

Recon 2.2: from reconstruction to model of human metabolism

Neil Swainston^{1,2,3} · Kieran Smallbone³ · Hooman Hefzi^{4,5} · Paul D. Dobson³ ·
Judy Brewer^{6,7} · Michael Hanscho^{8,9} · Daniel C. Zielinski⁴ · Kok Siong Ang^{10,11} ·
Natalie J. Gardiner² · Jahir M. Gutierrez^{4,5} · Sarantos Kyriakopoulos¹¹ ·
Meiyappan Lakshmanan¹¹ · Shangzhong Li^{4,5} · Joanne K. Liu¹³ · Veronica S. Martínez¹² ·
Camila A. Orellana¹² · Lake-Ee Quek¹² · Alex Thomas^{5,13} · Juergen Zanghellini⁹ ·
Nicole Borth^{8,9} · Dong-Yup Lee^{10,11} · Lars K. Nielsen¹² · Douglas B. Kell^{1,14} ·
Nathan E. Lewis^{5,15} · Pedro Mendes^{1,3,16}

Received: 29 March 2016 / Accepted: 27 May 2016 / Published online: 7 June 2016
© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract

Introduction The human genome-scale metabolic reconstruction details all known metabolic reactions occurring in humans, and thereby holds substantial promise for studying complex diseases and phenotypes. Capturing the whole human metabolic reconstruction is an on-going task and since the last community effort generated a consensus reconstruction, several updates have been developed.

Objectives We report a new consensus version, Recon 2.2, which integrates various alternative versions with significant additional updates. In addition to re-establishing a consensus reconstruction, further key objectives included providing more comprehensive annotation of metabolites and genes, ensuring full mass and charge balance in all reactions, and developing a model that correctly predicts ATP production on a range of carbon sources.

Methods Recon 2.2 has been developed through a combination of manual curation and automated error checking. Specific and significant manual updates include a respecification of fatty acid metabolism, oxidative phosphorylation

Electronic supplementary material The online version of this article (doi:10.1007/s11306-016-1051-4) contains supplementary material, which is available to authorized users.

✉ Neil Swainston
neil.swainston@manchester.ac.uk

¹ Manchester Centre for Synthetic Biology of Fine and Speciality Chemicals (SYNBIOCHEM), Manchester Institute of Biotechnology, The University of Manchester, Manchester M1 7DN, UK

² Faculty of Life Sciences, The University of Manchester, Manchester M13 9PL, UK

³ School of Computer Science, The University of Manchester, Manchester M13 9PL, UK

⁴ Department of Bioengineering, University of California, San Diego, La Jolla, CA, USA

⁵ Novo Nordisk Foundation Center for Biosustainability, University of California, San Diego School of Medicine, La Jolla, CA, USA

⁶ Harvard Extension School, 51 Brattle St., Cambridge, MA 02138, USA

⁷ Computer Science and Artificial Intelligence Lab, Massachusetts Institute of Technology, 32 Vassar St., Cambridge, MA 02139, USA

⁸ Department of Biotechnology, University of Natural Resources and Life Sciences, Vienna, Austria

⁹ Austrian Centre of Industrial Biotechnology, Vienna, Austria

¹⁰ Department of Chemical and Biomolecular Engineering, National University of Singapore, 4 Engineering Drive 4, Singapore 117585, Singapore

¹¹ Bioprocessing Technology Institute, Agency for Science, Technology and Research (A*STAR), 20 Biopolis Way, #06-01, Centros, Singapore 138668, Singapore

¹² Australian Institute for Bioengineering and Nanotechnology (AIBN), The University of Queensland, Corner College and Cooper Roads (Bldg 75), Brisbane, QLD 4072, Australia

¹³ Bioinformatics and Systems Biology Program, University of California, San Diego, La Jolla, CA, USA

¹⁴ School of Chemistry, The University of Manchester, Manchester M13 9PL, UK

¹⁵ Department of Pediatrics, University of California, San Diego, La Jolla, CA, USA

¹⁶ Center for Quantitative Medicine, UConn Health, 263 Farmington Avenue, Farmington, CT 06030-6033, USA

and a coupling of the electron transport chain to ATP synthase activity. All metabolites have definitive chemical formulae and charges specified, and these are used to ensure full mass and charge reaction balancing through an automated linear programming approach. Additionally, improved integration with transcriptomics and proteomics data has been facilitated with the updated curation of relationships between genes, proteins and reactions.

