



Genome-scale modeling and human disease: an overview

Matthew A. Oberhardt^{1*†} and Erwin P. Gianchandani^{2*†‡}

¹ Department of Molecular Microbiology and Biotechnology, Faculty of Life Sciences, School of Computer Sciences, and Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel

² Division of Computer and Network Systems, United States National Science Foundation, Arlington, VA, USA

*Correspondence: mattoby@gmail.com; egiancha@nsf.gov

† These authors have contributed equally to this work.

‡ The views expressed in this article are strictly those of the author and not representative of those of the US National Science Foundation.

Edited by:

Guillermo A. Cecchi, IBM Watson Research Center, USA

Reviewed by:

Pablo Meyer, IBM Watson Research Center, USA

Keywords: genome-scale models, systems analysis, genome-scale metabolic reconstruction, human diseases, systems biology

A commentary on

Genome-scale modeling and human disease

by Gianchandani, E. P., and Oberhardt, M. A.

The last several decades have seen extraordinary progress in the biomedical sciences. The explosion of sequencing and high-throughput data is both welcome and daunting for the study of human disease: while human disease is increasingly understood to be multi-factorial and systemic, the sheer complexity of the data being generated makes unaided interpretation nearly impossible. Meanwhile, genome-scale modeling (GSM) has emerged as a major scaffold and toolkit for contextualizing rich data, and one especially suited to the thousands-of-datapoints-per-measurement reality of contemporary disease research.

The archetypal genome-scale model is the genome-scale metabolic reconstruction (GENRE), a predictive network model that contains up to several thousand metabolic reactions, as well as associated genes and enzymes (but not kinetics, due to the scale) (Oberhardt et al., 2009). Recently available GENREs of human metabolism have opened up enormous avenues in disease research (Duarte et al., 2007; Ma et al., 2007; Thiele et al., 2013), especially when integrated with high-throughput data [for an extensive review, see in this topic: (Blazier and Papin, 2012)]. These models rely on

extensive manual curation, and annotating understudied or ambiguous parts of metabolism is critical for improving their predictive power. In an effort to address one of the most difficult-to-annotate areas of metabolism, researchers involved in the human metabolic reconstruction efforts have provided for this topic a large analysis of membrane transporters in human metabolism, including a discussion of how transport impacts multiple human diseases (Sahoo et al., 2014). GENREs are contributing to many areas of disease research, as detailed below, and their scope and influence will increase as a result of such contributions.

Systemic metabolic disorders such as obesity and diabetes exact a huge toll in the US and worldwide, and GSMs are increasingly being used for their study. Large-scale models of mitochondria, for example, have helped examine obesity-associated aberrations in mitochondrial fatty acid degradation (Van Eunen et al., 2013) and many other aspects of energy metabolism as reviewed in this topic: (Sangar et al., 2012). Similarly, the human GENRE has been used in a number of studies relevant to metabolic diseases [e.g., building a model of human adipocyte—(Mardinoglu et al., 2013); determining biomarkers for inborn errors of metabolism—(Shlomi et al., 2009)], as extensively reviewed here: (Varemo et al., 2013). GENREs are obvious choices for studying metabolically-based diseases, and will likely be relied on more in the future.

Another area of increasing interest in human disease is the impact of the microbial organisms that cohabit our bodies, collectively known as our “microbiome.” The gut microbiome, for example, has been shown to alter the metabolism of many drugs (Kang et al., 2013), and to be a causative factor in maintaining obese or healthy states (Turnbaugh et al., 2006). GENREs have been used to examine prominent members of the gut microbiota (Heinken et al., 2013), to understand interactions between gut microbes (Shoae et al., 2013), and to explore interactions between gut microbes and epithelial cells (Sahoo and Thiele, 2013). GSMs are still severely limited in this arena due to challenges in community microbial modeling. However, large-scale microbiome modeling efforts will likely have increasing impact as they mature in the coming years, both by driving new knowledge of complex community phenotypes (e.g., Freilich et al., 2011 and reviewed generally in Greenblum et al., 2013) and by including so-far neglected areas such as the oral microbiome, as reviewed in this topic: (Dimitrov and Hoeng, 2013).

