analyze the models' phenotypes. I will discuss a recently developed method that combines dynamic Flux Balance Analysis (dFBA) with a global constraint on cellular protein fraction [3]. We apply our method to a recent *E. coli* genome-scale reconstruction, finding that our method outperforms existing dFBA formulation. The globally-constrained dFBA method makes it possible to investigate cellular proteome composition in response to changing nutrient conditions.

- [1] E. Karlsen, C. Schulz and E. Almaas. Submitted.
 [2] N. Martyushenko and E. Almaas. Submitted.
 [3] E. Karlsen and E. Almaas. In preparation.

Incorporating Flux Sampling into a Minimal Assumption Dynamic Flux Balance Analysis Algorithm.

St. Elmo Wilken¹, Linda Petzold², and Michelle O'Malley¹
 (1)Chemical Engineering, UCSB, Santa Barbara, CA, (2) Computer Science, UCSB, Santa Barbara, CA

Recent advances in sequencing technology have contributed to the rapid development of several highly predictive genome-scale models. Techniques like flux balance analysis (FBA) allow for relatively simple, mechanistically based, interrogation of these models. A drawback of FBA is the assumption of steady state. Dynamic flux balance analysis (DFBA) is an extension of this technique to non-steady state systems; a primary advantage of this approach is the low extra cost associated with developing dynamic models that are still mechanistically valid.

DFBA simulators typically solve successive FBA problems, and use the resultant fluxes to move the dynamic states of the system forward in time. A challenge associated with DFBA is the proper way of handling the degenerate solution space resulting from FBA. We propose a novel algorithm that takes advantage of the degenerate solutions of FBA to more closely match *in vivo* conditions. Indeed, it is well known that noisy gene expression results in variable enzyme concentrations between clonal cells; therefore, one expects the metabolic fluxes experienced by each cell to be different. Our algorithm models this behavior by sampling from the solution space, and uses these sampled fluxes to move the states of the system forward in aggregate. The resultant time varying flux distributions represent an extension of flux variability analysis to the dynamic setting. Our algorithm stands in contrast to the current state of the art simulators that utilize sequential linear optimization to find a unique flux for each state that is integrated. While this approach is mathematically robust, it requires further assumptions that might not be valid, or could be hard to justify. If the predicted transient flux distributions associated with our approach prove to accurately capture the *in vivo* behavior of an organism, it would allow experimentalists the ability to carefully interrogate proposed engineering strategies in the dynamic setting.

Creation and Analysis of Biochemical Constraint-Based Models: The Cobra Toolbox v3.0.

Ronan M.T. Fleming

Leiden Academic Centre for Drug Research, Leiden University, Leiden, Netherlands

COntstraint-Based Reconstruction and Analysis (COBRA) provides a molecular mechanistic framework for integrative analysis of experimental molecular systems biology data and quantitative prediction of physicochemically and biochemically feasible phenotypic states. The COBRA Toolbox is a comprehensive desktop software suite of interoperable COBRA methods. It has found widespread applications in biology, biomedicine, and biotechnology because its functions can be flexibly combined to implement tailored COBRA protocols for any biochemical network. This protocol is an update to the COBRA Toolbox 1.0 and 2.0. Version 3.0 includes new methods for quality controlled reconstruction, modelling, topological analysis, strain and experimental design, network visualisation as well as network integration of chemoinformatic, metabolomic, transcriptomic, proteomic, and thermochemical data. New multi-lingual code integration also enables an expansion in COBRA application scope via high-precision, high-performance, and nonlinear numerical optimisation solvers for multi-scale, multi-cellular and reaction kinetic modelling, respectively. This protocol overviews all of these new features and can be adapted to generate and analyse constraint-based models in a wide variety of scenarios. The COBRA Toolbox 3.0 provides an unparalleled depth of constraint-based reconstruction and analysis methods.

POSTER ABSTRACTS

Towards a Metabolic and Expression Model of the Metabolically Versatile Bacterium *Rhodopseudomonas palustris*.

Adil Alsiyabi¹, Cheryl Immethun², and Rajib Saha³

(1)Chemical and Biomolecular Engineering, University of Nebraska - Lincoln, Lincoln, NE, (2)Chemical & Biomolecular Engineering, University of Nebraska, Lincoln, NE, (3)Department of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE
Rhodopseudomonas palustris is a metabolically versatile Purple Non-Sulfur Bacterium (PNSB).

Depending on growth conditions, *R. palustris* can operate in either one of four different forms of metabolism: photoautotrophic, photoheterotrophic, chemoautotrophic, and chemoheterotrophic. *R. palustris* is also a facultative anaerobe and is capable of fixing nitrogen. This metabolic flexibility suggests that *R. palustris* would be an ideal model organism for studying the network of interacting reactions, and how it's altered in response to changing conditions. Moreover, exploration into *R. palustris*' industrial applications has just begun with research being generated on the production of hydrogen, methane, and polyhydroxybutyrate (PHB). This study aims to combine systems and synthetic biology by reconstructing *R. palustris*' metabolic and expression (ME-) network and establishing a constitutive promoter library to gain a mechanistic understanding of the organism's functional capabilities under different conditions.

In this study, a genome-scale metabolic model was first constructed following a well-developed framework established in the lab. The metabolic model currently contains 925 reactions, 1220 metabolites, and 1134 genes. An expression matrix (E-matrix), containing transcription, translation, and post-transcriptional modifications, is currently being reconstructed. After the reconstruction of this matrix, the two networks will be integrated and previously established 'coupling constraints' will be implemented to link the synthesis of macromolecular molecules to catalysis of metabolic components. Moreover, preliminary experimental results showed that the bacterium is able to utilize a host of different carbon sources, including a host of lignin-derived compounds. An IPTG induced promoter was also established to express the CRISPRi components. This system will be used to verify and iteratively improve model predictions, resulting in a well-established mechanistic understanding of the organism's cellular functions. The predictive power of this work will expand as more omics datasets become available, facilitating the establishment of this metabolically versatile, non-model microorganism as a biotechnology platform.

Genetic Optimization Algorithm for Metabolic Engineering Revisited.

Tobias B. Alter, Lars M. Blank, and Birgitta E. Ebert
iAMB - Institute of Applied Microbiology, ABBt - Aachen Biology and Biotechnology, RWTH Aachen University, Aachen, Germany

To date, several independent methods and algorithms exist exploiting constraint-based stoichiometric models to find metabolic engineering strategies that optimize microbial production performance. Optimization procedures based on metaheuristics facilitate a straightforward adaption and expansion of engineering objectives as well as fitness functions, while being particularly suited for solving problems of high complexity. Regarding the uprising of multi-scale models and a need for solving advanced engineering problems, we strive to advance genetic algorithms, which stand out due to their intuitive optimization principles and proven usefulness in this field of research. A drawback of genetic algorithms is that premature convergence to sub-optimal solutions easily occurs if the optimization parameters are not adapted to the specific problem. Therefore, we conducted comprehensive parameter sensitivity analyses focusing on the impact on finding optimal strain designs. We further demonstrate the capability of genetic algorithms to simultaneously handle (i) multiple, non-linear engineering objectives (ii) the identification of gene target sets according to logical gene-protein-reaction associations, (iii) minimization of the number of network perturbations, and (iv) the insertion of non-native reactions, while employing genome-scale metabolic models. This framework adds a level of sophistication in terms of strain design robustness, which is exemplarily tested on succinate overproduction in *Escherichia coli*.

New Methods to Constrain Genome Scale Models with Stable Isotope Labeling Data.

Tyler Backman¹, David E. Ando², Jay D. Keasling³, and Hector Garcia Martin²

(1)Joint BioEnergy Institute, Emeryville, CA, (2)Joint BioEnergy Institute (JBEI), Emeryville, CA, (3)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby, Denmark

Here we present new methods to directly constrain genome scale models with ¹³C stable isotope labeling data. By reducing the assumptions required for flux modeling, these methods can infer fluxes in engineered organisms with unusual phenotypes. First, we present a method to systematically calculate flux bounds for any specified "core" of a genome-scale model so as to satisfy

the bow tie approximation that carbon tends to flow outward from core metabolism. Second, we automatically identify an updated set of core reactions that can satisfy this approximation more efficiently via Simulated Annealing. Lastly, we present a workflow to quantify the uncertainty of fluxes constrained with stable isotope labeling data at the genome scale using Markov Chain Monte Carlo (MCMC).

Mgpipe: A Toolbox for Investigating Metagenomics Data through Personalized Microbiota Metabolic Models.

Federico Baldini¹, Almut Heinken¹, Laurent Heirendt¹, Ronan M.T. Fleming², and Ines Thiele¹

(1)Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg,

(2)Leiden Academic Centre for Drug Research, Leiden University, Leiden, Netherlands

mgPipe: A toolbox for investigating metagenomics data through personalized microbiota metabolic models

Background/Purpose: Predictive, personalized medicine requires tools that can account for individual variability, such as the genome, life style, and the microbiome. Recently, the gut microbiome has received great attention for its potential to modulate human metabolism and its role in health and disease. The aim of this project is the development of an easy-to-use tool (mgPipe (1)) that enables researchers to investigate *in silico* the functional implications associated with changes in gut microbial composition in healthy individuals and patients.

Methods: mgPipe combines published algorithms, which map metagenomics data onto a reference database to determine microbial identity and abundance (2-3). Subsequently, this microbial information is combined with genome-scale metabolic reconstructions of more than 800 gut microbes (4) to generate personalized microbiome metabolic models. Using established constraint-based modeling methods from the COBRA toolbox (5) and machine learning approaches, mgPipe enables the *in silico* interrogation of the different functional, metabolic properties of these microbiome models.

Results and conclusion: mgPipe is an efficient tool to integrate metagenomic data with constraint based modeling enabling creation, simulation, and analysis of personalized microbiome metabolic models.

References

- Baldini F, Heinken A, et al. The Microbiome Modeling Toolbox: from microbial interactions to personalized microbial communities. *bioRxiv*. 2018:318485.
- Bolger AM, et al. *Bioinformatics*. 2014;btu170.
- Li H, et al. *Bioinformatics*. 2009;25(14):1754-60.

- Magnúsdóttir S, et al. Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. *Nature Biotechnology*. 2016.
- Heirendt L, Arreckx S, et al. Creation and analysis of biochemical constraint-based models: the COBRA Toolbox v3. 0. *arXiv preprint arXiv:171004038*. 2017.

Systems Approaches to Identify Metabolite Signatures in Placental Development.

Priyanka Baloni¹, Alison Paquette¹, Heather Brockway², Yoel Sadovsky³, Louis Muglia², and Nathan D. Price¹

(1)Hood-Price Lab, Institute for Systems Biology, Seattle, WA, (2)Cincinnati Children's Hospital Medical Center, Cincinnati, OH, (3)Magee-Womens Research Institute, Pittsburgh, PA

The human placenta is essential for successful pregnancy as it performs critical functions such as transfer of nutrients and oxygen to the fetus, removing harmful waste products, and providing mechanical, hormonal and immune support. Alterations to placental function can have profound impacts on the health of the developing fetus and the outcome of the pregnancy. In this study, we analyzed transcriptomic data of 200 term placentae and identified expression levels for metabolic genes. We used the most comprehensive human metabolic reconstruction, Recon 3D, and integrated transcriptomics data with the model to make it placenta-specific. The draft reconstruction of placenta consists of about 5525 reactions, more than 5500 metabolites and captures about 80% reactions in subsystems such as Vitamin K metabolism, cholesterol metabolism, citric acid cycle, chondroitin synthesis, nucleotide metabolism and glycerophospholipid metabolism compared with Recon 3D model. This is the first *in silico* metabolic network of the human placenta and will serve as a platform to integrate longitudinal transcriptomic and metabolomic data throughout pregnancy. We have also developed a pipeline to integrate a placental transcriptional regulatory network with this metabolic network, thus making it possible to predict transcription factors that regulate metabolic changes in placenta. Placenta metabolic reconstruction provides insights into normal placental development and can also be used to predict metabolic signatures in pregnancy complications. We aim to develop a predictive methodology that combines information from various omics measurements with metabolic reconstruction that can be used to predict a metabolic trajectory across pregnancy.

POSTER ABSTRACTS

Cobrapy: A Solid Foundation for a Diverse Python Metabolic Modeling Community.

Moritz E. Beber¹, João G.R. Cardoso², Christian Diener³, Mátyás Fodor⁴, Ali Kaafarani¹, Zachary A. King⁵, Matthias König⁶, Svetlana Kutuzova¹, Christian Lieven¹, Alba Lopez¹, Michael Mundy⁷, Vivek Rai⁸, Henning Redestig⁹, Nikolaus Sonnenschein¹, Peter C. St. John¹⁰, and Sascha Steinbüß¹¹

(1)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark, (2)Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark, (3)National Institute for Genomic Medicine, Mexico City, Mexico, (4)Genialis, Inc, Ljubljana, Slovenia, (5)Bioengineering, University of California, San Diego, La Jolla, CA, (6)Institute of Theoretical Biology, Humboldt University Berlin, Berlin, Germany, (7)Center for Individualized Medicine, Mayo Clinic, Rochester, MN, (8)Department of Computational Medicine & Bioinformatics, University of Michigan, Ann Arbor, MI, (9)Genencor International, DuPont, Leiden, Netherlands, (10)Biosciences Center, National Renewable Energy Laboratory, Golden, CO, (11)Debian Project, New York, NY

When the COBRApy project began, its goals were defined as providing an intuitive, object-oriented interface for constraint-based metabolic modelling in Python that 1) can represent complex biological processes, 2) meets the computational challenges of next generation stoichiometric constraint-based models and omics data sets, 3) promotes related research via freely available software, and 4) provides an interoperable alternative to the COBRA toolbox for Matlab.

Today, we can safely say that COBRApy has achieved all of those goals and has become an established tool that both academic and industry users rely on. We want to raise awareness for the maturity of COBRApy and how it has enabled productive constraint-based modeling by publishing a stable 1.0 release. In this work, we not only present the many improvements that have happened since the start of the project, but also highlight the vibrant and diverse community of Python packages built on top of COBRApy that together provide a plethora of methods related to constraint-based reconstructions and modeling.

While navigating a diverse software landscape akin to Bioconductor or the Python package index can be daunting at first, the benefits of distributed development are more publicly available tools due to the creative freedom provided, fast access to implementations of latest methods, and more direct contact between users and authors. Last but not least, a healthy open source community that encourages contributions from anyone can help mitigate the problem of abandoned academic

software. Over the course of its six years of existence, COBRApy has been under the stewardship of several people and has seen contributions from almost 30 authors without whom it would not be where it is today. Here, we aim to provide a solid overview of the packages and capabilities within the community built around COBRApy.

Probabilistic Metabolic Network Reconstruction: A Modular Approach to Integrate Omics Data and Represent Uncertainty.

David Bernstein^{1,2}, Shany Ofaim^{2,3}, Luca Zoccarato⁴, Daniel Sher⁵, and Daniel Segre^{1,2,3,6,7}

(1)Department of Biomedical Engineering, Boston University, Boston, MA, (2)Biological Design Center, Boston University, Boston, MA, (3)Bioinformatics Program, Boston University, Boston, MA, (4)Department of Experimental Limnology, Leibniz-Institute of Freshwater Ecology and Inland Fisheries, Neuglobsow, Germany, (5)Department of Marine Biology, University of Haifa, Haifa, Israel, (6)Department of Biology, Boston University, Boston, MA, (7)Department of Physics, Boston University, Boston, MA

Metabolic network reconstruction is central to many aspects of metabolic analysis and engineering, and is essential for ongoing efforts towards mechanistic microbial ecosystem modeling. In general, the reconstruction process involves translating genomic information into a metabolic network that can provide insight into cell-wide metabolic fluxes and other phenotypes. A number of well-curated reconstructions exist for different model organisms, but a major challenge in the field is to reconstruct metabolic networks for diverse environmentally and medically relevant organisms. Various data sources including genomic, metabolomic, and phenotypic data can be used to guide metabolic network reconstruction. Additionally, there are several methods available for automatically generating draft reconstructions, including by using probabilistic genome annotations. However, few methods have directly attempted a comprehensive, automated integration of multiple data types, while at the same time explicitly dealing with the inherent uncertainty in the reconstruction process. We have developed a Markov chain Monte Carlo method to sample metabolic networks from a probability space shaped by the available data. Our method formalizes the data integration process and clarifies the representation of uncertainty by generating a natural ensemble of metabolic network reconstructions. As a proof of concept, we implemented this method on simulated genomic and metabolomic data from the *E. coli* iJO1366 metabolic network. Moving forward, we

are working on including additional data types, such as phenotypic assays, with the goal of applying this method to a large group of environmentally-relevant marine microbes.

Essential Metabolism for a Minimal Cell.

Marian Breuer

Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL

The question of the core requirements of cellular life led to the construction of the “minimal cell” JCVI-syn3A: A cell where practically all genes were removed that were not essential for robust growth in a stress-free laboratory environment. With only 493 genes in a 543 kbp genome, JCVI-syn3A has a genome smaller than that of any independently-replicating cell found in nature. It provides a versatile platform to study the basics of cellular life and is small enough that a complete description of all cellular functions can be pursued. Here, we present an extensively curated metabolic reconstruction and flux balance analysis (FBA) model of this minimal cell, using the vast amount of experimental information available on its natural precursor, *Mycoplasma mycoides capri*. The model, featuring 337 reactions involving 304 metabolites, is near-complete with 98 % of enzymatic reactions justified through gene assignments and/or experimental evidence, and agrees well with gene essentiality data from transposon mutagenesis experiments. The 155 genes included in the reconstruction have a high in vivo essentiality or quasi-essentiality of 92 %, compared to 79 % in silico essentiality—underscoring the minimality of the network. The reconstruction itself and the comparison of in vivo and in silico essentialities lead to new hypotheses on particular metabolic functions, suggesting specific experiments. Thus, the model provides a solid foundation for further experimental and computational studies on the minimal cell.

A Distance Measure for Heterogeneity Using Genome Scale Metabolic Networks.

Andrea Cabbia¹, Peter A. J. Hilbers¹, and Natal A. W. van Riel^{1,2}

(1)Biomedical Engineering, Eindhoven University of Technology, Eindhoven, Netherlands, (2)Academic Medical Center, University of Amsterdam, Amsterdam, Netherlands

Physiological differences in the aging process are inherently present in a population, and increase with age, affecting the risk of developing disabilities and age-related diseases [1]. Patient-Derived Genome-Scale Metabolic Models (PD-GSMM) are built from human

GSMM and experimental data, mostly transcriptomics and proteomics, belonging to single individuals. Personalized genome scale models have recently been used to plan individualized anti-cancer therapies [2], and to address the variability among cancer patients, identifying key genes involved in tumour growth [3]. Despite their success in cancer metabolism, is still not clear the extent to which PD-GSMMs are representations of individual metabolic features in physiological conditions, and how successful such models are in capturing inter-individual heterogeneity when dealing with subtler phenotypes such as ageing. Starting from microarray datasets of younger and older adults’ skeletal muscle gene expression, we developed the first collection of patient-derived genome scale metabolic models of ageing individuals’ myocytes, and used a data science approach to define a distance metric and assess the variability between metabolic models. This research is part of the PANINI project (Physical Activity and Nutrition INfluences in Aging), and has received funding from the European Union’s Horizon2020 programme, under the Marie Skłodowska-Curie grant agreement 675003.

[1] Yang Y and Lee L.C.

Dynamics and Heterogeneity in the Process of Human Frailty and Aging: Evidence From the U.S. Older Adult Population
J Gerontol B Psychol Sci Soc Sci. 2010 Mar; doi: 10.1093/geronb/gbp102

[2] Agren R. et al.

Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling.
Mol Syst Biol. 2014 Mar; doi: 10.1002/msb.145122.

[3] Uhlen M. et al.

A pathology atlas of the human cancer transcriptome.
Science. 2017 Aug 18; doi: 10.1126/science.aan2507.

Thermodynamics As an Optimization Goal for Metabolism: Prediction of Metabolite Levels, Rate Constants and Post-Translational Regulation.

William R. Cannon¹, Jeremy D. Zucker¹, Douglas J. Baxter², Neeraj Kumar¹, Jennifer M. Hurley³, Scott E. Baker⁴, and Jay C. Dunlap⁵

(1)Computational Biology, Pacific Northwest National Laboratory, Richland, WA, (2)Research Computing, Pacific Northwest National Laboratory, Richland, WA, (3)Department of Biological Sciences, Rensselaer Polytechnic Institute, Troy, NY, (4)Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, (5)Department of Molecular and Systems Biology, Geisel School of Medicine at Dartmouth, Hanover, VT

Nature selects those organisms that can reproduce the fastest while maintaining fitness and extracting the least

amount of energy from their environment. This problem can be formulated as a maximum entropy production rate problem that includes experimental and physical constraints. We report the application this approach for modeling biological systems in lieu of having *in vivo* rate constants. The method is applied in four steps: (1) a new constrained optimization approach based on Marcelin's 1910 mass action equation is used to obtain the maximum entropy distribution, (2) the predicted metabolite concentrations are compared to those generally expected from experiment using a loss function from which post-translational regulation of enzymes is inferred, (3) the system is re-optimized with the inferred regulation from which rate constants are determined from the metabolite concentrations and reaction fluxes, and finally (4) a full ODE-based, mass action simulation with rate parameters and allosteric regulation is obtained. From the last step, the power characteristics and resistance of each reaction can be determined. The method is applied to the central metabolism of *Neurospora crassa* and the flow of material through the three competing pathways of upper glycolysis, the non-oxidative pentose phosphate pathway, and the oxidative pentose phosphate pathway are evaluated as a function of the NADP/NADPH ratio. It is predicted that regulation of phosphofructokinase (PFK) and flow through the pentose phosphate pathway are essential for preventing an extreme level of fructose 1, 6-bisphosphate accumulation. Such an extreme level of fructose 1,6-bisphosphate would otherwise result in a glassy cytoplasm with limited diffusion, dramatically decreasing the entropy and energy production rate and, consequently, biological competitiveness.

Comparative Genomics and Network Modeling of Parasites.

Maureen A. Carey¹, Michal Stolarczyk², Ana Untariou³, Gregory L. Medlock⁴, Jennifer L. Guler², and Jason A. Papin⁴

(1)Microbiology, Immunology, and Cancer Biology, University of Virginia, Charlottesville, VA, (2)Biology, University of Virginia, Charlottesville, VA, (3)Biomedical Engineering, University of Virginia, Charlottesville, VA, (4)Department of Biomedical Engineering, University of Virginia, Charlottesville, VA

Eukaryotic parasites cause diseases with large global burdens, resulting in over one million deaths annually. To develop novel antiparasitic drugs, experimental model systems often use alternative parasite species that are easier to manipulate genetically or to observe *in vivo* in mice, extrapolating these observations to disease-causing parasites. However, characterization of functional differences between parasite species is limited to *post*

hoc and single-target studies, limiting the utility of these experimental models of disease and historic data. Each parasite genome encodes unique enzyme annotations; however, it is unclear whether these differences arise from divergent functions or incomplete genome annotation.

