REVIEW ARTICLE

An analysis of a 'community-driven' reconstruction of the human metabolic network

Neil Swainston · Pedro Mendes · Douglas B. Kell

Abstract Following a strategy similar to that used in baker's yeast (Herrgård et al. Nat Biotechnol 26:1155–1160, 2008). A consensus yeast metabolic network obtained from a community approach to systems biology (Herrgård et al. 2008; Dobson et al. BMC Syst Biol 4:145, 2010). Further developments towards a genome-scale metabolic model of yeast (Dobson et al. 2010; Heavner et al. BMC Syst Biol 6:55, 2012). Yeast 5—an expanded reconstruction of the *Saccharomyces cerevisiae* metabolic network (Heavner et al. 2012) and in *Salmonella typhimurium* (Thiele et al. BMC Syst Biol 5:8, 2011). A community effort towards a knowledge-base and mathematical model of the human pathogen *Salmonella typhimurium LT2* (Thiele et al. 2011), a recent paper (Thiele et al. Nat Biotechnol 31:419–425, 2013). A community-driven global reconstruction of human metabolism (Thiele et al. 2013) described a much improved 'community consensus' reconstruction of the human metabolic network, called Recon 2, and the authors (that include the present ones) have made it freely available via a database at http://humanmetabolism.org/ and in SBML format at Biomodels (http://identifiers.org/biomodels.db/MODEL1109130000). This short analysis summarises the main findings, and suggests some approaches that will be able to exploit the availability of this model to advantage.

Keywords Metabolism · Modelling · Systems biology · Networks · Metabolic networks

N. Swainston · D. B. Kell
School of Chemistry, The University of Manchester,
Oxford Road, Manchester M13 9PL, UK

D. B. Kell
e-mail: dbk@manchester.ac.uk

N. Swainston (✉) · P. Mendes · D. B. Kell
Manchester Institute of Biotechnology, The University
of Manchester, Princess Street, Manchester M1 7DN, UK
e-mail: neil.swainston@manchester.ac.uk

P. Mendes
e-mail: pedro.mendes@manchester.ac.uk

P. Mendes
School of Computer Science, The University of Manchester,
Oxford Road, Manchester M13 9PL, UK

P. Mendes
Virginia Bioinformatics Institute, Virginia Tech, Washington St.
0477, Blacksburg, VA 24060, USA

1 Main findings of the Recon 2 paper

A highly curated consensus reconstruction of the human metabolic network, termed Recon 2, was recently released (Thiele et al. 2013). The development of the network followed a community 'jamboree' approach (Herrgård et al. 2008; Dobson et al. 2010; Heavner et al. 2012; Thiele and Palsson 2010), exploiting both genomic and literature data to expand upon existing reconstructions (Duarte et al. 2007; Ma et al. 2007; Gille et al. 2010) to produce a 'basal' network that contains 7,440 reactions, 5,063 metabolite pools and 2,626 unique metabolites. The reconstruction is extensively semantically annotated (Kell and Mendes 2008; Courtot et al. 2011), fully compliant with the MIRIAM standard (Le Novère et al. 2005), unambiguously identifying cellular compartments, metabolites, genes and enzymes with publicly available, external database terms (Krause et al. 2011). Thus, cellular compartments are annotated with Gene Ontology (GO) terms, while metabolites are annotated with terms from resources such as Chemical Entities of Biological Interest (ChEBI) (Hastings

et al. 2013) as well as using IUPAC International Chemical Identifier (InChI http://www.iupac.org/home/publications/e-resources/inchi.html) terms (Coles et al. 2005) where possible. Reactions are curator-validated and annotated with PubMed literature references, standardized GO evidence codes, and a confidence scoring system ranging from 0 (no evidence) to 4 (biochemical evidence). Metabolic reactions were checked to ensure correct stoichiometry, (ir)reversibility, the correct assignment of gene association and enzyme rules, and mass and charge balancing. Appropriate transport reactions were also included and these followed the same level of annotation as reactions.

In contrast to existing resources such as KEGG (Kanehisa and Goto 2000) or the Human Metabolome Database (HMDB) (Wishart et al. 2007), Recon 2 acts as both a knowledgebase and a predictive model, amenable to constraint-based analysis approaches such as flux balance analysis (Orth et al. 2010). To demonstrate this utility, the Recon 2 authors focused on five analyses of immediate interest. First they defined a metabolic task as a nonzero flux through a reaction or through a pathway leading to the production of a metabolite B from a metabolite A; 354 such metabolic tasks were defined and all carried out successfully in silico. Secondly, they established whether known mutations producing 'inborn errors of metabolism' (IEMs) did have the predicted effect on biomarkers (54 reported biomarkers for 49 different IEMs, with an accuracy of 77 %; see also Shlomi et al. 2009). Thirdly, they showed that Recon 2 should predict a large fraction of metabolites that are excreted (the 'metabolic footprint' (Allen et al. 2003) or 'exometabolome' (Kell et al. 2005)). Fourthly, based on expression profiling data from the Human Protein Atlas (Uhlén et al. 2010), they generated 65 draft cell-type-specific models, and fifthly they found (notwithstanding the rather promiscuous behaviour of many drugs (Hopkins 2008, 2009; Kell et al. 2013)) that they could map 1,290 drugs to 308 enzyme and enzymatic complexes.