Cancer is a complex and multifaceted disease, and a hallmark for huge data collection efforts. As such, it is a natural target for systems modeling [for a general review of systems biology approaches, see in this topic: (Hernandez Patino et al., 2012)]. Metabolic deregulation in cancer has generated considerable interest within the genome-scale metabolic modeling community, resulting in a number of

cancer-related metabolic reconstructions being recently published [see a review in this topic: (Lewis and Abdel-Haleem, 2013)]. Models of specific cancer subtypes are now being built based on the generic human GENRE (Jerby et al., 2010; Agren et al., 2012), and in a few cases, they have revealed insights with therapeutic potentiality (Frezza et al., 2011; Agren et al., 2014).

Due to their lack of kinetic parameters, GENREs alone cannot predict dynamic cell states, nor, surprisingly, can they integrate metabolite concentration data into basic kinetics or allosteric regulation. Since kinetic parameters are difficult to measure and can vary between conditions or cells, ensemble modeling was recently used to estimate kinetic models of human colorectal adenocarcinoma cell lines, and to reveal potential synthetic lethal interactions that could yield new drug targets [see in this topic: (Khazaei et al., 2012)]. Cancer is also a disease marked by the evolutionary

process that the cancerous cells undergo. Genomic data and increasingly sophisticated population models are now enabling elucidation of these processes, which are critical for establishing the basis for cures [see a review in this topic: (Stransky and De Souza, 2012)]. These areas have gained a lot of interest, and we expect many more systems-level studies of cancer in the near future.

By contrast, neurological disorders constitute a set of diseases that have not received as much attention in the GSM community, despite the significant impacts illustrated in **Figure 1**. Early attention focused on genome-wide expression analyses and gene-interaction networks, often using yeast pathways conserved in humans and implicated in neurodegenerative diseases such as Parkinson's, Alzheimer's, and Huntington's (Petranovic and Nielsen, 2008; Noorbakhsh et al., 2009; Wall et al., 2009). More recently, efforts have

begun to employ GSM with success. For example, (Lewis et al., 2010) integrated gene expression data, proteomics data, and literature-based manual curation to model brain energy metabolism and recapitulate the metabolic interactions between astrocytes and various neuron types relevant to Alzheimer's disease. In addition, transcriptomic data from Alzheimer's patients were integrated with a genome-scale computational human metabolic model to characterize the altered metabolism in the diseased state, and metabolic modeling methods were employed to predict metabolic biomarkers and drug targets (Stempler et al., 2014). We expect interest in neurological illnesses to continue to rise.

While much of this short review focuses on GENRE-based analyses, GENREs are by no means the only genome-scale models of note. Many alternative topology-based methods for pathway analysis are available and have been reviewed here: (Mitrea