Results Recon 2.2 now represents the most predictive model of human metabolism to date as demonstrated here. Extensive manual curation has increased the reconstruction size to 5324 metabolites, 7785 reactions and 1675 associated genes, which now are mapped to a single standard. The focus upon mass and charge balancing of all reactions, along with better representation of energy generation, has produced a flux model that correctly predicts ATP yield on different carbon sources.

Conclusion Through these updates we have achieved the most complete and best annotated consensus human metabolic reconstruction available, thereby increasing the ability of this resource to provide novel insights into normal and disease states in human. The model is freely available from the Biomedels database (<http://identifiers.org/biomodels.db/MODEL1603150001>).

Keywords Human · Metabolism · Modelling · Reconstruction · Model · Systems biology

1 Introduction

Metabolic processes are implicated in many important aspects of human health. Models are critical in the exploration and understanding of the complexity underlying human metabolism. Genome-scale models of human metabolism [Recon 1 (Duarte et al. 2007) and Recon 2 (Thiele et al. 2013)] have been used to predict biomarkers of inborn errors of metabolism (Shlomi et al. 2009), to identify drug targets (Frezza et al. 2011) and off-target drug effects (Thiele et al. 2013), to study cancer metabolism (Lewis and Abdel-Haleem 2013) and to improve understanding of microbial interactions with the host organism (Bordbar et al. 2010; Heinken and Thiele 2015). Furthermore, while human metabolic reconstructions are of obvious utility in the medical and pharmacological fields, it is worth noting that reconstructions of mammalian biochemical networks also act as blueprints for modelling systems of biotechnological significance, such as Chinese Hamster Ovary (CHO; Kaas et al. 2014) and Human Embryonic Kidney (HEK; Quek et al. 2014) cells. As these networks improve, their applications across basic, clinical and biotechnological research will continue to expand.

The core of a predictive metabolic model is a reconstruction of the underlying reaction network, which catalogues all metabolic reactions encoded within the genome. (One can consider a reconstruction to be a comprehensive knowledge base of biochemistry, covering metabolic reactions and their enzymes, and a metabolic model to be a mathematical representation of the reconstruction.) Naturally, the reconstruction should be both accurate and complete. The most recent human metabolic reconstruction, Recon 2, was published following a large international community effort to develop a consensus reconstruction from existing resources. This reconstruction was of a considerably larger scale than its predecessors and therefore represented an important advancement. Since its publication, Recon 2 has served as a valuable resource for many studies. As is common for such reconstructions, new knowledge about cell metabolism is regularly discovered and these insights must be added to the reconstruction. Furthermore, the use of reconstructions in modelling studies identifies necessary corrections. Thus, for Recon 2, updates and refinements could strengthen the accuracy of its predictions (Swainston et al. 2013; Quek et al. 2014; Kell and Goodacre 2014).

Following the release of Recon 2, several updates were published (Quek et al. 2014; Sahoo et al. 2014, 2015). These updates provided better definition of transport proteins, a wider consideration of drug metabolism, and a number of error corrections. An additional update, Recon 2.1 (Smallbone 2013) focused upon improving carbon balancing but did not cover full stoichiometric mass and charge balancing of every reaction. As such, there is a need for identifying imbalances to ensure accurate predictions of energy metabolism.

In this update, a number of errors are corrected and various improvements introduced to capture human metabolism more accurately and completely. Extensive manual curation has increased the reconstruction size, which now contains 5324 compartmentalised metabolites (of which 2652 are unique chemical species), 7785 reactions, and 1675 associated genes (Table 1). The focus upon mass and charge balancing of all reactions, along with better representation of energy generation, has produced a model that correctly predicts ATP flux on different carbon sources. Thus, through these updates, we have achieved the most complete and best-annotated consensus human metabolic reconstruction available. We demonstrate that Recon 2.2 is, to our knowledge, the first mammalian metabolic model to predict (free) energy production correctly, based upon carbon availability.