2 Some known shortcomings of Recon 2

While Recon 2 represents the 'state of the art' of public human metabolic network reconstructions, it should be acknowledged that it does have some known shortcomings, including the fact that a number of known metabolites and reactions (including those involving unliganded iron (Hower et al. 2009; Kell 2009, 2010; Chifman et al. 2012; Funke et al. 2013)) have still to be included, and there are increasing numbers of 'unexpected' metabolite-protein reactions that are being discovered (Li et al. 2010; Li and Snyder 2011; Kell 2011; Kell et al. 2013). These are thus mainly 'false negatives' (Broadhurst and Kell 2006), and dealing with them is clearly one of the goals that will

remain in any continuing curation process. It is recognised that the network reconstruction process is iterative (Reed and Palsson 2003), and the metabolomics and systems biology communities are encouraged to contribute to this ongoing effort. Following an approach that has been applied successfully in the generation of subsequent iterations of the yeast consensus model (Herrgård et al. 2008; Dobson et al. 2010; Heavner et al. 2012), suggested updates and amendments can be e-mailed to network.reconstruction @manchester.ac.uk.

3 What the Recon 2 network will allow us to do or to do better

3.1 General benefits of network models

As discussed previously (Kell 2006a; b), the availability of a systems biology model of a metabolic network allows one to effect a variety of analyses, some of which are illustrated in Fig. 1. We here mention just a few that are likely to be of most interest to the metabolomics community.

3.2 Improved predictions of metabolic fluxes, including in biotechnology

Given the topology of a network, and the stoichiometric and thermodynamic constraints under which metabolic networks must operate (Palsson 2006; Kell 2006a; b), it is possible to use generalised kinetics to predict metabolic fluxes (Liebermeister and Klipp 2006; Smallbone et al. 2007; Smallbone and Simeonidis 2008; Smallbone et al. 2010). The accuracy of these predictions can of course be enhanced by the use of known kinetic rate equations (Li et al. 2010), and even by expression profiles alone (Lee et al. 2012). Such an approach has been applied, exploiting both transcriptomics and fluxomics data, to constrain models derived from a precursor of Recon 2 in order to elucidate and validate new drug targets in renal-cell cancer (Frezza et al. 2011).

The use of network biology in predicting fluxes (and how to change them), as well as in parameter optimisation (Mendes and Kell 1998; Moles et al. 2003; Adams et al. 2013), has enjoyed particular success in biotechnology where it is usually the fluxes to external products that are of interest (Park et al. 2007; Lee et al. 2012; Park et al. 2010; Becker et al. 2011). In this area, the ongoing development of a systems biology toolkit for Chinese Hamster Ovary (CHO) cells, which will be increasingly utilised for biotechnological production of pharmaceutical proteins (Kildegaard et al. 2013), will be aided by the development of Recon 2, which can act as a template for development of a

Fig. 1 A summary of some of the intellectual areas in which we can create and exploit the contents of systems biology models as encoded in SBML

detailed metabolic reconstruction of CHO. Additionally, one may anticipate the importance of predictions of metabolic fluxes in understanding nutrition and regulation in health and disease.

3.3 Understanding and incorporating knowledge of drugs that use known transporters

As part of the need to incorporate 'new' proteins and their interactions with small molecules, one particular feature that has become increasingly apparent in recent years is that pharmaceutical drugs do not normally cross membranes 'passively' through any phospholipid bilayer portions that they may contain, but hitchhike on the carriers that participate in the transmembrane transport of

intermediary metabolites (Al-Awqati 1999; Dobson and Kell 2008; Dobson et al. 2009; Dobson et al. 2009; Kell and Dobson 2009; Giacomini et al. 2010; Burckhardt and Burckhardt 2011; Kell et al. 2011; Lanthaler et al. 2011; DeGorter et al. 2012; Kell and Goodacre 2013). It is likely that these kinds of issues contribute significantly to the dreadful attrition rates still seen in drug development (van der Greef and McBurney 2005; Kola and Landis 2004; Kola 2008; Empfield and Leeson 2010; Leeson and Empfield 2010; Kwong et al. 2011). The availability of Recon 2 and its tissue-specific versions will now make it much easier to correlate drug disposition with transporter expression, and thereby determine (with suitable machine learning analyses (Kell et al. 2001)) the roles of the different transporters in effecting the cellular uptake and

efflux of particular drugs. Incorporating this kind of knowledge into subsequent iterations of Recon 2 is an urgent priority.