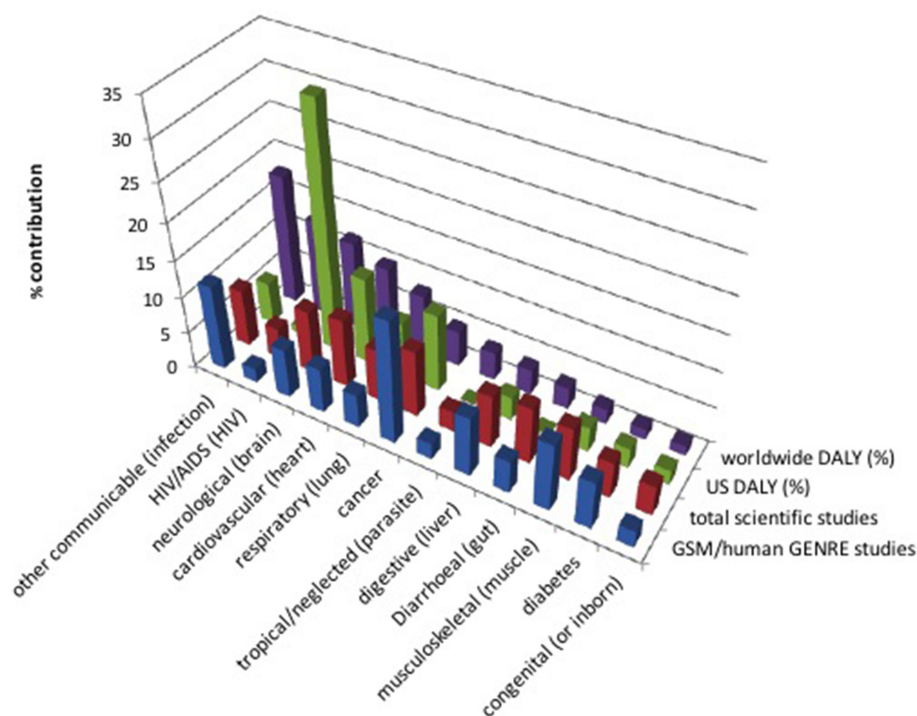


FIGURE 1 | Publications in different disease areas vs. morbidity rates. Age standardized disability adjusted life years (DALY), a measure of years of life lost due to disease, is reported as percent worldwide and in the US. Injury (11.3%) and "other non-communicable disease" (7.4%) are not listed due to lack of related search terms. The count of scientific studies found in Google Scholar searches [normalized to 83% (=100% - 11.3% - 7.4%)] of disease-related search terms (listed in parentheses, if different than the disease name) are

also shown, along with studies specifically related to genome-scale models or to the human GENRE [the numbers of hits were averaged between hits including the searchterm AND citing (Duarte et al., 2007), and hits including the searchterm AND including the phrase "genome scale model," before normalization]. DALY data were taken from the World Health Organization (WHO) report for the year 2002 (which is the most updated WHO source of DALYs of which these authors are aware) (Mathers et al., 2005).

et al., 2013). We also include in this topic a promising new Boolean-based model for somatic cell reprogramming: (Flottmann et al., 2012). Somatic cell reprogramming is a new and highly promising field—it first emerged in 2006 with the landmark paper (Takahashi and Yamanaka, 2006)—that could lead to novel therapeutic approaches, such as growing organs from skin cells for self-transplant.

GSM-based analysis is now a key asset in studying disease. The works in this topic reflect trends in the biomedical sciences at large, including areas of intense interest (e.g., cancer) as well as those that have been labeled as neglected diseases (e.g., few models have been built for studying parasitic tropical diseases or HIV/AIDS). Although eukaryote reconstructions are more challenging due to genome sizes, knowledge coverage, and the multitude of cellular compartments (Thiele and Palsson, 2010), we expect the successes described in this overview to continue to mount—with a particular focus in coming years on clinical problems with translatable outcomes, in which models will help identify new drug targets or alternate cures. This is already evident from recent DREAM Challenges, which have sought to foster collaboration and build communities around fundamental questions at the intersection of systems biology and translational medicine [see, for example, (Margolin et al., 2013)]. To help guide and contextualize disease study, we have included a chart of the most devastating diseases, along with the amount of focus in GSM studies as well as in science at large toward addressing them (Figure 1). Shifting focus toward neglected areas is a worthy goal to which we hope this mini-review will contribute.

