Genes Nutr (2008) 3:107–113
DOI 10.1007/s12263-008-0090-5

REVIEW

The challenges for molecular nutrition research 4: the "nutritional systems biology level"

Ben van Ommen · Duccio Cavallieri · Helen M. Roche · Ulla I. Klein · Hannelore Daniel

Received: 31 May 2008 / Accepted: 8 September 2008 / Published online: 30 September 2008

Abstract Nutritional systems biology may be defined as the ultimate goal of molecular nutrition research, where all relevant aspects of regulation of metabolism in health and disease states at all levels of its complexity are taken into account to describe the molecular physiology of nutritional processes. The complexity spans from intracellular to inter-organ dynamics, and involves iterations between mathematical modelling and analysis employing all profiling methods and other biological read-outs. On the basis of such dynamic models we should be enabled to better understand how the nutritional status and nutritional challenges affect human metabolism and health. Although the achievement of this proposition may currently sound unrealistic, many initiatives in theoretical biology and biomedical sciences work on parts of the solution. This review provides examples and some recommendations for the molecular nutrition research arena to move onto the systems level.

Keywords Systems biology of nutrition · Modelling · Networks · Challenges

All authors are member of the European Nutrigenomics Organisation (http://www.nugo.org).

B. van Ommen (✉)
Department of Biosciences, TNO-Quality of Life,
Zeist, The Netherlands
e-mail: ben.vanommen@tno.nl

D. Cavallieri
Department of Pharmacology, University of Firenze,
Florence, Italy
e-mail: duccio.cavalieri@unifi.it

H. M. Roche
UCD Conway Institute, University College Dublin,
Dublin, Ireland
e-mail: helen.roche@ucd.ie

U. I. Klein · H. Daniel
Molecular Nutrition Unit, Nutrition and Food Research Center,
Technische Universität München, Munich, Germany

Introduction

Let's start with a quote: "Perhaps surprisingly, a concise definition of systems biology that most of us can agree upon has yet to emerge" (Ruedi Aebersold, Ph.D., faculty member of the Institute for Systems Biology, Seattle, USA—http://www.systemsbiology.org). Since there is no unifying concept of what systems biology is, we may take the liberty here to define what we consider as Nutritional Systems Biology (NSB).

"NSB covers all approaches targeted towards understanding the key processes that define nutrition in the context of regulation of transcription, protein synthesis and turn-over, metabolism and genome stability in health and disease states at all levels of its complexity, to simulate these processes and to predict the outcome of any alteration (whether genetic or nutritional) of the system. NSB starts at the molecular and cellular levels and takes the regulatory processes onto the level of the different organs and the inter-organ flow of information, nutrients and metabolites to describe most comprehensively the organism response and the processes quantitatively in mathematical terms".

The "measure everything period"

In the initial euphoria of the functional genomics and nutrigenomics research, a rather simplified idea of

"describing the biological system by measuring everything" was applied. Investigations were stimulated by new technologies enabling with ease to quantify the effects of nutritional manoeuvres at the whole transcriptome and/or proteome levels. However, this "measuring everything" approach was taken without the appropriate tools for meaningful evaluation and most likely 99% of the generated data were left untouched. Moreover, unlike other biomedical, pharmacological and toxicological treatments, it was anticipated that any nutrient would produce a multitude of mainly minor changes in gene expressions, protein and metabolite concentrations and therefore one better should look for everything. We now realize that this type of inventory studies does not necessarily leads to a deeper understanding of the biological processes. For this it is inevitable to combine "omics" with firm functional assays and mechanistic studies as part of a comprehensive phenotyping.

What are the necessary measurements to move nutrigenomics research onto the systems level?