3.4 Other approaches to mining the metabolic network

An important recognition (Herrgård et al. 2008), continued in Recon 2, was that of the utility of the methods of cheminformatics (Gasteiger 2003) in providing chemically accurate and database-independent descriptions of the structures of metabolites that allowed models (such as those encoded in SBML (Hucka et al. 2003; Hucka et al. 2004)) to be interrogated computationally. In a similar vein, there is an increasing trend towards automated reasoning about the content of scientific papers from a systems biology point of view (Hakenberg et al. 2004; King et al. 2005; Ananiadou et al. 2006; Kell and Mendes 2008; Ananiadou et al. 2010; Ray et al. 2010; Miwa et al. 2012, 2013), including about their metabolomes (Knox et al. 2007; Attwood et al. 2009; Attwood et al. 2010; Nobata et al. 2011; Zhou et al. 2012; Hastings et al. 2013). The availability of Recon 2 will allow one to ask questions such as, "how many metabolites with a given substructure are present in the network?" or "which metabolites are common (or different) between these two networks?" or to plot out the distributions of various kinds of properties that may be of interest (Dobson et al. 2009); one such plot, simply showing the distribution of molecular masses, is given in Fig. 2.

### 3.5 Comparison with the experimental metabolome

Analysis of the metabolome as encoded in biochemical networks also leads one to recognise the importance of comparing systems biology models with the experimental metabolome (i.e. the concentrations of small molecules) measured in different circumstances ('comparative metabolomics', Raamsdonk et al. 2001; Levandi et al. 2008). As with Recon 2, the experimental metabolome of, for example, human serum consists of several thousands of reproducibly detectable metabolites (O'Hagan et al. 2005; 2007; Begley et al. 2009; Zelena et al. 2009). While some of these experimentally observed metabolites will certainly originate from nutrients or the gut microbiome (Goodacre 2007; Li et al. 2008; Wikoff et al. 2009; Zhao and Shen 2010; Wang et al. 2011; Bennett et al. 2013; Collino et al. 2013; Heinken et al. 2013), Recon 2 allows one to set down those that are at least encoded in the human genome sequence, and compare these with the contents of the various metabolome databases (Brown et al. 2009; Kamp et al. 2012; van Ravenzwaay et al. 2012; Sawada et al. 2012; Steinbeck et al. 2012; Tautenhahn et al. 2012; Wishart 2012; Zhou et al. 2012; Guo et al. 2013; Hastings et al. 2013; Haug et al. 2013; Li et al. 2013; Salek et al. 2013; Sakurai et al. 2013; Wishart et al. 2013). Data standards such as SBRML (Dada et al. 2010) allow a straightforward comparison of network models with experimental data encoded in a compatible format.

3.6 The importance of semantic annotation

Many of the application areas described above are dependent upon the semantic awareness of Recon 2, and the incorporation of thousands of unique, persistent, unambiguous semantic annotations that allow for software-driven analyses of the knowledgebase and derived models. By representing both the network and its semantic information using community-driven standards such as SBML (Hucka

Fig. 2 An assessment of the distribution of molecular masses of the metabolites in Recon 2
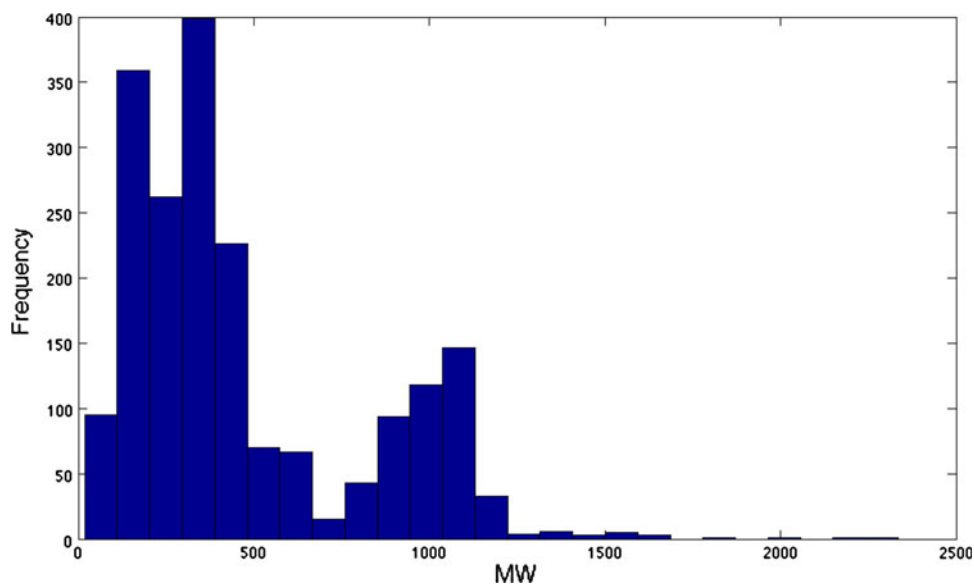
et al. 2003, 2004) and MIRIAM (Le Novère et al. 2005), software producers are able to develop against a given standard, decoupling the network model from the techniques used in its various more specialized analyses.

There is increasing community interest in the development of tissue- and condition-specific models, and this task is dependent upon the integration of large-scale 'omics data. Methods to perform such integration are many-fold and are in constant development (Mo et al. 2009; Jerby et al. 2010; Wang et al. 2012), but all are reliant on automated approaches, given that the size of the datasets involved renders manual integration impossible. Recon 2 is annotated with third-party identifiers across numerous scales, from genomics through to transcriptomics, proteomics and metabolomics, all of which can be mapped to appropriate web services, allowing for their automated interpretation (Swainston and Mendes 2009) and integration of multi-omics data (Li et al. 2008, 2008; Hyduke et al. 2013). Additionally, the definition of metabolites in structural terms permits the exploitation of cheminformatics tools such as the Chemistry Development Toolkit (Steinbeck et al. 2003) and Open Babel that allow one to translate the various encodings or mappings of chemical structures (O'Boyle et al. 2011).