The model is freely available from the Biomedels database (Chelliah et al. 2015), under the identifier MODEL1603150001 (<http://identifiers.org/biomodels.db/MODEL1603150001>).

2 Materials and methods

Recon 2.2 is an extension of Recon 2.1 (Smallbone, 2013). A series of manual curation steps led to the development of an interim version, Recon 2.1.5. Following this, a semi-automated curation approach was implemented, in which model updates were specified in simple, human-readable text files. These text files were interpreted by an updated version of the SuBliMinaL Toolbox (Swainston et al. 2011), built on libChEBI (Swainston et al. 2016), in order to automate the production of Recon 2.2 from Recon 2.1.5. All models, text files and software used to build Recon 2.2 are freely available from <https://github.com/mcisb/mcisb-recon> along with instructions on their use.

Semi-automatic mapping of NCBI identifiers resolved HGNC identifiers but also revealed duplicates, pseudogenes, ESTs and other non-gene records, which have been removed. The resulting reconstruction contains 1675 genes. All gene associations are expressed using disjunctive normal form: (A and B) or (A and C) rather than A and (B or C). This consistency facilitates the writing of parsers, but more importantly is the correct representation of the underlying biology, explicitly specifying genes in terms of the complexes their products form.

Beta-oxidation of fatty acids in the mitochondria and peroxisome was expanded to account explicitly for intermediary fatty acyl-CoA moieties and the full suite of enzymes necessary for complete oxidation of fatty acids. By breaking down lumped beta-oxidation reactions into the component two-carbon cycles, the differing specificities of the enzymes catalysing the first dehydrogenase/oxidase step can be resolved. We additionally clarified the gene-protein-reaction (GPR) relationships for these reactions to include the enzymes required for the full beta-oxidation cycle (acyl-CoA dehydrogenase/oxidase, enoyl CoA hydratase, 3-hydroxyacyl-CoA dehydrogenase, and β -keto-thiolase) as well as enzymes utilized for unsaturated fatty acid beta-oxidation (dienoyl-CoA reductase and/or enoyl-CoA isomerase).

The modelling utility of these improvements is demonstrated via a suite of tests. The test suite was developed in Python 2.7, and is available at <https://github.com/mcisb/mcisb-recon-analysis>. The test suite checks reaction balancing, ATP production under a range of nutrient sources, and the size of the reconstruction and its constituent elements. It can be run against any model that adheres to the COBRA convention (Schellenberger et al. 2011) of SBML formatting, and—similar to the concept of unit testing in software development (Yoo and Harman 2012)—can therefore be used to validate incremental updates to Recon 2.2 as the model develops further.

3 Results and discussion

The original goal of Recon 2 was to create a consensus from existing reconstructions. Recently published updates resulted in different versions of Recon 2, and these have all been incorporated into Recon 2.2 to create a new consensus. These include updates to transporter reactions (Sahoo et al. 2014), drug effects and metabolism (Sahoo et al. 2015), and the corrections published by Quek et al. (2014).

Due to their necessity for the development of tissue-specific models, the accurate definition of GPR relationships are of particular importance in multicellular, mammalian reconstructions. GPRs allow the development of models implementing constraints based upon experimentally measured expression data (Lee et al. 2012; Pornputapong et al. 2015; Uhlén et al. 2015). It follows that the accuracy of such models is directly dependent upon the accuracy of the gene associations in the original reconstruction. Gene association updates from a previous Recon 2 iteration (Recon 2.04, <http://vmh.uni.lu>) and further manual corrections and updates have also been incorporated into the present version. Previous human reconstructions used a variety of identifiers, mostly from NCBI, to denote genes. In Recon 2.2 genes are now represented using HGNC identifiers, the worldwide authority for assignment of standardised nomenclature to human genes.