Identifying the cellular "gatekeepers"

At the *cellular* level, we need to identify the pathways by which the intracellular nutrient-sensitive targets synchronize the cellular capability to changes in homeostasis upon stimulation by extracellular signals such as insulin, glucagon, catecholamines and other hormones, cytokines or mediators. It becomes apparent that the cellular response to an external hormone or other signalling molecule almost always involves an intracellular sensing mechanism of the available energy or nutrient supply status. Well known in this respect is the ATP/AMP ratio for the activity of the AMP-kinase [19]. In a similar way the mammalian target of rapamycin (mTOR) is activated in response to high nutrient status and availability of free amino acids [17]. mTOR is a conserved serine–threonine protein kinase that serves as a central regulator for cell growth and protein synthesis by integrating signals from nutrients and growth factors. The combined extracellular and intracellular signals merge in common signalling pathways for the cells final decision on execution of alterations in cell functions (Fig. 1). Many reviews address the mechanisms underlying these signalling processes that control prime metabolic functions such as the use of different energy sources (glucose, fatty acids, amino acids) for ATP production, protein synthesis or stress response. These essential nutritional processes are now brought into focus of many researchers who are interested in diabetes and other metabolic-stress related disorders. The incidence of these diseases increases dramatically and urgently requires new medical treatments and therefore these nutritional processes have also moved into the centre of pharma research. Recent genome wide screens have been performed to identify and annotate the downstream gene targets for example of the peroxisomal proliferating-activating receptors (PPARs) with natural as well as pharmacological ligands (for example the thiazolidinediones). It can be expected that there will be a growing number of these studies as well as corresponding databases with these "gatekeepers" of metabolic control. The prime focus will be in particular the nuclear receptors that are directly activated by nutrients or their derivatives like the retinoid X receptor (RXR)/retinoid acid receptor (RAR), the liver X receptor (LXR), vitamin D receptor (VDR), PPAR's and others and other transcription factors that are downstream targets of extracellular signals like the forkhead box class O factors (FoxO). The small subfamily of FoxO transcription factors plays an important role in longevity and tumor suppression and integrate a variety of environmental signals including insulin, growth factors, nutrients and oxidative stress [3, 8]. The next step includes mainly the screening of existing knowledge bases to extract the relevant information from nutrient-sensitive gene expression analysis studies.

Changes in gene expression, however, are per se not predictive enough to account for all metabolic changes. The complexity of biological organization and the multitude of interactions between the cellular proteins need to be assessed as well. The latter are the basis of interactome models, describing the static relationship between

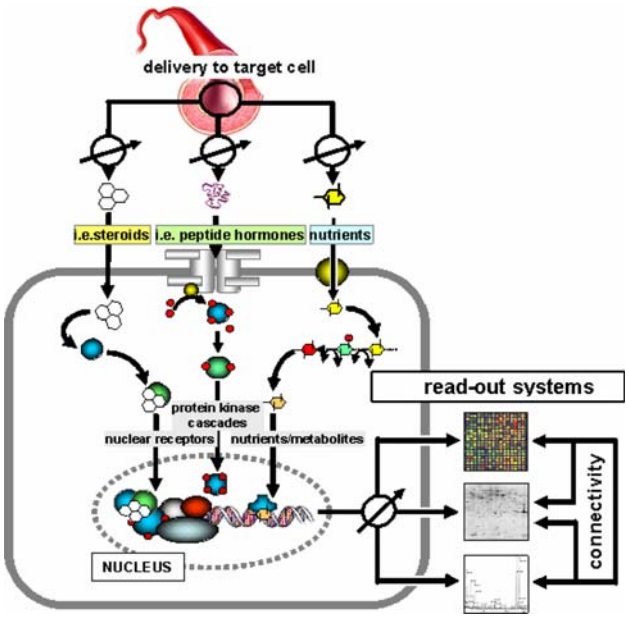

Fig. 1 The gatekeeper concept that involves intracellular synchronisation of external hormone signals with the cellular status that translate into changes in the transcriptome, proteome and metabolome