Furthermore, the specification of metabolites, enzymes and reactions in unambiguous terms facilitates the development of knowledgebases and models of related organisms, through automated or semi-automated means (Henry et al. 2010; Swainston et al. 2011; Agren et al. 2013). Recon 2 can therefore act as a template for the development of metabolic reconstructions of related model organisms, facilitating comparative studies and simulation of metabolism between human and other model systems (Sigurdsson et al. 2010).

4 Concluding remarks

The availability of Recon 2 allows a great many computational analyses to be performed. We have purposely rehearsed these at a rather general level, since particular analyses, that might be relevant to particular diseases, for instance, are simply implementations of the more general approaches. One new approach that will depend on the existence of such a network as a necessary resource is personalised medicine (Hood and Flores 2012). There one will develop models of metabolism calibrated for each specific individual, in large part using metabolomics methods, to be used as bases for diagnostics and decisions on course of treatment. Recon 2 is a very significant step towards such a map, where such measurements have to be anchored for various types of modelling that will underpin personalised treatment decisions.

The development of tissue- and condition-specific models has been demonstrated with Recon 2 and its predecessors (Jerby et al. 2010; Frezza et al. 2011; Wang et al. 2012). As subsequent iterations of Recon 2 develop, it is hoped that the scope of the knowledgebase, and the predictive power of derived models, will increase to keep pace with advancements in the community knowledge of human metabolism, many of which will be driven by the discipline of metabolomics.

References

Adams, R., Clark, A., Yamaguchi, A., et al. (2013). SBSI: an extensible distributed software infrastructure for parameter estimation in systems biology. *Bioinformatics, 29*, 664–665.

Agren, R., Liu, L., Shoaie, S., Vongsangnak, W., Nookaew, I., & Nielsen, J. (2013). The RAVEN toolbox and its use for generating a genome-scale metabolic model for *Penicillium chrysogenum*. *PLoS Computational Biology, 9*, e1002980.

Al-Awqati, Q. (1999). One hundred years of membrane permeability: does overton still rule? *Nature Cell Biology, 1*, E201–E202.

Allen, J. K., Davey, H. M., Broadhurst, D., et al. (2003). High-throughput characterisation of yeast mutants for functional genomics using metabolic footprinting. *Nature Biotechnology, 21*, 692–696.

Ananiadou, S., Kell, D. B., & Tsujii, J. (2006). Text mining and its potential applications in systems biology. *Trends in Biotechnology, 24*, 571–579.

Ananiadou, S., Pyysalo, S., Tsujii, J., & Kell, D. B. (2010). Event extraction for systems biology by text mining the literature. *Trends in Biotechnology, 28*, 381–390.

Attwood, T. K., Kell, D. B., McDermott, P., Marsh, J., Pettifer, S. R., & Thorne, D. (2009). Calling International rescue: knowledge lost in literature and data landslide! *Biochemical Journal, 424*, 317–333.

Attwood, T. K., Kell, D. B., McDermott, P., Marsh, J., Pettifer, S., & Thorne, D. (2010). Utopia documents: linking scholarly literature with research data. *Bioinformatics, 26*, i568–i574.

Becker, J., Zelder, O., Häfner, S., Schröder, H., & Wittmann, C. (2011). From zero to hero–design-based systems metabolic engineering of *Corynebacterium glutamicum* for L-lysine production. *Metabolic Engineering, 13*, 159–168.

Begley, P., Francis-McIntyre, S., Dunn, W. B., et al. (2009). Development and performance of a gas chromatography-time-of-flight mass spectrometry analysis for large-scale non-targeted metabolomic studies of human serum. *Analytical Chemistry, 81*, 7038–7046.

Bennett, B. J., Vallim, T. Q., Wang, Z., et al. (2013). Trimethylamine-*N*-oxide, a metabolite associated with atherosclerosis, exhibits

complex genetic and dietary regulation. *Cell Metabolism, 17*, 49–60.

Broadhurst, D., & Kell, D. B. (2006). Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics, 2*, 171–196.

Brown, M., Dunn, W. B., Dobson, P., et al. (2009). Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics. *Analyst, 134*, 1322–1332.

Burckhardt, G., & Burckhardt, B. C. (2011). In vitro and in vivo evidence of the importance of organic anion transporters (OATs) in drug therapy. *Handbook of Experimental Pharmacology, 201*, 29–104.

Chifman, J., Kniss, A., Neupane, P., et al. (2012). The core control system of intracellular iron homeostasis: a mathematical model. *Journal of Theoretical Biology, 300*, 91–99.

Coles, S. J., Day, N. E., Murray-Rust, P., Rzepa, H. S., & Zhang, Y. (2005). Enhancement of the chemical semantic web through the use of InChI identifiers. *Organic & Biomolecular Chemistry, 3*, 1832–1834.