A major goal for Recon 2.2 was to improve the simulation of energy metabolism. Towards this end, both mitochondrial and peroxisomal fatty acid oxidation were redefined and expanded by replacing previously lumped reactions with constituent two carbon cycle reactions (e.g. an n carbon fatty acid to an $n-2$ carbon fatty acid), for both saturated and unsaturated fatty acids. Genes associated with fatty acid oxidation have also have been expanded.

Significant improvements also were made by redefining the representation of oxidative phosphorylation. Specifically, this involved the definition of a new compartment, the mitochondrial intramembrane space, and the use of this compartment to define an electrochemical proton gradient. Introducing this specific pool of transmembrane protons enforces the coupling between the reactions of the electron transport chain with that of ATP synthase, and thus the coupling of mitochondrial NADH oxidation and O_2 reduction with ATP production (Martínez et al. 2014). While there may be inevitable simplifications of the representation of the mitochondrial intramembrane space in the model, its addition greatly improves the results of ATP flux predictions and will act as an incentive for further subsequent curation. The updated reactions are given in Supplementary Information: Table S1.

Recon 2 included reactions with incomplete mass and charge balancing (Table 1). Despite well over 90 % of

Table 1 Reconstruction scope

	Recon 2.2	Quek	Recon 2.1	Recon 2.04	Recon 2	Recon 1
Metabolites	5324	4962	5056	5063	5063	2766
Reactions	7785	7327	8089	7440	7440	3742
Balanced	7780	6761	7916	6948	6948	431
Unbalanced	0	211	51	94	94	
Unknown	5 ^a	355	122	398	398	3311
Genes	1675 ^b	2167	2193	2140	2191	

Recon 2.2 has increased the number of metabolites and reactions catalogued in the reconstruction relative to Recon 2, and the reactions have now been verified and corrected for mass balance. Note that boundary metabolites are excluded from this table (boundary metabolites being those outside of the system). The values for metabolites and reactions therefore correspond to the size of the stoichiometry matrix

^a The 5 unbalanced reactions remaining in Recon 2.2 are those relating to the biomass objective function

^b There appears to be fewer genes in Recon 2.2 than in previous reconstructions, due to the movement to representing genes by HGNC identifiers, which group alternative transcripts or splice variants into a single entity

reactions being correctly balanced, the presence of even a small number of incorrectly balanced reactions is sufficient for leaks (the erroneous or ‘alchemical’ creation of mass) to occur, which can lead to unreliable flux predictions. To address this in Recon 2.2, an automated reaction balancing method, originally introduced in the SuBliMinaL Toolbox software suite (Swainston et al. 2011), was extended and applied. The original algorithm employed linear programming to check and correct reaction stoichiometry based upon element and charge counts of the reaction participants. It was also able to add ‘missing’ protons and water molecules, which are commonly absent from reaction definitions. This algorithm has been extended here to balance reactions involving non-specific metabolites, that is, those containing generic R-groups (‘Markush structures’), or those containing repeating units [e.g. (CH₂)_n]. The use of R-groups is especially prevalent in Recon 2 in defining lipid metabolism, where in the interests of simplicity, multiple reactions involving fatty acids of differing chain lengths were condensed into a single reaction. The R-groups that remain are those representing conserved moieties such as acyl-carrier protein (ACP), which cannot be represented as a defined chemical formula but whose presence do not affect the mass and charge balancing of reactions.

Reversible reactions that thermodynamics suggest should be unidirectional under typical physiological conditions can also impact the accuracy of model predictions. Many of these have been discovered through multiple rounds of manual curation, driven by the requirement to predict realistic ATP yields. This iterative process involved performing a flux-balance analysis (FBA) test, inspecting the resulting flux pattern for anomalous reactions, and correcting their directionality based on literature searches and definitions in pathway databases.

Previous work has highlighted the advantages of augmenting metabolic reconstructions with unambiguous, publicly available identifiers mapping elements to entries in persistent, external data resources (Herrgård et al. 2008). Due to its breadth and accuracy, ChEBI (Hastings et al. 2016) has become a de facto standard for the annotation of metabolic species in systems biology models. Recon 2.2 has been further curated to expand the number of metabolites that are annotated to ChEBI entries. Additionally, metabolites that are not currently in the ChEBI database have been submitted to the ChEBI curators with the intention of expanding the database, and incorporating these new ChEBI identifiers into a subsequent iteration. In the interim, Recon 2.2 metabolites have been annotated with InChI string representations of molecular structure.