molecular components, and protein–protein interaction maps have been derived for bacteria, yeast or *C. elegans* [26]. Despite the fact that protein-protein interactions are known to be essential for execution of cellular functions, they are usually not considered to be affected by nutrition. However, knowing the role for example of trace elements (Zn, Cu, Iron, Se, etc.) as protein structure determinants or the reaction of pyridoxalphosphate and other vitamins with protein entities, there is a layer of biological interactions that could involve numerous diet-dependent alterations that have been almost completely neglected so far. Although proteome analysis as performed in nutrition research has yielded interesting findings in view of biomarker discovery or even with respect to post-translational modifications, the current methods available are far from the state were a comprehensive picture of the proteome of a cell can be obtained. The same holds true for the cellular metabolome. Regardless of the recent technological advances and achievements in mass spectrometry and nuclear magnetic resonance (NMR) techniques for identifying and quantifying hundreds of metabolites simultaneously, they all cover only those intermediates with higher concentrations in the biological sample. Nevertheless, even with this limited number of metabolites, it is a challenge to bring the changes in metabolic profiles back into the cellular biochemistry and to the pathways in which they are generated or used. In particular the research on bacteria, yeast and plants is currently most advanced concerning metabolome analysis in the context of cell biology. The approach was even extended by incorporation of radiotracer or stable isotopes to describe the fluxes of nutrients and their metabolites in distinct metabolic pathways in time and space [20, 25]. Despite the fact that nutritional science has a long tradition in employing tracers to characterize for example protein synthesis, amino acid oxidation rates or lipid metabolisms, flux studies have not yet been incorporated into systemic approaches in the nutrigenomics era.

Modelling nutritional processes at the cellular level

It is the intrinsic concept of systems biology that the biological information is compiled and a model is constructed. The model is then used for predictions on the cell behaviour and experimentally validated with quantitative measures of the cells response and those data are used to revise the model in an iterative approach (Fig. 2).

Good models need good experts to filter the available knowledge. Despite the wonderful tools for example for automatic text mining, expert interpretation and annotation are essential elements for model generation. Model construction typically involves the translation of the knowledge base into reactants and reactions and the transformation into a network diagram with a set of linked differential equations [1]. There are numerous commercial and non-commercial software tools available meanwhile that based on graphical approaches allow to draw gene-regulatory and biochemical networks [2] with graphical notation systems as proposed by Kitano [14]. It needs to be emphasized that the generation of a predictive model is based on a precise data model [18]. The data model is a descriptor that annotates all the possible interactions and in a way represents a language that the scientists use to describe the different components of the system in a computationally amenable form and in a widely accepted format. The most relevant language is currently the systems biology markup language (SBML; http://www.sbml.org), a standard for representing models of biochemical and gene-regulatory networks [11]. Other languages mainly used for annotations of reactions are the web ontology language (OWL) and the biological pathways exchanging language (Biopax; [24]. A free database containing numerous SBML models is available [16].

Model verification and validation is essential to finally prove that the equations comprising a model predict correctly the outcome. When mechanisms are complex and the data sets used to generate the model rely on input from different species, the models are prone to errors. Currently the only way to verify the validity of the model is to measure and correct to get the fit to data as close as possible. The remaining challenge is to experimentally collect enough high quality data to effectively redesign the model. Most biological pathway models are currently poorly calibrated and validated. One of the recently introduced concepts in model generation is the robustness [13]. This means from a systems perspective that any living cell is fairly robust in its reaction to a perturbation of the environment whether physical or nutritional. The fundamentals of this robustness are complex feedback control loops, the functional redundancy in many gene products, the modularity of the systems and the structural stability of most of the biological entities. It is essential to build in these robust features into the models that describe nutritional processes in mathematical terms when nutritional systems biology is the goal. In nutrition—which means the supply (or lack) of energy substrates and nutrient substrates that varies in time and space like no other environmental factor—robustness is intrinsically a crucial determinant in metabolic control. The robustness has conceptually been discussed in the context of the metabolic syndrome by Kitano [15] with a map for the interaction of adipocytes, hepatocytes, skeletal muscle cells and pancreatic beta-cells in the metabolic syndrome.

How can we promote systems approaches in nutrition research at the cellular level and what steps need to be taken?