Collino, S., Martin, F. P., & Rezzi, S. (2013). Clinical metabolomics paves the way towards future healthcare strategies. *British Journal of Clinical Pharmacology, 75*, 619–629.

Courtot, M., Juty, N., Knüpfer, C., et al. (2011). Controlled vocabularies and semantics in systems biology. *Molecular Systems Biology, 7*, 543.

Dada, J. O., Spasic, I., Paton, N. W., & Mendes, P. (2010). SBRML: a markup language for associating systems biology data with models. *Bioinformatics, 26*, 932–938.

DeGorter, M. K., Xia, C. Q., Yang, J. J., & Kim, R. B. (2012). Drug transporters in drug efficacy and toxicity. *Annual Review of Pharmacology and Toxicology, 52*, 249–273.

Dobson, P. D., & Kell, D. B. (2008). Carrier-mediated cellular uptake of pharmaceutical drugs: an exception or the rule? *Nature Reviews Drug Discovery, 7*, 205–220.

Dobson, P. D., Lanthaler, K., Oliver, S. G., & Kell, D. B. (2009a). Implications of the dominant role of transporters in drug uptake by cells. *Current Topics in Medicinal Chemistry, 9*, 163–184.

Dobson, P. D., Patel, Y., & Kell, D. B. (2009b). "Metabolite-likeness" as a criterion in the design and selection of pharmaceutical drug libraries. *Drug Discovery Today, 14*, 31–40.

Dobson, P. D., Smallbone, K., Jameson, D., et al. (2010). Further developments towards a genome-scale metabolic model of yeast. *BMC Systems Biology, 4*, 145.

Duarte, N. C., Becker, S. A., Jamshidi, N., et al. (2007). Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proceedings of the National Academy of Sciences USA, 104*, 1777–1782.

Empfield, J. R., & Leeson, P. D. (2010). Lessons learned from candidate drug attrition. *IDrugs, 13*, 869–873.

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

Funke, C., Schneider, S. A., Berg, D., & Kell, D. B. (2013). Genetics and iron in the systems biology of Parkinson's disease and some related disorders. *Neurochemistry International, 62*, 637–652.

Gasteiger, J. (Ed.). (2003). *Handbook of chemoinformatics: from data to knowledge*. Weinheim: Wiley-VCH.

Giacomini, K. M., Huang, S. M., Tweedie, D. J., et al. (2010). Membrane transporters in drug development. *Nature Reviews Drug Discovery, 9*, 215–236.

Gille, C., Bölling, C., Hoppe, A., et al. (2010). HepatoNet1: a comprehensive metabolic reconstruction of the human hepatocyte for the analysis of liver physiology. *Molecular Systems Biology, 6*, 411.

Goodacre, R. (2007). Metabolomics of a superorganism. *Journal of Nutrition, 137*, 259S–266S.

Guo, A. C., Jewison, T., Wilson, M., et al. (2013). ECMDB: the *E. coli* Metabolome Database. *Nucleic Acids Research, 41*, D625–D630.

Hakenberg, J., Schmeier, S., Kowald, A., Klipp, E., & Leser, U. (2004). Finding kinetic parameters using text mining. *OMICS: A Journal of Integrative Biology, 8*, 131–152.

Hastings, J., de Matos, P., Dekker, A., et al. (2013). The ChEBI reference database and ontology for biologically relevant chemistry: enhancements for 2013. *Nucleic Acids Research, 41*, D456–D463.

Haug, K., Salek, R. M., Conesa, P., et al. (2013). MetaboLights-an open-access general-purpose repository for metabolomics studies and associated meta-data. *Nucleic Acids Research, 41*, D781–D786.

Heavner, B. D., Smallbone, K., Barker, B., Mendes, P., & Walker, L. P. (2012). Yeast 5—an expanded reconstruction of the *Saccharomyces cerevisiae* metabolic network. *BMC Systems Biology, 6*, 55.

Heinken, A., Sahoo, S., Fleming, R. M., & Thiele, I. (2013). Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. *Gut Microbes, 4*, 28–40.

Henry, C. S., DeJongh, M., Best, A. A., Frybarger, P. M., Linsay, B., & Stevens, R. L. (2010). High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nature Biotechnology, 28*, 977–982.

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

Hood, L., & Flores, M. (2012). A personal view on systems medicine and the emergence of proactive P4 medicine: predictive, preventive, personalized and participatory. *New Biotechnology, 15*, 613–624.

Hopkins, A. L. (2008). Network pharmacology: the next paradigm in drug discovery. *Nature Chemical Biology, 4*, 682–690.

Hopkins, A. L. (2009). Predicting promiscuity. *Nature, 462*, 167–168.

Hower, V., Mendes, P., Torti, F. M., et al. (2009). A general map of iron metabolism and tissue-specific subnetworks. *Molecular BioSystems, 5*, 422–443.

Hucka, M., Finney, A., Bornstein, B. J., et al. (2004). Evolving a *lingua franca* and associated software infrastructure for computational systems biology: the systems biology markup language (SBML) project. *Systematic Biology, 1*, 41–53.

Hucka, M., Finney, A., Sauro, H. M., et al. (2003). The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics, 19*, 524–531.