To address this challenge, we generated metabolic reconstructions from 160 parasite genomes, representing 38 genera and 111 species. Unsurprisingly, one of the largest genomes (*Chromera velia*) has the most unique reactions (35); however, the smallest genome (*Plasmodium billcollinsi*) has six unique reactions. Reconstruction and genome sizes are correlated, but even small networks contain unique features. We next applied our novel automated curation approach to leverage manual curation to improve reconstructions for related organisms. We used these semi-curated reconstructions to compare metabolic capacity and pathway utilization and generate hypotheses about species-specific functions. Here, we focus on the most lethal malaria parasite, *Plasmodium falciparum*, and the mouse model of severe malaria, *P. berghei*. For example, while *P. falciparum* grows in anucleated host erythrocytes, *berghei* prefers nucleated reticulocytes; *in silico*, *P. falciparum* has a reduced demand for host purines when compared to *P. berghei*, perhaps explaining this host cell preference and highlighting a pathway for which *P. berghei* is a poor model organism for *P. falciparum* antimalarial development. By performing these analyses with all 160 reconstructions, we developed a framework to identify metabolic discrepancies and commonalities between genera and species, facilitating comparison of experimental findings and optimizing model system selection to develop therapeutics for parasitic diseases.

Models Constrained with Transcriptomics or Proteomics Data Generate Discordant Predictions.

Maureen A. Carey¹, Ana Untariou², and Jason A. Papin³

(1)Microbiology, Immunology, and Cancer Biology, University of Virginia, Charlottesville, VA, (2)Biomedical Engineering, University of Virginia, Charlottesville, VA, (3)Department of Biomedical Engineering, University of Virginia, Charlottesville, VA

Modern high-throughput technologies have enabled detailed characterizations of an organism's phenotype; this large-scale molecular data allows for a systems-level analysis of biological conditions. Moreover, genome-scale metabolic reconstructions (GENREs) can be leveraged to interpret these data by integrating 'omics into a GENRE, resulting in a condition-specific models and the generation condition-specific predictions. Transcriptomic data have been most frequently used for data integration due to the

datasets' capacity for genome-wide coverage. However, regulation occurs at the translational level as well as transcriptional level, so proteomics-derived models may be used to generate more accurate condition-specific predictions than transcriptomic-derived models.

To investigate the functional differences between expression integration using either transcriptomic and proteomic data, matched datasets from five biological conditions (GSE65209 and PXD001659) were integrated into a GENRE (from Bosi, *et al.*, doi: 10.1073/pnas.1523199113) using GIMME. As transcriptomic datasets have higher coverage than many proteomics datasets, the effect of data coverage was investigated by random sampling subsets of the data. Predictions (specifically, reaction essentiality) from models generated with matched transcriptomic and proteomic data (*i.e.* collected from the same biological condition) and mismatched datasets (*i.e.* from different conditions) were compared; surprisingly, these matched predictions were no more similar than predictions generated from mismatched models. Transcriptomic-derived models predicted more essential reactions than proteomics models, even when controlling for scope (*i.e.* number of enzymes constrained), indicating transcriptomic-derived models are more constrained than proteomics-derived models. These results emphasize the importance of interpreting predictions as condition-, datatype-, and approach-specific results. Future work will explore the effects of alternative integration approaches and datasets.

Toward a Whole-Cell Model of Human Embryonic Stem Cells: A Composable Model of Metabolism Built on a Scalable Framework for Integrating Multiple Pathway Submodels.

Yin Hoon Chew, Andrew Sundstrom, and Jonathan R. Karr

Department of Genetics & Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York City, NY

Stem cell behaviors such as self-renewal and differentiation result from complex interactions among signaling, gene regulation, metabolism, and other pathways. To understand how these pathways collectively determine cell behavior, we aim to develop the first whole-cell computational model of a human cell, focusing on the well-characterized H1 human embryonic stem cell (hESC) line.

Our approach is to build and integrate submodels of each pathway. To begin, we have developed a scalable framework for integrating multiple pathway submodels and built a genome-scale submodel of single-cell metabolism.

The integration framework consists of state variables that represent the instantaneous configuration of a cell such as its mass, volume, and composition including the abundance and compartmental location of each molecular component. We developed the initial conditions for the framework by integrating extensive data including the DNA sequence of H1-hESCs; the location of each transcript, exon, and protein lifted from the hg19 reference; protein-complex compositions; macromolecular contents; small-metabolite concentrations; RNA expression; protein abundances; and the stoichiometry, catalysis, rate law, and kinetic data of each reaction.

We built the metabolism submodel by: (1) expanding Recon 2.2 to produce and recycle the metabolites that are needed and generated by other pathways, respectively; (2) refining the submodel for integration into a stochastic dynamical model of individual cells; and (3) using multi-omics and biochemical data to contextualize the submodel to represent H1-hESCs. The metabolism submodel correctly predicts the essentiality of 97% of metabolism-associated genes in hESCs. 87.5% of the false negative have non-metabolic functions, highlighting the need for integrated models that describe multiple pathways.

Going forward, we will create submodels of other pathways, integrate the submodels, and use the integrated model to investigate stem cell self-renewal. Ultimately, we believe whole-cell models have the potential to transform basic science and medicine through unprecedented *in silico* experiments personalized to each patient.

Metabolic Analysis of Nitrogen Fixation in the Soybean Micro-Symbiont *Sinorhizobium fredii* Using Constraint-Based Modeling.

Carolina A. Contador¹, Qianwen Wang¹, Siu Hung Joshua Chan², and Hon-Ming Lam¹

(1)School of Life Sciences and Center for Soybean Research of the Partner State Key Laboratory of Agrobiotechnology, Chinese University of Hong Kong, Hong Kong, China, (2)Department of Chemical and Biological Engineering, Colorado State University, Fort Collins, CO

Rhizobia are soil bacteria known for being able to establish symbiosis with diverse host plants. In this context, *Sinorhizobium fredii* is a soil bacterium able to form nitrogen-fixing nodules with diverse legumes such as soybean. In particular, *S. fredii* strain CCBAU45436 represents one of the dominant sublineage of *S. fredii* that nodulates soybeans in alkaline-saline soils in the Huang-Huai-Hai Plain region in China. Here, we present

POSTER ABSTRACTS

a genome-scale metabolic model for the soybean micro-symbiont *Sinorhizobium fredii* strain CCBAU45436. The model was manually curated based on physiological, biochemical information and integration of high-throughput data. A symbiosis reaction was defined to represent the specific exchange of nutrients between the bacteroid and soybean. The symbiosis reaction together with the nitrogen fixation reaction defines the objective function to describe the stage of nitrogen fixation. The model developed can be used to systematically analyze symbiotic factors and predict essential genes for nitrogen fixation. In addition, the model can be expanded to represent different life stages of the bacteria to try to understand niche adaptation. No reconstructions are currently available for *S. fredii* strains despite of their potential benefits to sustainable agriculture and ecosystem.

This work is supported by the Hong Kong Research Grants Council GRF (14108014) and Area of Excellence Scheme (AoE/M-403/16) awarded to H.-M.L.

Metabolic Network Reconstruction for the Pan-Genome: A Scalable Method to Get High-Quality Metabolic Models across the Tree of Life.

Kevin Correia and Radhakrishnan Mahadevan
Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON, Canada

Background

Genome-scale metabolic models (GEM's) are a valuable tool used to study the metabolism of organisms with metabolic simulations, but they also provide a scaffold for 'omics data integration. The drop in cost for genome sequencing has lead to an exponential increase in the diversity of sequenced genomes, but the number of curated GEM's has not kept pace. This gap hinders our ability to study physiology across the tree of life. Furthermore, our review of published GEM's has found that metabolic models contain significant commission and omission errors in central metabolism. To address these quantity and quality GEM issues, we propose open and transparent efforts to curate the pan-genome, pan-reaction, and pan-metabolome of larger groups of organisms by research communities, rather than for a single species. We outline our approach for budding yeasts.

Method

We created a consensus metabolic network from 13 fungi/yeasts GEM's spanning seven species. We added additional reactions to the network to account for growth substrates and non-natural enzyme activities in the fungal literature that were not present in biochemical databases.

Reactions were annotated with ortholog ID's from AYbRAH, our fungal ortholog database, to generate 33 fungi/yeasts GEM's from the subkingdom Dikarya.

Results

The fungal pan-GEM contains 2224 reactions, 2696 metabolites, 1543 orthologs, and 10 compartments. The gene coverage in the species models created from our pan-GEM is higher than published GEM's. Pathways not previously included in published GEM's, such as degradation of alkanes and aromatics, are captured using our approach.

Conclusion

Maintaining a metabolic network reconstruction annotated with a curated ortholog database at a higher pan-genome level improves the quality and quantity of GEM's. Capturing accurate ortholog-reaction-protein associations minimizes commission and omission errors, and the ability to scale GEM's to more branches of the tree of life.

Reconciling Barseq Datasets for *Z. Mobilis* Using a Genome-Scale Metabolic Model.

Dylan Courtney¹, Wai Kit Ong¹, Shu Pan¹, and Jennifer Reed²

(1)Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI, (2)Chemical and Biological Engineering, University of Wisconsin Madison, Madison, WI

Zymomonas mobilis is an industrially relevant, gram negative, ethanologen known for very high glycolytic fluxes, high ethanol production, and low biomass yields. Development of engineered strains of *Z. mobilis* has been relatively slow, and future efforts stand to benefit from being informed and directed by updated genome-scale metabolic models. Here we present iZM4_472, an updated model of *Z. mobilis* ZM4, and use it to analyze and reconcile BarSeq datasets from pooled mutant fitness experiments [Deutschbauer et al. *Journal of bacteriology* (2014)]. iZM4_472 contains 757 metabolic and transport reactions (of which 611 have GPR associations), 472 genes, and 632 unique metabolites, making it the largest and best annotated model of *Z. mobilis* to date. Data from reported pooled mutant experiments was used to assess the accuracy of gene essentiality predictions and identify genes associated with gap-filled reactions in the model. Several discrepancies between gene essentiality predictions and experimental results were caused by gene duplication, polar effects, or mis-mapped barcodes in the mutant library; thus, highlighting potential challenges with applying these high-throughput datasets to improve metabolic models. Correlations between transposon

mutant fitness across 492 experiments were used to identify gene candidates for gap-filled reactions involved in histidine biosynthesis. Additional genes for reactions involved in tyrosine, biotin, ubiquinone, and pyridoxal 5'-phosphate synthesis were identified and confirmed through a combination of *E. coli* mutant complementation experiments, and our previously developed model-enabled gene search (MEGS) approach [Pan et al. *Journal of Biological Chemistry* (2017)]. Finally, flux coupling analysis of iZM4_472 was used to analyze the fitness scores of all 492 experiments in the reported BarSeq datasets to identify metabolic modules where mutant phenotypes were poorly correlated. These modeling efforts help pinpoint where additional metabolic knowledge gaps of *Z. mobilis* metabolism reside.

Developing and Evaluating Integrated Metabolic Regulatory Models for Microbial Life.

Molly Creagar¹, Thomas Kamp², Ruby Fore³, Derek Friend⁴, Craig Disselkoen⁵, Christopher S. Henry⁶, Matthew DeJongh⁷, Aaron Best⁸, William Lindsey², and Nathan Tittle⁹

(1)University of San Francisco, San Francisco, CA, (2)Dordt College, Sioux Center, IA, (3)Brown University, Providence, RI, (4)University of Nevada, Reno, Reno, NV, (5)University of California, San Diego, San Diego, CA, (6)Argonne National Laboratory, Argonne, IL, (7)Computer Science, Hope College, Holland, MI, (8)Biology, Hope College, Holland, MI, (9)Mathematics, Dordt College, Sioux Center, IA

A number of papers have described methods to include transcriptional regulatory networks (TRNs) in the development of metabolic models. However, in general, these models do not allow for statistical uncertainty in TRN interactions. Furthermore, methods which do allow for uncertainty do so in a non-rigorous manner which, effectively, downweights prior transcriptomics data to the point where it has little impact on the resulting model. To combat these limitations, we developed an integrated metabolic regulatory model (iMRM). This novel approach explicitly models the statistical uncertainty in gene state activity inferences by using gene activity state estimates, and the iMRM process measures the flux through reactions for a sole single-cell organism. This allows for benefits such as differential thresholds for gene activity for different genes, a stronger correlation with experimental flux data, and the application of a non-Boolean confidence metric. Furthermore, this is the first such metabolic model to be grounded on a statistically rigorous foundation.

Constructing a High Quality Genome-Scale Metabolic Network for *Streptomyces lividans* TK24.

Wouter Daniels¹, Kjerstin De Winter¹, Justien Ghesquière¹, Kenneth Simoens¹, Tobias Busche², Christian Rückert², Bart Lievens³, Tomer Shlomi⁴, Jörn Kalinowski², and Kristel Bernaerts¹

(1)Bio- and Chemical Systems Technology, Reactor Engineering, and Safety (CREaS), Department of Chemical Engineering, KU Leuven, University of Leuven, Leuven, Belgium, (2)Center for Biotechnology (CeBiTec), University of Bielefeld, Bielefeld, Germany, (3)Laboratory for Process Microbial Ecology and Bioinspirational Management, Department of Microbial and Molecular Systems, KU Leuven, University of Leuven, Sint-Katelijne Waver, Belgium, (4)Computer Science, Technion IIT, Haifa, Israel

Streptomyces lividans is increasingly attracting attention as a potential industrial host for heterologous production of proteins and secondary metabolites. The current interest in *S. lividans*, combined with the increasing relevance of constraint-based modeling for strain improvement, urges for the creation of a high-quality genome-scale metabolic network (GSMN).

The most recent GSMN of model organism *S. coelicolor*—with high genomic similarity to *S. lividans*—is taken as a starting point for model construction. Comparative genomics is used for the conversion into *S. lividans* gene-reaction relationships. Reactions which lack *S. lividans* enzymes are removed when this did not result in dead-end reactions with expressed (RNA-seq) enzymes or loss of growth on known growth substrates did not occur. Furthermore, KEGG reaction identifiers are added to the model, and KEGG metabolite identifiers and Enzyme Commission numbers are re-evaluated.

The model is extended using genomic, metabolomic and phenomic data. From the genome, metabolic reactions catalyzed by predicted enzymes not yet included in the model are added, as well as their substrates if they are not yet included as metabolites. Furthermore, additional metabolites, detected through LC-MS, are linked to the model by addition of metabolic reactions whenever this is reasonable with respect to the in silico phenotype and literature. Biolog Phenotype MicroArrays are used for testing respiration of *S. lividans* TK24 on 348 substrates and subsequent gap-filling of the model. To maintain high quality, the library for gap-filling is limited to 13 selected published bacterial GSMNs. Gap-filling was performed by searching for minimal sets of reactions that restore growth on a given substrate through mixed integer linear programming, and selecting the most appropriate solution. [Research funded by EU FP7-KBBE-2013-7 StrepSynth (grant n°613877).]

Construction of a Dynamic Flux Balance Analysis Model for *Streptococcus Gordonii*.

Kjerstin De Winter¹, Wim Teughels², Steffen Waldherr³, and Kristel Bernaerts⁴

(1)Department of Chemical Engineering, KU Leuven, University of Leuven, Leuven, Belgium, (2)Periodontology and Oral Microbiology, Department of oral health science, KU Leuven, University of Leuven, Leuven, Belgium, (3)Bio- and Chemical Systems Technology, Reactor Engineering and Safety (CREaS), Department of Chemical Engineering, KU Leuven, University of Leuven, Leuven, Belgium, (4)Bio- and Chemical Systems Technology, Reactor Engineering, and Safety (CREaS), Department of Chemical Engineering, KU Leuven, University of Leuven, Leuven, Belgium

Periodontitis is a chronic and inflammatory disease of the tooth supporting tissues. Environmental changes can bring oral biofilms from a healthy state to dysbiosis, which potentially evolves to periodontitis. Overall, we aim to construct a biofilm community model that brings more insight in dysbiosis of the oral biofilm and the influence of environmental factors. We develop a dynamic flux balance analysis (DFBA) model for *Streptococcus gordonii* which forms the first step in the development of an oral biofilm community model linked to periodontitis.

S. gordonii produces H_2O_2 and lactate, inhibiting pathogen outgrowth and enabling cross-feeding in the biofilm, respectively. First, the automated reconstruction taken from the AGORA database^[1] was curated using literature information and databases (e.g., KEGG, RHEA, BIGG). FBA was used to evaluate growth rates, and to analyze carbon routing through the central carbon metabolism and amino acid biosynthesis pathways, resulting in a GSMN with 1278 reactions and 1078 metabolites, reproducing realistic FBA predictions. Secondly, the GSMN was transformed into a planktonic model in DFBA^{lab}^[2]. Simulations with glucose as main C-source predict the expected coupled production of lactate and H_2O_2 , and show that protoheme depletion diverts carbon resources to lactate, thus triggering lactate production. Protoheme is an essential compound that cannot be synthesized by *S. gordonii*. H_2O_2 production is predicted with and without protoheme limitation, although concentrations are low. Analyses of growth on amino acids showed enhanced production of H_2O_2 for growth on glucose and arginine, without influencing lactate production. Finally, DFBA simulations in a *S. gordonii* biofilm are established. Simulations still show the coupled production of lactate and H_2O_2 . In a following stage, additional community members are added to the model.

^[1]S. Magnúsdóttir et al., Nat. Biotechnol., vol. 35, no. 1, 2017.

^[2]J. A. Gomez, K. Höffner, and P. I. Barton, BMC Bioinformatics,

vol. 15, no. 1, 2014.

Identification and Analysis of Bacterial Genomic Metabolic Signatures in the Department of Energy's Systems Biology Knowledgebase.

Matthew DeJongh¹, Megan Grimes¹, Lauren Satkiewicz¹, Nathan Tintle², and Aaron Best³

(1)Computer Science, Hope College, Holland, MI, (2) Mathematics, Dordt College, Sioux Center, IA, (3)Biology, Hope College, Holland, MI

With continued rapid growth in the number and quality of fully sequenced and accurately annotated bacterial genomes, we have an unprecedented opportunity to understand metabolic diversity. In previous work, we selected 101 diverse and representative completely sequenced bacteria and implemented a manual curation effort to identify 846 unique metabolic variants present in these bacteria. The presence or absence of these variants act as a metabolic signature for each of the bacteria, which can then be used to cluster them into a metabolic tree and analyze similarities and differences between and across bacterial groups. In our current work, we are developing an application to automate the assignment of metabolic variants to bacterial genomes in the Department of Energy's Systems Biology Knowledgebase (KBase). This will enable us to expand the set of bacteria to which this approach is applied and use the resulting tree to test broad questions about metabolic diversity and complexity across the bacterial tree of life.

This work is funded by the National Science Foundation awards MCB-1716285 to Hope College and MCB-1715211 to Dordt College.

References:

Nathaniel Bowerman, Nathan Tintle, Matthew DeJongh, Aaron A Best. Identification and Analysis of Bacterial Genomic Metabolic Signatures. Pacific Symposium on Biocomputing 2017, pp. 3-14. 2017.

AP Arkin, RW Cottingham, CS Henry et al. KBase: The United States Department of Energy Systems Biology Knowledgebase. Nature Biotechnology, Vol 36, pp. 566-569, 2018.

Untangling Dynamic Protein Behaviour with Rule-Based Modelling and Sensitivity Analysis.

Robin Denhardt-Eriksson

EPFL - LCSB, Lausanne, Switzerland

The endoplasmic reticulum (ER) is key to synthesizing mature, functional proteins. Its complex structure consisting of large thin sheets densely packed around the nucleus and elongated tubules reaching throughout the cell is key to its function, and is dynamically regulated in

response to cellular signals. CLIMP-63 is the only known protein to play a role in shaping the ER, while also being the target of palmitoylation, one of the few reversible post-translational lipid modifications. In addition to its role in ER structure regulation, CLIMP-63 plays a role as a receptor at the cell surface, where it is known to bind to several ligands. These multiple roles make untangling the regulation of CLIMP-63 abundance and localisation a non-trivial task. We employed rule-based modelling to build a dynamical representation of CLIMP-63 birth, dimerisation, palmitoylation, trafficking and degradation. Experimental data in the form of time-dependent radioactive labelling was used to calibrate the model with heuristic optimisation methods. Multiple predictions and hypothesis were generated as to the regulation of CLIMP-63 levels and localisation. Sensitivity Analysis was used on the calibrated model to maximise the extraction of useful knowledge and hypotheses.

Towards a Genome-Wide Transport Systems Encoding Genes Tracker.

Lagoa D. (1), Liu F. (2), Ferreira E.C. (1), Faria J. (2), Henry C. (2), **Dias O. (1)**

(1) CEB - Centre of Biological Engineering, University of Minho

(2) Mathematics and Computer Science Division, Argonne National Laboratory

The Transport Systems Tracker (*TranSyT*) is a tool to identify transport systems and the compounds carried across membranes by those systems, annotating the proteins with the Transporter Classification Database (TCDB)¹ families numbers. In addition, this tool also generates the respective transport reactions.

TranSyT's internal database contains every possible reaction for all compounds described in each TCDB entry, as well as their hierarchical children metabolites.

TranSyT starts by assigning identifiers to metabolites, as TCDB does not provide cross links to any database. These metabolites are identified by searching a BioDB, a database developed in-house, for all possible names and synonyms. The relationships between compounds (hierarchical ontology) was also determined using BioDB, which combines information retrieved from several sources such as ModelSEED, KEGG, MetaCyc and BiGG.

The second step involves retrieving the TC families information from TCDB, which allows to find the suitable transport type for each metabolite (e.g. symport and antiport) and direction (in/out or out/in).

All generated information is then stored in a Neo4j graph database, which allows *TranSyT* to retrieve the required reactions for the genome in study, saving time and resources.

All sequences of the genome in study are aligned against the whole set of records retrieved from TCDB, containing the sequences of well-known transport systems, using BLAST+.

After the alignments, only genes having evidences of encoding transport systems are associated to transport reactions. The compartmentalization of such reactions is also possible, as the direction of the reaction was beforehand defined.

TranSyT was developed in Java and is a the next iteration of the previously developed TRIAGE² tool.

References:

- 1 Saier, M. H. et al. Nucleic Acids Res. 34, D181–D186 (2006)
- 2 Dias, O. et al. IEEE/ACM Trans. Comput. Biol. Bioinforma. XX, 1–1 (2016)

Predicting Shifts in Cardiomyocyte Metabolism during Heart Failure.

Bonnie Dougherty¹, Kristopher Rawls¹, Glynis Kolling¹, Vinnakota Kalyan², Venkat Pannala², Anders Wallqvist², and Jason A. Papin¹

(1)Department of Biomedical Engineering, University of Virginia, Charlottesville, VA, (2)Department of Defense Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Fort Detrick, MD

Contextualizing the human metabolic network to model the metabolic states of specific tissue- and cell-types has generated novel hypotheses and tremendous insight into the role of metabolism in disease. Here, we use a recently published human GENRE, iHsa, to model cardiomyocyte metabolism through the integration of proteomics data available in the Human Protein Atlas (HPA). We validate the cardiomyocyte model through completion of pre-defined cardiomyocyte relevant tasks as well as evaluation of general tasks published with the original human metabolic network reconstruction. Both sets of tasks have provided a starting point for manual curation of the cardiomyocyte model. Next, we use the cardiomyocyte model to identify reactions, pathways, and genes differentially affected in heart failure through the utilization of different carbon sources for ATP production, an established phenotype in heart failure. The cardiomyocyte model demonstrates enrichment for carbon-source specific pathways, demonstrating the ability of the model

to connect genotype to phenotype using the phenotypic changes observed in carbon substrate utilization that occur in heart failure.

Metabolic Network Reconstruction of *Treponema Pallidum* Spp. *Pallidum*.

