

## REVIEWS

### Box 3 | Physiological modelling tools

Physiological modelling involves spatiotemporal systems being represented by partial differential equations (PDEs); solving PDEs is more involved than for ordinary differential equations (ODEs). In most cases, PDEs are solved by the Finite Element Method (FEM), a numerical technique that provides approximate solutions for PDEs. FEM is supported in tools like ANSYS, FreeFEM++ and OpenFEM. MATLAB can be also used to solve PDEs with the PDE Toolbox.

Several simulation software tools can be used for physiological modelling. These include: JSim, a Java-based simulation system that can solve first-degree, one-dimensional PDEs; OpenCell, an environment for working with CellML; Virtual Physiology (produced by cLabs), a set of tissue-specific simulators, such as for neurons, skin senses and muscle, that can also be used as educational tools; and Flexible Large-scale Agent-based Modelling Environment (FLAME), an agent-based modelling and simulation framework.

Many simulation tools are also being developed for specific areas. These include: STEPS, a stochastic engine for pathway simulation that is used in molecular modelling; GENESIS, a simulation platform for multiple levels of neural systems, from subcellular biochemical reactions to networks of neurons; Neuron, a simulation environment for computational models of neurons and networks of neurons; SimHeart and Heart Simulator, which are used in cardiology; and AnyBody, a full-body musculoskeletal simulation tool.

Although many standards have been defined for data and model representations, they only ensure that data and models that comply with these standards can be used by software that support these standards; they do not ensure that multiple software tools can be used seamlessly<sup>100</sup>. When software tools are developed by independent research groups or companies without an explicit agreement as to how they can be integrated, this can cause problems when forming a workflow of multiple tools. This is because the tools are likely to be inconsistent in their operating procedures and their use of various non-standardized data formats. Thus, users often have to convert data formats, to learn operating procedures for each tool, and sometimes even to adjust operating environments. This impedes productivity, undermines the flexibility of the workflow and is prone to errors.

As an example workflow, a researcher working on modelling an oncogenic MAPK pathway may wish to quickly access, by one click, the sequences of genes that are involved in this pathway to see the mutations that are associated with a specific subgroup of cancer patients. They might then search a protein structure database for the three-dimensional structures of proteins that are encoded by these mutated genes to see how the mutations might affect the three-dimensional structures. Next, they might explore possible docking interactions of candidate kinase inhibitors (using virtual docking simulations). Then, using advanced text-mining tools, the researcher could search for experimental and clinical articles that have reported possible effects of the compound and similar compounds on the cell line of interest or a cell line with similar mutations. Finally, the researcher could modify the original model to incorporate possible differences in networks owing to the mutations and could run dynamical simulations to see the effects on the cellular responses to specific compounds. Currently, this workflow requires multiple separate software tools, and there is no transfer of retrieved information among software tools. A successful, integrated

software tool would enable such a workflow in a few clicks, so that users could concentrate on science rather than on software operation.

Recently, several initiatives have been launched to move towards software integration. The US Department of Energy is initiating the Systems Biology Knowledgebase project for building an integrated environment for data, knowledge and tools as part of their Genomes to Life programme. Another example is the Garuda Alliance, which was formed with the aim of creating a platform and a set of guidelines to achieve a highly productive and flexible software and data environment; that is, a one-stop service for systems biology and bioinformatics. The aim is to have a high level of interoperability among software in a language-agnostic manner, to provide consistent user experiences and to offer a broader accessibility of tools and resources. To achieve these objectives, the Garuda Core will provide defined and comprehensive APIs, a wide range of program and widget parts, and a series of design guidelines. Developers of tools will be able to use the provided APIs to make their own tools operational through the Garuda Core. Garuda-compliant software would need to adopt user-interface guidelines so that researchers could use a range of tools without the need for additional learning. Initially, software — such as CellDesigner, the Panther pathway database<sup>101</sup>, bioCompendium, PathText<sup>102</sup>, the Edinburgh SBSI tools and others — will be provided as Garuda-compliant software. The intention is to host increasing numbers of software and data or knowledge resources. Achieving a smooth workflow is still a long way off, but these efforts are certainly the first step.