Hyduke, D. R., Lewis, N. E., & Palsson, B. Ø. (2013). Analysis of omics data with genome-scale models of metabolism. *Molecular BioSystems, 9*, 167–174.

Jerby, L., Shlomi, T., & Ruppin, E. (2010). Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism. *Molecular Systems Biology, 6*, 401.

Kamp, H., Strauss, V., Wiemer, J., et al. (2012). Reproducibility and robustness of metabolome analysis in rat plasma of 28-day repeated dose toxicity studies. *Toxicology Letters, 215*, 143–149.

Kanehisa, M., & Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research, 28*, 27–30.

Kell, D. B. (2006a). Metabolomics, modelling and machine learning in systems biology: towards an understanding of the languages of cells. The 2005 Theodor Bücher lecture. *FEBS Journal, 273*, 873–894.

Kell, D. B. (2006b). Systems biology, metabolic modelling and metabolomics in drug discovery and development. *Drug Discovery Today, 11*, 1085–1092.

Kell, D. B. (2009). Iron behaving badly: inappropriate iron chelation as a major contributor to the aetiology of vascular and other

progressive inflammatory and degenerative diseases. *BMC Medical Genomics, 2*, 2.

Kell, D. B. (2010). Towards a unifying, systems biology understanding of large-scale cellular death and destruction caused by poorly liganded iron: parkinson's, Huntington's, Alzheimer's, prions, bactericides, chemical toxicology and others as examples. *Archives of Toxicology, 577*, 825–889.

Kell, D. B. (2011). Metabolites do social networking. *Nature Chemical Biology, 7*, 7–8.

Kell, D. B., Brown, M., Davey, H. M., Dunn, W. B., Spasic, I., & Oliver, S. G. (2005). Metabolic footprinting and systems biology: the medium is the message. *Nature Reviews Microbiology, 3*, 557–565.

Kell, D. B., Darby, R. M., & Draper, J. (2001). Genomic computing: explanatory analysis of plant expression profiling data using machine learning. *Plant Physiology, 126*, 943–951.

Kell, D. B., & Dobson, P. D. (2009). The cellular uptake of pharmaceutical drugs is mainly carrier-mediated and is thus an issue not so much of biophysics but of systems biology. In M. G. Hicks & C. Kettner (Eds.), *Proc Int Beilstein Symposium on Systems Chemistry* (pp. 149–168). Berlin: Logos Verlag.

Kell, D. B., Dobson, P. D., Bilsland, E., & Oliver, S. G. (2013). The promiscuous binding of pharmaceutical drugs and their transporter-mediated uptake into cells: what we (need to) know and how we can do so. *Drug Discovery Today, 18*, 218–239.

Kell, D. B., Dobson, P. D., & Oliver, S. G. (2011). Pharmaceutical drug transport: the issues and the implications that it is essentially carrier-mediated only. *Drug Discovery Today, 16*, 704–714.

Kell, D.B., Goodacre, R. (2013). Metabolomics and systems pharmacology: why and how to model the human metabolic network for drug discovery. *Drug Discovery Today*. (Submitted).

Kell, D. B., & Mendes, P. (2008). The markup is the model: reasoning about systems biology models in the semantic web era. *Journal of Theoretical Biology, 252*, 538–543.

Kildegaard, H.F., Baycin-Hizal, D., Lewis, N.E., Betenbaugh, M.J. (2013). The emerging CHO systems biology era: harnessing the 'omics revolution for biotechnology. *Current Opinion in Biotechnology*, doi: 10.1016/j.copbio.2013.02.007. (Epub ahead of print).

King, R. D., Garrett, S. M., & Coghill, G. M. (2005). On the use of qualitative reasoning to simulate and identify metabolic pathways. *Bioinformatics, 21*, 2017–2026.

Knox, C., Shrivastava, S., Sothard, P., Eisner, R., & Wishart, D. S. (2007). Biospider: a web server for automating metabolome annotations. *Pacific Symposium on Biocomputing, 12*, 145–156.

Kola, I. (2008). The state of innovation in drug development. *Clinical Pharmacology and Therapeutics, 83*, 227–230.

Kola, I., & Landis, J. (2004). Can the pharmaceutical industry reduce attrition rates? *Nature Reviews Drug Discovery, 3*, 711–715.

Krause, F., Schulz, M., Swainston, N., & Liebermeister, W. (2011). Sustainable model building the role of standards and biological semantics. *Methods in Enzymology, 500*, 371–395.

Kwong, E., Higgins, J., & Templeton, A. C. (2011). Strategies for bringing drug delivery tools into discovery. *International Journal of Pharmaceutics, 412*, 1–7.

Lanthaler, K., Bilsland, E., Dobson, P., et al. (2011). Genome-wide assessment of the carriers involved in the cellular uptake of drugs: a model system in yeast. *BMC Biology, 9*, 70.

Le Novère, N., Finney, A., Hucka, M., et al. (2005). Minimum information requested in the annotation of biochemical models (MIRIAM). *Nature Biotechnology, 23*, 1509–1515.