Silvia Morini^{1,2}, Isabella Casini^{3,4}, Thomas M. Hamm^{1,5}, Kay Nieselt^{5,6}, and **Andreas Dräger**^{1,5}

(1)Computational Systems Biology of Infection and Antimicrobial-Resistant Pathogens, Center for Bioinformatics Tübingen (ZBIT), Tübingen, Germany, (2)Department of biotechnologies and biosciences, Università degli Studi di Milano - Bicocca, Milano, Italy, (3)Biological and Environmental Engineering, Max Planck Institute for Developmental Biology, Tübingen, Germany, (4)Department of Geosciences, University of Tübingen, Tübingen, Germany, (5)Department of Computer Science, University of Tübingen, Tübingen, Germany, (6)Integrative Transcriptomics, Center for Bioinformatics Tübingen (ZBIT), Tübingen, Germany

Background: Since the discovery, made in 1905 by Schaudinn and Hoffmann, of *Treponema pallidum* ssp. *pallidum* as the etiologic agent of syphilis, medicine has made significant progress against this disease. Yet, despite the availability of diagnostic tests and a therapy based on antibiotics, the world has not stopped being burdened by syphilis, that has been re-emerging globally over the last few decades (WHO Report, 2008), for which no vaccine is still available and which, moreover, in its early stages enhances the transmission of HIV. Continuous in vitro culture of this organism has still not been achieved, imposing a substantial roadblock to its experimental inspection, and even the sequencing of its genome (Fraser et al., 1998) did not yield an obvious solution to the cultivation problem. While much has already been tried on the laboratory bench, this pathogen has still not (to our knowledge) been tackled using a systems biology approach.

Results: Here, we present a first manually curated draft reconstruction of the metabolic network in *Treponema pallidum* ssp. *pallidum* towards a genome-scale metabolic model (GEM). At this time, the model iSM161 comprises 161 genes of 1,039 predicted open reading frames that are responsible for 239 reactions of 277 metabolites. The model is still under development and steadily updated. For the reconstruction, COBRApy has been used, where subsystem information is added and parsed as SBML groups extension using libSBML.

Discussion: Using this reconstruction, together with the application of COBRA methods, we anticipate to gather new insights into the pathogen's physiology and

pathology, and in how this spirochete, which has earned the designation of “stealth pathogen,” succeeds in making a living and eluding human's immune defenses as well as cultivation attempts. It is planned to make the model available to the community in SBML format.

Modeling of Potentially Virulence-Associated Metabolic Pathways in *Pseudomonas Aeruginosa* PA14 Including Experimental Verification.

Alina Renz^{1,2}, Erwin Bohn³, Monika Schütz³, and **Andreas Dräger**^{1,2}

(1)Computational Systems Biology of Infection and Antimicrobial-Resistant Pathogens, Center for Bioinformatics Tübingen (ZBIT), Tübingen, Germany, (2) Department of Computer Science, University of Tübingen, Tübingen, Germany, (3)Institute for Medical Microbiology and Hygiene, University Hospital Tübingen, Tübingen, Germany

According to the report of the ‘Antimicrobial resistance surveillance in Europe (2015),’ *P. aeruginosa* is an opportunistic, human pathogen that causes many infections in hospitalized patients with immune defects or impairments. Since it is difficult to control *P. aeruginosa* in hospitals, it can cause hospital-acquired pneumonia [1]. Some strains of *P. aeruginosa* seem to be associated with higher mortality than others. With the help of a published genome-scale model of PA14 [2] and sequencing data of a highly pathogenic patient strain that was recently isolated in Tübingen, metabolic differences between the laboratory and patient strain shall be identified and subsequently verified in a laboratory experiment.

First, differences in the sequences of the two strains were identified by performing SNP analysis. These differences were then used to find metabolic alterations affecting the virulence. By using FBA and *in-silico* gene knock-out, three genes were identified that could be responsible for the difference in pathogenicity between the laboratory and the highly pathogenic patient strain. These genes affect metabolic reactions that are associated with virulence in *P. aeruginosa*. Gene knock-outs of these three genes are momentarily being performed in a laboratory experiment to verify their metabolic relevance in virulence.

However, only a fraction of the genes of *P. aeruginosa* is included in the published model. Many additional genes are associated with virulence, differ between the sequenced laboratory and patient strains, or both. Therefore, the model is being extended to increase its predictive value.

European Centre for Disease Prevention and Control.
Antimicrobial resistance surveillance in Europe 2015. Annual
Report of the European Antimicrobial Resistance Surveillance

Network (EARS-Net). Stockholm: ECDC; 2017. <https://ecdc.europa.eu/en/publications-data/antimicrobial-resistance-surveillance-europe-2015#no-link>
 Bartell JA, Blazier AS, Yen P, Thøgersen JC, Jelsbak L, Goldberg JB, and Papin JA. Reconstruction of the metabolic network of *Pseudomonas aeruginosa* interrogate virulence factor synthesis, Nature Communications (2017). doi:10.1038/ncomms14631

Representation of ME-Models in SbmL

Marc Alexander Voigt^{1,2}, Colton J. Lloyd³, Laurence Yang³, Zachary A. King³, Oliver Kohlbacher^{2,4,5,6}, Kay Nieselt^{2,7}, and Andreas Dräger^{1,2}

(1)Computational Systems Biology of Infection and Antimicrobial-Resistant Pathogens, Center for Bioinformatics Tübingen (ZBIT), Tübingen, Germany, (2)Department of Computer Science, University of Tübingen, Tübingen, Germany, (3)Bioengineering, University of California, San Diego, La Jolla, CA, (4)Max Planck Institute for Developmental Biology, Tübingen, Germany, (5)Center for Personalized Medicine (ZPM), University of Tübingen, Tübingen, Germany, (6)Quantitative Biology Center (QBiC), University of Tübingen, Tübingen, Germany, (7)Integrative Transcriptomics, Center for Bioinformatics Tübingen (ZBIT), Tübingen, Germany

Metabolism and Expression models (ME-models) [1] are a constraint-based modeling approach, which explicitly accounts for the cost of macromolecular biosynthesis. This allows ME-models to investigate genotype-phenotype relationships with a quantitative incorporation of ‘-omics’ data. Common standards are a prerequisite for interoperability of systems biology tools. Novel approaches often require additional or changed data structures that cannot be represented in existing standards. This has been the case for ME-models. ME-models are powerful tools, but a lack of standards for encoding and creation, and the increased complexity of these models prevented widespread use.

SBMLme is an extension of current model encoding standards that enables SBML representations of ME-models that were reconstructed using COBRAme [2]. A prototype of the extension has been created in Java together with a standalone, bi-directional converter, between this extension and COBRAme’s model storage format. The converter showed that SBMLme could fully and correctly encode a COBRAme model.

The use of SBMLme enables ME-models to be shared more efficiently, and a wider variety of tools to access ME-models. This should promote the propagation of ME-models. SBMLme may be used as a proof-of-concept for an official SBML package for ME-models. The standardization process requires ongoing discussion, the

identification of a consensus SBML draft standard, and the implementation of the existing validation rules.

SBMLme is freely available at <https://github.com/draeger-lab/SBMLme> under the terms of the MIT license.

Ines Thiele, Ronan M. T. Fleming, Richard Que, Aarash Bordbar, Dinh Diep, and Bernhard O. Palsson. Multiscale Modeling of Metabolism and Macromolecular Synthesis in *E. coli* and Its Application to the Evolution of Codon Usage. PLOS ONE, 7(9):1-18, 09 2012

Colton J Lloyd, Ali Ebrahim, Laurence Yang, Zachary Andrew King, Edward Catolui, Edward J O'Brien, Joanne K Liu, and Bernhard O Palsson. COBRAme: A Computational Framework for Models of Metabolism and Gene Expression. bioRxiv, 2017.

Automated Reconstruction, Evaluation and Comparison of Diverse Genome Scale Fungal Models.

Janaka N Edirisinghe^{1,2}, José P. Faria¹, Sara Calhoun³, Filipe Liu¹, Igor Grigoriev³, and Christopher S. Henry^{1,2}

(1)Argonne National Laboratory, Argonne, IL, (2)Computation Institute, University of Chicago, Chicago, IL, (3)Lawrence Berkeley National Laboratory, Berkeley, CA

Genome-scale metabolic models offer an efficient mechanism for predicting phenotypes across many environmental conditions. These models are also key tools in understanding interactions between species in the environment, including microbial, plant, and fungal species. While methods have emerged for the automated reconstruction of microbial and plant models from genomic data, the construction of high-quality fungal models remains a challenge. Here we introduce a new method for constructing genome-scale fungal models in an automated fashion based on a curated set of reactions and gene associations that are derived from 14 published fungal metabolic models. As the basis for our method, we produced a fungal model template that encompasses the biochemistry data from the published fungal models and the structural annotations from the associated fungal genomes. Using structural annotations of a fungal genome we compute a set of orthologous proteins against the curated fungal template in order to assert the presence or absence of specific biochemical reactions and pathways. Once the orthologous protein families are determined, the related biochemistry data is propagated to construct a new draft metabolic model. This method is deployed in the KBase (<https://narrative.kbase.us/>) as an app called “Build Fungal Model”. We have built tools in KBase to produce high-quality metabolic models on prokaryotes and plants and new approaches to identify trophic interactions between species. Using these methods, we are able to study how fungal-bacterial, plant-fungal and microbiomes evolve as a sustainable community.

POSTER ABSTRACTS

We applied “Build Fungal Model” app to build draft models for more than 100 fungal genomes imported from the Joint Genome Institute (JGI) MycoCosm resource (<https://genome.jgi.doe.gov/programs/fungi/index.jsf>).

We compared these models side-by-side, exploring how each genome overlaps with our curated model template and plotting model variance along the phylogenetic tree of fungal genomes. All draft fungal models are available for download via a KBase Narrative.

Metabolic Supply and Demand Constrain Optimal Regulatory Strategies: Structural Evidence from the E coli Small Molecule Regulatory Network.

Christian Euler¹ and Radhakrishna Mahadevan²

(1)Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON, Canada, (2)Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON, Canada

Balancing supply and demand is a life or death-or-task for bacteria. In resource-limited environments, they must be able to precisely coordinate metabolic flux distributions to balance the demands required for optimal growth without wasting costly metabolites. To this end, they have evolved complex, layered regulatory systems which respond to changing external contexts. While the structure and function of hierarchical regulatory mechanisms (transcription and translation) have been characterized extensively, regulation at the metabolic level remains less well understood. To address this, we reconstructed a genome-scale small-molecule regulatory network (SMRN) for E coli to understand the role small-molecule regulation plays in balancing supply and demand. Structural analysis of our SMRN revealed a dense network of interactions with layers of organization. We found that at the local level, SMR interactions are organized into well-understood feedforward loop motifs that provide specific input-output relationships between fluxes through relatively distant metabolic reactions. Functionalization of these motifs demonstrates that the SMRN has the capacity to influence supply and demand balancing in two ways: by coordinating signals within the core metabolism and effecting tertiary fluxes based on these signals; and by directly balancing fluxes of metabolites which are essential for biomass production. We explore the constraints under which this regulatory system is likely to be active to show that it can work synergistically with transcriptional regulation and that protein degradation rate is the dominant factor which determines the optimal balance of control burden across these layers of regulation.

Improving High Throughput Genome-Scale Metabolic Model Reconstruction and Validation with Tnseq Data Using Modelseed 2.

José P. Faria¹, Janaka N Edirisinghe¹, Filipe Liu², Samuel M.D. Seaver¹, James G. Jeffries¹, Qizh Zhang¹, Pamela Weisenhorn¹, Boris Sadkhin¹, Nidhi Gupta¹, Tian Gu¹, and Christopher S. Henry¹

(1)Argonne National Laboratory, Argonne, IL, (2)Centre of Biological Engineering, University of Minho, Braga, Portugal

The Department of Energy Systems Biology Knowledgebase (KBase) is a platform designed to solve the grand challenges of Systems Biology. KBase has implemented bioinformatics tools that allow for multiple workflows including genome annotation, comparative genomics, and metabolic modeling. We selected a phylogenetically diverse set of approximately 1000 genomes and constructed draft genome-scale metabolic models (GSMMs) using the ModelSEED pipeline implemented in KBase. We used these 1000 genomes as a test set to improve the quality of models produced by the ModelSEED. First, we curated our mapping of RAST functional roles to biochemistry by reconciling with data mined from KEGG and published metabolic models; we corrected errors in our reaction reversibility assertions to improve overall model constraints; we applied a new method to predict auxotrophy across all 1000 genomes to predict improved gapfilling media; we refined our gapfilling procedure to prevent draft models from our pipeline from overproducing ATP; and we process all models through the Memote pipeline, accompanying complete reconstructions with Memote reports. We show how all of our pipeline improvements increase the number of gene associations, decrease the number of gapfilled reactions, improve the accuracy of growth and ATP production yield predictions, and decrease the number of blocked reactions across all models. The addition of Memote to our pipeline enables us to provide a measure for model quality that is consistent across reconstruction platforms. We show how auxotrophy, and pathway presence varies across our 1000 training-set genomes along the phylogenetic tree. We also plot model quality across the phylogenetic tree, identify taxa where model quality is lower. Finally, we select five specific genomes for which comprehensive TN-seq data is available, and we compare model predictions of all data with experimental results, showing significant improvement in accuracy between models generated by the original ModelSEED and models from ModelSEED 2.

A Deep Dive into False Positive Gene Essentiality Predictions with Modeling and Multi-Omics Analysis.

Adam M. Feist

Bioengineering, University of California, San Diego, La Jolla, CA

Essentiality assays are commonly practiced as important tools for the discovery of gene functions. Growth/no growth screens of single gene knockout strain collections are often utilized to test the predictive power of COBRA models. False positive predictions occur when computational analysis predicts a gene to be non-essential, however experimental screens deem the gene to be essential. In this study, we explored the definition of conditional essentiality from a phenotypic and genomic perspective. Gene-deletion strains associated with false positive predictions of gene essentiality on defined minimal medium for *E. coli* were targeted for extended growth tests followed by population sequencing. Of the twenty false positive strains available and confirmed from the Keio knock-out collection, 11 strains were shown to grow with longer incubation periods making these actual true positives. These strains grew reproducibly with a diverse range of growth phenotypes. It was found that 9 out of 11 of the false positive strains that grew acquired mutations in at least one replicate experiment and the types of mutations ranged from SNPs and small indels associated with regulatory or metabolic elements to large regions of genome duplication. Comparison of the detected adaptive mutations and modeling predictions of alternate pathways and isozymes suggested agreement for the observed growth phenotype for 6 out of the 9 cases where mutations were observed, which was further validated by analysis of the transcriptomes of selected strains. In conclusion, longer-term growth experiments followed by whole genome sequencing and transcriptome analysis can provide a better understanding of conditional gene essentiality and mechanisms of adaptation to such perturbations. Compensatory mutations are largely reproducible mechanisms and are in agreement with genome-scale modeling predictions to loss of function gene deletion events.

A Multiobjective Strain Design Platform for Modular Cell Engineering.

Sergio Garcia¹ and **Cong T. Trinh²**

(1)Chemical and Biomolecular Engineering, University of Tennessee, Knoxville, TN, (2)Department of Chemical and Biomolecular Engineering, University of Tennessee, Knoxville, TN

Metabolic engineering has enabled the use of microbial cell factories for industrial production of biochemicals.

However, developing an optimal strain for synthesis of one product with the conventional strategy is laborious and expensive. To accelerate and reduce the cost of strain engineering, we formulate the modular cell (MODCELL) design principles by exploiting the modular organization of metabolic networks and combinatorial possibilities of metabolic modules that enable the synthesis of a large space of biochemicals. Using multiobjective optimization methods, we develop novel algorithms to implement the MODCELL design for genome-scale metabolic networks and the associated software ModCell2.0. We demonstrate ModCell2.0 for design, construction, and validation of an *E. coli* modular cell for combinatorial synthesis of biochemicals, e.g., alcohols and bioesters from fermentable sugars and organic wastes. We envision MODCELL will provide a useful platform for modular cell engineering.

Development of an Accelerated Workflow for Parameterizing Kinetic Models of Metabolism.

Satyakam Dash, Saratram Gopalakrishnan, and Costas D. Maranas

Department of Chemical Engineering, The Pennsylvania State University, University Park, PA

Kinetic models can simulate an organism's metabolism under genetic perturbations by capturing the changes in metabolite pools and the impact on reaction fluxes in the form of regulations. The challenge of estimating kinetic parameters can be resolved using the Ensemble modeling (EM) approach which samples parameters consistent with a training dataset. However, the EM approach implemented using GA is computationally expensive. Thus, we have described here a gradient-based kinetic parameterization workflow that exploits the strengths and overcomes the shortcomings of the EM approach. First, the kinetic parameters are anchored to the steady-state metabolic flux distribution in the wild-type strain by exploiting the rate laws governed by mass-action kinetics relating fluxes through elementary reactions to sampled enzyme levels and kinetic parameters. Following this, a genetic perturbation is introduced by adjusting the total abundance of specific enzymes relative to the wild-type strain so as to reflect the *in vivo* genetic perturbation in the form of up/down-regulations or gene deletions. The steady-state fluxes in the mutant conditions are evaluated by decomposing the underlying system of bilinear equations into two linear sub-problems and iterating between them until convergence is achieved. Based on the deviation of predicted fluxes from measured data, kinetic parameters are updated using a Newtonian step until optimality is achieved. This accelerated procedure has been applied to a *C. thermocellum* core model

POSTER ABSTRACTS

(106 reactions, 72 metabolites, 674 elementary kinetic parameters, and 4 mutant conditions) and an *E. coli* near genome-scale model (319 reactions, 276 metabolites, 2536 elementary kinetic parameters and 7 mutant conditions). Parameterization was completed in 12 hours for *C. thermocellum* and 48 hours for *E. coli*, compared to a computational time requirement of 1 week and 6 weeks for *C. thermocellum* and *E. coli*, respectively using the GA-based procedure.

MLProScape: Machine Learning (ML) Based Method for Engineering Enzymes Faster By Modeling Protein Fitness Landscape (ProScape).

Sanjan T.P. Gupta^{1,2}, Evan Glasgow^{2,3}, Brian G. Fox^{2,3}, Parmeswaran Ramanathan⁴, and Jennifer Reed^{1,2}

(1)Chemical and Biological Engineering, University of Wisconsin Madison, Madison, WI, (2)Great Lakes Bioenergy Research Center, Madison, WI, (3)Biochemistry, University of Wisconsin Madison, Madison, WI, (4)Electrical and Computer Engineering, University of Wisconsin Madison, Madison, WI

Protein engineering aims to improve the functional properties of a protein, such as thermo-stability, binding affinity, and/or catalytic activity and has become a vital step in developing industrial enzymes. This work describes a machine learning (ML) based method called 'MLProScape' that builds an accurate model of the protein fitness landscape (ProScape) and then, uses the model to design synthetic protein designs with superior functional properties. Unlike approaches using directed evolution, which requires experimentally screening millions of protein variants, MLProScape requires fewer protein variants (on the order of tens to hundreds) to be tested experimentally, owing to the power of statistical inference.

As a proof-of-concept, MLProScape was applied to enhance the catalytic activity of glycoside hydrolases – a key enzyme used to degrade lignocellulosic biomass for biofuel production. Experimentally measured specific activities for a diverse set of glycoside hydrolases were used to train the ML models. The resulting elastic net regression models have a high predictive power (with correlation coefficient and R^2 values as high as 0.896 and 0.714, respectively, between the predicted and experimentally measured specific activities using a 5-fold cross validation). Moreover, by using position specific features, amino acid positions distal to the active site that might play a key role in modulating the activity level can be identified. MLProScape is also capable of modeling complex design criteria, such as engineering the catalytic activity of an enzyme towards multiple substrates

simultaneously, as well as, to account for other desirable traits such as high stability and better *in vivo* expression.

Development of methods like MLProScape will complement and add value to the current growing repertoire of *in-silico* pathway engineering tools as it will enable metabolic engineers to alleviate bottleneck steps en route target chemical of interest.

Key words: machine learning, enzyme engineering, sequence-to-function, experimental design

Constraint-Based Modeling Reveals the Distinct Metabolic Potential in Gut Microbiomes of Inflammatory Bowel Disease Patients with Dysbiosis.

Almut Heinken¹, Laurent Heirendt¹, Federico Baldini¹, Ronan M.T. Fleming², and Ines Thiele¹

(1)Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg, (2)Leiden Academic Centre for Drug Research, Leiden University, Leiden, Netherlands

The human gut microbiome plays an important role in human health. It is well known that inflammatory bowel diseases (IBD) are associated with microbial dysbiosis and altered fecal metabolomics profiles, however, the etiology of this complex, multifactorial disease remains poorly understood.

Recently, we published AGORA, a resource of curated genome-scale metabolic reconstructions for 773 common gut microbial strains. Moreover, we have published a COBRA Toolbox extension that enables the creation of personalized microbiome models from metagenomics data, as well as a high-performance implementation of constraint-based modeling in Julia. Here, we systematically profile the metabolic capabilities of gut microbiomes in health and disease states.

We retrieved strain-level relative abundances from metagenomics data from a cohort of pediatric Crohn's Disease patients with and without dysbiosis and healthy control children (108 samples in total) and constructed a personalized microbiome model for each sample using the corresponding AGORA reconstructions. Subsequently, we predicted for each microbiome model the quantitative biosynthesis potential for all secreted metabolites as well as the strain-level contributions to each metabolite in each individual microbiome.

In total, 182 metabolites from diverse subsystems could be secreted by at least one individual microbiome, of which 124 had statistically significantly different production potential in dysbiotic IBD patient microbiomes. The production potential for potentially detrimental

metabolites such as lactate, sulfide, trimethylamine, and acetaldehyde was increased in dysbiotic IBD microbiomes. These metabolites were mainly produced by Proteobacteria representatives in this group but not in non-dysbiotic IBD or healthy microbiomes. In contrast, dysbiotic IBD microbiomes were depleted in production potential of butyrate, and B vitamins due to lower abundances of the contributing taxa.

In summary, we present an efficient computational approach to systematically interrogate individual microbiome models. This approach enables us to mechanistically link disease-relevant metabolites with gut microbial taxa known to play a role in IBD.

COBRA.jl - Gearing up for the huge scale.

Laurent Heirendt¹, Sylvain Arreckx¹, Venkata P. Satagopam¹, Reinhard Schneider¹, Ines Thiele¹, and Ronan M.T. Fleming²

(1)Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg, (2)Leiden Academic Centre for Drug Research, Leiden University, Leiden, Netherlands

Ever larger biochemical networks are being developed by the COBRA community. Reconstructions with millions of biochemical reactions are within reach, and standard implementations of COBRA software written in MATLAB, Python, or C, hit their limit. Existing implementations are not suitable for simultaneously analyzing thousands of these large networks and simulations are months-long.

Accelerating existing and new COBRA methods bears the challenges of short development times and speeding up analyses significantly, all while assuring scalability and parallelism across multiple computing nodes and without recurring to complex message-passing interfaces. The Julia language [1] satisfies this need. It is set to become the language of choice to accelerate analyses of large (< 500k biochemical reactions) and huge-scale biochemical networks (> 500k reactions).

A step towards gearing up for high-dimensional COBRA modeling is the release of the high-level, high-performance, and open-source COBRA.jl package (git.io/COBRA.jl). DistributedFBA.jl [2], part of COBRA.jl, allows performing a flux balance analysis and many related analysis types efficiently, especially on large and huge-scale models. Most COBRA analyses are based on the COBRA Toolbox [3], a comprehensive software suite of interoperable methods written in MATLAB, which has found widespread applications in biology, biomedicine, and biotechnology. PALM.jl allows launching analyses

across several computing nodes simultaneously, and this for thousands of models at once.

Through the open-source high-performance COBRA.jl package, the analysis capabilities of the COBRA community are lifted to another level, as reconstruction and analysis of large and huge-scale models are made possible and accelerated.