A comparison of maximum ATP yields per unit of carbon source was calculated for a number of existing models, under both aerobic and anaerobic conditions. The results show that, in contrast to previous versions, Recon 2.2 is able to correctly predict maximum ATP fluxes. The results are given in Table 2 and Supplementary Information: Table S2.

To summarise, Recon 2.2 compiles updates from the various different updates to Recon 2.0 that have been published. In addition, hundreds of novel manual updates have been included, and semi-automatic checks have been conducted to create a new consensus human metabolic reconstruction and associated model. Importantly, Recon 2.2 has eliminated all mass leaks from improperly balanced reactions that resulted in previous models being able to simulate growth without a carbon source. ATP synthesis is also now coupled to carbon availability. Simulation of growth and energy metabolism using Recon 2.2 gives biologically realistic results.

Annotations in Recon 2.2 have been improved by increasing the use of ChEBI identifiers for metabolites, and

Table 2 Comparison of maximum ATP yields under different carbon sources and oxygenation

Carbon source	ATP yield					
	Theoretical ^a	Recon 2.2	Quek	Recon 2.04	Recon 2	Recon 1
Aerobic						
glc, aerobic	31	32	∞	∞	∞	∞
fru, aerobic	31	32	∞	∞	∞	∞
C4:0, aerobic	21.5	22	∞	∞	∞	∞
C6:0, aerobic	35.25	36	^b	^b	^b	^b
C8:0, aerobic	49	50	∞	∞	∞	∞
C10:0, aerobic	62.75	64	∞	∞	∞	^b
C12:0, aerobic	76.5	82.5	∞	∞	∞	^b
C14:0, aerobic	90.25	92	∞	∞	∞	∞
C16:0, aerobic	104	106.75	∞	∞	∞	∞
C18:0, aerobic	117.75	120.0	∞	∞	∞	∞
C20:0, aerobic	131.5	134	∞	∞	∞	∞
C22:0, aerobic	145.25	147.25	∞	∞	∞	^b
C24:0, aerobic	159	160.5	∞	∞	∞	∞
C26:0, aerobic	172.75	170.75	∞	∞	∞	∞
Anaerobic						
glc, anaerobic	2	2	∞	∞	∞	∞
fru, anaerobic	2	2	∞	∞	∞	∞
C4:0, anaerobic	0	0	∞	∞	∞	∞
C6:0, anaerobic	0	0	^b	^b	^b	^b
C8:0, anaerobic	0	0	∞	∞	∞	∞
C10:0, anaerobic	0	0	∞	∞	∞	^b
C12:0, anaerobic	0	0	∞	∞	∞	^b
C14:0, anaerobic	0	0	∞	∞	∞	∞
C16:0, anaerobic	0	0	∞	∞	∞	∞
C18:0, anaerobic	0	0	∞	∞	∞	∞
C20:0, anaerobic	0	0	∞	∞	∞	∞
C22:0, anaerobic	0	0	∞	∞	∞	^b
C24:0, anaerobic	0	0	∞	∞	∞	∞
C26:0, anaerobic	0	0	∞	∞	∞	∞

Glc, glucose; fru, fructose; Cn:0, a saturated fatty acid of length *n*

^a Theoretical yields are calculated according to Salway (2003)

^b Carbon source was not present in the model

standardising gene annotations to HGNC. During this process numerous misannotations were removed and new annotations incorporated.

The development of Recon 2.2 followed a semi-automated approach. The introduction of this approach provides full traceability of the updates implemented, and will facilitate and accelerate the process of developing subsequent updates (to both the human and any other constraint-based metabolic model). Regarding simplicity, updates are supplied in simple text files that are parsed and interpreted by custom software. By providing the facility to update models in such an automated manner, the potential user base for the model building process expands beyond those that have an intimate understanding of software, and of

formats such as SBML and the COBRA Toolbox. The benefits in terms of promoting reproducible science are also clear: all changes made to a given model to produce a new version are catalogued in text files, which essentially act as a *diff*¹ between versions. Finally, once the underlying software to interpret the text files has been written, the process of further developing future iterations of the model rests solely in writing updated text files. This approach allows the model developer to focus on the *content* of the updates, rather than the technical means of implementing them.