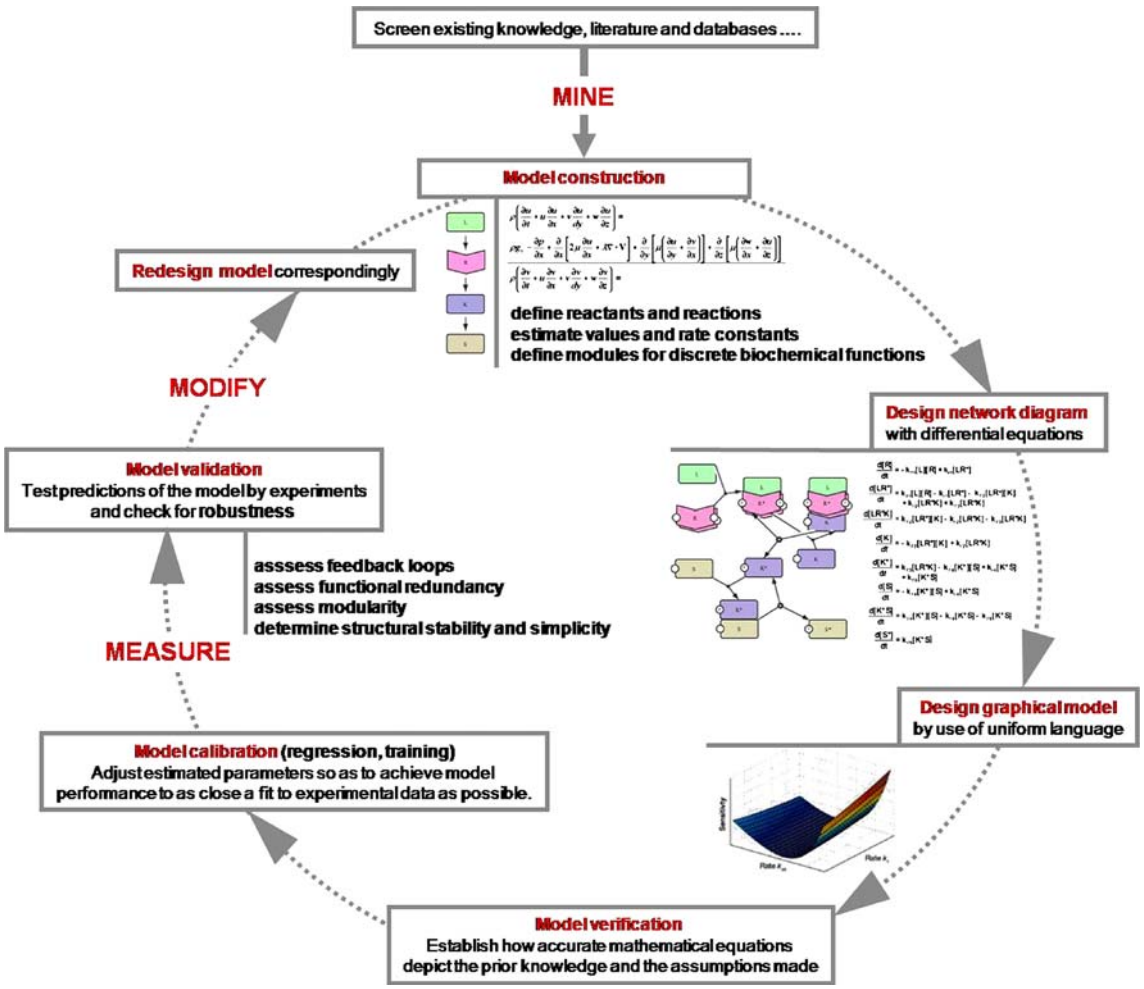

Fig. 2 The cycle of mining, modelling, measuring and remodelling in a system biology approach (modified according to the concept of [1]