Lee, J. W., Na, D., Park, J. M., Lee, J., Choi, S., & Lee, S. Y. (2012a). Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nature Chemical Biology, 8*, 536–546.

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

Leeson, P. D., & Empfield, J. R. (2010). Reducing the risk of drug attrition associated with physicochemical properties. *Annual Reports in Medicinal Chemistry, 45*, 393–407.

Levandi, T., Leon, C., Kaljurand, M., Garcia-Cañas, V., & Cifuentes, A. (2008). Capillary electrophoresis time-of-flight mass spectrometry for comparative metabolomics of transgenic versus conventional maize. *Analytical Chemistry, 80*, 6329–6335.

Li, P., Dada, J. O., Jameson, D., et al. (2010a). Systematic integration of experimental data and models in systems biology. *BMC Bioinformatics, 11*, 582.

Li, X., Gianoulis, T. A., Yip, K. Y., Gerstein, M., & Snyder, M. (2010b). Extensive in vivo metabolite-protein interactions revealed by large-scale systematic analyses. *Cell, 143*, 639–650.

Li, L., Li, R., Zhou, J., Zuniga, A., et al. (2013). MyCompoundID: using an Evidence-based Metabolome Library for Metabolite Identification. *Analytical Chemistry, 85*, 3401–3408.

Li, P., Oinn, T., Soiland, S., & Kell, D. B. (2008a). Automated manipulation of systems biology models using libSBML within Taverna workflows. *Bioinformatics, 24*, 287–289.

Li, X., & Snyder, M. (2011). Metabolites as global regulators: a new view of protein regulation: systematic investigation of metabolite-protein interactions may help bridge the gap between genome-wide association studies and small molecule screening studies. *BioEssays, 33*, 485–489.

Li, M., Wang, B., Zhang, M., et al. (2008b). Symbiotic gut microbes modulate human metabolic phenotypes. *Proceedings of the National Academy of Sciences USA, 105*, 2117–2122.

Liebermeister, W., & Klipp, E. (2006). Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theoretical Biology and Medical Model, 3*, 41.

Ma, H., Sorokin, A., Mazein, A., et al. (2007). The Edinburgh human metabolic network reconstruction and its functional analysis. *Molecular Systems Biology, 3*, 135.

Mendes, P., & Kell, D. B. (1998). Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. *Bioinformatics, 14*, 869–883.

Miwa, M., Ohta, T., Rak, R., et al. (2013). A method for integrating and ranking the evidence for biochemical pathways by mining reactions from text. *Bioinformatics*. (In press).

Miwa, M., Thompson, P., McNaught, J., Kell, D. B., & Ananiadou, S. (2012). Extracting semantically enriched events from biomedical literature. *BMC Bioinformatics, 13*, 108.

Mo, M. L., Palsson, B. O., & Herrgård, M. J. (2009). Connecting extracellular metabolomic measurements to intracellular flux states in yeast. *BMC Systems Biology, 3*, 37.

Moles, C. G., Mendes, P., & Banga, J. R. (2003). Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome Research, 13*, 2467–2474.

Nobata, C., Dobson, P., Iqbal, S. A., et al. (2011). Mining metabolites: extracting the yeast metabolome from the literature. *Metabolomics, 7*, 94–101.

O'Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open babel: an open chemical toolbox. *Journal of Cheminformatics, 3*, 33.

O'Hagan, S., Dunn, W. B., Broadhurst, D., et al. (2007). Closed-loop, multi-objective optimisation of two-dimensional gas chromatography (GCxGC-tof-MS) for serum metabolomics. *Analytical Chemistry, 79*, 464–476.

O'Hagan, S., Dunn, W. B., Brown, M., Knowles, J. D., & Kell, D. B. (2005). Closed-loop, multiobjective optimisation of analytical instrumentation: gas-chromatography-time-of-flight mass spectrometry of the metabolomes of human serum and of yeast fermentations. *Analytical Chemistry, 77*, 290–303.

Orth, J. D., Thiele, I., & Palsson, B. Ø. (2010). What is flux balance analysis? *Nature Biotechnology, 28*, 245–248.

Palsson, B. Ø. (2006). *Systems biology: properties of reconstructed networks*. Cambridge: Cambridge University Press.

Park, J. M., Kim, T. Y., & Lee, S. Y. (2010). Prediction of metabolic fluxes by incorporating genomic context and flux-converging pattern analyses. *Proceedings of the National Academy of Sciences USA, 107*, 14931–14936.

Park, J. H., Lee, K. H., Kim, T. Y., & Lee, S. Y. (2007). Metabolic engineering of *Escherichia coli* for the production of L-valine based on transcriptome analysis and in silico gene knockout simulation. *Proceedings of the National Academy of Sciences USA, 104*, 7797–7802.

Raamsdonk, L. M., Teusink, B., Broadhurst, D., et al. (2001). A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nature Biotechnology, 19*, 45–50.

Ray, O., Whelan, K., & King, R. (2010). Automatic revision of metabolic networks through logical analysis of experimental data. *Inductive Logic Programming, 5989*, 194–201.

Reed, J. L., & Palsson, B. Ø. (2003). Thirteen years of building constraint-based in silico models of Escherichia coli. *Journal of Bacteriology, 185*, 2692–2699.