[1] Bezanson et al., *Julia: A Fresh Approach to Numerical Computing*, *SIAM Review*, 59(1):65–98, 2017. [2] Heirendt et al., *DistributedFBA.jl: high-level, high-performance flux balance analysis in Julia*, *Bioinformatics*, 33(9):1421–1423, 2017. [3] Heirendt & Arreckx et al., *Creation and analysis of biochemical constraint-based models: the COBRA Toolbox v3.0*, *Nature Protocols (accepted)*, 2018.

Genome-Scale Flux Elucidation in the Fast-Growing Cyanobacterium *Synechococcus Elongatus UTEX 2973*.

John I Hendry¹, Saratram Gopalakrishnan¹, Himadri B. Pakrasi², Yinjie J. Tang³, and Costas D. Maranas¹

(1)Department of Chemical Engineering, The Pennsylvania State University, University Park, PA, (2)Department of Biology, Washington University, St. Louis, MO, (3) Department of Energy, Environmental and Chemical Engineering, Washington University, St. Louis, MO

A short doubling time of 2h makes *Synechococcus elongatus* UTEX 2973 a promising platform for solar based production of biofuels and other bio-chemicals. Successful metabolic engineering strategies can be informed by the knowledge of intracellular flux distribution in this organism. In this meta-analysis, the intracellular flux distribution in a genome-scale metabolic model of *Synechococcus elongatus* UTEX 2973 was elucidated using isotopic non-stationary ¹³C-metabolic flux analysis using the experimentally measured labeling dynamics of 13 central carbon metabolites reported in an earlier study. To this end, a genome scale carbon mapping model, *imSyu593*, was constructed using the existing mapping model for *Synechocystis* sp PCC 6803, *imSyn617*, as the starting point to trace the flow of carbons through 593 reactions spreading across the central carbon metabolism, amino acid metabolism and peripheral pathways. Flux elucidation revealed a near complete routing (>96%) of the assimilated carbons to biomass formation. This high biomass yield is afforded by the ability of this organism to reincorporate carbons oxidized in anabolic and photorespiratory pathways while increasing the flux through non-decarboxylating reactions. The ability to reincorporate oxidized carbons sustains a higher flux through the photorespiratory C2 cycle to meet the glycine and serine demand for the biomass synthesis. In contrast with other cyanobacteria, Malic enzyme is found to be dispensable

POSTER ABSTRACTS

with pyruvate being synthesized via Pyruvate Kinase. And instead of using the lower glycolysis and Pyruvate Dehydrogenase reaction, Acetyl-CoA was synthesized using the carbon efficient Phosphoketolase pathway. These findings suggest the existence of a highly efficient metabolism in *Synechococcus elongatus* UTEX 2973, which, in conjunction with fast growth rate, supports the development of this organism into an ideal photoautotrophic production platform.

Miren: An Optimization Tool for Data-Driven Discovery of Global Regulatory Phenomena Used to Elucidate the Heat Stress Response Mechanism in Rice Seed.

Mohammad Mazharul Islam¹, Jaspreet Sandhu², Harkamal Walia², and Rajib Saha³

(1)Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE, (2)Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE, (3)Department of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE

Plants use a suite of strategies to respond to abiotic stresses, including changing the abundance of stress-responsive genes/proteins that ultimately lead to the large-scale changes in gene expression levels, protein abundances, and metabolites. Complex gene-protein-reaction associations as well as regulatory mechanisms constitute a challenge to elucidate stress response mechanisms in plants. This research aims to understand the heat stress response mechanisms in developing rice seed using temporal transcriptomic analyses in control and stress conditions and *in silico* metabolic analyses.

Rice plants were exposed to heat stress at 12 and 36 hours after fertilization with a 16h-light/8h-dark cycle, and young developing seeds were collected from control and stressed plants. Total RNA isolated from developing seeds was used for differential gene expression analysis, which yielded ~7000 significantly stress-responsive genes. Clustering analysis was used to develop a minimal gene interaction network and identify global regulators. The highly connected “hub” genes included previously-identified MADS-box genes as well as a large number of novel regulatory genes. MiReN, an MILP optimization-based tool was developed to decipher the minimal regulatory network using the time-series transcriptomic data. MiReN identified important regulatory relationships for stress-responsive rice transcription factors (e.g., OsMYB, OsbZIP, OsMADS etc.) and predicted the minimal global regulatory network for rice seed in control and stress conditions. A comparative analysis of the network

topology reveals the shift in regulatory mechanisms in presence of stressors and allows for integration of transcriptomic data with a genome-scale metabolic model of rice seed. Work on other rice tissues and modeling the interactions between them using multi-level and multi-objective modeling frameworks to develop a robust plant-scale rice model is underway. Our predictive mathematical model will identify biologically important and non-intuitive solutions to questions related to stress response mechanisms and accelerate the development of tolerant plant varieties in an efficient and accurate fashion.

Leveraging a Clostridium Difficile Genre and Metagenomics to Identify Candidate Probiotic Bacterial Strains.

Matthew Jenior¹ and Jason Papin²

(1)Biomedical Engineering, University of Virginia, Charlottesville, VA, (2)Department of Biomedical Engineering, University of Virginia, Charlottesville, VA

Clostridium difficile is a Gram-positive, sporulating anaerobe that has become the leading causing of hospital-acquired infection. Exposure to antibiotics sensitizes hosts to colonization of the gastrointestinal tract by *C. difficile* through disruption of the healthy resident bacterial community, known as the gut microbiota. Previous studies have strongly supported that colonization resistance to *C. difficile* is driven by competition for growth nutrients by members of the gut microbiota. As *C. difficile* is known to colonize numerous host species and is capable of catabolizing a large array of possible growth substrates, it is difficult to determine the bacterial groups in a resistant community that confer this property. In order to more closely study nutrient utilization by *C. difficile* during infection across distinctly susceptible environments within the complex milieu of the gut microbiome, we employed Flux-Balance Analysis in an *in vivo* contextualized GENRE of *C. difficile* strain 630. In consideration of large deficits in nutrient utilization pathways and in an effort to address several challenges with existing reconstructions of *C. difficile*, we generated a *de novo* reconstruction of *C. difficile* 630 followed by extensive manual curation of its core metabolism and nutrient acquisition systems. *In silico* predictions in both rich and minimal medias, as well as transcriptomic and metabolomic data contextualization collected from an animal model of infection, closely mimic experimental characterization across multiple *in vitro* and *in vivo* studies. Furthermore, correlations between predicted *C. difficile* substrate utilization across distinct infection conditions with metagenome-enabled metatranscriptomic results from those microbiotas reveal discrete groups of

bacteria which share distinct metabolic functionalities with the pathogen when compared to resistant communities. Ultimately, by identifying metabolic functionalities of resistant/sensitive communities we can design targeted probiotic therapies with the desired community phenotype.

Metabolic Chaos during Microbial Stress Responses.

Paul A. Jensen

Bioengineering Department and Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL

Bacteria mount a physiologic stress response to survive hostile or stringent conditions. The genes that comprise the stress response can be defined in two ways: a transcriptional network composed of genes with expression changes during stress, and a phenotypic network of genes whose deletion lowers the bacteria's fitness during stress. On first thought, one would expect these two networks to be the same, or at least have significant overlap. The genes important for surviving stress should be up-regulated, and genes not essential for tolerating the stress should be down-regulated. However, studies from multiple microbes and conditions have discovered that the transcriptional and phenotypic stress responses are nearly disjoint; genes with expression changes during stress can be deleted without changing fitness, and phenotypically important genes are rarely differentially expressed. Although several studies have described this striking phenomenon, it remains unclear why bacteria appear to have two separate stress response networks.

We discovered that fitness and expression changes can be linked using COBRA models. During nutrient depletion, genes with expression changes are often adjacent to phenotypically important genes. Even the magnitude of the fitness or expression changes are linked — the largest fitness changes are closest to the largest expression changes, and vice versa. The expression changes are much closer to the fitness changes than expected if both sets of genes were randomly distributed throughout the network. By contrast, the response to antibiotic stress is largely uncoordinated. Changes in fitness and expression appear randomly distributed across the network, and the magnitudes of the nearest fitness and expression changes are uncorrelated. Using a novel “enzyme coupling” framework, we show how modularity and redundancy in the metabolic network enable the split between transcriptional and phenotypic stress responses – in both

the coherent starvation response and the incoherent antibiotic response.

Fine Tuning Thresholds to Facilitate Integration of Transcriptomics Data.

Chintan Joshi, Song-Min Schinn, Austin Chiang, Anne Richelle, and Nathan Lewis

University of California, San Diego, La Jolla, CA

Integration of omics datatypes is becoming increasingly commonplace in the field of metabolic modeling. Among all of the omics datatypes, transcriptomics is the most chosen for integration with genome-scale metabolic models (GEMs) resulting in a wide array of algorithms. These algorithms are often preceded by preprocessing the transcriptomics data to generate list of high confidence genes or reactions. Multiple studies in the past have identified parameters related to preprocessing to be important, specifically thresholding of gene expression profile. However, it is not clear how thresholds should be decided. Here, in our analysis we look at tissue-specific transcriptomic data from HPA and GTex and provided novel insights on what factors could determine thresholds and what are the largest sources of false positive and false negative predictions. Our analysis also provides novel insights on gene expression patterns of tissue-specific and housekeeping genes. Further, we also suggest improvements, validate using proteomics data and metabolic tasks. We also applied our thresholding method to cancer cell-line transcriptomics data sets. Thus, we demonstrate better applicability of thresholds across different data sets.

A Multi-Omics Investigation Unveiling the Tiered Regulation of Breast Cancer Cell Metabolism.

Rotem Katzir

University of Maryland, College Park, MD

Altered metabolism is a hallmark of cancer, but little is still known about its regulation. In this study, we measure transcriptomic, proteomic, phospho-proteomic and fluxomics data in a breast cancer cell-line(MCF7) across three different growth conditions. Integrating these multiomics data within a genomescale human metabolic model in combination with machine learning we systematically chart the different layers of metabolic regulation in breast cancer cells, predicting which enzymes and pathways are regulated at which level. We distinguish between two types of reactions, directly or indirectly regulated. Directly-regulated reactions include those whose flux is regulated by transcriptomic

POSTER ABSTRACTS

alterations (~890) or via proteomic or phospho-proteomics alterations (~140) in the enzymes catalyzing them. Potentially indirectly regulated reactions are those that currently lack evidence for direct regulation (~930). Many metabolic pathways are predicted to be regulated at different levels, and those may change at different media conditions. Remarkably, we find that the flux of predicted indirectly regulated reactions is strongly coupled to the flux of the predicted directly regulated ones, uncovering a hierarchical organization of breast cancer cell metabolism. Furthermore, the predicted indirectly regulated reactions are predominantly reversible. Taken together, this architecture may facilitate the rapid and efficient metabolic reprogramming in response to the varying environmental conditions incurred by the tumor cells. The approach presented lays a conceptual and computational basis for a more complete mapping of metabolic regulation in different cancers with incoming additional data.

Constraint-based modeling of allelic variation mechanistically links genome-wide associations

Erol Kavvas

Bioengineering, UCSD, La Jolla, CA

Mechanistic genotype-phenotype maps are difficult to construct using existing approaches. Here we present a method that integrates microbial genomics data, constraint-based metabolic modeling, and machine learning to infer a systemic bridge linking genetic variation and phenotypic variation. Our method, named Game4all (game for alleles), formulates the problem as a coordination game in which the alleles (players) must coordinate their constraints (actions) such that the computed outputs of the strain-specific genome-scale models are predictive (payoff) of the measured phenotypes. We apply Game4all to 1,595 drug-tested *M. tuberculosis* strains to construct a mechanistic landscape of drug resistance evolution. Evaluation of the genotype-system map infers mechanistic effects of causal alleles and uncovers epistatic interactions. Analysis of the system-phenotype map identifies metabolic states differentiating resistant and susceptible *M. tuberculosis* strains, emphasizing redox and bioenergetic homeostasis as a determinant of antibiotic efficacy. Constraint-based modeling can thus contribute to the study and analysis of genetic factors underlying observed phenotypic variation, a field previously dominated by statistical methods.

Laboratory Evolution Reveals a Two-Dimensional Rate-Yield Tradeoff in Microbial Metabolism.

Zachary A. King¹ and **Chuankai Cheng²**

(1)Bioengineering, University of California, San Diego, La Jolla, CA, (2)University of California, San Diego, La Jolla, CA

Growth rate (μ) and yield (Y) are fundamental features of microbial growth. However, we lack a mechanistic and quantitative understanding of the μ - Y relationship. Studies pairing computational predictions with experiments have shown the importance of maintenance energy and proteome allocation in explaining rate-yield tradeoffs and overflow metabolism. Recently, adaptive evolution experiments of *Escherichia coli* reveal a phenotypic diversity beyond what has been explained using a simple μ - Y relationship. Here, we identify a two-dimensional μ - Y tradeoff in adapted *E. coli* strains where the dimensions are (A) a tradeoff between μ and Y and (B) a tradeoff between substrate uptake rate (q_{glc}) and Y . We employ a multi-scale modeling approach, combining a previously reported small-scale proteome allocation model with a genome-scale model of metabolism and gene expression (ME-model), to develop a quantitative description of this two dimensional μ - Y relationship using data for *E. coli* K-12 MG1655. The analysis provides a mechanistic explanation of the two-dimensional μ - Y tradeoff. Furthermore, the analysis identifies modifications to the P/O ratio as a potential mechanism that enables the q_{glc} - Y tradeoff. Thus, the μ - Y tradeoffs that govern microbial adaptation to new environments are more complex than previously reported, and they can be understood in mechanistic detail using a multi-scale modeling approach.

A Metabolic Reconstruction of Lactobacillus Reuteri and Analysis of Its Potential As a Cell Factory.

Thordis Kristjansdottir¹, **Elleke F. Bosma²**, **Filipe Branco dos Santos³**, **Emre Özdemir⁴**, **Alex Toftgaard Nielsen⁴**, and **Steinn Gudmundsson¹**

(1)Centre For Systems Biology, University Of Iceland, Reykjavik, Iceland, (2)DTU Biosustain, (3)University of Amsterdam, Amsterdam, Netherlands, (4)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark

Lactic acid bacteria (LAB) have long been used by humans in the food industry and is *Lactobacillus reuteri* one of them. Additionally, *L. reuteri* is also known to produce several different compounds, with potential biotechnological interest, such as; the antimicrobial compound reuterin, vitamin B12 and 1,3-propanediol, which can be used as a starting material for plastic production. In an attempt to further explore *L. reuteri* as a cell factory for production of valuable compounds, a genome-scale metabolic model was reconstructed and

thoroughly curated and validated using data from the literature and several experimental data sets, including knockout strains. The biomass reaction was built from species specific data when available and alternatively from *L. plantarum* and *L. lactis*.

L. reuteri metabolizes glucose via the phosphoketolase pathway. Genome data indicates the absence of a phosphofructokinase (PFK), which results in an inactive EMP pathway. Model simulations show that PFK does not provide a significant advantage in terms of growth.

A pathway prediction algorithm was used to identify several heterologous pathways for production of 1-propanol. Theoretical yields of 1-propanol were evaluated for multiple carbon sources. The most promising pathways were further optimized *in silico* by identifying targets for overexpression and knockouts for increased production of 1-propanol.

Metabolic Modeling of a Model Diatom.

Manish Kumar¹, Shawn W. Polson², Lisa Zeigler Allen³, and Karsten Zengler⁴

(1)Department of Pediatrics, University of San Diego, San Diego, CA, (2)Department of Computer and Information Sciences, University of Delaware, Newark, DE, (3)Department of Microbial and Environmental Genomics, J. Craig Venter Institute, La Jolla, CA, (4)Department of Pediatrics, University of California, San Diego, CA

Diatoms are unicellular photosynthetic microalgae that are ubiquitous in fresh water and marine and environments. These microalgae are amenable for ~20% of the global carbon fixation. Diatoms are highly diverse and have substantially higher productivity and capabilities to accumulate lipids compared to other classes of microalgae. These properties make diatoms ideal candidates for advanced production of biofuels and chemicals. However, metabolism of diatoms is poorly characterized both in laboratory settings as well as in the environment, where diatoms interact. Previous studies revealed a very high abundance of marine viruses, some of which have been shown to infect diatoms along with other phytoplankton cells, and thus altering their metabolism. To gain in-depth knowledge about diatom metabolism and to address how viral infection modifies diatom physiology, we applied constraint-based modeling to the model diatom *Cylindrotheca* at different physiological states during its life-cycle. A genome-scale metabolic model of *Cylindrotheca* has been generated using annotated genome and three previously developed metabolic reconstructions of photosynthetic

organisms (*Phaeodactylum tricornutum*, *Chlamydomonas reinhardtii*, and *Chlorella vulgaris*) as template database of biochemical reactions. The model is validated with growth/no growth physiological data at several growth conditions. Furthermore, the model is constrained using single-cell RNA-sequencing data with and without viral infection to elucidate the involved mechanisms in viral infection to *Cylindrotheca*. Metabolic model integrated with time-course gene expression data is employed to explore the physiological characteristics, such as growth and lipid biosynthesis at different stages, and to determine how viruses affect the metabolism of the diatom host at these states. The outcome of this work will underpin future studies to advance our understanding of diatom metabolism and to elucidate their role on the ecology and biogeochemistry of the ocean.

Computational Modeling of *Ostreococcus Tauri* Central Metabolism Based on Statistical Thermodynamics.

Neeraj Kumar¹, William R. Cannon¹, James E. Evans², and Jeremy D. Zucker³

(1)Computational Biology, Pacific Northwest National Laboratory, Richland, WA, (2)Environmental Molecular Sciences Laboratory, Pacific Northwest National Lab, Richland, WA, (3)Pacific Northwest National Laboratory, Richland, WA

Developing predictive metabolic model based on the fundamental principles of physics is very critical to understand how cell structure affects function and subsequently investigate cellular dynamics relevant for bioenergy application. New computational tools are needed for quantitative modeling of cell that predict metabolite levels in high throughput manner, characterize thermodynamics and kinetics of individual reactions and energetics requirement of most likely metabolic pathways. In this talk, I will discuss our newly developed ODE-based optimization approach based on statistical thermodynamics for developing a central metabolic model of *Ostreococcus tauri*, the smallest known photoautotrophic eukaryote. We used the maximum entropy production principle to derive fluxes through a central metabolic network of *Ostreococcus tauri*. The predicted metabolite concentrations produced from maximum entropy production rate solution are compared to those typically expected from experiment using a loss function from which post-translational regulation of enzymes is inferred. Subsequently, we re-optimize the system with the inferred regulation and determine optimal rate constants for the metabolic network from the metabolite concentrations and reaction fluxes. At

POSTER ABSTRACTS

the end, I will discuss biophysical insights obtained from the modeling of *Ostreococcus tauri* that are relevant for biofuels production.

BOFdat: Generating Biomass Objective Function for Constraint-Based Metabolic Models from Experimental Data.

Jean-Christophe Lachance

Biology, Université de Sherbrooke, Sherbrooke, QC, Canada

The computation of growth phenotypes by genome-scale metabolic models (GEMs) requires the definition of an objective function. To simulate growth related phenotypes a biomass objective function (BOF) that includes key biomass components of the cell such as the major macromolecules (DNA, RNA, proteins), lipid composition, coenzymes, inorganic ions and species-specific components. While the accurate definition of the BOF was shown to provide precise phenotypic predictions, no standardized computational platform is available to generate species-specific data-driven BOFs. To fulfill this gap in the software ecosystem, we implemented BOFdat: a Python package for the definition of Biomass Objective Function from experimental data. BOFdat is a 3-step process. At each step, different macromolecular categories are added along with their macromolecular weight fractions, facilitating the mass balancing of the objective function. The first step adds the main macromolecules and lipids using both omics data and macromolecular weight fraction of each category to calculate the stoichiometric coefficients of each metabolite. The second step determines coenzymes by performing a connectivity analysis of the entire metabolic network, and inorganic ions by comparing against a list of ions found in all previously generated metabolic models in the BiGG Models database. The third step is a novel, unbiased approach using a genetic algorithm to find the combination of metabolites that best match experimental gene essentiality data, and clusters the results into sets of metabolic end goals. We used BOFdat to reconstruct the BOF of the *Escherichia coli* model iML1515 and found that the BOF generated by BOFdat had more similar biomass composition, growth rate, essentiality to the original predictions than other available methods. By providing an unbiased, data-driven approach to defining biomass objective functions, BOFdat has the potential to improve the quality of new genome-scale models and also greatly decrease the time required to generate a new reconstruction.

Elucidating the Metabolic Rewiring of Pluripotent Stem Cells during Differentiation into Adult Progenitors Using an Updated and Enzyme-Constrained Human Genome-Scale Model.

Meiyappan Lakshmanan¹, Dong-Yup Lee^{2,3}, and Lokanand Koduru⁴

*(1)Bioinformatics Group, Bioprocessing Technology Institute, Singapore, Singapore, (2)Bioprocessing Technology institute, A*STAR, Singapore, Singapore, (3)School of Chemical Engineering, Sungkyunkwan University, Gyeonggi-do, Korea, Republic of (South), (4) National university of Singapore, Singapore, Singapore*

Pluripotent stem cells (PSCs) offer new potentials for treating multiple diseases via gene therapy and regenerative medicine due to their unique regenerative abilities. PSCs differentiate into many distinct germ layer cells and produce various human organs via various transcriptional, epigenetic and metabolic changes. While the transcriptional and epigenetic changes during stem cell differentiation is widely researched, the metabolic changes of PSCs during differentiation remain largely uncovered, except the well-known switch from aerobic glycolysis into oxidative phosphorylation for energy generation. Therefore, we utilize a systematic approach combining transcriptomic data and the human genome-scale model (GEM) to examine the metabolic signatures of pluripotent stem cells, germ layer cells, and adult stem cells in comparison to various mature cells. Here, it should be highlighted that although the availability of several human GEMs has substantially improved our understanding on the molecular basis of various diseases and phenotypes, we still observed the variability in internal metabolic fluxes predicted by the models, across conditions due to the limited constraints typically used in constraints-based flux analysis. To address this critical issue, here, we present the most comprehensive and biochemically consistent human GEM to-date, Recon3E, and included the necessary kinome information to enable its ready use during the simulations, thereby improving the phenotype predictions further. Recon3E accounts for 3463 open reading frames, 7182 reactions corresponding to 987 unique enzymes, and 3023 unique metabolites. Importantly, Recon3E contains the kinetic information, i.e. k_{cat} , for 672 of the total 987 enzymes where 74% of these values are human-specific. Using Recon3E, we were able to predict the metabolic rewiring which occurs during the differentiation of PSCs into adult progenitors, emphasizing the role of unsaturated lipids, amino acid derivatives and one-carbon metabolites in maintaining self-renewal and pluripotency, in addition

to the commonly cited glycolysis and oxidative phosphorylation. (Funding: SSAC# PJ01334605, Republic of Korea)

Implementation and Application of a New Combined Stoichiometric and Thermodynamic Flux Balance Analysis.