REFERENCES

- Agren, R., Bordel, S., Mardinoglu, A., Pornputtpong, N., Nookaew, I., and Nielsen, J. (2012). Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using INIT. *PLoS Comput. Biol.* 8:e1002518. doi: 10.1371/journal.pcbi.1002518
- Agren, R., Mardinoglu, A., Asplund, A., Kampf, C., Uhlen, M., and Nielsen, J. (2014). Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling. *Mol. Syst. Biol.* 10, 721. doi: 10.1002/msb.145122
- Blazier, A. S., and Papin, J. A. (2012). Integration of expression data in genome-scale metabolic network reconstructions. *Front. Physiol.* 3:299. doi: 10.3389/fphys.2012.00299
- Dimitrov, D. V., and Hoeng, J. (2013). Systems approaches to computational modeling of the oral microbiome. *Front. Physiol.* 4:172. doi: 10.3389/fphys.2013.00172
- 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. *Proc. Natl. Acad. Sci. U.S.A.* 104, 1777–1782. doi: 10.1073/pnas.0610772104
- Flottmann, M., Scharp, T., and Klipp, E. (2012). A stochastic model of epigenetic dynamics in somatic cell reprogramming. *Front. Physiol.* 3:216. doi: 10.3389/fphys.2012.00216
- Freilich, S., Zarecki, R., Eilam, O., Segal, E. S., Henry, C. S., Kupiec, M., et al. (2011). Competitive and cooperative metabolic interactions in bacterial communities. *Nat. Commun.* 2, 589. doi: 10.1038/ncomms1597
- 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. doi: 10.1038/nature10363
- Greenblum, S., Chiu, H. C., Levy, R., Carr, R., and Borenstein, E. (2013). Towards a predictive systems-level model of the human microbiome: progress, challenges, and opportunities. *Curr. Opin. Biotechnol.* 24, 810–820. doi: 10.1016/j.copbio.2013.04.001
- Heinken, A., Sahoo, S., Fleming, R. M., and Thiele, I. (2013). Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. *Gut Microbes* 4, 28–40. doi: 10.4161/gmic.22370
- Hernandez Patino, C. E., Jaime-Munoz, G., and Resendis-Antonio, O. (2012). Systems biology of cancer: moving toward the integrative study of the metabolic alterations in cancer cells. *Front. Physiol.* 3:481. doi: 10.3389/fphys.2012.00481
- Jerby, L., Shlomi, T., and Ruppim, E. (2010). Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism. *Mol. Syst. Biol.* 6:401. doi: 10.1038/msb.2010.56
- Kang, M. J., Kim, H. G., Kim, J. S., Oh Do, G., Um, Y. J., Seo, C. S., et al. (2013). The effect of gut microbiota on drug metabolism. *Expert Opin. Drug Metab. Toxicol.* 9, 1295–1308. doi: 10.1517/17425255.2013.807798
- Khazaei, T., McGuigan, A., and Mahadevan, R. (2012). Ensemble modeling of cancer metabolism. *Front. Physiol.* 3:135. doi: 10.3389/fphys.2012.00135
- Lewis, N. E., and Abdel-Haleem, A. M. (2013). The evolution of genome-scale models of cancer metabolism. *Front. Physiol.* 4:237. doi: 10.3389/fphys.2013.00237
- Lewis, N., Schramm, G., Bordbar, A., Schellenberger, J., Andersen, M., Cheng, J., et al. (2010). Large-scale *in silico* modeling of metabolic interactions between cell types in the human brain. *Nat. Biotechnol.* 28, 1279. doi: 10.1038/nbt.1711
- Ma, H., Sorokin, A., Mazein, A., Selkov, A., Selkov, E., Demin, O., et al. (2007). The Edinburgh human metabolic network reconstruction and its functional analysis. *Mol. Syst. Biol.* 3, 135. doi: 10.1038/msb4100177
- Mardinoglu, A., Agren, R., Kampf, C., Asplund, A., Nookaew, I., Jacobson, P., et al. (2013). Integration of clinical data with a genome-scale metabolic model of the human adipocyte. *Mol. Syst. Biol.* 9. doi: 10.1038/msb.2013.5
- Margolin, A., Bilal, E., Huang, E., Norman, T., Ottestad, L., Mecham, B., et al. (2013). Systematic analysis of challenge-driven improvements in molecular prognostic models for breast cancer. *Sci. Transl. Med.* 5, 181re1. doi: 10.1126/scitranslmed.3006112
- Mathers, C. D., Fat, D. M., Inoue, M., Rao, C., and Lopez, A. D. (2005). Counting the dead and what they died from: an assessment of the global status of cause of death data. *Bull. World Health Organ.* 83, 171–177. Available online at: http://www.scielo.org/scielo.php?script=sci_arttext&pid=S0042-96862005000300009&lng=en&nrn=iso&tng=en
- Mitrea, C., Taghavi, Z., Bokanizad, B., Hanoudi, S., Tagett, R., Donato, M., et al. (2013). Methods and approaches in the topology-based analysis of biological pathways. *Front. Physiol.* 4:278. doi: 10.3389/fphys.2013.00278
- Noorbakhsh, F., Overall, C., and Power, C. (2009). Deciphering complex mechanisms in neurodegenerative diseases: the advent of systems biology. *Trends Neurosci.* 32, 88. doi: 10.1016/j.tins.2008.10.003
- Oberhardt, M. A., Palsson, B. O., and Papin, J. A. (2009). Applications of genome-scale metabolic reconstructions. *Mol. Syst. Biol.* 5, 320. doi: 10.1038/msb.2009.77
- Petranovic, D., and Nielsen, J. (2008). Can yeast systems biology contribute to the understanding of human disease? *Trends Biotechnol.* 26, 584. doi: 10.1016/j.tibtech.2008.07.008
- Sahoo, S., Aurich, M. K., Jonsson, J. J., and Thiele, I. (2014). Membrane transporters in a human genome-scale metabolic knowledgebase and their implications for disease. *Front. Physiol.* 5:91. doi: 10.3389/fphys.2014.00091
- Sahoo, S., and Thiele, I. (2013). Predicting the impact of diet and enzymopathies on human small intestinal epithelial cells. *Hum. Mol. Genet.* doi: 10.1093/hmg/ddt119
- Sangar, V., Eddy, J. A., Simeonidis, E., and Price, N. D. (2012). Mechanistic modeling of aberrant energy metabolism in human disease. *Front. Physiol.* 3:404. doi: 10.3389/fphys.2012.00404
- Shlomi, T., Cabili, M. N., and Ruppim, E. (2009). Predicting metabolic biomarkers of human inborn errors of metabolism. *Mol. Syst. Biol.* 5, 263. doi: 10.1038/msb.2009.22
- Shoae, S., Karlsson, F., Mardinoglu, A., Nookaew, I., Bordel, S., and Nielsen, J. (2013). Understanding the interactions between bacteria in the human gut through metabolic modeling. *Sci. Rep.* 3:2532. doi: 10.1038/srep02532
- Stempler, S., Yizhak, K., and Ruppim, E. (2014). Integrating transcriptomics with metabolic modeling predicts biomarkers and drug targets for Alzheimer's Disease. *PLoS ONE* 9:e105383. doi: 10.1371/journal.pone.0105383
- Stransky, B., and De Souza, S. J. (2012). Modeling tumor evolutionary dynamics. *Front. Physiol.* 3:480. doi: 10.3389/fphys.2012.00480
- Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and