¹ From Wikipedia: *In computing, the diff utility is a data comparison tool that calculates and displays the differences between two files.*

Acknowledgments K.S. thanks Brandon Barker, James Eddy, Emanuel Gonçalves and Neema Jamshidi for their changes incorporated to Recon 2.1. N.S. thanks Ines Thiele and Ronan Fleming for maintaining updates of Recon 2. As ever, this work is indebted to the inexhaustible well of inspiration provided by Michael Howard. This is a contribution from the Manchester Centre for Synthetic Biology of Fine and Speciality Chemicals (SYNBIOCHEM).

Funding We acknowledge support for this work from the Biotechnology and Biological Sciences Research Council (BBSRC) Grant “Enriching Metabolic PATHwaY models with evidence from the literature” (EMPATHY; BB/M006891/1) [N.S., P.D.D., D.B.K., P.M.]; from the BBSRC Grant “Continued development of ChEBI towards better usability for the systems biology and metabolic modelling community” (BB/K019783/1) [N.S., P.M.]; from the BBSRC Grant “Centre for synthetic biology of fine and specialty chemicals (SYNBIOCHEM)” (BB/M017702/1) [N.S., D.B.K., P.M.]; from the University of Manchester Faculty of Life Sciences Quantitative Biology initiative awarded to Dr. Natalie Gardiner, “Modelling and sensitivity analysis of metabolic networks in diabetic neuropathy” [N.G., K.S., P.D.D.]; from the National Institutes of Health (GM080219) [P.M.]; from the NUS Synthetic Biology for Clinical and Technological Innovation (SynCTI) and Next-Generation Bio-Green 21 Program [SSAC, No. PJ01109405] [K.S.A., D.Y.L.]; from the Federal Ministry of Science, Research and Economy (BMFWF), the Federal Ministry of Traffic, Innovation and Technology (bmvit), the Styrian Business Promotion Agency SFG, the Standortagentur Tirol, the Government of Lower Austria and ZIT—Technology Agency of the City of Vienna through the COMET-Funding Program managed by the Austrian Research Promotion Agency FFG [M.H., J.Z., N.B.]; and from the Novo Nordisk Foundation that had been provided to the Center for Biosustainability at the Technical University of Denmark [H.H., N.E.L.].

Compliance with ethical standards

Conflict of interest All authors declare no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Bordbar, A., Lewis, N. E., Schellenberger, J., Palsson, B. Ø., & Jamshidi, N. (2010). Insight into human alveolar macrophage and *M. tuberculosis* interactions via metabolic reconstructions. *Molecular Systems Biology*, *6*, 422.

Chelliah, V., Juty, N., Ajmera, I., Ali, R., Dumousseau, M., Glont, M., et al. (2015). BioModels: Ten-year anniversary. *Nucleic Acids Research*, *43*, D542–D548.

Duarte, N. C., Becker, S. A., Jamshidi, N., Thiele, I., Mo, M. L., Vo, T. D., et al. (2007). Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 1777–1782.

Frezza, C., Zheng, L., Folger, O., Rajagopalan, K. N., MacKenzie, E. D., Jerby, L., et al. (2011). Haem oxygenase is synthetically lethal with the tumour suppressor fumarate hydratase. *Nature*, *477*, 225–228.

Hastings, J., Owen, G., Dekker, A., Ennis, M., Kale, N., Muthukrishnan, V., et al. (2016). ChEBI in 2016: Improved services and an expanding collection of metabolites. *Nucleic Acids Research*, *44*, D1214–D1219.

Heinken, A., & Thiele, I. (2015). Systematic prediction of health-relevant human-microbial co-metabolism through a computational framework. *Gut Microbes*, *6*, 120–130.