Simeon Leupold, Bastian Niebel, and Matthias Heinemann

Molecular Systems Biology, University of Groningen, Groningen, Netherlands

We recently uncovered that cellular metabolism is geared towards growth maximization but constrained by an upper rate of cellular Gibbs energy dissipation. Exploiting this principle in flux balance analysis generated excellent predictions of exchange and intracellular fluxes across a wide range of conditions and carbon sources, and even including some metabolite concentrations. This method requires a thermodynamically consistent formulation of cellular processes including Gibbs energies of reaction for every metabolic process, and a Gibbs energy balance, which states that the Gibbs energy dissipated by all cellular processes is equal to the Gibbs energy exchanged with the environment. We here present a detailed workflow, exemplified for a genome-scale metabolic model of *Escherichia coli*, how to implement and apply this method starting from any stoichiometric metabolic reconstruction. This workflow encompasses the formulation of such a model, the parameterization using experimental data together with regression analysis and the application in flux balance analysis. Furthermore, we present strategies how to solve the required nonlinear and nonconvex optimizations. Given the limited amount of required input data, and the precision and extent of the model predictions, we consider this method a valuable addition to current flux balance analysis approaches.

Implementing and Evaluating a Gaussian Mixture Framework for Identifying Gene Function from Tnseq Data.

Kevin Li¹, Rachel Chen², William Lindsey³, Aaron Best⁴, Matthew DeJongh⁵, Christopher S. Henry⁶, and Nathan Tintle⁷

(1)Columbia University, New York, NY, (2)North Carolina State University, Raleigh, NC, (3)Dordt College, Sioux Center, IA, (4)Biology, Hope College, Holland, MI, (5)Computer Science, Hope College, Holland, MI, (6)Argonne National Laboratory, Argonne, IL, (7)Mathematics, Dordt College, Sioux Center, IA

The rapid acceleration of microbial genome sequencing increases opportunities to understand bacterial gene function. Unfortunately, only a small proportion of genes have been studied. Recently, TnSeq has been proposed as a cost-effective, highly reliable approach to predict gene functions as a response to changes in a cell's fitness before-after genomic changes. However, major questions remain about how to best determine whether an observed quantitative change in fitness represents a meaningful change. To address the limitation, we develop a Gaussian mixture model framework for classifying gene function from TnSeq experiments. In order to implement the mixture model, we present the Expectation-Maximization algorithm and a hierarchical Bayesian model sampled using Stan's Hamiltonian Monte-Carlo sampler. We compare these implementations against the frequentist method used in current TnSeq literature. From simulations and real data produced by *E.coli* TnSeq experiments, we show that the Bayesian implementation of the Gaussian mixture framework provides the most consistent classification results.

Sbmltools: A Sbml Integration Tool for Kbase.

Filipe Liu¹, José P. Faria¹, Qizhi Zhang¹, Miguel Rocha², Isabel Rocha², and Christopher S. Henry¹

(1)Argonne National Laboratory, Argonne, IL, (2)Centre of Biological Engineering, Department of Biological Engineering, University of Minho, Braga, Portugal

Since the release of the first version of System Biology Markup Language (SBML) it become the de facto exchange format to store and share Genome Scale Metabolic Models (GSM). However, due to limitations in earlier versions of the SBML users adopted their own strategies to represent important attributes regarding to GSMs: gene-protein-reaction rules (GPR), metabolite/reaction annotation and model constraints. While most of these features were addressed in the latest version of SBML it left a heterogenous community of SBML models.

The SBMLTools provides methods to import and integrate existing SBML models into the Department of Energy Systems Biology Knowledgebase (KBase). The purpose of the SBMLTools is to reshape SBML models to fit the KBase demands: separation of the exchange reactions for the media formulation, integration of the metabolic network with the ModelSEED biochemistry, integration of GPR genes to match with a compatible genome and normalize the compartments with the vocabulary in KBase.

The tool compiles an internal in-house built integrated biochemistry database that compiles existing compound

POSTER ABSTRACTS

and reaction databases (ModelSEED, KEGG, MetaCyc, BiGG, etc) for annotation purposes. It also takes advantage of the KBase genome catalog that hosts more than 70,000 RefSeq bacterial genomes, indexing several attributes regarding to their genes (e.g., gene aliases).

While the main purpose of the SBMLTools is to improve the compatibility of the existing tools in KBase with external built models, it can also be used by researchers to normalize SBML models regarding to identifiers (compounds and reactions), exchange media, compartments and genes.

The genetic basis for adaptation of model-designed syntrophic co-cultures.

Colton J. Lloyd¹, Ali Ebrahim¹, Laurence Yang¹, Zachary A. King¹, Edward Catoiu¹, Edward J. O'Brien², Joanne Liu³, and Bernhard O. Palsson¹

(1)Bioengineering, University of California, San Diego, La Jolla, CA, (2)Bioinformatics, University of California, San Diego, La Jolla, CA, (3)University of California, San Diego, La Jolla, CA

Understanding the fundamental characteristics of microbial communities has far reaching implications for human health and applied biotechnology. However, the genetic basis of viable community formation of synthetic communities has not been studied in detail. By pairing auxotrophic mutants in co-culture, it has been demonstrated that viable nascent *E. coli* communities can be established where the mutant strains are metabolically coupled. A novel algorithm, OptAux, was constructed to design 61 unique multi-knockout *E. coli* auxotrophic strains that require significant metabolite uptake to grow. These predicted knockouts included a diverse set of novel non-specific auxotrophs that result from inhibition of major biosynthetic subsystems. Three OptAux predicted non-specific auxotrophic strains—with diverse metabolic deficiencies—were co-cultured with an L-histidine auxotroph and optimized via adaptive laboratory evolution (ALE). Time-course sequencing revealed the genetic changes employed by each strain to achieve higher community growth rates and provided insights into mechanisms for adapting to the syntrophic niche. A community model of metabolism and gene expression was utilized to predict the relative community composition and fundamental characteristics of the evolved communities. This work presents new insight into the genetic strategies underlying viable nascent community formation and a novel computational method to elucidate metabolic changes that empower the creation of cooperative communities.

Exploring the Metabolic Role of Multidrug-Resistance Efflux Pumps in *Salmonella Typhimurium*.

Sean G. Mack¹, Xuan Wang-Kan², Laura J. V. Piddock², and Daniel J. Dwyer^{1,3,4,5}

(1)Dept. of Chemical and Biomolecular Engineering, University of Maryland, College Park, MD, (2)Inst. of Microbiology and Infection, University of Birmingham, Birmingham, United Kingdom, (3)Dept. of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD, (4)Inst. for Physical Science and Technology, University of Maryland, College Park, MD, (5)Dept. of Bioengineering, University of Maryland, College Park, MD

The surge in antibiotic resistance poses a serious threat to human health. In Gram-negative pathogens, diverse resistance-nodulation-division (RND) efflux pumps provide broad-spectrum resistance and therefore greatly contribute to antibiotic treatment failure [1]. While the structure and protective roles of these pumps has been established [2], their homeostatic contributions are poorly understood. Recent work has suggested a broader metabolic role for the AcrB efflux pump in *Salmonella enterica Typhimurium*, with AcrB being the most common member of the RND family. The loss of AcrB efflux, but not expression, led to a loss in virulence and the downregulation of various pathogenicity factors [3]. Given that AcrB is highly conserved across all Enterobacteriaceae [2], further exploration of its potential metabolic role is clearly needed.

To address this gap in understanding, we have developed a computational framework to predict and compare the metabolic behaviors of wild-type and the loss-of-function AcrB mutant of *S. Typhimurium*. Specifically, we analyzed RNAseq data from *S. Typhimurium* WT and mutant strains grown in minimal media for differentially expressed genes. Using a modified version of the metabolic transformation algorithm (MTA) [5], transcriptomic data were integrated with a genome-scale metabolic model for *S. Typhimurium* [6] to generate metabolic flux predictions for each strain. From these flux profiles, we identified genetic perturbations that would shift WT metabolism closer to mutant metabolism (and vice versa). Our findings suggest the loss in AcrB efflux leads to increased fatty acid biosynthesis which may contribute to the overall loss in virulence. To validate these predictions, we are currently testing our top targets via metabolomics and infection assays.

Blair, J. M. A. et al. Nature Reviews Microbiology (2015).

Du, D. et al. Biological Chemistry (2015).

Wang-Kan, X. et al. mBio (2017).

Yizhak, K. et al. Nature Communications (2013).

Thiele, I. et al. BMC Systems Biology (2011).

Thermodynamic Metabolic Flux Analysis with Consistent Handling of Errors in the Estimates from the Component Contribution Method.

Vishnuvardhan Mahamkali¹, Kaspar Valgepea¹, Tim McCubbin¹, Esteban Marcellin¹, and Lars K. Nielsen^{1,2}

(1)AIBN, The University of Queensland, Brisbane, Australia, (2)DTU Biosustain, Technical University of Denmark, Kongens Lyngby, Denmark

Thermodynamic metabolic flux balance analysis (TMFA) is used extensively in genome scale metabolic modeling to obtain thermodynamically consistent flux and metabolic concentration profiles. TMFA involves the use of thermodynamic constraints to avoid thermodynamically infeasible solutions and provides information about free energy change of reactions as well as the range of metabolite activity. The current best method for estimating standard Gibbs energy of reaction ($\Delta_r G^0$) uses the component contribution method combining reactant and group contribution methods. The estimates of $\Delta_r G^0$ are associated with an error, which is calculated from errors on the individual reactant and group contributions.

Recent implementations of TMFA allow for the error in the estimates by allowing each individual $\Delta_r G^0$ to vary independently within its estimated confidence. However, this is not the correct treatment as the individual $\Delta_r G^0$ are not independent; they share reactant and group contributions. Allowing them to vary independently will inevitably result thermodynamic inconsistencies, where closed loops have non-zero ΔG . This problem was resolved by returning to the original TMFA formulation, where the error was allowed in the individual groups and then identically propagated to all $\Delta_r G^0$. Using the component method, this approach is here expanded by allowing for errors in both reaction and group contributions (each is allowed with their respective confidence ellipsoid) and then propagating the adjusted estimates correctly to all $\Delta_r G^0$. The method is made available as an OpenCobra function.

Development and Application of Constraint-Based Modeling Methods to Study Vulnerabilities Associated to Lipid Metabolism in Prostate Cancer.

Igor Marín de Mas¹, Esther Aguilar², Carmen Bedia³, Timothy M. Thomson⁴, Romà Tauler³, Balazs Papp⁵, Marta Cascante², and Lars K. Nielsen¹

(1)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby,

Denmark, (2)Departamento de Bioquímica I Biología Molecular, University of Barcelona, Barcelona, Spain, (3)IDAEA, CSIC, Barcelona, Spain, (4)Department of Environmental Chemistry, Molecular Biology Institute of Barcelona, National Research Council (IBMB-CSIC), (5)Department of Biochemistry and Molecular Biology, Biological Research Centre of the Hungarian Academy of Sciences, Szeged, Hungary

It is increasingly clear that significant alterations in the lipid profile of cancer cells accompany tumor progression and metastasis. These changes are induced by a metabolic reprogramming that enhances the malignant phenotype of cancer cells. In this context, genome-scale metabolic models (GEM) have emerged as a valuable platform to integrate different omic data to study cancer metabolism from a holistic perspective.

Here we present a constraint-based model-driven approach integrating transcriptomic data to study the metabolic profile of two clonal sub-populations from a prostate cancer cell line (PC-3): PC-3/M and PC-3/S, representing pre and post Epithelial-Mesenchymal-Transition stages. This model-driven analysis and experimental validations unveiled a marked metabolic reprogramming in long-chain fatty acids metabolism.

While PC-3/M cells showed an enhanced entry of long-chain fatty acids into the mitochondria, PC-3/S cells used this pool as precursors of eicosanoid metabolism. This metabolic reprogramming endows PC-3/M cells with augmented energy metabolism for fast proliferation and PC-3/S cells with increased eicosanoid production impacting angiogenesis, cell adhesion and invasion.

These findings highlight the relevance of lipid metabolism in cancer development and progression. However, lipid-associated pathways are poorly annotated in human GEMs which limits the use of this tool. To overcome this limitation, we developed an algorithm-based metabolic network building method to enrich existing GEMs with lipid-associated pathways. This algorithm was applied in the study of the metabolic profile of another prostate cancer cell line (DU145) in response to a chronic exposure to different endocrine disruptors inducing malignancy and alterations in the lipid profile. The resulting lipid-enriched GEM covered 98% of the altered lipids compared to only 5% in the original model. Thus, this approach improves lipidomic data integration into GEM reconstructions, enabling a more in-depth study of the mechanisms underlying diseases with a strong metabolic component such as cancer.

Modelling of Cell Metabolism to Reduce the Lactate Generation in CHO Cell Cultures.

Iván Martínez-Monge^{1,2}, Svetlana Volkova¹, Igor Marín de Mas¹, Hooman Hefzi³, Pere Comas², Nathan Lewis³, Martí Lecina^{2,4}, Jordi Joan Cairó², and Lars Keld Nielsen¹

(1)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kongens Lyngby, Denmark, (2)Department of Chemical, Biological and Environmental Engineering, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Spain, (3)Department of Bioengineering, University of California, San Diego, CA, (4)Bioengineering Department, IQS-Universitat Ramon Llull, Barcelona, Spain

Chinese Hamster Ovary cells (CHO) display Warburg metabolism characterized by high glucose consumption and high lactate production under aerobic conditions. Lactate is a by-product widely reported to inhibit cell growth in culture, imposing an important burden to industrial processes. In order to reduce lactate secretion and thus its inhibitory effects, two different approaches have been applied: i) Bioprocess Engineering, controlling extracellular conditions in the bioreactor to trigger concomitant glucose/lactate consumption, and ii) Synthetic Biology, to generate non-lactate producer mutant cells knocking out lactate dehydrogenase and the regulators responsible for inhibiting pyruvate conversion to acetyl-CoA.

In the current study, the metabolic profile of CHO cells was investigated in batch bioreactor cultures performed under three conditions: normal WT, controlled WT, and zero-lactate CHO (CHO ZeLa). Exometabolomic data was integrated into a reduced genome-scale metabolic model using Flux Balance Analysis (FBA) during the mid-exponential phase, as well as Dynamic FBA (DFBA) to capture the dynamic changes occurring over time in CHO cell metabolism. Model reduction was performed through a novel semi-automated reduction protocol to generate fully functional metabolic models.

FBA showed that in wild-type cell line, lactate is produced to fulfill the NADH regeneration requirements in the cytoplasm and only a small amount of pyruvate is introduced into TCA through Acetyl-CoA. When concomitant glucose and lactate consumption was triggered, as well as in CHO ZeLa, glucose uptake was significantly reduced and a balance between glycolysis and TCA cycle fluxes was reached, yielding a more efficient substrate consumption. Moreover, DFBA illuminated the metabolic mechanisms by which wild-type CHO switches from a Warburg (glucose consumption/

lactate production) phenotype to a glucose/lactate co-consumption phenotype.

Our approach enabled us to explore the mechanisms underlying dynamic metabolic response of wild-type CHO and CHO ZeLa with potential implications in the industry of bioproducts.

Systematic Reduction of Genome-Scale Models for the Study of Metabolic Phenotypes of Human Cells.

Maria Masid¹, Meriç Ataman¹, and Vassily Hatzimanikatis²

(1)Laboratory of Computational Systems Biotechnology (LCSB), EPFL, Lausanne, Switzerland, (2)Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland

In the last years, the analysis of cellular metabolism has sparked new interest in systems biology and metabolic modelling. In particular, modelling the phenotypes of healthy and diseased cells will help to understand the main metabolic characteristics of disease development and progression. It will also be key to design more effective therapies.

The reconstruction of genome-scale models (GEMs) enables the computation of phenotypic traits based on the genetic composition of a target organism. To overcome the well-known challenges when working with large networks, we generate systematically reduced models around specific subsystems. Within this framework, we curate the GEMs to include the thermodynamic properties of the network metabolites and reactions. Furthermore, we consider the composition and utilization of the extracellular-medium metabolites, and the synthesis of the biomass precursor metabolites. The reduced models can be used for a broad range of applications extending from omics data integration to kinetic models.

We demonstrate here that the study of reduced human GEMs can provide insight and a systematic framework to compare different versions of GEMs (Recon 2 versus Recon 3) as well as different cellular physiologies, such as, diseased versus healthy phenotypes.

Elementary Mass Action Stoichiometric Simulation Models Predict Non-Negligible Concentrations of Enzyme-Bound Metabolites.

Marta R. A. Matos¹, Daniel C. Zielinski², Lars K. Nielsen³, and Nikolaus Sonnenschein⁴

(1)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark,

Kgs Lyngby, Denmark, (2)Bioengineering, University of California, San Diego, CA, (3)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Lyngby, Denmark, (4)The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark

Improvement of cell factory performance, to the point where production of chosen molecules becomes economically profitable, requires extensive metabolic modification. Metabolic models have been used to both provide new insights into the inner workings of metabolism and new directions for strain engineering. Kinetic models in particular are key to model the dynamics of metabolism and substrate-level enzyme regulation.

Here, we use the elementary Mass Action Stoichiometric Simulation (eMASS) framework to build a prototype kinetic model ensemble for glycolysis in *E. coli*. Following eMASS, we first build an individual kinetic model for each reaction by decomposing it into elementary reactions, according to the respective reaction mechanism. Next, the associated elementary rate constants are fitted to kinetic data obtained from the literature, e.g. k_{cat} , K_m , K_i , while satisfying the respective Haldane relationship. Finally, we assemble the system's kinetic model by combining the individual reaction models and integrating fluxomics and metabolomics data. Since we model both free and enzyme-bound metabolite concentrations explicitly, we drop the common assumption that enzyme-bound metabolite concentrations are negligible ($x_{tot} \approx x_{free}$) and predict the free metabolite concentrations.

Our preliminary results show that the percentage of enzyme-bound metabolite vs. total metabolite can be as high as 40%, i.e. the amount of enzyme-bound metabolite may not be negligible even in glycolysis. These results derive from the fact that in some cases the concentration of metabolites is in the same order of magnitude as the enzymes with which they interact.

Medusa: A Software Package for Construction and Analysis of Ensembles of Metabolic Networks.

Gregory L. Medlock and Jason A. Papin

Department of Biomedical Engineering, University of Virginia, Charlottesville, VA

Genome-scale metabolic network reconstructions (GENREs) are powerful tools for predicting the metabolic behavior of organisms across broad environmental contexts and in the presence of drugs or genetic manipulation. Many applications of GENREs, such as contextualization of 'omics data and network expansion

through integration of phenotypic data, lead to multiple feasible network structures which are all equally likely. Common practice is to choose a single network and perform downstream analyses without consideration of other equally-likely networks. Recently, maintaining these alternative network structures and performing analysis in an ensemble fashion has demonstrated improved predictive performance.

Here, we present Medusa, a software package for the construction and analysis of ensembles of GENREs. Medusa is an open-source python software package that performs constraint-based reconstruction and analysis (COBRA) methods at ensemble-scale. The underlying data structures implemented in Medusa extend cobrapy, a python package for analysis of GENREs using COBRA methods.

In this presentation, we highlight three key capabilities of Medusa. The first is generation and analysis of ensembles of GENREs from phenotypic data. The second is generation of ensembles composed of alternative GENRE states created using integration of transcriptomic data. The third key capability is identification of network components associated with variable predictions throughout an ensemble using supervised and unsupervised machine learning. Using this workflow of ensemble generation and uncertainty quantification, Medusa can be used to select experiments that maximally resolve uncertainty in GENRE structure.

Simulations performed using Medusa are probabilistic in that each ensemble generates a distribution of predictions. We hope that wide adoption of ensemble analysis methods will reduce investment in GENRE-based predictions that are spurious (e.g. only one ensemble member makes the prediction) and help guide experiments to quickly improve the quality and predictive power of GENREs.

A Systematic Assessment of Current Genome-Scale Metabolic Reconstruction Tools.

Sebastián N. Mendoza, Brett G. Olivier, Douwe Molenaar, and Bas Teusink

Systems Bioinformatics, Amsterdam Institute for Molecules, Medicines and Systems, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

Several genome-scale metabolic reconstruction software platforms have been developed and continuously updated during the last fifteen years. These tools have been widely applied to reconstruct metabolic models for hundreds of microorganisms ranging from important human pathogens

POSTER ABSTRACTS

to species of industrial relevance. So far, these platforms have not been systematically evaluated with respect to software quality, best potential uses and intrinsic capacity to generate high-quality genome-scale metabolic models. In consequence, this lack of assessment keeps users away from selecting the tool that best fits the purpose of their research. In this work, we reviewed the current genome-scale reconstruction software platforms and we performed a systematic assessment of the most promising tools. We defined a list of features for assessing software quality related to genome-scale reconstruction which we hope it would be useful for potential users. In addition, we compared the output networks generated using each of these tools with the high-quality manually curated models of *Lactobacillus plantarum* and *Bordetella pertussis*, representatives of gram-positive and gram-negative bacteria, respectively. We showed that none of the tools outperform the others in all the studied features and that users should carefully choose one or another, or even combining them, depending on the particularities of their projects.

Genome-Scale Metabolic Model of Chromochloris, an Emerging Model Organism for Sustainable Fuel Production.

Alexander Metcalf and Nanette R. Boyle

Chemical and Biological Engineering, Colorado School of Mines, Golden, CO

One of the main challenges of developing alternative energy sources is the ability to compete economically with fossil fuels. One approach to make bioproduction more economical is to engineer organisms which can produce fuels as well as value added molecules. *Chromochloris zofingiensis*, a green alga, can accumulate up to 40% of their dry weight as triacylglycerols making it an excellent candidate for biofuels production. It also accumulates high amounts of astaxanthin, a high value ketocarotenoid that can be used as a pharmaceutical, nutraceutical, cosmetic, or as food or feed supplements¹⁻³. Naturally derived astaxanthin has a market price of approximately \$7,000 per kilogram. The co-production of this molecule product makes the economic feasibility of biodiesel production by *C. zofingiensis* much more attainable. In order to investigate the metabolic capacity of this organism for both fuel and astaxanthin production, we have reconstructed the metabolic network from the published genome. We will present our work to date and results of simulations to maximize the production of both products. (Research supported by grant DE-SC0018301 from the Department of Energy.)

References

- 1 Hussein, G., Sankawa, U., Goto, H., Matsumoto, K. & Watanabe, H. Astaxanthin, a Carotenoid with Potential in Human Health and Nutrition. *Journal of Natural Products* 69, 443-449, doi:10.1021/np050354+ (2006).
- 2 Yuan, J. P., Peng, J., Yin, K. & Wang, J. H. Potential health-promoting effects of astaxanthin: A high-value carotenoid mostly from microalgae. *Molecular Nutrition & Food Research* 55, 150-165, doi:10.1002/mnfr.201000414 (2011).
- 3 Liu, J. et al. *Chlorella zofingiensis* as an Alternative Microalgal Producer of Astaxanthin: Biology and Industrial Potential. *Marine Drugs* 12, 3487-3515, doi:10.3390/md12063487 (2014).

Exploring Biosynthetic Pathways for the Production of Five Methyl Ethyl Ketone Precursors.