- We need as a community an agreement to adapt a uniform language (like SBML) and converters to migrate existing information into this language This can easily be achieved as standards are ready to be absorbed from other science communities. The nutrigenomics community should play an active part in standardization and integration of life science databases which are a prerequisite for any successful application of e-Science in nutrigenomics research.
- We need to create and collect case examples of nutritional systems biology models. Because no model will currently serve all needs and as many parameters will be lacking, a pragmatic approach is needed. In constructing the framework of such a model as a collaborative effort missing parameters as collected from cell, animal or human studies and bioinformatics simulations/fittings can be integrated any time for adjusting the models.
- We need to establish a "nutritional interactome project" as a show case, describing all essential parts in a relevant area of nutrition such as energy metabolism from the macroscopic level of inter-organ flow of energy substrates to mitochondrial respiration, or the "micronutrient interactome", describing all processes and mechanisms related to micronutrient function and their role in the biological response. The nutritional interactome would (among others) be the compilation of all components that show up to be relevant for interpretation of "omics" data as described in the first paragraph.
- Eventually, major metabolic and regulatory processes linked to nutritionally relevant disorders (like the metabolic syndrome) may be approached from a modelling perspective, where all sub-processes can be described individually but are connected to the overarching (disease) phenotype, and thus become a "top-down" modelling approach. This type of approach is commercially developed and exploited [12]. Open source academic initiatives should be initiated but need a strict coordination and major funding. A framework

architecture where the various models could be inserted, may be constructed where the various nutritional models may be interlinked with other depositories like the biomodels database [16].

Moving to the next layer of complexity

The ultimate objective for nutritional systems biology is to understand the whole organism rather than a single cell type. At the *organ* and *inter-organ level*, we need to define and quantify the relative contribution of the different cell types in an organ to the overall organ function and, on the next layer, to assess the different organs contribution to the metabolic adaptations found at the organism level. This should be optimized first in animal studies, wherein organs are accessible for detailed analysis of the changes in the transcriptome, proteome and metabolome. Furthermore animal models tend to have less inherent variability than observed between human subjects. Only a few examples are found so far in literature in which the different "omics" technologies are used to define the metabolic status of a whole organ. Most recently, valuable data on mice liver from a cross breading of a diabetes-resistant and a diabetes-susceptible strain were derived by simultaneously determining about 80 metabolites as well as transcript profiles that could also be linked to quantitative trait loci [6]. Such an "organ centric" profiling approach to a nutritional challenge, such as a high-fat diet, is also embedded into the currently ongoing "proof of principle study" by the European Nutrigenomics Organisation (NuGO) (see http://www.nugo.org).

In human studies access to organs or tissue samples is highly limited (except for easily accessible cells such as peripheral blood mononuclear cells) and therefore we will rely almost completely on biofluids to assess the nutritional status or on imaging techniques such as magnetic resonance imaging (MRI) or NMR for obtaining non-invasively information on organ composition. With respect to biofluids as a source of information on the metabolic state we need surrogate markers which can only be derived from well designed human investigations in which the nutritional environment is controlled. Moreover, animal studies with embedded biofluid profiling are essential to accurately assess organ specific responses under well defined feeding regimens. This should provide a better definition of the robustness of surrogate markers in plasma or urine.

However, it is more than feasible to assume that nutritional "omics" studies will lead to the identification of various subsets of biological processes modulated by diet and relevant to nutrition-dependent diseases. These processes, most of them involving all of the above levels (cell, organ and organism) need to be characterized in terms of their normal physiology and of those aberrations that may lead to or finally characterize a disorder. One of the challenges which nutrition research faces is how to describe early disease onset, in order to design nutritional prevention strategies. These early changes will be subtle and likely to be hidden in the day-to-day fluctuations of maintaining homeostasis in a constantly changing environment. Also, a large number of possible processes and parameters may initially be involved in reaching the same (disease) endpoint, and this is subject to strong inter-individual variation. We need longitudinal studies with multiple observations and analyses along the time-course of disorder development (or health improvement). This issue strongly relates to the "nutritional phenotype" topic. Bioinformatics approaches need to be developed that can handle and interpret longitudinal "omics" data. Statistical approaches (correlational networks, parallel factor analysis, cluster analysis, etc.) need to be tailored towards exploiting the time-line of events. The importance of analysing the time-dependence of gene expression changes to define the role of "gatekeepers" is exemplified by studies in murine lymphocytes [5]. In this large-scale array-based gene expression analysis transcript data were aggregated into large gene groups with related behaviour (megamodules) and time-dependent changes were characterized. The authors observed for example a medium-term critical global transcriptional dependence on ATP-generating genes in mitochondria and a long-term dependence glycolytic genes. Gene expression profiles of circulating leukocytes in response to bacterial endotoxin infusion have also provided the basis for a network-based analysis of systemic inflammation in humans [4]. In the end, input from biostatistics and bioinformatics is needed for the best design and evaluation of time-series analysis in nutritional studies.