### A vision for the future

Creating and making the best use of software and data resources will facilitate an in-depth understanding of biological systems. However, the impact of creating a widely accepted software platform may go far beyond productivity improvements in each research group because the platform could potentially connect research groups globally. Although international collaboration in scientific projects is common, determining how best to create a successful open collaboration is still a challenge. For example, creating and maintaining a comprehensive and in-depth model of a biological system is often beyond the scope of any single research group. Maintaining, updating and improving pathway databases — such as Reactome, KEGG, and the Panther pathway database — requires continuous funding. Also, such databases are not sufficiently in-depth for many complex pathways, especially when compared with interaction maps that have been developed by a few dedicated researchers who are focused on specific pathways<sup>66</sup>.

Some alternative approaches have been proposed that use Web2.0 services, as Wikipedia does. There are several such attempts, including Wikipathways<sup>74</sup>, Wikigenes<sup>103</sup> and Gene wiki<sup>104</sup>. However, many of these efforts are struggling<sup>105</sup>. One possible reason is the lack of incentive for scientists to contribute their knowledge and data. Why would somebody spend time to share their knowledge when such a contribution is not properly



acknowledged<sup>106</sup>? Although there are discussions of schemes to systematically acknowledge such efforts, it is yet to be seen whether these schemes can change social dynamics and hence the motivations of potential contributors. There may be a great opportunity to enhance our scientific productivity when a 'network of intelligence' or 'wisdom of crowds'<sup>107</sup> approach can become a reality because everyone could gain from the ideas and experiences of others. However, we do not know yet how best to achieve this in reality. Web2.0 approaches are often suggested by computer science-based researchers because of the success of such approaches in their field. However, there are cultural differences in biological research, and overcoming these differences may be a substantial challenge and may also require the involvement of a broader range of experts, such as sociologists and psychologists.

Our vision is that the increased capability to navigate and relate various data and knowledge resources using integrated platforms would enable researchers to enjoy a higher level of productivity and a greater potential for innovation. Connecting genomics, molecular networks and physiology will provide us with a deeper understanding of how individual differences in the genome affect physiological processes through alterations in molecular networks. The current reality is that there are various software tools that can be used for a broad range of systems biology research, and these tools are being increasingly integrated owing to standardization and alliance efforts. Emerging comprehensive, consistent and community-wide software platforms enable us to promote systems biology research today, and also to think about what comes next.