Milenko Tokic¹, Noushin Hadadi², Meric Ataman³, Dário Neves⁴, Birgitta E. Ebert⁴, Lars M. Blank⁴, **Ljubisa Miskovic³**, and Vassily Hatzimanikatis⁵

(1)Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), 1015 Lausanne, Switzerland, (2)Swiss Institute of Bioinformatics (SIB), 1015 Lausanne, Switzerland, (3) Laboratory of Computational Systems Biotechnology (LCSB), EPFL, Lausanne, Switzerland, (4)iAMB - Institute of Applied Microbiology, ABBt - Aachen Biology and Biotechnology, RWTH Aachen University, Aachen, Germany, (5)Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland

Limited reserves of oil and natural gas and the establishment of new environmental policies sparked off intensive research towards sustainable production of the 2nd generation of biofuels, with Methyl Ethyl Ketone (MEK) being one promising fuel candidate. MEK is a commercially valuable petrochemical with an extensive application as a solvent. However, as of today, a sustainable and economically viable production of MEK has not yet been achieved despite several attempts of introducing biosynthetic pathways in industrial microorganisms. We used BNICE.ch as a retrobiosynthesis tool to discover all novel pathways around MEK. Out of 1'325 identified compounds connecting to MEK with one reaction step, we selected 3-oxopentanoate, but-3-en-2-one, but-1-en-2-olate, butylamine, and 2-hydroxy-2-methyl-butanenitrile for further study. We reconstructed 3'679'610 novel biosynthetic pathways towards these 5 compounds. We then embedded these pathways into the genome-scale model of *E. coli*, and a set of 18'622 were found to be most biologically feasible ones based on thermodynamics and their yields. For each novel reaction in the viable pathways, we proposed the most similar KEGG reactions,

with their gene and protein sequences, as candidates for either a direct experimental implementation or as a basis for enzyme engineering. Through pathway similarity analysis we classified the pathways and identified the enzymes and precursors that were vital for the production of the target molecules. These retrobiosynthesis studies demonstrate the potential of BNICE.ch for discovery, systematic evaluation, and analysis of novel pathways in synthetic biology and metabolic engineering studies.

Predicted Responses of a Large-Scale *Pseudomonas Putida* KT2440 Kinetic Metabolic Model to Several Single-Gene Knockouts Are Consistent with Experimental Observations.

Milenko Tokic¹, Ljubisa Miskovic², and Vassily Hatzimanikatis³

(1)Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), 1015 Lausanne, Switzerland, (2)Laboratory of Computational Systems Biotechnology (LCSB), EPFL, Lausanne, Switzerland, (3)Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland

P. putida emerged as one of the most promising production hosts for a wide range of chemicals, due to its fast growth with a low nutrient and cellular energy demand, considerable metabolic versatility, ability to grow in wide range of chemicals, suitability for genetic manipulations and its robustness and high flexibility to adapt and counteract different stresses. One of the main advantages of *P. putida* compared to commonly used industrial hosts such as *E. coli* is its superior tolerance to toxic compounds such as benzene, toluene, ethylbenzene, xylene and other hydrocarbons. In this work, we developed a large-scale kinetic model of *P. putida* to predict the metabolic phenotypes and design metabolic engineering interventions for the production of biochemicals. We first performed a gap-filling and thermodynamic curation of the genome-scale iJN1411 model of *P. putida* KT2440. The redGEM and lumpGEM algorithms for the systematic reduction of stoichiometric genome-scale models are then applied to the curated iJN1411 to derive a consistently reduced large-scale stoichiometric model of *P. putida*. Using this model as a scaffold, we next employed the ORACLE framework to generate a population of large-scale kinetic models around the experimentally observed steady state. To illustrate the predictive capabilities of these models, we performed two studies. First, for a wild-type strain of *P. putida* KT2440 growing on glucose under aerobic

conditions, we computed metabolic responses to several single-gene knockouts, and the developed kinetic models successfully captured the experimentally observed phenotypes. In the second study, we proposed metabolic engineering interventions for improved robustness of this organism to stress conditions. Overall, the results from these studies suggest that the developed models of *P. putida* metabolism can successfully be used for metabolic engineering design.

Genome-Scale Metabolic Modeling of Actinomycetes for Secondary Metabolites Production.

Omkar Mohite¹, Tilmann Weber¹, Hyun Uk Kim², and Sang Yup Lee^{1,2}

(1)Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Kgs. Lyngby, Denmark, (2)Department of Chemical and Biomolecular Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, Republic of (South)

Actinomycetes are an important source for the discovery of medically useful secondary metabolites, such as anticancers and antibiotics. Major challenges of producing secondary metabolites include optimization of metabolic fluxes towards the biosynthesis of a target secondary metabolite and identification of an appropriate heterologous production host. Systems biology approaches including reconstruction of genome-scale metabolic models (GEMs) are increasingly applied to explore the dependencies between primary and secondary metabolism to predict gene manipulation targets for the overproduction of secondary metabolites. GSM (Genome-scale metabolic Modeling with Secondary Metabolism) pipeline (unpublished) allows automatic reconstruction of GEMs with secondary metabolites biosynthetic pathways. Here, we exemplify the strength of GSM by reconstructing GEMs of important actinobacteria such as *Streptomyces venezuelae* and *Streptomyces collinus*. Growth prediction performance of GEMs under a variety of substrate conditions was compared with and fine-tuned using these two organisms' phenotype microarray data. The resulting GEMs can provide an important tool for applying systems metabolic engineering strategies for the optimal production of secondary metabolites in various actinomycetes as a heterologous production host. In this regard, GSM can serve as a useful platform to explore the potential of various actinomycetes by generating a collection of GEMs to produce diverse secondary metabolites.

POSTER ABSTRACTS

Genome Scale Metabolic Analysis of the *Lactobacillus* Genus with a Focus on Probiotic Strains.

Thomas J. Moutinho Jr., Benjamin C. Neubert, and Jason A. Papin

Department of Biomedical Engineering, University of Virginia, Charlottesville, VA

The *Lactobacillus* genus has, primarily, dominated the probiotic industry with a great deal of success in conferring health benefit as shown in a variety of rigorous clinical trials. In this study, we generate over 200 genome-scale metabolic reconstructions from a broad array of *Lactobacillus* species, including many of the currently sequenced species that are typically found in marketed probiotics, resident strains from human isolation sites, and fermented foods. We have conducted a strain level comparison of the reactomes for each genome, demonstrating that the metabolic reactome contributes to clustering of the genomes based on human health related isolation sites, including the gut, oral cavity, and vagina. Strains resident to the vaginal microbiota cluster separately from strains resident to the gastrointestinal tract, while there is a large overlap in metabolic potential between gut and oral residents. A random forest analysis was used to identify reactions specific to resident strains to the gastrointestinal tract when compared to other *Lactobacillus* strains. Carbon and amino acid source utilization was predicted using the generated metabolic reconstructions, our data indicate that there is a great deal of variation among *Lactobacillus* strains with some correlation to isolation site. We present a phylogenetic analysis based on metabolic capabilities, demonstrating the similarities among *Lactobacillus* strains indicating which groups are most likely to compete in a microbial community. The metabolic reconstructions were generated using ModelSEEDs universal format, the PATRIC database, Mackinac, CobraPy, and Probannopy. This study provides a public resource of systematically generated genome scale metabolic reconstructions for a wide range of *Lactobacillus* strains for further curation and analysis of probiotics relevant to human oral, gut, and vaginal health. Additionally, this work lays the ground work for personalize design, selection, and delivery of *Lactobacillus* based probiotics.

Multiple Reconstructions of *P. Putida* for Metabolic Engineering Applications.

Joshua Mueller

Chemical and Biomolecular Engineering, University of Nebraska - Lincoln, Lincoln, NE; Biological Engineering, University of California, San Diego, San Diego, CA

Pseudomonas putida is an industrially relevant species of bacteria. It has Generally recognized as safe (GRAS) status, is amenable to genetic manipulations, and has a robust metabolism and tolerance to several solvents used in industry. These attributes give it the potential to be a versatile platform for the bio-industry. Since 2003 multiple different strains of *P. putida* with varying capabilities have had their genomes fully sequenced. Previously detailed models of the strain KT2440 have been created, but little has been done about other strains. By creating a multi-strain reconstruction of more than 20 strains of this species, new pathways and usages can be elucidated. A pan- and core-genome can be identified which will aid in identifying central pathways as well as peripheral pathways that can be manipulated to add new capabilities and improve output.

System-Level Examination of Metabolism in Rhizosphere Grown Burkholderiales.

Ali Navid¹, Jennifer Pett-Ridge², Evan Star³, Jillian F. Banfield⁴, Mary Firestone⁵, and Erin Nuccio⁶

(1)Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, Livermore, CA, (2)Physical and Life Sciences Directorate, Lawrence Livermore National Laboratory, (3)2Department of Plant & Microbial Biology, University of California, Berkeley, CA, (4)University of California, Berkeley, CA, (5)Department of Environmental Science, Policy & Management, University of California, Berkeley, CA, (6)Lawrence Livermore National Laboratory, Livermore, CA

It is well known that rhizosphere (in the vicinity of plant roots) microbes have a significant effect on plant health and growth rate. However, our knowledge of the system-level biochemical interactions between these microbes and their surroundings is incomplete. To gain a better understanding of these interactions, we have sequenced the soil metagenomes and constructed metagenome assembled genomes (MAGs) of active organisms growing in rhizosphere and bulk soil. We have developed a genome-scale model of metabolism for one of these MAGs (a member of the Burkholderiales family, SM_S39) and used the model to explore time-resolved gene-expression data for samples collected within and outside of the influence of roots.

Our results show that SM_S39 greatly upregulates its import and metabolism of inorganic and complex nitrogen sources such as amino acid dimers, chitosamine when growing in the rhizosphere. In addition, this organism's rhizosphere CAZyme transcripts were upregulated for enzymes potentially involved in the decomposition

of cellulose, hemicellulose, cell membranes, lipids, polysaccharides, proteins, starches, and biopolymers.

We also identified pathways that were significantly upregulated when comparing the bulk soil to the rhizosphere soil environment. We found that pathways related to generation and consumption of fatty metabolites and those involved in biosynthesis and metabolism of amino acids (especially aromatic, branched-chain, as well as some common nitrogen carrying amino acids aspartate and arginine) were strongly upregulated. These results indicate that although it is intuitive to think that exchange of carbon between the root and the microbes is the primary driver of the metabolic interactions between the plant and the soil microbes like SM_S39, microbial nitrogen metabolism might be affected even more significantly by the presence of roots.

The Strain Optimization API: A Flexible Strain Design Formalism for an Automated High-Throughput Industrial Pipeline.

Chiam Yu Ng

Bioinformatics and Software Engineering, Amyris, Inc, Emeryville, CA

As a pioneer in making molecules in a sustainable way by strain engineering and industrial fermentation, Amyris has invested heavily in technology to develop strains and bring valuable molecules to the market at unprecedented speed. In the past three years, we have developed the Automated Scientist, a suite of mathematical models and machine learning algorithms that can iteratively perform the Design-Build-Test-Learn cycle for automated strain design and improvement. Thus far, the Automated Scientist has successfully created > 50,000 strains producing over 190 molecules. The design module of the Automated Scientist incorporates genome-scale metabolic models, an in-house knowledge store of design elements annotated with public data, and proprietary design algorithms to create pathway and strain designs. The pathway design to a target molecule can be visualized in the open-source tool Escher (first extended for industrial application at Amyris in 2014 by the author of the code) for route curation and additional flux balance simulation by biologists. To facilitate the exchange of knowledge between human scientists and the Automated Scientist, we have recently developed the Strain Optimization API. The API allows human scientists to specify the intended strain designs at a high level of abstraction, to validate their designs in silico, and to visualize the genotype-phenotype relationships. The Automated Scientist can parse the high-level input

and automatically design the genetic architecture of the strains. In addition, the API can be accessed by Amyris' high-throughput omics pipeline to identify strain-specific gene/protein/metabolite targets for testing up to 5,000 samples/day, and thus seamlessly accelerate the strain re-design process.

Genome Scale Metabolic Modeling and Analysis of *Clostridium Difficile*.

Charles J. Norsigian, Bernhard Ø. Palsson, and Jonathan M. Monk

Bioengineering, University of California, San Diego, La Jolla, CA

Clostridium difficile is a pathogen of high clinical interest due to its persistence as a hospital-borne infection. Previous studies suggest that differences in metabolic capabilities are contributing factors to resistance and virulence. As a means of systematically understanding the linkage between genotype and phenotype of this organism we constructed a new genome-scale reconstruction of *C. difficile* strain 630 that builds and improves upon previous efforts. The model recapitulates experimental gene knockout predictions with 91% accuracy. We deploy constraint-based reconstruction and analysis along with flux-balance analysis to investigate metabolic capabilities by systematically predicting growth capabilities in over 180 nutrient environments. We utilize the reconstruction to build strain-specific models of 32 isolates including both symptomatic and asymptomatic strains and identified unique differentiating metabolic capabilities. We experimentally validate the models by generating Biolog phenotypic microarray data for all 32 of these strains and compared results to model predictions.

An Integrated COBRA-PBPK Model to Study Interactions between the Gut Microbiome and the Brain in Autism.

Meghana Palukuri¹, Shruti Shivakumar², Swagatika Sahoo³, and Raghunathan Rengaswamy³

(1)Institute for Computational Engineering and Sciences, The University of Texas at Austin, Austin, TX, (2)College of Computing, Georgia Institute of Technology, Atlanta, GA, (3)Indian Institute of Technology Madras, Chennai, India

Autism, also ASD (Autism Spectrum Disorder) is a complex neurological condition characterized by social, behavioural and developmental issues such as impaired communication and repetitive actions. Gastro-intestinal disturbances in autistic patients have motivated experiments that identified abnormal concentrations of certain bacteria in the gut. To understand the role

played by the most reported gut bacteria associated with autism, namely *Bacteroides vulgatus*, *Clostridium perfringens*, *Desulfovibrio desulfuricans*, *Bifidobacterium longum* and *Lactobacillus acidophilus*, a hybrid modeling approach was adopted. Genome-scale metabolic models of these bacteria were integrated with the human small intestinal enterocyte model. Simultaneously, a permeability-limited, two sub-compartment, whole body PBPK transport model was incorporated to determine the distribution of gut-bacteria derived toxins throughout organs. A novel framework was formulated for the development and analysis of an integrated gut-PBPK-brain model and was applied to the study of autism for the first time. The integrated model predicted that intake of probiotics (*L. acidophilus* and *B. longum*) leads to reduced oxidative stress in the brain, resulting in relief from autistic symptoms. Further, investigating the effect of various dietary schemes indicated that a typical western diet worsens symptoms in autistic individuals. Importantly, potential biomarkers for autism were predicted in terms of bacterial secretion products and affected pathways in the genome-scale metabolic model of the brain. Thus, the integrated model breaks new ground in developing a quantitative model for studying gut-brain interactions, with the potential to understand and remedy several neurological problems with links to the gut, such as autism, depression and anxiety. More generally, the framework for integration and analysis of constraint-based and PBPK models can be applied to study dynamically, the interactions between systems in various fields. Its applications are demonstrated by providing a mechanistic basis for autism pathogenesis and by proposing dietary changes and intake of probiotics for the alleviation of autistic symptoms.

Genome-Scale Constraint-Based Metabolic Modeling and Analysis of *Cryptococcus Curvatus*.

Nhung Pham, Maarten Reijnders, Maria Suarez-Diez, Bart Nijssse, Peter J Schaap, and Vitor A. P. Martins Dos Santos
Laboratory of Systems and Synthetic Biology, Wageningen University and Research, Wageningen, Netherlands

Cryptococcus curvatus, also known as *Cutaneotrichosporon curvatus*, an oleaginous fungus, is gaining attention due to its ability to use a wide range of carbon sources and its ability to accumulate more than 60% lipid in its biomass. In this study we constructed a genome-scale metabolic model for *C. curvatus* that can help to understand and achieve higher yield of lipid production.

A draft model was constructed using the scaffold-based method by Loira et al (2012). The draft model was extensively curated based on published data and literature with special emphasis on lipid synthesis pathways. The model was used to simulate growth on different carbon sources. RNAseq data were obtained at C:N of 16:1 and 8:5 with two replicates for each condition. These expression data were integrated into the model to study lipid synthesis in *C. curvatus*.

The current version of the model has 831 genes, 1436 metabolites and 1679 reactions distribute in 11 compartments. The growth predictions are highly consistent with experimental data. However, the initial model was not able to predict differences in lipid synthesis at varying nitrogen concentrations. Integration of expression data will lead to improve predictions on lipid production at nitrogen abundant and nitrogen depleted conditions. These results illustrate that the model can be further used to study and guide metabolic engineering for optimization of *C. curvatus* as a cell factory for lipid production.

References

Loira, N., Dulermo, T., Nicaud, J.-M., & Sherman, D. (2012). A genome-scale metabolic model of the lipid-accumulating yeast *Yarrowia lipolytica*. *BMC Systems Biology*, 6(1), 35. <https://doi.org/10.1186/1752-0509-6-35>

Identifying Orphan Enzymes in *Pseudomonas Putida* Using Cobra-Based Methods.

Brandon Phan¹, Shu Pan², Christine Vera Colon³, Nathaniel Bennett³, and Jennifer L. Reed³

(1)Cellular and Molecular Biology, University of Wisconsin-Madison, Madison, WI, (2)Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI, (3)Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI

Genome-scale metabolic models have a wide variety of applications. While these models capture gene-protein-reaction relationships, they often are incomplete because they include misidentified reactions or metabolic gaps due to inaccurate genome annotations. Identifying genes responsible for missing reactions will provide a more complete understanding of metabolic networks and genotype-phenotype relationships. Previously, we developed a model-enabled gene search (MEGS) method to identify missing genes in a metabolic network associated with gap-filled reactions using functional selection experiments. MEGS involved creating a genomic library for a species with gap-filled reactions, and transforming the library into a recipient strain,

such as *Escherichia coli*, whereby, the missing gap-filled reaction is coupled to growth [1]. Here, we extend MEGS to find genes associated with orphan enzymes in *Pseudomonas putida*. Databases exist that list orphan enzymes from many different organisms, whose metabolic activity has been shown experimentally but for which no associated gene sequences are known (including ~17% of enzymes with EC numbers) [2]. To verify some of the orphan enzymes in *P. putida*, we used COBRA methods to design *E. coli* recipient strains, which can only grow if they express genes associated with the appropriate orphan enzyme from *P. putida*. By finding sequences for orphan enzymes, we can build new GPR associations, thereby improving model predictions not only for *P. putida* but for other organisms with similar genes as well.

References:

- Pan, S., Nikolakakis, K., Adamczyk, P. A., Pan, M., Ruby, E. G., & Reed, J. L. (2017). Model-enabled gene search (MEGS) allows fast and direct discovery of enzymatic and transport gene functions in the marine bacterium *Vibrio fischeri*. *Journal of Biological Chemistry*, 292(24), 10250–10261.
- Shearer, A. G., Altman, T., & Rhee, C. D. (2014). Finding sequences for over 270 orphan enzymes. *PLoS ONE*, 9(5).

Enzyme Multicollinearity in Genome-Scale Metabolic Models Revealed By an Efficient Coupling Algorithm

Dikshant Pradhan and **Paul A. Jensen**

University of Illinois at Urbana-Champaign Bioengineering Department

Numerous reactions control bacterial behavior and homeostasis. Dimensionality reduction is often necessary to analyze COBRA models, especially non-convex models that require intensive computations. One method to reduce networks is to identify reactions with colinear, or fully coupled, fluxes. In a COBRA model, coupled reactions can be identified using the Flux-Coupling Finder (FCF) algorithm. However, identification of coupled reaction sets is not as useful as identification of coupled enzyme sets, since enzymes are targets of metabolic engineering and drug discovery efforts. To calculate coupling between enzymes, we developed a mathematical transformation linking reaction fluxes and enzyme activity while preserving the nonlinear mapping between genes, proteins, and reactions. This framework, called Flux and Activity Linked Constraints (FALCON) expands the original linear program to a larger mixed-integer linear program (MILP). Although solving MILPs requires significantly more computational power, we developed an efficient variant of FCF (cached-FCF) that drastically reduces the runtime required to identify couplings in both COBRA and FALCON models.

We tested cached-FCF on the FALCON model for the bacterium *Pseudomonas aeruginosa*. We discovered that the enzymes in coupled sets showed significantly higher correlation of expression and fitness than sets of enzymes created from sets of coupled reactions. The coupled enzyme sets form the functional units of the organism's metabolic network. We further tested the effects of perturbing the metabolic network with single-gene knockouts. We observed that gene deletions force additional couplings among enzyme sets. These additional couplings link enzyme sets with similar expression or fitness patterns.

Model-Guided Engineering of Cyanobacteria for Improved Biofuel Production.

Hugh M. Purdy and **Jennifer L. Reed**

Department of Chemical and Biological Engineering, University of Wisconsin-Madison, Madison, WI

Cyanobacteria are promising hosts for fuel and chemical production as their main metabolic inputs are minimal and renewable (i.e. light, carbon dioxide, water, and a nitrogen source). However, cyanobacteria engineered to produce compounds of interest frequently suffer from low yields and/or are genetically unstable. As such, metabolically coupling chemical production to growth may be important in engineering cyanobacteria because it can generate robust production characteristics. Additionally, a growth-coupled production approach may help to circumvent gaps in our knowledge about cyanobacterial metabolic regulation, which likely factors into the poor production characteristics that are often reported relative to heterotrophic strains. In this work, a genome-scale metabolic model is being used to guide engineering of the cyanobacterium *Synechococcus* sp. PCC 7002 for growth-coupled overproduction of short- to mid-chain alcohols (e.g. n-butanol) that are potential biofuels. PCC 7002 is of particular interest as a cyanobacterial production strain due to its relatively rapid growth-rate and tolerance to saline and high-light conditions. Metabolic engineering strategies for improving the production of various alcohols in PCC 7002 were investigated via the metabolic model iSyp708, which was previously developed by our lab. Using this approach, a potential strategy for high-yield growth-coupled production of n-butanol and 2-methyl-1-butanol in PCC 7002 was successfully identified. This strategy hinges on coupling production of these alcohols to nitrate assimilation by rewiring PCC 7002's native NADH-cycling pathways. Strains are currently being constructed to test this approach. If a strongly growth-

POSTER ABSTRACTS

coupled production strain is successfully engineered, it will be adaptively evolved to further improve its alcohol production characteristics and genetic changes in these evolved strain(s) will be identified and analyzed. This study demonstrates that genome-scale metabolic models are useful in guiding cyanobacterial metabolic engineering and will help identify genetic alterations for optimizing the production of biofuel compounds in cyanobacteria.

Kinetic Model of *Clostridium Beijerinckii* Based on Phenotypic States Superposition.

Marcelo Rivas-Astroza, Ivan Paredes, German Aroca, and Raul Conejeros

School of Biochemical Engineering, Pontificia Universidad Catolica de Valparaiso, Valparaiso, Chile

Growth kinetics follow lag, exponential, and stationary phases, with their transitions marked by drastic changes in biomass growth rates. This may be due to shifts in transcriptional activity, where high substrate concentrations promote increasing ribosome production to support the exponential phase. Under this hypothesis, there are smooth transitions from low to high growth rates on a phenotypically uniform population of cells. However, this is not the case for *Clostridium beijerinckii* which, when grown in batches, present various phenotypes coexisting. It starts as a vegetative phenotype, where fast substrate consumption is paired by fast biomass growth and production of acidic products. Acidification promotes a fraction of cells to morph into a clostridial phenotype. These are non-dividing cells whose metabolism is geared towards sporulation, a process where glucose and the previously released acidic species in the medium are consumed, producing in turn solvents. Released spores may later germinate into vegetative cells if conditions became favorable.

We propose a kinetic model where all three stages of *C. beijerinckii* are present at any given time. The dynamics between phenotypic states are described by differential equations within dynamic FBA. To model the vegetative-clostridial and spore-vegetative transitions, we used the concentration of acidic species as arguments to a logistic equation whose value can be interpreted as the fraction of the vegetative cells transiting to clostridial state. Likewise, we used a second logistic equation, this time as a function of the concentration of solvents, to model the fraction of clostridial cells that change into spores.

Using public RNA-seq databases we determined divergently expressed genes in each phenotype. We pruned the metabolic network of *C. beijerinckii* accordingly

to create phenotype-specific metabolic networks. To train our model we measured changes in substrates, and acidic and solventogenic species along batch cultures of *C. beijerinckii* under various conditions.