How can we improve systems approaches in nutrition research when focusing on human studies and what steps need to be taken?

- We need to create biological trajectories of the transition from optimal metabolic function through impaired homeostasis to dysfunction for nutrition relevant processes. This will also help to define the robustness of the biological systems and should promote the markers associated with the early onset of diet-related diseases.
- We need to define how "overarching processes" like the nutritional status (postprandial vs. fasting vs. starvation) or certain states of metabolic and/or oxidative stress and inflammation interact over time and which markers sets best describe the perturbation.
- We need to assess to which extent early life nutrition (intra-uterine and post-partum) determines the time

course of nutrition related disorders in a humans life span addressing metabolic programming and epigenomics as general principals affecting any systems approach.

Whatever type of studies in NSB are planned or done in the future it seems most important that we agree upon a set of minimal requirements for study design, analytical procedures, data storage and sharing. This exercise will result in the definition of a "minimal parameter set" to be reported and quantified for a given type of dietary intervention study. This minimal parameter set will have a relatively high density of regulatory elements, receptors, signal transduction pathways, rate determining metabolic enzymes etc., all related to nutritional processes. Finally, a data warehouse with knowledge system allowing the query of these studies (results coupled to their study description) will be useful, as this will allow optimal comparison and integration of the acquired results. Such a system would facilitate for example the extraction of all gene expression changes in liver related to PPAR-alpha modulation by a high fat diet, combined with the changes in inflammation related proteins in plasma. To turn this concept into reality, detailed planning of the data-capture strategy is essential so that sufficient flexibility and power is available for the query options. Data storage standards have been developed or are in development for all of the "omic" technologies. However, capturing all of the relevant associated study details remains a significant challenge, particularly in nutrition studies using various cell systems, animal models and human subjects. Also, where studies move beyond the examination of single compounds to whole foods and diets, data capturing of the relevant details becomes progressively more complex. The NutriBase concept put forward by NuGO forms the basis for this database effort.

Conclusion

Nutritional systems biology is in its infancy and only a very few examples exist for the description of nutritional processes on the basis of models and in mathematical terms. These examples include the kinetic analysis and even multi-compartment models of lipoprotein metabolism [7, 21], or even computational models to analyse human energy metabolism during semi-starvation and in cancer cachexia [9, 10]. Examples to describe mammalian nutrient metabolism in quantitative kinetic models have been put forward for folate and most recently for glutathione [22, 23]. Despite the lack of examples it seems that nutritional science is ready to move forward to the systems approaches. We have learned to handle the "omics" technologies and now it is time to go back to where we started and this is metabolism. We all had to learn the biochemical pathways and they are the principles of metabolic adaptation. Good nutritionists should turn into good systems experts and start modelling the processes. Modelling in nutrition research as well as tracer techniques employing stable isotopes are well established and usually describe compartments and kinetics of distinct metabolic processes—however so far more at the macroscopic scale. With the new generation of nutrition researchers that now work on the molecular and cellular level, that study gene expression, protein functions and regulatory processes in all details, the road is paved to move on to the systems level.