- Kitano, H. Systems biology: a brief overview. *Science* **295**, 1662–1664 (2002).
- Kitano, H. Computational systems biology. *Nature* **420**, 206–210 (2002).
- Kitano, H. Perspectives on systems biology. *New Generation Computing* **18**, 199–216 (2000).
- Ideker, T., Galitski, T. & Hood, L. A new approach to decoding life: systems biology. *Annu. Rev. Genomics Hum. Genet.* **2**, 343–372 (2001).
- Tyson, J. J., Chen, K. & Novak, B. Network dynamics and cell physiology. *Nature Rev. Mol. Cell Biol.* **2**, 908–916 (2001).
- Novak, B. & Tyson, J. J. Numerical analysis of a comprehensive model of M-phase control in *Xenopus* oocyte extracts and intact embryos. *J. Cell Sci.* **106**, 1153–1168 (1993).
- Chen, K. C. *et al.* Integrative analysis of cell cycle control in budding yeast. *Mol. Biol. Cell* **15**, 3841–3862 (2004).  
**A pioneering study using computational modelling and analysis of the budding yeast cell cycle. The model computationally reproduced the phenotypes of various gene deletion mutants.**
- Aoki, K., Yamada, M., Kunida, K., Yasuda, S. & Matsuda, M. Processive phosphorylation of ERK MAP kinase in mammalian cells. *Proc. Natl Acad. Sci. USA* **108**, 12675–12680 (2011).
- Schoeberl, B. *et al.* An ErbB3 antibody, MM-121, is active in cancers with ligand-dependent activation. *Cancer Res.* **70**, 2485–2494 (2010).
- Schoeberl, B. *et al.* Therapeutically targeting ErbB3: a key node in ligand-induced activation of the ErbB receptor-PI3K axis. *Sci. Signal.* **2**, ra51 (2009).
- Evans, D., Hagiu, A. & Schmalensee, R. *Invisible Engines: How Software Platforms Drive Innovation and Transform Industries*. (MIT Press, 2006).  
**An easy-to-read introduction to the concept of software platforms in industries.**
- Lee, T. L. Big data: open-source format needed to aid wiki collaboration. *Nature* **455**, 461 (2008).
- Brown, F. Saving big pharma from drowning in the data pool. *Drug Discov. Today* **11**, 1043–1045 (2006).
- Kröger, P. & Bry, F. A computational biology database digest: data, data analysis, and data management. *Distributed and Parallel Databases* **13**, 7–42 (2003).
- Field, D., Tiwari, B. & Snape, J. Bioinformatics and data management support for environmental genomics. *PLoS Biol.* **3**, e297 (2005).
- Keator, D. B. Management of information in distributed biomedical laboratories. *Methods Mol. Biol.* **569**, 1–23 (2009).
- Van Deun, K., Smilde, A. K., van der Werf, M. J., Kiers, H. A. & Van Mechelen, I. A structured overview of simultaneous component based data integration. *BMC Bioinformatics* **10**, 246 (2009).
- Brazma, A., Krestyaninova, M. & Sarkans, U. Standards for systems biology. *Nature Rev. Genet.* **7**, 593–605 (2006).
- Brazma, A. *et al.* Minimum information about a microarray experiment (MIAME)—toward standards for microarray data. *Nature Genet.* **29**, 365–371, (2001).
- Taylor, C. F. *et al.* The minimum information about a proteomics experiment (MIAPE). *Nature Biotech.* **25**, 887–893 (2007).
- Martens, L., Palazzi, L. M. & Hermjakob, H. Data standards and controlled vocabularies for proteomics. *Methods Mol. Biol.* **484**, 279–286 (2008).
- Taylor, C. F. *et al.* Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. *Nature Biotech.* **26**, 889–896 (2008).
- Saltz, J. *et al.* caGrid: design and implementation of the core architecture of the cancer biomedical informatics grid. *Bioinformatics* **22**, 1910–1916 (2006).
- Oinn, T. *et al.* Taverna: a tool for the composition and enactment of bioinformatics workflows. *Bioinformatics* **20**, 3045–3054 (2004).
- Lee, S., Wang, T. D., Hashmic, N. & Cummings, M. P. Bio-STEER: A semantic Web workflow tool for Grid computing in the life sciences. *Future Generation Computer Systems* **23**, 497–509 (2007).
- Giardine, B. *et al.* Galaxy: a platform for interactive large-scale genome analysis. *Genome Res.* **15**, 1451–1455 (2005).
- Schadt, E. E., Friend, S. H. & Shaywitz, D. A. A network view of disease and compound screening. *Nature Rev. Drug Discov.* **8**, 286–295 (2009).
- van't Veer, L. J. *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530–536 (2002).
- Altschuler, D., Daly, M. J. & Lander, E. S. Genetic mapping in human disease. *Science* **322**, 881–888 (2008).