Etfl: A Formulation of ME-Models Accounting for Expression and Thermodynamics in Constraint-Based Models.

Pierre Salvy¹ and Vassily Hatzimanikatis²

(1)Laboratory of Computational Systems Biotechnology (LCSB), EPFL, Lausanne, Switzerland, (2)Laboratory of Computational Systems Biotechnology (LCSB), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland

Since flux balance analysis (FBA) has been introduced in Systems Biology, several attempts have been made at supplementing it with expression data [1-3]. With the notable exception of O'Brien et al. [1], none of them account directly for the enzyme production as a part of the mathematical programming formulation of the problem. Furthermore, O'Brien et al. propose a bilinear programming formulation that is computationally heavy.

We propose a top-down model formulation, from metabolism to RNA synthesis, that allows the simulation of thermodynamics-compliant intracellular fluxes, as well as enzyme and mRNA levels. The formulation results in a mixed integer linear problem (MILP) which enables both relative and absolute metabolite, proteins, and mRNA concentration integration. The proposed formulation does not require any bilinear solver. We present here the results obtained using the *Escherichia coli* model iJO1366 [4]. We show that the formulation is able to reproduce proteome-limited growth, which flux balance analysis cannot [1], as well as predict feasible mRNA and enzyme concentrations in the cell.

O'Brien, E.J., et al., Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Mol Syst Biol*, 2013. **9**: p. 693.

Sánchez, B.J., et al., Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Molecular Systems Biology*, 2017. **13**(8).

Beg, Q.K., et al., Intracellular crowding defines the mode and sequence of substrate uptake by *Escherichia coli* and constrains its metabolic activity. *Proc Natl Acad Sci U S A*, 2007. **104**(31): p. 12663-8.

Orth, J.D., et al., A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism-2011. *Molecular Systems Biology*, 2011. **7**.

Effective Visualization for Investigating Elementary Flux Modes in Genome-Scale Metabolic Models.

Chaitra Sarathy¹, Michael Lenz¹, Martina Kutmon^{1,2}, Chris T. Evelo^{1,2}, and Ilja C.W. Arts^{1,3}

(1)Maastricht Centre for Systems Biology (MaCSBio), Maastricht University, Maastricht, Netherlands, (2) Department of Bioinformatics - BIGCaT, NUTRIM School of Nutrition and Translational Research in Metabolism, Maastricht University, Maastricht, Netherlands, (3) Department of Epidemiology, CARIM School for Cardiovascular Diseases, Maastricht University, Maastricht, Netherlands

Elementary Flux Modes (EFMs) are an indispensable tool for constraint-based modeling and metabolic network analysis¹. EFMs allow exploration of cellular metabolism beyond conventional pathway definitions and integration with high-throughput data. Since the field has been driving towards improving algorithms to enumerate all (or subsets) of EFMs, their visualization has been limited to mapping only the flux data in standalone tools. However, flux mapping is just one component of systems studies, where several molecular datatypes are combined. There is a need for an efficient visual analysis approach that complements current efforts integrating omics data with genome-scale metabolic models. In this study, we developed a simple, efficient and MATLAB-based workflow for graphically visualizing EFMs as a network of reactions, metabolites and genes. Our flexible workflow seamlessly integrates COBRA² or RAVEN³ with the open-source tool, Cytoscape⁴. Cytoscape not only enables data visualization and advanced network analysis, but also network extension with other molecular interaction datatypes, all within a single framework. Once processed through our workflow, SBGN layout is automatically applied on the visualized network requiring little-to-no manual effort by the user. Furthermore, combining existing tools avoids additional installation of flux-mapping tools, used previously, thus making it both user-friendly and time-saving. We illustrate our workflow using a subset of EFMs generated from both small and large models and map gene expression data on the visualized EFMs, thereby, demonstrating that efficient, integrative visualization can allow for a better understanding of biological processes.

References: [1] Klamt, S. *et al.*, DOI:10.1371/journal.pcbi.1005409. [2] Schellenberger J. *et al.*, DOI:10.1038/nprot.2011.308. [3] Agren, R. *et al.*, DOI:10.1371/journal.pcbi.1002980. [4] Shannon, P. *et al.*, DOI:10.1101/gr.1239303.

Deconvoluting Independent Regulatory Signals in the Escherichia coli Transcriptome.

Anand Sastry¹, Ye Gao², Richard Szubin¹, Ying Hefner¹, Sibe Xu¹, Donghyuk Kim^{1,3}, Kumari Sonal Choudhary¹, Laurence Yang¹, Zachary A. King¹, and Bernhard O. Palsson¹

(1)Bioengineering, University of California, San Diego, La Jolla, CA, (2)Biological Sciences, University of California, San Diego, San Diego, CA, (3)Department of Genetic Engineering and Graduate School of Biotechnology, Kyung Hee University, Yongin, Korea, Republic of (South)

Underlying cellular responses is a transcriptional regulatory network (TRN) that modulates gene expression. Current descriptions of the TRN cannot quantitatively deconvolute the transcriptome into the relative contributions of individual transcriptional regulators. Here, we applied unsupervised learning to a compendium of high-quality Escherichia coli RNA-seq datasets (median $R^2 = 0.99$ between biological replicates) to identify 71 statistically independent regulatory signals. Summation of the 71 signals explained over 80% of expression variation across 115 unique experimental conditions. Of these, 50 signals were directly linked to characterized transcriptional regulators. Condition-specific signal strengths were validated by exposure to new environmental conditions, confirming 76% of predicted signal activations. The resulting decomposition of the transcriptome provided: 1) a quantitative, mechanistic explanation of responses to environmental and genetic perturbations, 2) a guide to gene function discovery, 3) characterization of a novel pyruvate-responsive transcription factor, and 4) a basis for comparing transcriptional regulation across closely related strains. Thus, we find that signal summation forms an underlying principle that can describe the composition of the prokaryotic transcriptome.

Characterizing and Ranking Computed Metabolic Engineering Strategies.

Philipp Schneider and Steffen Klamt

Analysis and Redesign of Biological Networks, Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany

Computational strain design, i.e. the calculation of metabolic intervention strategies from a mathematical model, is a key component of an integrated metabolic engineering approach. A broad range of methods has been developed for this task, including bilevel optimization approaches (e.g., OptKnock and RobustKnock) and the framework of Minimal Cut Sets (MCSs). Most of these

POSTER ABSTRACTS

algorithms enforce the coupling between growth and the synthesis of the desired product. Some of them, such as the MCSEnumerator, may return a large pool of thousands of possible intervention strategies from which the most suitable strategy can then be selected.

This work focuses on how to characterize and rank, in a meaningful way, a given pool of intervention strategies calculated for growth-coupled product synthesis. Some criteria are straightforward and can be assessed easily, for example, the number of necessary reaction or gene cuts, the maximal growth rate and the guaranteed (minimum) product yield. Additionally, we present more extensive criteria that are worth considering when picking the 'best' intervention strategy. Among others we investigate the existence of metabolites that may disrupt the growth coupling when accumulated or secreted and check whether the interventions interrupt pathways at their origin or (less preferred) at later steps. We also assess the maximum thermodynamic driving force of the pathway(s) favored by the intervention strategy and take into account whether a specific intervention strategy relies on flux re-routing within the central metabolism or on other pathways with possibly minor throughputs. Furthermore, strategies that have significant overlap with alternative solutions are also ranked higher because they provide flexibility in implementation.

We applied the presented ranking approach to several sets of intervention strategies for growth-coupled synthesis of native and heterologous products computed from the genome-scale *E. coli* model iJO1366 using the MCSEnumerator algorithm.

Optfill: A Novel Optimization-Based Tool to Automate the Gapfilling of Genome-Scale Metabolic Models.

Wheaton Schroeder¹ and **Rajib Saha²**

(1)Chemical and Biomolecular Engineering, University of Nebraska - Lincoln, Lincoln, NE, (2)Department of Chemical and Biomolecular Engineering, University of Nebraska-Lincoln, Lincoln, NE

Computational modeling of metabolism is now an indispensable tool to drive the processes of understanding, discovering, and redesigning of biological systems. By defining the metabolic space, genome-scale metabolic (GSM) models can assess allowable cellular phenotypes and explore metabolic potential under specific environmental and/or genetic conditions. GSM model curation processes typically involve gleaning information on gene annotations and reactions from major public databases; however,

incomplete gene annotations and system knowledge leaves gaps in any GSM reconstruction. Gapfill (and numerous similar tools) automates addressing of these gaps, applying Mixed Integer Linear Programming (MILP) and utilizes functionalities from related organisms or changing the direction of existing reactions to fill gaps on a per-gap basis. This is done without consideration of thermodynamically infeasible cycle (TIC) avoidance. Hence, Gapfill makes redundant changes and increases the number of TICs in GSM models, which ultimately require further manual scrutiny. To address the limitations of current automated gapfilling procedures, introduced here is an improved method, namely OptFill, which performs automated gapfilling. OptFill applies two MILP problems in series, which addresses the fixes needed on a per-GSM model basis. The first optimization problem (single-level MILP) seeks to identify all TICs which could be created by adding new functionalities to the GSM model. The second optimization problem (multi-level MILP) seeks to maximize the number of gap metabolites fixed subject to minimizing the number of new functionalities added. Included in the second problem is an integer cut based on the results of the first problem, which prevents any solution containing TICs from being viable solutions. OptFill has thus far been successfully applied to metabolic models of a generic test system, *Saccharomyces cerevisiae* (baker's yeast), *Exophiala dermatitidis*. Thus, OptFill provides a distinct advantage over the traditional Gapfill approach in the extent of automation and needed manual curation.

Sammi: An Interactive, Semi-Automated Tool for the Illustration and Visualization of Metabolic Networks.

Andre Schultz and **Rehan Akbani**

Bioinformatics and Computational Biology, University of Texas MD Anderson Cancer Center, Houston, TX

The visualization of large metabolic networks is a challenging process. Bottom-up, manual construction of such maps is extremely laborious, while automated methods can yield useful but often muddled networks. Therefore, the development of semi-automated network visualization tools, which automatically generate large-scale metabolic maps while at the same time allowing user-defined manual curation, are necessary. Here we present SAMMI, an interactive tool for the illustration and visualization of metabolic networks. SAMMI draws metabolic networks as dynamic bipartite graphs. Force-directed node orientation allows for the constant repositioning of network nodes, assisting in the map design and visualization process. SAMMI also provides

a wide array of customization features that allow users to efficiently build metabolic maps. These functionalities include positioning, fixing, grouping, splitting, and coloring of network nodes, temporarily removing secondary metabolites, adding and removing metabolites and reactions, adding annotation text and shapes, and parsing the model into subgraphs based on model annotation (e.g. reaction subsystem or metabolite compartment) or user-defined input (e.g. lists of reactions). The implementation of graph theory algorithms also allow for the fast selection and re-arrangement of relevant network nodes. SAMMI illustrations can be exported in SAMMI specific or ESCHER compatible formats, as well as PDF images suitable for publications. We believe SAMMI is a powerful tool for the illustration of metabolic network maps, as well as the visualization of multiple data types.

Genome-Scale Metabolic Reconstructions for Phylogeny?

Christian Schulz and **Eivind Almaas**

Dept. of Biotechnology, NTNU - Norwegian University of Science and Technology, Trondheim, Norway

Constraint-based modeling is a widely used and powerful methodology to assess the metabolic phenotypes and capabilities of an organism. Genome-scale metabolic models, based on the genome and including all reactions of the organism, are mostly created for that purpose. However, that does not cover their full potential.

Since all reactions are included in the models, it is possible to compare these models and thereby the species directly by creating a consolidated reconstruction of several organisms. This consolidated reconstruction including every reaction occurring at least once allows for identifying cross-organism wide conserved reactions, linked to a gene of the source-organism.

AutoKEGGRec is an automated tool that creates first draft reconstructions of single organisms, community and a consolidated reconstruction for a list of organisms or strains. AutoKEGGRec is developed in Matlab and works seamlessly with the COBRA Toolbox v3, and it is based on only using the KEGG database as external input. The first draft reconstructions consist of all reactions for a KEGG organism ID and corresponding linked genes. This provides a comprehensive starting point for further refinement and curation using the host of COBRA toolbox functions or other preferred tools. Furthermore, additional functionalities are provided, like the organisms-reaction-genes matrix allowing for advanced cross-organism reaction engineering.

Here, a consolidated model for organisms across Domains was created using the innovative capabilities of AutoKEGGRec. The metabolic reconstructions were used to create a phylogenetic tree of life based on the basis of the reaction networks. Both the direct gene sequence as well as the underlying protein sequence were used.

This tree of “metabolic networks” can be used to compare organisms focusing on the whole metabolome instead of using small parts of the DNA. Furthermore, it enlightens the metabolic engineering possibilities in between species throughout their maybe similar reactions for similar living environments in distinct points.

A Computational Knowledge-Base Elucidates the Response of Staphylococcus Aureus to Different Media Types.

Bernhard O. Palsson¹, Yara Seif², Jon Monk³, Nathan Mih⁴, Saugat Poudel², Jared T. Broddrick¹, Cristal Zuniga P⁵, Hannah Tsunemoto⁶, and Karsten Zengler⁷

(1)Bioengineering, University of California, San Diego, La Jolla, CA, (2)UCSD, San Diego, CA, (3)UCSD, La Jolla, CA, (4)UCSD, (5)Pediatrics, University of California, San Diego, CA, (6)University of California San Diego, San Diego, (7)Department of Pediatrics, University of California, San Diego, CA

S. aureus is classified as a serious threat pathogen and is a priority that guides the discovery and development of new antibiotics. It has been observed to exhibit differential antibiotic resistance profiles when cultured in different media. Despite growing knowledge of *S. aureus* metabolic capabilities, our understanding of its systems-level responses to different media types remains incomplete. Here, we develop a manually reconstructed genome-scale model (GEM-PRO) of metabolism with 3D protein structures for *S. aureus* USA300 str. JE2 containing 854 genes, 1,440 reactions, 1,327 metabolites and 673 3-dimensional protein structures. Computations were in 85% agreement with gene essentiality data from random barcode transposon site sequencing (RB-TnSeq) and 68% agreement with experimental physiological data. Comparison of computational predictions with experimental observations highlight: 1) cases of non-essential biomass precursors; 2) metal cofactor promiscuity; 3) genes subject to transcriptional regulation involved in Staphyloxanthin biosynthesis; and 4) gaps of knowledge in C2-carbon assimilation; 5) the essentiality of purine and amino acid biosynthesis in synthetic physiological medium; and 6) a switch to aerobic glycolysis upon exposure to extracellular glucose that was elucidated using a time-course of quantitative

POSTER ABSTRACTS

metabolomics data. An up-to-date GEM-PRO thus serves as a knowledge-based platform to elucidate *S. aureus*' metabolic response to its environment.

Describing the Glucose-Lactate Consumption Rate during Expansion and Osteogenic Differentiation of Human Bone Marrow Derived Mesenchymal Stem Cells: A Premise for Building a Systems Biology Model of Osteogenesis Using Metabolomics Analysis.

Póra B. Sigmarsdóttir¹, Sarah McGarrity², Óttar Rólfsson³, and Olafur Sigurjonsson⁴

(1)Reykjavík University, Reykjavík, Iceland, (2)Center for Systems Biology, University of Iceland, 101 Reykjavik, Iceland, (3)Centre For Systems Biology, University Of Iceland, Reykjavik, Iceland, (4)Reykjavik University

Transplant of cells derived from mesenchymal stem cells to treat diseases such as osteoporosis is tantalisingly close. The production of cells for transplant requires further study and currently there is lack of knowledge on metabolomic changes that occur during expansion and osteogenic differentiation (OD) of MSCs. It has been shown that mesenchymal stem cell osteogenic differentiation is accompanied by metabolic shifts, especially the utilisation of glucose. In order to better understand this process and expand the knowledge of other metabolic pathways that interact with stem cell fate this project is building constraint based metabolic models to integrate transcriptomic and metabolomics data.

Publically available transcriptomic data and new glucose and lactate measurements were combined in the CobraToolbox. This created initial models of the first stages of expansion and osteogenic differentiation based. Functional assays confirmed the process of osteogenesis in vitro.

Glucose and lactate measurements from six donors during expansion and osteogenic differentiation indicate multiple stages to these processes. . The final stage of expansion shows an increase in the use of glucose, without an increase in lactate production. In osteogenesis the average glucose consumption and lactate production measures higher over the later 14 days, giving an indication of the possibility of two phenotypes.

The models were able to recreate known metabolic features of mesenchymal stem cells during expansion and osteogenesis. These include relative levels of oxidative phosphorylation to glycolysis and production of kynurenine from tryptophan. These models also suggest glycan production as well as more central metabolism

as areas key to the shift from expansion to osteogenic differentiation.

Genome scale metabolic modelling offers a way of combining multiple types of omics data in order to better understand and manipulate the features of a biological system.

Mighty Morphing Metabolic Models: Leveraging Manual Curations for Automatic Metabolic Reconstruction of Clades.

Evangelos Simeonidis, Brendan King, Matthew A. Richards, and Nathan D. Price

Hood-Price Lab, Institute for Systems Biology, Seattle, WA

Genome-scale metabolic models have been employed with great success for phenotypic studies of organisms over the last two decades. The most difficult step in the reconstruction of these models is manual curation; a time-intensive and laborious process of literature review that is nevertheless essential for a high-quality network. Although various methods have been developed to accelerate automated reconstruction, much effort must still be expended to supplement an automatically-generated draft model with manually curated information from the published literature to achieve the quality needed for successful simulation of the metabolic processes of the organism. Here, we are utilizing our tool for likelihood-based gene annotation (King et al., 2018) to create a method that “morphs” a manually curated metabolic model to a draft model of a closely related organism. Our method combines genes from the original, manually curated model with genes from an annotation database to create a final structure that contains gene-associated reactions from both sources. The benefits of such an approach are twofold: 1) the effort and accumulated knowledge that has gone into the construction of the original model is leveraged to create a metabolic model for a closely related organism; 2) starting from an already completed and functioning model allows the user to run simulations at every step as necessary, offering the ability to predict how modifications will affect the performance of the model. We demonstrate our method by applying it on our model of *Methanococcus maripaludis* (Richards et al., 2016), to create morphed models of three related methanogenic archaea. Model morphing offers a viable alternative to other automated reconstruction methods, particularly for organisms that are evolutionarily dissimilar to those that form the foundation of annotation databases, and provides a faster method to reconstruct a clade of metabolic models for related organisms from a manually curated representative.

Utilizing RNA-Seq Data in Bayesian Estimation of Gene Activity States.

Eldon Sorensen¹, Jaden Boehme², Angela Gasdaska³, Matthew DeJongh⁴, Aaron Best⁵, William Lindsey⁶, and Nathan Tintle⁷

(1)Pacific Lutheran University, Tacoma, WA, (2)Oregon State University, Corvallis, OR, (3)Emory University, Atlanta, GA, (4)Computer Science, Hope College, Holland, MI, (5)Biology, Hope College, Holland, MI, (6)Dordt College, Sioux Center, IA, (7)Mathematics, Dordt College, Sioux Center, IA

Recently, there has been interest in exploring how to infer gene activity states (e.g., whether a gene is active or inactive in a particular condition of interest) from genome-wide transcriptomic data. This knowledge is useful in many downstream applications, including the potentially improved use of transcriptomics data to improve flux predictions in metabolic models. Recently, a rigorous Bayesian approach (MultiMM) to classifying gene activity states was proposed that leverages a priori knowledge of operon structure as well as genome-wide transcriptomics data from multiple conditions in order to classify gene activity states. However, the MultiMM approach was developed for use on microarrays, and only evaluated on a very large set of over 900 E. coli arrays. Here, we extend the Bayesian model to RNA sequencing data and then evaluate its performance. Importantly, we evaluate performance in situations with both large (100s) and small (10s) of conditions, and provide intuition on necessary sample sizes for robust performance.

Bayesian Inference of Metabolic Kinetics from Genome-Scale Multiomics Data.

Peter C. St. John¹, Jonathan Strutz², and Yannick J. Bomble¹

(1)Biosciences Center, National Renewable Energy Laboratory, Golden, CO, (2)Chemical and Biological Engineering, Northwestern University, Evanston, IL

Optimizing the metabolism of microorganisms for maximum yields and titers is a critical step in improving bioprocess economics. With the growing availability of transcriptomic, proteomic, metabolomic, and fluxomic analysis techniques, characterization of engineered strains has become increasingly detailed. However, utilizing multiomics data to make informed decisions about future strain improvements remains a major challenge in modern bioengineering. Parameterizing kinetic models from indirect, in vivo data is typically infeasible at the genome-scale. There is therefore a need for mechanistic

modeling frameworks that can consume the large amount of data generated through multiomics experiments to yield actionable insight for strain engineers.

Metabolic ensemble modeling has emerged in recent years as an effective tool for estimating confidence intervals in kinetic metabolic models from observable steady-state flux and concentration measurements. The method operates by sampling feasible kinetic parameters and filtering to sets of parameters that match experimental observations. However, these methods rely on an explicit numerical integration of large kinetic models, and therefore scale poorly both in the size of the network that can be modeled, as well as the amount of data that can be used to constrain the parameters.

We demonstrate the coupling of linear-logarithmic kinetics to a Bayesian inference framework. The resulting method therefore represents a scalable, flexible method for modeling omics data. Specifically, linear-logarithmic kinetics enable steady-state fluxes to be predicted linearly from kinetic parameters, removing the computational burden associated with solving for steady-state flux. Additionally, since linear solutions permit the easy determination of likelihood gradients, advanced Bayesian inference techniques can be applied to reliably estimate posterior parameter distributions even for high-dimensional models. We demonstrate the method through a number of case studies, from simple in vitro reaction systems to complex, medium-scale metabolic models with hundreds of metabolites and reactions parameterized with modern multi-omic data.

Dynamic FBA with Time-Course Transcriptomics.

Snorre Sulheim^{1,2}, Alexander Wentzel¹, and Eivind Almaas³

(1)Dept. of Biotechnology and Nanomedicine, SINTEF Industry, Trondheim, Norway, (2)Dept. of Biotechnology and Food Science, NTNU - Norwegian University of Science and Technology, Trondheim, Norway, (3)Dept. of Biotechnology, NTNU - Norwegian University of Science and Technology, Trondheim, Norway

Genome-scale metabolic models provide a framework for the computational prediction of a wide range of phenotypes, such as growth in different nutrient environments, and effects of genetic manipulations, as well as aiding in the interpretation of 'omics data. Dynamic flux balance analysis (dFBA) is a method where the dynamics of growth is incorporated into the framework of steady-state flux balance analysis (FBA) by updating environmental conditions at regular time-intervals. Here,

POSTER ABSTRACTS

we present tt-dFBA by extending the dFBA methodology through incorporating time-course transcriptome data to account for changes in gene expression levels during batch fermentation.

We apply our framework to the case of batch fermentation of *Streptomyces coelicolor*, using the recent reconstruction iKS1317, to explain the observed metabolic switching and onset of secondary metabolite production when the organism faces phosphate depletion. Our results provide insights into the metabolic flux rerouting causing the observed metabolic changes.

Determination of CHO Biomass Composition.