- adult fibroblast cultures by defined factors. *Cell* 126, 663–676. doi: 10.1016/j.cell.2006.07.024
- Thiele, I., and Palsson, B. (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* 5, 93. doi: 10.1038/nprot.2009.203
- 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. *Nat. Biotechnol.* doi: 10.1038/nbt.2488
- Turnbaugh, P. J., Ley, R. E., Mahowald, M. A., Magrini, V., Mardis, E. R., and Gordon, J. I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444, 1027–1031. doi: 10.1038/nature05414
- Van Eunen, K., Simons, S. M., Gerding, A., Bleeker, A., Den Besten, G., Touw, C. M., et al. (2013). Biochemical competition makes fatty-acid beta-oxidation vulnerable to substrate overload. *PLoS Comput. Biol.* 9:e1003186. doi: 10.1371/journal.pcbi.1003186
- Varemo, L., Nookaew, I., and Nielsen, J. (2013). Novel insights into obesity and diabetes through genome-scale metabolic modeling. *Front. Physiol.* 4:92. doi: 10.3389/fphys.2013.00092
- Wall, D., Esteban, E., DeLuca, T., Huyck, M., Monaghan, T., Velez de Mendizabal, N., et al. (2009). Comparative analysis of neurological disorders focuses genome-wide search for autism genes. *Genomics* 93, 120. doi: 10.1016/j.ygeno.2008.09.015
- Conflict of Interest Statement:** There are no commercial or financial conflicts of interest. However, the views expressed in this article are strictly those of EPG and not representative of those of the US National Science Foundation.
- Received: 28 August 2014; accepted: 23 December 2014; published online: 23 January 2015.
- Citation: Oberhardt MA and Gianchandani EP (2015) Genome-scale modeling and human disease: an overview. *Front. Physiol.* 5:527. doi: 10.3389/fphys.2014.00527
- This article was submitted to *Computational Physiology and Medicine*, a section of the journal *Frontiers in Physiology*.
- Copyright © 2015 Oberhardt and Gianchandani. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.