- Dewan, A. *et al.* *HTRA1* promoter polymorphism in wet age-related macular degeneration. *Science* **314**, 989–992 (2006).
- Yang, Z. *et al.* A variant of the *HTRA1* gene increases susceptibility to age-related macular degeneration. *Science* **314**, 992–993 (2006).
- Chester, E. J. *et al.* Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. *Nature Genet.* **37**, 233–242 (2005).
- Monks, S. A. *et al.* Genetic inheritance of gene expression in human cell lines. *Am. J. Hum. Genet.* **75**, 1094–1105 (2004).
- Morley, M. *et al.* Genetic analysis of genome-wide variation in human gene expression. *Nature* **430**, 743–747 (2004).
- Zhu, J. *et al.* An integrative genomics approach to the reconstruction of gene networks in segregating populations. *Cytogenet. Genome Res.* **105**, 363–374 (2004).
- Zhu, J. *et al.* Increasing the power to detect causal associations by combining genotypic and expression data in segregating populations. *PLoS Comput. Biol.* **3**, e69 (2007).
- Zhu, J. *et al.* Integrating large-scale functional genomic data to dissect the complexity of yeast regulatory networks. *Nature Genet.* **40**, 854–861 (2008).
- Margolin, A. A. *et al.* ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinformatics* **7** (Suppl. 1), S7 (2006).
- Faith, J. J. *et al.* Large-scale mapping and validation of *Escherichia coli* transcriptional regulation from a compendium of expression profiles. *PLoS Biol.* **5**, e8 (2007).
- Shen-Orr, S. S., Milo, R., Mangan, S. & Alon, U. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature Genet.* **31**, 64–68 (2002).
- Alon, U. Network motifs: theory and experimental approaches. *Nature Rev. Genet.* **8**, 450–461 (2007).
- Fadda, A. *et al.* Inferring the transcriptional network of *Bacillus subtilis*. *Mol. Biosyst.* **5**, 1840–1852 (2009).
- Cho, B. K. *et al.* The transcription unit architecture of the *Escherichia coli* genome. *Nature Biotech.* **27**, 1043–1049 (2009).
- Mendoza-Vargas, A. *et al.* Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. *PLoS ONE* **4**, e7526 (2009).
- Lemmens, K. *et al.* DISTILLER: a data integration framework to reveal condition dependency of complex regulons in *Escherichia coli*. *Genome Biol.* **10**, R27 (2009).
- De Smet, R. & Marchal, K. Advantages and limitations of current network inference methods. *Nature Rev. Microbiol.* **8**, 717–729 (2010).
- Ferrucci, D. *et al.* Building Watson: an overview of the DeepQA project. *AI Magazine* **31**, 3 (2010).
- Oda, K. & Kitano, H. A comprehensive map of the toll-like receptor signaling network. *Mol. Syst. Biol.* **2**, 2006.0015 (2006).
- Oda, K., Matsuoka, Y., Funahashi, A. & Kitano, H. A comprehensive pathway map of epidermal growth factor receptor signaling. *Mol. Syst. Biol.* **1**, 2005.0010 (2005).
- Caron, E. *et al.* A comprehensive map of the mTOR signaling network. *Mol. Syst. Biol.* **6**, 453 (2010).
- Kaizu, K. *et al.* A comprehensive molecular interaction map of the budding yeast cell cycle. *Mol. Syst. Biol.* **6**, 415 (2010).
- Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).
- Joshi-Tope, G. *et al.* Reactome: a knowledgebase of biological pathways. *Nucleic Acids Res.* **33**, D428–D432 (2005).
- Mi, H. *et al.* The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Res.* **33**, D284–D288 (2005).
- Cerami, E. G. *et al.* Pathway Commons, a web resource for biological pathway data. *Nucleic Acids Res.* **39**, D685–D690 (2011).
- Karp, P. D. *et al.* Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. *Nucleic Acids Res.* **33**, 6083–6089 (2005).
- Hucka, M. *et al.* The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* **19**, 524–531 (2003).  
**An original paper on SBML that triggered various standardization efforts in systems biology.**
- Demir, E. *et al.* The BioPAX community standard for pathway data sharing. *Nature Biotech.* **28**, 935–942 (2010).
- Le Novère, N. *et al.* The Systems Biology Graphical Notation. *Nature Biotech.* **27**, 735–741 (2009).
- Le Novère, N. *et al.* Minimum information requested in the annotation of biochemical models (MIRIAM). *Nature Biotech.* **23**, 1509–1515 (2005).