Diana Szeliova^{1,2}, **David Ruckerbauer**^{2,3}, **Sarah Galleguillos**^{1,2}, **Stephan Hann**², **Michael Hanscho**^{1,2}, and **Nicole Borth**^{1,2}

(1)The Austrian Centre of Industrial Biotechnology, Vienna, Austria, (2)University of Natural Resources and Life Sciences, Vienna, Austria, (3)Austrian Centre of Industrial Biotechnology, Vienna, Austria

Chinese hamster ovary (CHO) cells are the primary host organism for the production of protein biopharmaceuticals. Significant improvements in product yield and cell growth were achieved in the past years by bioprocess and media optimization, directed evolution and targeted genetic engineering. However, a deeper understanding of the underlying processes in the cells is still limited. Recently, a CHO-specific genome scale metabolic model was created in a large community effort. This model is a comprehensive resource of CHO metabolism. Using COBRA methods, we are now starting to get valuable insights into the cells' metabolism, their protein production capabilities and potential limitations. One essential input for the model is biomass composition. It has been shown that using strain and condition specific biomass together with bioprocess data improves the model's predictions. Currently, however, the model uses estimates and literature values, since comprehensive data about CHO cell composition, specifically of individual cell lines or strains, are lacking. In this work, methods for the determination of CHO biomass composition (proteins and amino acids, lipids, DNA, RNA and cell dry mass) were established. These include chromatography and mass spectrometry determination of amino acids, fluorimetric and spectrophotometric quantification of nucleic acids and gravimetric quantification of cell dry mass. For the first time, biomass compositions of various CHO host and producer cell lines in exponential phase were accurately characterized in media with or without glutamine. The goal was to assess variability of the biomass composition

among different strains and conditions to see which components are the most variable and therefore should be measured every time and which could be generalized for all CHO cells.

Grid-Based Computational Methods for the Design of Constraint-Based Parsimonious Chemical Reaction Networks to Simulate Metabolite Production: Gridprod.

Takeyuki Tamura

Bioinformatics Center, Institute for Chemical Research, Kyoto University, Uji, Kyoto, Japan

Constraint-based metabolic flux analysis of knockout strategies is an efficient method to simulate the production of useful metabolites in microbes. Owing to the recent development of technologies for artificial DNA synthesis, it may become important in the near future to mathematically design minimum metabolic networks to simulate metabolite production. Accordingly, we have developed a computational method where parsimonious metabolic flux distribution is computed for designated constraints on growth and production rates which are represented by grids. When the growth rate of this obtained parsimonious metabolic network is maximized, higher production rates compared to those noted using existing methods are observed for many target metabolites. The set of reactions used in this parsimonious flux distribution consists of reactions included in the original genome scale model iAF1260. The computational experiments show that the grid size affects the obtained production rates. Under the conditions that the growth rate is maximized and the minimum cases of flux variability analysis are considered, the developed method produced more than 90% of metabolites, while the existing methods produced less than 50%. Mathematical explanations using examples are provided to demonstrate potential reasons for the ability of the proposed algorithm to identify design strategies that the existing methods could not identify. The source code is freely available, and is implemented in MATLAB and COBRA toolbox. This study has been accepted by BMC Bioinformatics.

Identification of the Enzymes Responsible for Diverse Phenotypic States of Yeast Lipid Metabolism Using Comprehensive Mechanistic Models.

Sophia Tsouka¹ and **Vassily Hatzimanikatis**²

(1)Laboratory of Computational Systems Biotechnology (LCSB), EPFL, Lausanne, Switzerland, (2)Laboratory of Computational Systems Biotechnology (LCSB), Swiss

Federal Institute of Technology (EPFL), Lausanne, Switzerland

Lipids carry an important role in cell structure and function, as well as in the physiopathology of multiple diseases. Maintenance of the lipid profiles should be tightly regulated for preserving membrane permeability, cell integrity and several other functions.

Large-scale kinetic models of metabolic networks are essential in order to capture and predict such behaviors of cellular systems when subject to perturbations. To this end, we developed a detailed model of the lipid metabolism, based on genome-scale metabolic models of *S. cerevisiae*, in order to identify how the stoichiometric and kinetic coupling determines lipid homeostasis and its regulation.

We curated this model using thermodynamic data as well as lipidomic measurements and we used the Optimization and Risk Analysis of Complex Living Entities (ORACLE) framework to generate populations of parametrized kinetic models that are consistent with the given physiology, while satisfying the stoichiometric and thermodynamic constraints and accounting for the parametric uncertainty.

The model encompasses 1143 reactions and 799 metabolites across 6 cellular compartments (cytosol, mitochondria, peroxisomes, endoplasmic reticulum, Golgi and nucleus), and includes the following lipid-related subsystems: biosynthesis, elongation, and degradation of fatty acids, biosynthesis and esterification of sterols, biosynthesis of phospholipids, sphingolipids, and cardiolipin, triacylglyceride decomposition, dolichol biosynthesis and the mevalonate pathway. It also includes several key parts of yeast metabolism such as glycolysis, citric acid cycle, oxidative phosphorylation etc.

Using the distributions of the computed kinetic models' parameters, we constructed the dynamic mass balances of the species, in order to simulate the dynamic evolution of concentration profiles in response to small perturbations of enzyme activities, and to identify the enzymes that control the distributions of fluxes and metabolic concentrations at a representative steady state. We can further use this analysis to identify the changes in specific enzyme activities that are responsible for given mutant phenotypes.

Selective Metabolic Vulnerabilities in Multiple Myeloma.

Luis Vitores Valcárcel^{1,2}, Raquel Ordoñez², Cem Meydan³, Iñigo Apaolaza¹, Ari Melnick³, Xabier Agirre², Felipe Prosper², and **Francisco J. Planes**¹

(1)TECNUN, University of Navarra, San Sebastian, Spain, (2)CIMA, University of Navarra, Pamplona, Spain, (3) Cornell University, New York, NY

Multiple myeloma (MM) is a hematological cancer characterized by a heterogeneous clinical presentation and an abnormal accumulation of clonal plasma cells in the bone marrow. MM is the second most common cancer worldwide and accounts more than 30,000 new diagnoses cases in 2016 just in United States. Immunomodulatory drugs and proteasome inhibitors are the main existing treatments against MM, which significantly improves the life expectancy of patients; however, MM remains an incurable disease.

In order to identify novel therapeutic strategies in MM, we systematically search for selective metabolic vulnerabilities using the concept of genetic Minimal Cut Sets (gMCSs), recently introduced in Apaolaza et al. 2017. To that end, we first enumerated 20,000 gMCSs from Recon3D (Brunk et al. 2018), the most recent reconstruction of the human metabolism. Secondly, RNA-seq data from the CoMMpass (Relating Clinical Outcomes in Multiple Myeloma to Personal Assessment of Genetic Profile) project, the largest study performed to date by the Multiple Myeloma Research Foundation (MMRF), was mapped onto the gMCSs to discover metabolic targets in MM. Third, in order to identify selective target for MM cells, we also analyzed an unpublished RNA-seq study that includes healthy samples for different cell types arising from the B-cell differentiation: bone marrow plasma cells, tonsil plasma cells, memory B cells, centrocytes, centroblasts, naïve B cells. Fourth, the same analysis was repeated in different MM cell lines in order to identify suitable candidates for in-vitro experimental validation. Main results and future directions are presented and discussed.

[1] Apaolaza, Iñigo, et al. "An in-silico approach to predict and exploit synthetic lethality in cancer metabolism." *Nature communications* 8.1 (2017): 459.

[2] Brunk, Elizabeth, et al. "Recon3D enables a three-dimensional view of gene variation in human metabolism." *Nature biotechnology* 36.3 (2018): 272.

Computational Pathway Design for Funneling Lignin Intermediates to Aromatic Products.

Lin Wang¹, Gregg T. Beckham², and Costas D. Maranas¹

(1)Department of Chemical Engineering, The Pennsylvania State University, University Park, PA, (2)National Bioenergy Center, National Renewable Energy Laboratory, Golden, CO

POSTER ABSTRACTS

The heterogeneity of the aromatic products from lignin catalytic depolymerization is one of the major challenges for lignin valorization. Microbes have evolved many catabolic pathways to tackle the challenge by funneling such heterogeneous intermediates to a few central aromatic products, which lead further to intra- or extradiol ring opening to produce value-added chemicals. However, such funneling pathways are only well-characterized in a few organisms such as *Sphingobium* sp. SYK-6 and *Pseudomonas putida* KT2440. Further effort to explore novel bacterial funneling pathways is necessary. Herein, we apply the de novo pathway design tool (novoStoic [1]) to computationally prospect all possible pathways for lignin-derived mono- and biaryls. First, we augment the MetRxn database with a new dataset of elementally balanced reaction operators using the automated atom mapping based reaction rule extraction procedure termed rePrime. For each reaction, the rePrime procedure identifies and captures as reaction rules the molecular graph topological changes underpinning the substrate to product graph conversion. A reaction rule captures the location of active reaction centers affected by the conversion of substrates to products. The reaction rules and known reactions are then operated upon a mixed integer linear programming algorithm (novoStoic) to identify a mass-balanced biochemical network that converts a source to a target metabolite while minimizing the number of de novo steps. We demonstrate the application of novoStoic in a few case studies such as designing shorter pathways from ferulate to protocatechuate and exploring C α -C β cleavage pathway of guaiacylglycerol- β -guaiacyl ether. By exploring the uncharted chemical space enabled by enzyme promiscuity, novoStoic paves the way to identify more carbon/energy efficient lignin funneling pathways with minimal heterologous enzymes to engineer, while exploring the organism's potential underground metabolism.

References

- [1] Kumar, A., et al. Pathway design using de novo steps through uncharted biochemical spaces. *Nature communications* 9.1 (2018): 184.

Template of Metabolic Reprogramming in Cancer and Healthy Cells for Inferring Oncogenes.

Wu-Hsiung Wu¹, Fan-Yu Li¹, Yi-Chen Shu¹, Jin-Mei Lai², Peter Mu-Hsin Chang³, Chi-Ying F. Huang⁵, and **Feng-Sheng Wang¹**

(1)Department of Chemical Engineering, National Chung Cheng University, Chiayi, Taiwan, (2)Department of Life Science, Fu-Jen Catholic University, New Taipei City,

Taiwan, (3)Faculty of Medicine, National Yang Ming University, Taipei, Taiwan, (4)Institute of Biopharmaceutical Sciences, National Yang Ming University, Taipei, Taiwan

Cancer cells exhibit unusual metabolic activity, characterized by high rates of glucose consumption and lactate production even under aerobic conditions, as well as increased glutamine catabolism and amino acid metabolism. Based on the behaviors of metabolic reprogramming, oncogenesis can be inferred from a genome-scale metabolic model of cancer cells. In this study, we incorporated the CORDA method [1] with data of the Human Protein Atlas database [2] and Recon 2.2 network [3] to reconstruct tissue-specific metabolic models for normal and cancer cells, respectively. These models were applied to assess the flux alternation template using flux-sum balance analysis. We developed a bilevel optimization formulation to infer oncogenes for the tissue-specific models. The objective function in the outer optimization problem is the similarity ratio of flux alternations of the dysregulated cell to the template for cancer or health models, or Warburg hypothesis from the literature. The inner optimization problem consisted of the flux balance model and uniform flux distribution model [4]. A nested hybrid differential evolution algorithm was introduced to solve the bilevel optimization problem. The head and neck tissue-specific model was used for a case study to infer oncogenes. Results show that TPK1 achieved the highest similarity ratio of 0.84 and three genes (GNPD1/GNPD2, THTPA, and PTEN) exhibited similarity ratio of 0.839. Literature shows that TPK1 is an oncogene for colon cancer [5] and PTEN is a tumor suppressor gene of squamous cell carcinomas of the head and neck [6].

Schulta et al., *PLOS Computational Biology*, 12(3): e1004808, 2016

<https://www.proteinatlas.org/>

Swainston1 et al., *Metabolomics*, 12:109, 2016

Wu et al., *PLOS Computational Biology*, 13(7): e1005618, 2017

Tiwana1 et al., *Oncotarget*, 6(8) : 5978–89, 2015

M. Poetsch et al., *Cancer Genetics and Cytogenetics*, 132(1):20–4, 2002

Gapseq: A Novel Approach for in silico Prediction and Analysis of Bacterial Metabolic Pathways and Genome-Scale Networks.

Johannes Zimmermann, Christoph Kaleta, and **Silvio Waschina**

Institute for experimental medicine, Christian-Albrechts-University Kiel (UKSH Campus), Kiel, Germany

In silico biochemical pathway analysis and metabolic phenotype prediction based on bacterial genome sequences became powerful research tools in biotechnology, medicine, and ecology. The accuracy of genome sequence-derived catalytic capability predictions and metabolic models are largely hindered by two factors: (i) False-positive and false-negative enzymatic reaction predictions due to incomplete knowledge of yet-unknown enzyme mechanisms, especially in non-model organisms, and (ii) inconsistencies in reaction databases, which frequently cause thermodynamically infeasible futile cycles and, hence, false phenotype predictions.

Here we present a new tool for pathway analysis and automated reconstruction of metabolic networks, called GapSeq, that addresses these obstacles. First, pathways are predicted by combining sequence homology search results to estimate the evidence of the presence of the pathways' individual reactions, pathway topologies, and information about the pathways' key reactions. Second, reactions predicted with high evidence scores are consolidated to construct draft genome-scale networks using a reaction database that has been manually curated and is completely futile cycle-free. This ensures, that also all possible reconstructions are free of thermodynamically infeasible substrate cycles. Third, a novel gap-filling algorithm was implemented that adds candidate reactions based on the reaction's evidence scores. Moreover, the gap-filling algorithm does not only enable biomass production under the defined media composition, but also fills gaps in alternative resource utilization pathways as well as anabolic pathways. This approach has the advantage that also gaps in the network are filled with candidate reactions that are only employed by the organism under specific environmental conditions or when interacting with different organisms.

We tested GapSeq on a large set of bacterial strains, which are physiologically well-described. We show that not only the presence of strain-specific pathways are correctly predicted, but also metabolic phenotypes, such as the production of specific fermentation products, are accurately predicted by GapSeq in combination with Flux Balance Analysis.

Erythrodb: A Multi-Omic Visual Knowledge Base for the Human Red Blood Cell.

James T. Yurkovich and **Bernhard O. Palsson**
Bioengineering, University of California, San Diego, La Jolla, CA

As the life sciences have embraced the information age, we have witnessed the birth of many outstanding resources such as BioCyc, UniProt, the Protein Databank, and KEGG. With the continued generation of -omics data, we are faced with challenges in contextualizing, visualizing, and extracting information from these data. Here, we present ErythroDB (<http://erythrodb.org>), a multi-omic visual knowledge base for the biochemistry and metabolism of the human red blood cell with applications for transfusion medicine. This resource contains three primary forms of content: (1) an interactive metabolic map enhanced with data from the literature and information from external databases, (2) an interactive "bibliome" that contains manually-curated literature on red blood cell metabolism and storage, and (3) a data repository that should be a valuable resource to the community and empower parameter-dependent systems biology research.

Multimics Integration for Prediction of Complex Metabolic Phenotypes.

Aleksej Zelezniak

Chalmers University of Technology, Gothenburg, Sweden

A key challenge in solving genotype to phenotype relationship is to predict a cell's metabolome. Here, we quantified the proteomes in 97 non-essential kinase knock-out strains from *Saccharomyces cerevisiae* and associated these to their metabolomes. We observed that kinase knockouts affect proteomes broadly but in a distinct manner; and that these are quantitatively dominated by expression changes in metabolic enzymes. Analysing these data in the context of kinetic modeling demonstrated that enzyme abundance affects metabolite concentrations through the redistribution of flux control, resulting in a many-to-many relationship between enzyme abundances and the metabolite concentrations. Machine learning enabled mapping these relationships, the prediction of experimentally measured metabolite concentrations, as well as to identify candidate genes important for the regulation of metabolism. Overall, our study suggests that hierarchical metabolism regulation acts predominantly through adjustment of broad expression patterns rather than over individual rate-limiting enzymes, and may account for more than half of metabolism regulation.

Genome Scale Metabolic Model Assisted Strain Designs for Itaconic Acid Production in Yeast.

Eric M. Young¹, **Zheng Zhao**², **Bianca Gielesen**², **Liang Wu**², **Ben Gordon**³, **Johannes A. Roubos**², and **Christopher A. Voigt**⁴

POSTER ABSTRACTS

(1)*Synthetic Biology Center, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA*, (2)*DSM Biotechnology Center, Delft, Netherlands*, (3)*The Foundry, MIT / Broad Institute, Cambridge, MA*, (4)*Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA*

Itaconic acid production by the natural producer *Aspergillus terreus* could be limited by difficult fermentation characteristics, such as high viscosity, shear stress, and low growth rate. An alternative host, such as *Saccharomyces cerevisiae*, might open new revenue for further process intensification. Various pathways for itaconic acid production in yeast, varying in the supply of AcCoA precursor, were designed using a genome scale metabolic model. By comparing the flux distribution between optimal growth and optimal production, additional targets were identified for overexpression and deletion. Four pathway variants were selected for 'build' and 'test'. The best tested variant was further improved in several iterative rounds using Design of Experiment methodologies. The full design-build-test-learn cycle delivered the highest itaconic acid titer in *S. cerevisiae* to date.

A Machine Learning-Empowered Kinetic Reconstruction of *E. coli* Metabolism.

Daniel C. Zielinski¹ and Bernhard O. Palsson²

(1)*Bioengineering, University of California, San Diego, CA*,
(2)*Bioengineering, University of California - San Diego, La Jolla, CA*

Constraint-based models of metabolism have been powerful tools for characterizing and predicting organism function with minimal parameter requirements. However, several important applications of metabolic models have emerged that depend upon enzyme kinetic parameters, which are scarce in the literature. In this work, we present a genome-scale kinetic reconstruction of the *E. coli* metabolic network. We curated the literature and existing databases to extract available kinetic data, including enzyme kinetic parameters, kinetic mechanisms, and protein structural information. To address the data scarcity issue, we fill gaps in measured parameters with statistical models for k_{cat} , K_m , and K_{eq} trained on literature data. We reconcile potential inconsistencies in data using a nonlinear regression approach that finds rate constants satisfying kinetic data for each enzyme to the extent possible. The resulting kinetic reconstruction should be a valuable resource to the community and empower parameter-dependent systems biology research.

AIChE® Volunteer and Meeting Attendee Conduct Guidelines

AIChE's volunteers are the core of the Institute and make all of its programs, conferences and educational efforts possible. These offerings provide excellent opportunities for AIChE members and meeting attendees to gain greater technical expertise, grow their networks, and enhance their careers. AIChE events provide engineers, scientists, and students a platform to present, discuss, publish and exhibit their discoveries and technical advances.

At all times, volunteers and meeting attendees should act in accordance with AIChE's Code of Ethics, upholding and advancing the integrity, honor and dignity of the chemical engineering profession. AIChE's Board of Directors has developed these guidelines to foster a positive environment of trust, respect, open communications, and ethical behavior. These guidelines apply to meetings, conferences, workshops, courses and other events organized by AIChE or any of its entities and also to volunteers who conduct other business and affairs on behalf of AIChE.

SPECIFICALLY:

1. Volunteers and meeting attendees should understand and support AIChE's Code of Ethics.
2. Volunteers and meeting attendees should contribute to a collegial, inclusive, positive and respectful environment for fellow volunteers and attendees, and other stakeholders, including AIChE staff.
3. Volunteers and meeting attendees should avoid making inappropriate statements or taking inappropriate action based on race, gender, age, religion, ethnicity, nationality, sexual orientation, gender expression, gender identity, marital status, political affiliation, presence of disabilities, or educational background. We should show consistent respect for colleagues, regardless of discipline, employment status, and organizations for which they work, whether industry, academia, or government.
4. Disruptive, harassing or other inappropriate statements or behavior toward other volunteers, members, another stakeholders, including AIChE staff, is unacceptable.
5. Volunteers and meeting attendees should obey all applicable laws and regulations of the relevant governmental authorities while volunteering or attending meetings. Volunteers and meeting attendees taking part in any AIChE event, including the Chem-E-Car Competition®, should also comply with all applicable safety guidelines.

Any violations of the foregoing should be reported to the President or the Executive Director of the Institute.



American Institute of Chemical Engineers
PO Box 4429 • Danbury, CT 06813-4429 USA

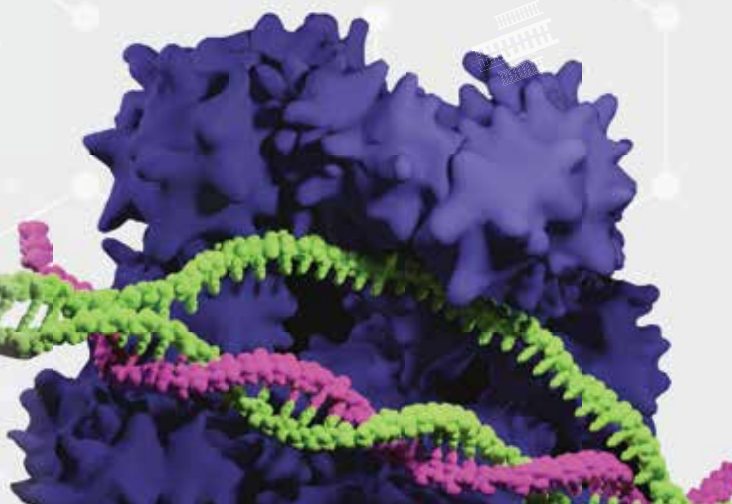
©2018 AIChE 3012_18 • 09.18

NOTES



CRISPR TECHNOLOGIES CONFERENCE

December 10-12, 2018
Bahia Resort Hotel | San Diego, CA



CALL FOR ABSTRACTS AND REGISTRATION NOW OPEN

The Society for Biological Engineering invites you to participate in the **International Conference on CRISPR Technologies**, December 10-12, 2018 in San Diego, CA.

This conference will bring together leaders and trainees from the cutting edge of CRISPR technologies and their application to genome editing and beyond. Academic, clinical, and industrial researchers are invited to share their recent discoveries to progress the field in CRISPR science and engineering.

Submit Abstracts for these sessions:

- CRISPR-Cas biology
- Genome editing and DNA repair
- Clinical applications of CRISPR
- Beyond editing: novel applications
- CRISPR-based gene drives
- Delivery and specificity of CRISPR-based tools
- Social, ethical, and economic implications of CRISPR-based technologies

To showcase your work, please contact
Keith Joseph at keitj@aiche.org.

Visit www.aiche.org/crispr today and
save on your conference registration.

CONFERENCE CHAIRS

- Samira Kiani, *Arizona State University*
- Sam H. Sternberg, *Columbia University*

KEYNOTE SPEAKER

- Fyodor Urnov, *Altius Institute for Biomedical Sciences*

INVITED SPEAKERS

- Omar Akbari, *University of California, San Diego*
- Chase Beisel, *Helmholtz Institute*
- Albert Cheng, *Jackson Laboratory*
- Wei Leong Chew, *Genome Institute of Singapore*
- Mo Ebrahimkhani, *Arizona State University/Mayo Clinic*
- Charles Gersbach, *Duke University*
- Patrick Hsu, *Salk Institute for Biological Studies*
- Yinan Kan, *eGenesis*
- Alexis Komor, *University of California, San Diego*
- Prashant Mali, *University of California, San Diego*
- Megan Palmer, *Stanford University*
- Shengdar Tsai, *St. Jude Children's Research Hospital*
- Harris Wang, *Columbia University*
- Yan Zhang, *University of Michigan*
- John Zuris, *Editas Medicine*

ORGANIZED BY



The illustration (header) is intended for promotional purposes only and is not an accurate representation of the B-DNA structure.



Institute for 
Systems Biology
Revolutionizing Science. Enhancing Life.