References

1. Aldridge BB, Burke JM, Lauffenburger DA, Sorger PK (2006) Physicochemical modelling of cell signalling pathways. Nat Cell Biol 8:1195–1203
2. Alves R, Antunes F, Salvador A (2006) Tools for kinetic modeling of biochemical networks. Nat Biotechnol 24:667–672
3. Calnan DR, Brunet A (2008) The FoxO code. Oncogene 27:2276–2288
4. Calvano SE, Xiao W, Richards DR et al (2005) A network-based analysis of systemic inflammation in humans. Nature 437:1032–1037
5. de Bivort B, Huang S, Bar-Yam Y (2004) Dynamics of cellular level function and regulation derived from murine expression array data. Proc Natl Acad Sci USA 101:17687–17692
6. Ferrara CT, Wang P, Neto EC et al (2008) Genetic networks of liver metabolism revealed by integration of metabolic and transcriptional profiling. PLoS Genet 4:e1000034
7. Gill JM, Brown JC, Caslake MJ et al (2003) Effects of dietary monounsaturated fatty acids on lipoprotein concentrations, compositions, and subfraction distributions and on VLDL apolipoprotein B kinetics: dose-dependent effects on LDL. Am J Clin Nutr 78:47–56
8. Gross DN, van den Heuvel AP, Birnbaum MJ (2008) The role of FoxO in the regulation of metabolism. Oncogene 27:2320–2336
9. Hall KD (2006) Computational model of in vivo human energy metabolism during semistarvation and refeeding. Am J Physiol Endocrinol Metab 291:E23–E37
10. Hall KD, Baracos VE (2008) Computational modeling of cancer cachexia. Curr Opin Clin Nutr Metab Care 11:214–221
11. Hucka M, Finney A, Sauro HM et al (2003) The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. Bioinformatics 19:524–531
12. Kansal AR (2004) Modeling approaches to type 2 diabetes. Diabetes Technol Ther 6:39–47
13. Kitano H (2007) Towards a theory of biological robustness. Mol Syst Biol 3:137
14. Kitano H, Funahashi A, Matsuoka Y, Oda K (2005) Using process diagrams for the graphical representation of biological networks. Nat Biotechnol 23:961–966
15. Kitano H, Oda K, Kimura T et al (2004) Metabolic syndrome and robustness tradeoffs. Diabetes 53(Suppl 3):S6–S15
16. Le Novere N, Bornstein B, Broicher A et al (2006) BioModels Database: a free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. Nucleic Acids Res 34:D689–D691

17. Lian J, Yan XH, Peng J, Jiang SW (2008) The mammalian target of rapamycin pathway and its role in molecular nutrition regulation. Mol Nutr Food Res 52:393–399
18. Luciano JS, Stevens RD (2007) e-Science and biological pathway semantics. BMC Bioinformatics 8(Suppl 3):S3
19. McGee SL, Hargreaves M (2008) AMPK and transcriptional regulation. Front Biosci 13:3022–3033
20. Miccheli A, Tomassini A, Puccetti C et al (2006) Metabolic profiling by 13C-NMR spectroscopy: [1, 2–13C2] glucose reveals a heterogeneous metabolism in human leukemia T cells. Biochimie 88:437–448
21. Nguyen MN, Chan DC, Dwyer KP et al (2006) Use of Intralipid for kinetic analysis of HDL apoC-III: evidence for a homogeneous kinetic pool of apoC-III in plasma. J Lipid Res 47:1274–1280
22. Nijhout HF, Reed MC, Budu P, Ulrich CM (2004) A mathematical model of the folate cycle: new insights into folate homeostasis. J Biol Chem 279:55008–55016
23. Reed MC, Thomas RL, Pavisic J et al (2008) A mathematical model of glutathione metabolism. Theor Biol Med Model 5:8
24. Strömbäck L, Lambrix P (2005) Representations of molecular pathways: an evaluation of SBML, PSI MI and BioPAX. Bioinformatics 21:4401–4407
25. Zamboni N, Kümmel A, Heinemann M (2008) anNET: a tool for network-embedded thermodynamic analysis of quantitative metabolome data. BMC Bioinformatics 9:199
26. Zhu X, Gerstein M, Snyder M (2007) Getting connected: analysis and principles of biological networks. Genes Dev 21:1010–1024