# REVIEWS

61. Kitano, H., Funahashi, A., Matsuoka, Y. & Oda, K. Using process diagrams for the graphical representation of biological networks. *Nature Biotech.* **23**, 961–966 (2005).
62. Klipp, E., Liebermeister, W., Helbig, A., Kowald, A. & Schaber, J. Systems biology standards — the community speaks. *Nature Biotech.* **25**, 390–391 (2007).
63. Sauro, H. M. *et al.* Next generation simulation tools: the Systems Biology Workbench and BioSPICE integration. *OMICS* **7**, 355–372 (2003).
64. van Iersel, M. P. *et al.* Presenting and exploring biological pathways with PathVisio. *BMC Bioinformatics* **9**, 399 (2008).
65. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
66. Bauer-Mehren, A., Furlong, L. I. & Sanz, F. Pathway databases and tools for their exploitation: benefits, current limitations and challenges. *Mol. Syst. Biol.* **5**, 290 (2009).
67. Calzone, L., Gelay, A., Zinovyev, A., Radvanyi, F. & Barillot, E. A comprehensive modular map of molecular interactions in RB/E2F pathway. *Mol. Syst. Biol.* **4**, 173 (2008).
68. Thiele, I. & Palsson, B. O. Reconstruction annotation jamborees: a community approach to systems biology. *Mol. Syst. Biol.* **6**, 361 (2010).  
**This paper discusses issues regarding community efforts to reconstruct comprehensive metabolic networks.**
69. Thiele, I. & Palsson, B. O. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* **5**, 93–121 (2010).
70. Feist, A. M., Herrgard, M. J., Thiele, I., Reed, J. L. & Palsson, B. O. Reconstruction of biochemical networks in microorganisms. *Nature Rev. Microbiol.* **7**, 129–143 (2009).  
**A review on the current state-of-the-art in data-driven genome-wide network reconstruction.**
71. Herrgard, M. J. *et al.* A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nature Biotech.* **26**, 1155–1160 (2008).
72. Wu, G., Zhu, L., Dent, J. E. & Nardini, C. A comprehensive molecular interaction map for rheumatoid arthritis. *PLoS ONE* **5**, e10137 (2010).
73. Matsuoka, Y., Ghosh, S., Kikuchi, N. & Kitano, H. Payao: a community platform for SBML pathway model curation. *Bioinformatics* **26**, 1381–1383 (2010).
74. Pico, A. R. *et al.* WikiPathways: pathway editing for the people. *PLoS Biol.* **6**, e184 (2008).
75. Wierling, C., Herwig, R. & Lehrach, H. Resources, standards and tools for systems biology. *Brief. Funct. Genomic. Proteomic.* **6**, 240–251 (2007).
76. Klipp, E. *et al.* *Systems Biology: A Textbook* (Wiley-VCH, 2009).  
**A text book with examples of modelling and computational analysis.**
77. Lopez-Aviles, S., Kapuy, O., Novak, B. & Uhlmann, F. Irreversibility of mitotic exit is the consequence of systems-level feedback. *Nature* **459**, 592–595 (2009).
78. McAdams, H. H. & Arkin, A. Stochastic mechanisms in gene expression. *Proc. Natl Acad. Sci. USA.* **94**, 814–819 (1997).
79. Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D. & van Oudenaarden, A. Regulation of noise in the expression of a single gene. *Nature Genet.* **31**, 69–73 (2002).
80. Arkin, A., Ross, J. & McAdams, H. H. Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics* **149**, 1633–1648 (1998).
81. Emonet, T., Macal, C. M., North, M. J., Wickersham, C. E. & Cluzel, P. AgentCell: a digital single-cell assay for bacterial chemotaxis. *Bioinformatics* **21**, 2714–2721 (2005).
82. Hofestadt, R. & Thelen, S. Quantitative modeling of biochemical networks. *Stud. Health Technol. Inform.* **162**, 3–16 (2011).
83. Blinov, M. L., Faeder, J. R., Goldstein, B. & Hlavacek, W. S. BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. *Bioinformatics* **20**, 3289–3291 (2004).
84. Swainston, N. *et al.* Enzyme kinetics informatics: from instrument to browser. *FEBS J.* **277**, 3769–3779 (2010).
85. Waltemath, D. *et al.* Minimum Information About a Simulation Experiment (MIASE). *PLoS Comput. Biol.* **7**, e1001122 (2011).
86. Dada, J. O., Spasic, I., Paton, N. W. & Mendes, P. SBRML: a markup language for associating systems biology data with models. *Bioinformatics* **26**, 932–938 (2010).
87. Hoops, S. *et al.* COPASI — a Complex Pathway Simulator. *Bioinformatics* **22**, 3067–3074 (2006).
88. Klipp, E., Herwig, R., Kowald, A., Wierling, C. & Lehrach, H. *Systems Biology in Practice: Concepts, Implementation and Application* (John Wiley & Sons, 2005).