Sakurai, N., Ara, T., Kanaya, S., et al. (2013). An application of a relational database system for high-throughput prediction of elemental compositions from accurate mass values. *Bioinformatics, 29*, 290–291.

Salek, R.M., Haug, K., Conesa, P., et al. (2013). The MetaboLights repository: curation challenges in metabolomics. *Database*, 2013, bat029.

Sawada, Y., Nakabayashi, R., Yamada, Y., et al. (2012). RIKEN tandem mass spectral database (ReSpect) for phytochemicals: a plant-specific MS/MS-based data resource and database. *Phytochemistry, 82*, 38–45.

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

Sigurdsson, M. I., Jamshidi, N., Steingrimsson, E., Thiele, I., & Palsson, B. Ø. (2010). A detailed genome-wide reconstruction of mouse metabolism based on human Recon 1. *BMC Systems Biology, 4*, 140.

Smallbone, K., & Simeonidis, E. (2008). Flux balance analysis: a geometric perspective. *Journal of Theoretical Biology, 258*, 311–315.

Smallbone, K., Simeonidis, E., Broomhead, D. S., & Kell, D. B. (2007). Something from nothing: bridging the gap between constraint-based and kinetic modelling. *FEBS Journal, 274*, 5576–5585.

Smallbone, K., Simeonidis, E., Swainston, N., & Mendes, P. (2010). Towards a genome-scale kinetic model of cellular metabolism. *BMC Systems Biology, 4*, 6.

Steinbeck, C., Conesa, P., Haug, K., et al. (2012). MetaboLights: towards a new COSMOS of metabolomics data management. *Metabolomics, 8*, 757–760.

Steinbeck, C., Han, Y., Kuhn, S., Horlacher, O., Luttmann, E., & Willighagen, E. (2003). The Chemistry Development Kit (CDK): an open-source Java library for Chemo-and Bioinformatics. *Journal of Chemical Information and Computer Sciences, 43*, 493–500.

Swainston, N., & Mendes, P. (2009). libAnnotationSBML: a library for exploiting SBML annotations. *Bioinformatics, 25*, 2292–2293.

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.

Tautenhahn, R., Cho, K., Uritboonthai, W., Zhu, Z. J., Patti, G. J., & Siuzdak, G. (2012). An accelerated workflow for untargeted metabolomics using the METLIN database. *Nature Biotechnology, 30*, 826–828.

Thiele, I., Hyduke, D. R., Steeb, B., et al. (2011). A community effort towards a knowledge-base and mathematical model of the human pathogen *Salmonella typhimurium LT2. BMC Systems Biology, 5*, 8.

Thiele, I., & Palsson, B. Ø. (2010). Reconstruction annotation jamborees: a community approach to systems biology. *Molecular Systems Biology, 6*, 361.

Thiele, I., Swainston, N., Fleming, R. M. T., et al. (2013). A community-driven global reconstruction of human metabolism. *Nature Biotechnology, 31*, 419–425.

Uhlén, M., Oksvold, P., Fagerberg, L., et al. (2010). Towards a knowledge-based Human Protein Atlas. *Nature Biotechnology, 28*, 1248–1250.

van der Greef, J., & McBurney, R. N. (2005). Rescuing drug discovery: in vivo systems pathology and systems pharmacology. *Nature Reviews Drug Discovery, 4*, 961–967.

van Ravenzwaay, B., Herold, M., Kamp, H., et al. (2012). Metabolomics: a tool for early detection of toxicological effects and an opportunity for biology based grouping of chemicals-from QSAR to QBAR. *Mutation Research, 746*, 144–150.

Wang, Y., Eddy, J. A., & Price, N. D. (2012). Reconstruction of genome-scale metabolic models for 126 human tissues using mCADRE. *BMC Systems Biology, 6*, 153.

Wang, Z., Klipfell, E., Bennett, B. J., et al. (2011). Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature, 472*, 57–63.

Wikoff, W. R., Anfora, A. T., Liu, J., et al. (2009). Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proceedings of the National Academy of Sciences USA, 106*, 3698–3703.

Wishart, D. S. (2012). Chapter 3: small molecules and disease. *PLoS Computational Biology, 8*, e1002805.

Wishart, D. S., Jewison, T., Guo, A. C., et al. (2013). HMDB 3.0–The Human Metabolome Database in 2013. *Nucleic Acids Research, 41*, D801–D807.

Wishart, D. S., Tzur, D., Knox, C., et al. (2007). HMDB: the Human Metabolome Database. *Nucleic Acids Research, 35*, D521–D526.

Zelena, E., Dunn, W. B., Broadhurst, D., et al. (2009). Development of a robust and repeatable UPLC-MS method for the long-term metabolomic study of human serum. *Analytical Chemistry, 81*, 1357–1364.

Zhao, L., & Shen, J. (2010). Whole-body systems approaches for gut microbiota-targeted, preventive healthcare. *Journal of Biotechnology, 149*, 183–190.

Zhou, B., Wang, J., & Ressom, H. W. (2012). MetaboSearch: tool for mass-based metabolite identification using multiple databases. *PLoS One, 7*, e40096.