Herrgård, M. J., Swainston, N., Dobson, P., Dunn, W. B., Arga, K. Y., Arvas, M., et al. (2008). A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nature Biotechnology*, *26*, 1155–1160.

Kaas, C. S., Fan, Y., Weilguny, D., Kristensen, C., Kildegaard, H. F., & Andersen, M. R. (2014). Toward genome-scale models of the Chinese hamster ovary cells: Incentives, status and perspectives. *Pharmaceutical Bioprocessing*, *2*, 437–448.

Kell, D. B., & Goodacre, R. (2014). Metabolomics and systems pharmacology: Why and how to model the human metabolic network for drug discovery. *Drug Discovery Today*, *19*, 171–182.

Lee, D., Smallbone, K., Dunn, W. B., Murabito, E., Winder, C. L., Kell, D. B., et al. (2012). Improving metabolic flux predictions using absolute gene expression data. *BMC Systems Biology*, *6*, 73.

Lewis, N. E., & Abdel-Haleem, A. M. (2013). The evolution of genome-scale models of cancer metabolism. *Frontiers in Physiology*, *4*, 237.

Martínez, V. S., Quek, L. E., & Nielsen, L. K. (2014). Network thermodynamic curation of human and yeast genome-scale metabolic models. *Biophysical Journal*, *107*, 493–503.

Pornputtapong, N., Nookaew, I., & Nielsen, J. (2015). Human metabolic atlas: An online resource for human metabolism. *Database*. doi:10.1093/database/bav068.

Quek, L. E., Dietmair, S., Hanscho, M., Martínez, V. S., Borth, N., & Nielsen, L. K. (2014). Reducing Recon 2 for steady-state flux analysis of HEK cell culture. *Journal of Biotechnology*, *184*, 172–178.

Sahoo, S., Aurich, M. K., Jonsson, J. J., & Thiele, I. (2014). Membrane transporters in a human genome-scale metabolic knowledgebase and their implications for disease. *Frontiers in Physiology*, *5*, 91.

Sahoo, S., Haraldsdóttir, H. S., Fleming, R. M., & Thiele, I. (2015). Modeling the effects of commonly used drugs on human metabolism. *FEBS Journal*, *282*, 297–317.

Salway, J. G. (2003). *Metabolism at a glance* (3rd ed.). Hoboken, NJ: Wiley-Blackwell.

Schellenberger, J., Que, R., Fleming, R. M., Thiele, I., Orth, J. D., Feist, A. M., et al. (2011). Quantitative prediction of cellular metabolism with constraint-based models: The COBRA Toolbox v2.0. *Nature Protocols*, *6*(9), 1290–1307.

Shlomi, T., Cabili, M., & Ruppin, E. (2009). Predicting metabolic biomarkers of human inborn errors of metabolism. *Molecular Systems Biology*, *5*(263), 2009.

Smallbone, K. (2013). *Striking a balance with Recon 2.1*. arXiv:1311.5696.

Swainston, N., Hastings, J., Dekker, A., Muthukrishnan, V., May, J., Steinbeck, C., & Mendes, P. (2016). libChEBI: An API for accessing the ChEBI database. *Journal of Cheminformatics*, *8*, 11.

Swainston, N., Mendes, P., & Kell, D. B. (2013). An analysis of a ‘community-driven’ reconstruction of the human metabolic network. *Metabolomics*, *9*, 757–764.

- Swainston, N., Smallbone, K., Mendes, P., Kell, D., & Paton, N. (2011). The SuBliMinaL Toolbox: Automating steps in the reconstruction of metabolic networks. *Journal of Integrative Bioinformatics*, 8, 186.
- Thiele, I., Swainston, N., Fleming, R. M., Hoppe, A., Sahoo, S., Aurich, M. K., et al. (2013). A community-driven global reconstruction of human metabolism. *Nature Biotechnology*, 31, 419–425.
- Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., et al. (2015). Tissue-based map of the human proteome. *Science*, 347, 1260419.
- Yoo, S., & Harman, M. (2012). Regression testing minimization, selection and prioritization: A survey. *Software Testing Verification and Reliability*, 22, 67–120.