89. Haefner, J. W. *Modeling Biological Systems: Principles and Applications* (Kluwer Academic Pub, 1996).
90. Kauffman, S. A. Metabolic stability and epigenesis in randomly constructed genetic nets. *J. Theor. Biol.* **22**, 437–467 (1969).
91. Zheng, J. *et al.* SimBoolNet — a Cytoscape plugin for dynamic simulation of signaling networks. *Bioinformatics* **26**, 141–142 (2010).
92. Iglesias, P. & Ingaalls, B. *Control Theory and Systems Biology* (MIT Press, 2009).  
**An excellent collection of introductory articles on how control theory can be applied to systems biology analysis.**
93. Chen, O. *et al.* Genetic basis and molecular mechanism for idiopathic ventricular fibrillation. *Nature* **392**, 293–296 (1998).
94. Noble, D. Modeling the heart — from genes to cells to the whole organ. *Science* **295**, 1678–1682 (2002).
95. Nomura, T. Towards integration of biological and physiological functions at multiple levels. *Front. Physiol.* **1**, 164 (2010).
96. Gleeson, P. *et al.* NeuroML: a language for describing data driven models of neurons and networks with a high degree of biological detail. *PLoS Comput. Biol.* **6**, e1000815 (2010).
97. Asai, Y. *et al.* Specifications of insilicoML 1.0: a multilevel biophysical model description language. *J. Physiol. Sci.* **58**, 447–458 (2008).
98. Plewczynski, D., La niewski, M., Augustyniak, R. & Ginalski, K. Can we trust docking results? Evaluation of seven commonly used programs on PDBbind database. *J. Comput. Chem.* **32**, 742–755 (2011).
99. Englebienne, P. & Moitessier, N. Docking ligands into flexible and solvated macromolecules. 4: are popular scoring functions accurate for this class of proteins? *J. Chem. Inf. Model.* **49**, 1568–1580 (2009).
100. Swertz, M. A. & Jansen, R. C. Beyond standardization: dynamic software infrastructures for systems biology. *Nature Rev. Genet.* **8**, 235–243 (2007).
101. Mi, H. & Thomas, P. PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. *Methods Mol. Biol.* **563**, 123–140 (2009).
102. Kemper, B. *et al.* PathText: a text mining integrator for biological pathway visualizations. *Bioinformatics* **26**, i374–i381 (2010).
103. Maier, H. *et al.* LitMiner and WikiGene: identifying problem-related key players of gene regulation using publication abstracts. *Nucleic Acids Res.* **33**, W779–W782 (2005).
104. Huss, J. W. *et al.* The Gene Wiki: community intelligence applied to human gene annotation. *Nucleic Acids Res.* **38**, D633–D639 (2010).
105. Callaway, E. No rest for the bio-wikis. *Nature* **468**, 359–360 (2010).
106. Kitano, H., Ghosh, S. & Matsuoka, Y. Social engineering for virtual 'big science' in systems biology. *Nat. Chem. Biol.* **7**, 323–326 (2011).  
**This paper discusses social issues in community-driven efforts in systems biology.**
107. Surowiecki, J. *The Wisdom of Crowds*. (Anchor, 2005).
108. Edwards, J. S. & Palsson, B. O. How will bioinformatics influence metabolic engineering? *Biotechnol. Bioeng.* **58**, 162–169 (1998).
109. Edwards, J. S., Ibarra, R. U. & Palsson, B. O. *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nature Biotech.* **19**, 125–130 (2001).
110. Smith, D. A. in *Metabolism, Pharmacokinetics and Toxicity of Functional Groups* 61–94 (Royal Society of Chemistry Publishing, 2010).

## Acknowledgements

This work is, in part, supported by funding from the HD-Physiology Project of the Japan Society for the Promotion of Science (JSPS) to the Okinawa Institute of Science and Technology (OIST). Additional support is from a Canon Foundation Grant, the International Strategic Collaborative Research Program (BBSRC-JST) of the Japan Science and Technology Agency (JST), the Exploratory Research for Advanced Technology (ERATO) programme of JST to the Systems Biology Institute (SBI) and from a strategic cooperation partnership between the Luxembourg Centre for Systems Biomedicine and SBI.

## Competing interests statement

The authors declare no competing financial interests.

## FURTHER INFORMATION

The Systems Biology Institute: <http://www.sbi.jp>  
BioCatalogue: <http://www.biocatalogue.org>  
BioModels.net: <http://biomodels.net>

## SUPPLEMENTARY INFORMATION

See online article: [S1 \(table\)](#)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF



