From genomes to systems: the path with yeast

Stephen G. Oliver*

*Faculty of Life Sciences, The University of Manchester, Michael Smith Building,
Oxford Road, Manchester M13 9PT, UK*

Metabolic Control Analysis (MCA) is a conceptual and mathematical formalism that models the relative contributions of individual effectors in a pathway to both the flux through the pathway and the concentrations of individual intermediates within it. To exploit MCA in an initial Systems Biology analysis of the eukaryotic cell, two categories of experiments are required. In category 1 experiments, flux is changed and the impact on the levels of the direct and indirect products of gene action is measured. We have measured the impact of changing the flux on the transcriptome, proteome and metabolome of *Saccharomyces cerevisiae*. In this whole-cell analysis, flux equates to growth rate. In category 2 experiments, the levels of individual gene products are altered, and the impact on the flux is measured. We have used competition analyses between the complete set of heterozygous yeast deletion mutants to reveal genes encoding proteins with high flux control coefficients. These genes may be exploited, in a top-down analysis, to build a coarse-grained model of the eukaryotic cell, as exemplified by yeast. More detailed modelling requires that ‘natural’ biological systems be identified. The combination of flux balance analysis with both genetics and metabolomics in the definition of metabolic systems is discussed.

Keywords: functional genomics; systems biology; yeast; *Saccharomyces cerevisiae*; metabolic control analysis; flux balance analysis

1. INTRODUCTION

Systems Biology is central to the post-genomic agenda and there are plans to construct complete mathematical models of unicellular organisms. However, such grand syntheses are still a long way off—not least because much of the quantitative data that will be required, if such models are to have predictive value and explanatory power, simply do not exist. Therefore, we will have to first construct models of smaller sub-systems (e.g. glycolysis, nucleotide biosynthesis, the cell cycle, etc.) and then integrate these component modules into a single construct, representing the entire cell. The problem, then, is to ensure that the modules can be joined up in a seamless manner to make a complete working model of a living cell that makes experimentally testable predictions and can be used to explain empirical data. There is a general awareness of this problem and there is much debate about the relative merits of ‘bottom-up’ and ‘top-down’ approaches in Systems Biology. In fact, both will be needed—but the foregoing discussion demonstrates that the top-down approach faces the larger conceptual problems.

It is difficult to construct an overarching framework for a model of, say, a yeast cell when one has no idea what the final model will look like. It would be very useful to build a very coarse-grained model of the yeast cell based on our current knowledge. However, this is a dangerous step since our current knowledge is very incomplete, with much relevant data being unavailable at present. Such a construct would very likely lead to us

being in a ‘cannot get there from here’ situation a few years down the road, where we would find it impossible to integrate specific models of individual sub-systems, which had been generated by bottom-up approaches, with one another in a comprehensive model of the yeast cell, patterned on this over-arching framework.

A coarse-grained model is certainly desirable, but it might be sensible to get the yeast cell to construct it for us, rather than make an imperfect attempt to construct it ourselves. We have chosen to use the formalism of Metabolic Control Analysis (MCA; Kacser & Burns 1973; Kacser 1995; Fell 1997) for this purpose. The three main principles of MCA theory can be summed up as follows: (i) in a metabolic pathway, the control of flux through the pathway is ‘shared’ between the different effectors (usually enzymes) that determine the individual steps in the pathway and is not exerted by a single enzyme catalysing some ‘rate-limiting step’. (ii) The relative control exerted by each effector (enzyme) is measured by the ‘flux control coefficient’, C_e^J . This can be defined as the ratio of the relative change in flux due to a change in the relative activity of an individual effector (enzyme) at steady state. (iii) The ‘summation theorem’ of MCA states that, for an unbranched pathway, the sum of the flux control coefficients of all the effectors is equal to one. (That is, a completely rate-determining enzyme will have a value of $C_e^J = 1$; while an enzyme that exerts no control over flux through the pathway at all will have $C_e^J = 0$. In practice, control is shared between all the enzymes in a pathway and flux control coefficients have values closer to 0 than 1.) We have suggested that MCA forms a general conceptual and mathematical framework for the discovery of gene function, and not just for those

*steve.oliver@manchester.ac.uk

One contribution of 15 to a Discussion Meeting Issue ‘Bioinformatics: from molecules to systems’.

genes whose products directly participate in or regulate metabolism (Oliver *et al.* 1998; Teusink *et al.* 1998; Raamsdonk *et al.* 2001).

Therefore, we must persuade yeast to identify those components of the eukaryotic cell system that exert the greatest degree of control over the pathways in which they participate, or which they regulate. In other words, we need to identify those components of a yeast cell that exert the greatest degree of control over its rate of growth and division. In the parlance of MCA, these components would be said to have high Flux Control Coefficients. In the following sections, I will first discuss bottom-up approaches to the identification of sub-systems within the yeast cell. I will then return to this top-down approach, based on MCA, for the construction of a coarse-grained model of the yeast cell. However, the concept of flux, its distribution and control, is intrinsic to both the bottom-up and top-down approaches to be described.

2. BOTTOM-UP APPROACHES

For the bottom-up approach, the initial problem is one of systems identification. While a lot of time is currently spent debating the question: 'what is Systems Biology?', that time would be much better employed in answering the question: 'what is a biological system?' The basic system that is studied in my laboratory is the eukaryotic cell, as exemplified by the budding yeast *Saccharomyces cerevisiae*. Among the properties that make *S. cerevisiae* a particularly suitable organism with which to study the system of the eukaryotic cell are its rapid growth, well-dispersed cells, and simple methods of cultivation under controlled and chemically defined conditions. Moreover, yeast's life cycle, with its ability to grow vegetatively in either the haploid or diploid phase (allowing the direct analysis of the products of meiosis) means that it represents an extremely well-defined genetic system for which facile techniques of exquisitely accurate genetic manipulation have been developed (Brown & Tuite 1998; Sherman 1998, 2002; Castrillo & Oliver 2004). While it may be premature to speak of the 'e-yeast' or the 'virtual *Saccharomyces*', if the goal of constructing a predictive and explanatory model of the eukaryotic cell can be achieved at all, then it ought to be possible to achieve it with yeast.

Why then (in an organism where we know so much about its biochemistry, physiology and cell biology) should it be a problem to identify the biological sub-systems that must be fully characterized and built into a comprehensive model of the eukaryotic cell? This problem arises because we have previously studied these biological systems in isolation and in a rigorously reductive fashion. Now, we must study them as parts of an integrated whole (see Kell & Oliver 2004; Oliver 2003). The problem is that our current view of, say, a metabolic or signal transduction pathway is often two-dimensional (rather than four-dimensional) and is frequently poorly integrated, if at all, with other cellular pathways. Thus our view of the network of metabolic pathways, represented by that chart taped to the door of the fridge in the laboratory, may not be the same as the yeast cell's. In order to gain a 'yeast's-eye view' of the different metabolic systems within the eukaryotic

cell, we have coupled flux balance analysis with both metabolomics and genetics. Although the initial aim of these approaches is the identification of the natural metabolic systems of yeast, the principles involved should be more widely applicable to the problem of biological systems identification.

(a) *Flux coupling analysis and metabolomics*

Saccharomyces cerevisiae is a rather specialized organism and its metabolic network is quite small, and probably better defined than that of any other eukaryote. Thus it has been possible to construct models, simply on the basis of stoichiometry (Holms *et al.* 1991) and the availability of a fully annotated genome sequence (Goffeau *et al.* 1996, 1997), of the complete metabolic network of the organism (Forster *et al.* 2003). Using such a model, simulations may be run that allow the calculation of the flux ratio between any pair of reactions at steady state. If this ratio is constant in all possible steady states, then the two reactions are said to be coupled; if it varies, they are said to be uncoupled (Burgard *et al.* 2004). Metabolomic analysis of either yeast cell extracts (the endometabolome or the metabolic fingerprint; Raamsdonk *et al.* 2001; Castrillo *et al.* 2003) or the culture supernatant (exometabolome or metabolic footprint; Allen *et al.* 2003) of mutants deleted for genes encoding either regulators or enzymes within a metabolic pathway may be used to check the predicted coupling of reactions (Bundy *et al.* in preparation). This combination of *in silico* flux coupling analysis and *in vivo* metabolomic analysis should be a powerful and efficient route to metabolic systems identification.

(b) *Flux balance analysis and synthetic phenotypes*

The recent completion of the human genome sequence (Lander *et al.* 2001; Venter *et al.* 2001) has provided a powerful demonstration that the increase in complexity of organisms up the evolutionary scale is not simply a consequence of an increase in gene number; rather, it is the result of an increased sophistication in the interactions between genes. Thus, to define the sub-systems within the eukaryotic cell, using yeast, we must develop a bottom-up strategy that allows us to reveal these gene interactions on a global level. This can readily be done by identifying synthetic interactions between genes. Thus, it may be that there is no overt phenotype due to deleting either gene A (A^-B^+) or gene B (A^+B^-) from the yeast genome (A^-B^+ or A^+B^-), but cells with *both* genes A and B deleted (A^-B^-) show a clear phenotype. The most obvious phenotype is death and screens are frequently made, on a gene-by-gene basis, for synthetic lethal interactions. Synthetic interactions between genes can be the consequence of gene redundancy (two genes encode the same protein) or of the occurrence of alternative pathways for the synthesis of the same product. A systematic attempt is being made in Canada to uncover all synthetic lethal interactions between yeast's 6000 protein-encoding genes by constructing all 12 million or so possible double mutants (Tong *et al.* 2004). So far, about 4% of the possible double-mutants

have been constructed and *ca* 0.6% have proved to be synthetically lethal on rich medium.

While this comprehensive analysis will provide a resource of incomparable value, it is possible (for genes involved in metabolism and its control) to take a short cut. We (Papp *et al.* in preparation) are using the stoichiometric model of yeast metabolism constructed by Palsson, Nielsen and colleagues (Forster *et al.* 2003) to carry out simulations (Segre *et al.* 2005) that allow us to predict synthetic interactions between metabolic genes. So far, construction of double-mutants based on these predictions has revealed synthetic interactions at a frequency of *ca* 30% (compared to 0.6% for the comprehensive approach). This suggests that the prediction and testing of synthetic interactions will be a relatively efficient way of uncovering partially redundant natural metabolic systems within the yeast cell.

3. TOP-DOWN APPROACHES WITHIN A METABOLIC CONTROL ANALYSIS FRAMEWORK

In order to exploit MCA in an initial top-down Systems Biology analysis of the eukaryotic cell, two categories of experiments are required: In category 1, flux is changed and the impact on the levels of the direct and indirect products of gene action is measured. In category 2, the levels of individual gene products are altered, and the impact on the flux is measured. The integration of the datasets produced by these two kinds of experiments should allow us to link the controllers with what they control in a truly comprehensive manner. This integration should permit the construction of a coarse-grained model of the eukaryotic cell that, in common with the bottom-up approaches described above, represents a yeast's-eye view. This coarse-grained model should, therefore, provide a suitable framework into which the sub-system models produced by bottom-up studies may be fitted, eventually permitting the 'grand synthesis' of a comprehensive and integrated model of the yeast cell, the e-yeast.

(a) *Category 1 experiments: changing the flux*

We have measured the impact of changing the flux on the transcriptome, proteome and metabolome of *S. cerevisiae*. In this whole-cell analysis, flux equates to growth rate, and we have exploited the technique of chemostat culture in order to set the growth rate (or flux) by controlling the supply of a growth-rate-determining nutrient. By setting growth rate to the same values in chemostats in which either glucose, ammonium, sulphate or phosphate represents the limiting nutrient, we have studied the control of the yeast transcriptome, proteome and metabolome in a manner that allows the dissection of growth-rate effects from nutritional effects (Castrillo *et al.* in preparation). Our previous experiments on transcriptome analysis (Hayes *et al.* 2002; Lim *et al.* 2003; Wu *et al.* 2004) had demonstrated that the use of chemostat culture removes many of the confounding variables that complicate the analysis of the data, and the same should be true for the proteome and metabolome. Therefore, the results generated by these experiments

should provide important kinetic information for the construction of dynamic models of the yeast cell in Systems Biology (Kitano 2005; Nurse 2003; Castrillo & Oliver 2005).

We have found that data on the exometabolome (that is, the totality of metabolites excreted from the cells into the growth medium), proteome and transcriptome all show a clear separation between the carbon-limited (i.e. glucose-limited) and the carbon-sufficient (i.e. ammonium-, sulphate- and phosphate-limited) cultures. Moreover, the proteome and transcriptome data show clear growth-rate-associated trends under all three nutrient limitations, but only the exometabolome data from the carbon-sufficient cultures shows such a trend. Since the exometabolome is the product of overflow metabolism (Kell *et al.* 2005), we infer that yeast more tightly regulates its metabolism with changes in flux (growth rate) under carbon-limited than under carbon-sufficient conditions, such that (in the former) the spectrum of metabolites excreted remains relatively unchanged.

The most detailed analyses of the impact of changing the flux have been made at the level of the transcriptome, where we were able to exploit the Gene Ontology (Ashburner *et al.* 2000) to place sets of genes significantly up- or downregulated with growth rate in their functional categories. We found that a large proportion of all growth-rate-regulated yeast genes share orthologues with other eukaryotes, including humans. The set of *ca* 500 genes significantly upregulated with increasing growth rate are frequently essential and encode evolutionarily conserved proteins of known function that participate in many protein-protein interactions. In contrast, more unknown, and fewer essential, genes are included in the set of *ca* 400 genes downregulated with increasing growth rate. The protein products of this downregulated set rarely interact with one another, but more frequently interact with the products of the upregulated genes. Presumably these interactions between proteins encoded by such differently regulated sets of genes are transient in nature and may indicate that the controls that mediate the switch-over between one set of genes and the other, as growth rate changes, may act at the post-transcriptional level.

The set of *ca* 500 genes of known function that are upregulated with growth rate shows a high proportion that are involved in the following biological processes: translation initiation, ribosome biogenesis and assembly, protein biosynthesis, nucleotide and nucleic acid metabolism. The same gene set showed the following biological functions of their protein products to be over represented: translation initiation factor activity and nucleic acid (RNA) binding, structural constituents of ribosome activity, ligase activity forming aminoacyl-tRNAs and DNA-directed RNA polymerase activity. These proteins that are upregulated with increasing growth rate occur in a variety of sub-cellular compartments (cytosol, exosome, nucleus) and complexes (e.g. eukaryotic translation initiation complexes, nucleolus and ribosome subunits).

Analysis of the set of *ca* 300 genes of known function whose transcription was significantly downregulated with increasing growth rate showed a high proportion

of these genes to be involved in the following biological processes: response to external stimulus, cell communication and signal transduction, autophagy, homeostasis, response to stress, vesicle recycling within Golgi. The predominant molecular functions categories for this gene set involve a variety of catalytic, signal transduction, transcription regulator and transport activities. The protein products of these growth-rate downregulated genes are components of the plasma membrane, vacuole and repairosome.

These genes that are downregulated with increasing growth rate are probably involved in maximizing the efficient utilization of cellular resources when nutrients are scarce. Our data indicate that this is a poorly understood aspect of the cell's economy since 30% of these genes is of as yet undetermined function. This is despite the fact that nutrient scarcity is likely to be a common situation for micro-organisms growing in their natural environment (Ferenci 1999). Among the genes of known function that are upregulated at low growth rates are those involved in mobilization and storage of available resources in the vacuole. Autophagy is a major system of bulk degradation of cellular components, and we find genes involved in this process are among those upregulated at low growth rates. Autophagy mediates a reduction in the pool of ribosomes, thus slowing cell growth when nutrients are limiting (Ohsumi 2004).

Other genes that are upregulated at low growth rates are those encoding transcriptional repressors whose action results in the activation of alternative routes for the assimilation of substrates as an adaptation to the environment. In all, the data on the downregulated genes present a picture of the slowly growing yeast cell activating pathways involved in the response to external stimuli, maintenance of homeostasis, vacuolar transport and storage and autophagy—the whole being directed towards a more efficient use of scarce resources.

(b) Category 2 experiments: changing the levels of gene products

We are determining which of yeast's 5000-plus protein-encoding genes display a haploinsufficiency phenotype by carrying out competition experiments between hemizygous, bar-coded, mutants in chemostat culture (Delneri *et al.* 2003). These studies have shown that the majority of genes have no impact on the fitness of yeast when their copy number is reduced from 2 to 1 in a diploid strain. However, there are a number of genes that cause a significant reduction of fitness (haploinsufficiency) or a significant improvement in fitness ('haploproficiency') when in the hemizygous state. Some of the genes exhibit a haploinsufficient or a haploproficient phenotype in all conditions studied, while others only exhibit an effect under specific environmental conditions. For instance, genes involved in vesicular transport show a haploinsufficient phenotype in all the environments examined so far (nitrogen-limited, carbon-limited and phosphorus-limited). In contrast, genes encoding proteins that form components of the proteasome complex, or are involved elsewhere in ubiquitin-mediated protein degradation, show a haploproficient phenotype in a nitrogen-limited

environment but not in conditions of carbon or phosphorus limitation.

If one integrates the data from the Category 1 ('change the flux') and Category 2 ('change the gene product concentrations') experiments, a striking result emerges. Virtually none of the genes that show a significant haploinsufficient or haploproficient phenotype under all three nutrient limitations examined has its expression regulated at the transcriptional level. Thus, in the parlance of MCA, very few yeast gene products with a high flux control coefficient for cell growth are, themselves, subject to regulation at the level of transcription. This may be a fundamental 'design rule' of eukaryotic cell systems (Kitano 2005; Nurse 2003; Castrillo & Oliver 2005), but such terms should be used with caution since living organisms are not the results of design; rather, they are the products of evolution. This fact represents a major challenge for the systems approach to biology.

4. THE GRAND SYNTHESIS: CHALLENGES FOR BIOINFORMATICS

The grand synthesis represents the task of integrating the data and models from both top-down and bottom-up investigations of the Systems Biology of the eukaryotic cell, as exemplified by studies on *S. cerevisiae*. At all costs, we must avoid a cannot get there from here situation in which the grand synthesis cannot be achieved because the over-arching framework used for top-down studies is inappropriate or inadequate, or the different sub-system models are incompatible with one another for either biological or computational reasons. The avoidance of these pitfalls will require considerable forward planning and cooperation on the part of the research community, and bioinformatics has a central role to play in solving the problems involved.

In preparing for the grand synthesis, we need to ensure comparability, compatibility and interoperability at both the experimental and informatic levels. A lot of this is about standards. The very word makes some researchers' hackles rise, because they think that the quality of their work is being judged. In truth, an improvement in research quality is often a consequence of the imposition of international data standards, but this is a fortunate side-product rather than the main aim. To avoid the 'we cannot get there from here' problem, researchers in yeast Systems Biology will have to agree a number of standards for their experiments—including the genetic background for the strains of *S. cerevisiae* used, the culture media on which those strains are grown, and a number of classical physiological transitions that can act as benchmarks for studies. Fortunately, the yeast research community has a long history of working in a cooperative and collaborative manner, which was largely due to the characters of a number of the founding fathers of the field—people like Herschel Roman & Bob Mortimer (Hall & Linder 1993). In recent times, there have been major world-wide collaborations between many laboratories in the yeast research community to achieve the sequencing of the *S. cerevisiae* genome (Goffeau *et al.* 1996, 1997) and the generation of

a comprehensive set of gene-deletion mutants and their functional analysis (Oliver 1996; Winzeler *et al.* 1999; Giaever *et al.* 2002). The Yeast Systems Biology Network has been established recently (Nielsen 2004) and, as with the Sequencing and Functional Analysis projects, funding from the European Commission looks likely to play a major catalytic role. It can only be hoped that the larger international Systems Biology scene (Kitano 2005) will emulate the characteristic positive cooperative spirit and effective self-organizing ability exhibited by the yeast researchers.

At the bioinformatic level, data standards have to do with ensuring that the results of experiments are stored in an accessible and interoperable manner, and that sufficient information about the way the experiments were performed (the so-called metadata—the data about the data) is also available to allow only appropriate and effective comparisons and integrations to be made. At this level, most progress has been made with the establishment of standards for transcriptomic data, and this effort has been ably led by another author of this issue, Alvis Brazma. Under his guidance, the MGED (www.mged.org) Consortium has both proposed what information must be captured about a microarray experiment (MIAME; <http://www.ebi.ac.uk/arrayexpress/>; Brazma *et al.* 2001), and developed XML representations for such data (MAGE-ML; Spellman *et al.* 2002). Developing agreed and effective representations for complex datasets originating in different labs, using different equipment, acting over diverse organisms is far from straightforward. Nevertheless, MGED has made significant progress in this direction (<http://www.ebi.ac.uk/arrayexpress/>). The development of standard representations for proteome and metabolome data is at a much earlier stage of development. However, the need for such standard representations is recognized and a credible start has been made for both proteomics (Orchard *et al.* 2003; Taylor *et al.* 2003; Garwood *et al.* 2004; <http://pedro.man.ac.uk/>) and metabolomics (Kell 2004; Jenkins *et al.* 2004; <http://www.armet.org/>). What is required now is for a universal ‘front end’ to be developed that allows experimentalists using multiple levels of omic analysis to capture the description of their experiment, and the associated metadata, just once. Again, this is an area of bioinformatics that is only just beginning to receive serious attention (Jones *et al.* 2004). A Functional Genomics Object Model (FuGE-OM; <http://fuge.sourceforge.net>) is under construction to enable common data standards for all levels of functional genomics to be developed.

We not only need compatibility and interoperability of the bioinformatics systems relating to the data gathered in functional genomics and Systems Biology experiments, we also need to ensure that the models that will be used increasingly to guide those experiments also have these characteristics. There has been an early awareness of this problem, and the development of the Systems Biology Mark-Up Language (Finney *et al.* 2001; Hucka *et al.* 2003) should permit the facile exchange of models between researchers. There has even been a demonstration that it is possible to completely automate the generation of hypotheses and the design and execution of functional genomics

experiments to test them (King *et al.* 2004). For all that, it will take a lot of good will and cooperation on the part of all-too-human researchers to ensure that ‘we can get there from here’ and that the virtual yeast will become a reality.

I am grateful to Kevin Brindle, Juan Castrillo, Daniela Delneri, and Balázs Papp for their critical reading of the manuscript and permission to refer to their unpublished data. Douglas Kell is thanked for many stimulating discussions. Work on the functional genomics and systems biology of yeast in my own laboratory has been supported by the BBSRC, NERC, EPSRC, EC, Wellcome Trust, DTI and AstraZeneca.

REFERENCES

- Allen, J., Davey, H. M., Broadhurst, D., Heald, J. K., Rowland, J. J., Oliver, S. G. & Kell, D. B. 2003 High-throughput classification of yeast mutants using metabolic footprinting. *Nat. Biotechnol.* **21**, 692–696. (doi:10.1038/nbt823)
- Ashburner, M. *et al.* 2000 Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29. (doi:10.1038/75556)
- Brazma, A. *et al.* 2001 Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nat. Genet.* **29**, 365–371. (doi:10.1038/ng1201-365)
- Brown, A. J. P. & Tuite, M. F. 1998 *Yeast gene analysis. Methods in Microbiol.* San Diego: Academic Press.
- Bundy, J., Papp, B., Burton, N., Harmston, R., Oliver, S. G., & Brindle, K. M. In preparation. Analysing metabolic networks using NMR-based metabolite profiling in conjunction with flux coupling analysis.
- Burgard, A. P., Nikolaev, E. V., Schilling, C. H. & Maranas, C. D. 2004 Flux coupling analysis of genome-scale metabolic network reconstructions. *Genome Res.* **14**, 301–312. (doi:10.1101/gr.1926504)
- Castrillo, J. I. & Oliver, S. G. 2004 Yeast as a touchstone in post-genomic research. Strategies for integrative analysis in functional genomics. *J. Biochem. Mol. Biol.* **37**, 93–106.
- Castrillo, J. I. & Oliver, S. G. 2005 Metabolomics and Systems Biology in *Saccharomyces cerevisiae*. In *The mycota. Vol. XIII. A comprehensive treatise on fungi as experimental systems for basic and applied research. Fungal genomics* (ed. Karl Esser). Berlin: Springer.
- Castrillo, J. I., Hayes, A., Mohammed, S., Gaskell, S. J. & Oliver, S. G. 2003 An optimised protocol for metabolome analysis in yeast using direct infusion electrospray mass spectrometry. *Phytochemistry* **62**, 929–937. (doi:10.1016/S0031-9422(02)00713-6)
- Castrillo J. I. *et al.* In preparation. Growth control of the eukaryote cell: a systems biology study in yeast.
- Delneri, D., Leong, H. S., Hayes, A., Davey, H. M., Kell, D. B. & Oliver, S. G. 2003 Assessing contributions to fitness of individual genes via genome-wide competition analysis. *Yeast* **20**, S337.
- Fell, D. A. 1997 *Understanding the control of metabolism*. London: Portland Press Ltd.
- Ferenci, T. 1999 Regulation by nutrient limitation. *Curr. Opin. Microbiol.* **2**, 208–213. (doi:10.1016/S1369-5274(99)80036-8)
- Finney, A. M., Hucka, M., Sauro, H. M., Doyle, J., Kitano, H. & Bolouri, H. 2001 The systems biology markup language. *Mol. Biol. Cell* **12**, 708S.
- Forster, J., Famili, I., Fu, P., Palsson, B. O. & Nielsen, J. 2003 Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res.* **13**, 244–253. (doi:10.1101/gr.234503)

- Garwood, K. *et al.* 2004 PEDRo: a database for storing, searching and disseminating experimental proteomics data. *BMC Genomics* **5**, 68. (doi:10.1186/1471-2164-5-68)
- Giaever, G. *et al.* 2002 Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**, 387–391. (doi:10.1038/nature00935)
- Goffeau, A. *et al.* 1996 Life with 6000 genes. *Science* **274**, 546–567. (doi:10.1126/science.274.5287.546)
- Goffeau, A. *et al.* 1997 The yeast genome directory. *Nature* **387**, 5–105.
- Hall, M. & Linder, P. (eds) 1993 *The early days of yeast genetics*. Plainview, NY: Cold Spring Harbor Laboratory Press.
- Hayes, A., Zhang, N., Wu, J., Butler, P. R., Hauser, N. C., Hoheisel, J. D., Lim, F., Sharrocks, A. D. & Oliver, S. G. 2002 Hybridization array technology coupled with chemostat culture: tools to interrogate gene expression in *S. cerevisiae*. *Methods* **26**, 281–290. (doi:10.1016/S1046-2023(02)00032-4)
- Holms, W. H., Hamilton, I. D. & Mousdale, D. 1991 Improvements in microbial productivity by analysis of metabolic fluxes. *J. Chem. Technol. Biotechnol.* **50**, 139–141.
- Hucka, M. *et al.* 2003 The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* **19**, 524–531. (doi:10.1093/bioinformatics/btg015)
- Jenkins, H. *et al.* 2004 A proposed framework for the description of plant metabolomics experiments and their results. *Nat. Biotechnol.* **22**, 1601–1606. (doi:10.1038/nbt1041)
- Jones, A., Hunt, E., Wastling, J. M., Pizarro, A. & Stoekert, C. J. 2004 An object model and database for functional genomics. *Bioinformatics* **20**, 1583–1590. (doi:10.1093/bioinformatics/bth130)
- Kacser, H. 1995 Recent developments beyond metabolic control analysis. *Biochem. Soc. Trans.* **23**, 387–391.
- Kacser, H. & Burns, J. A. 1973 The control of flux. *Symp. Soc. Exp. Biol.* **27**, 65–104.
- Kell, D. B. 2004 Metabolomics and systems biology: making sense of the soup. *Curr. Opin. Microbiol.* **7**, 296–307. (doi:10.1016/j.mib.2004.04.012)
- Kell, D. B. & Oliver, S. G. 2004 Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. *Bioessays* **26**, 99–105. (doi:10.1002/bies.10385)
- Kell, D. B., Brown, M., Davey, H. M., Dunn, W. B., Spasic, I., & Oliver, S. G. In press. Metabolic footprinting and Systems Biology: the medium is the message. *Nat. Rev. Microbiol.*
- King, R. D., Whelan, K. E., Jones, F. M., Reiser, P. G. K., Bryant, C. H., Muggleton, S. H., Kell, D. B. & Oliver, S. G. 2004 Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature* **427**, 247–255. (doi:10.1038/nature02236)
- Kitano, H. 2005 International alliances for quantitative modeling in systems biology. *Mol. Syst. Biol.* Published online 29 March 2005. (doi:10.1038/msb4100011)
- Lander, E. S. *et al.* 2001 Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921. (doi:10.1038/35057062)
- Lim, F.-L., Hayes, A., West, A. G., Pic, A., Morgan, B. A., Oliver, S. G. & Sharrocks, A. D. 2003 Mcm1p-Induced DNA bending regulates the formation of ternary transcription factor complexes. *Mol. Cell Biol.* **23**, 450–461. (doi:10.1128/MCB.23.2.450-461.2003)
- Nielsen, J. 2004 The first international workshop on systems biology of yeast, St Louis, USA, 9 November, 2003. *FEMS Yeast Res.* **4**, 757–758. (doi:10.1016/j.femsyr.2004.02.001)
- Nurse, P. 2003 Systems biology: understanding cells. *Nature* **424**, 883. (doi:10.1038/424883a)
- Ohsumi, Y. 2004 Autophagy: reversing cell growth. In *Cell growth. Control of cell size* (ed. M. N. Hall, M. Raff & G. Thomas) *Monograph* **42**, pp. 413–429. New York: Cold Spring Harbor Laboratory Press.
- Oliver, S. 1996 A network approach to the systematic analysis of yeast gene function. *Trends Genet.* **12**, 241–242. (doi:10.1016/0168-9525(96)30053-X)
- Oliver, S. G. 2003 Functional genomics: all the King's horses AND all the King's men CAN put Humpty together again. *Mol. Cell* **12**, 1343–1344. (doi:10.1016/S1097-2765(03)00501-X)
- Oliver, S. G., Winson, M. K., Kell, D. B. & Baganz, F. 1998 Systematic functional analysis of the yeast genome. *Trends Biotechnol.* **16**, 373–378. (doi:10.1016/S0167-7799(98)01214-1)
- Orchard, S., Hermjakob, H. & Apweiler, R. 2003 The proteomics standards initiative. *Proteomics* **3**, 1374–1376. (doi:10.1002/pmic.200300496)
- Papp, B., Harrison R., Pal, C., Delneri, D., & Oliver, S. G. In preparation. Predicting synthetic lethal interactions between yeast genes using flux balance analysis.
- Raamsdonk, L. M. *et al.* 2001 A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat. Biotechnol.* **19**, 45–50. (doi:10.1038/83496)
- Segre, D., DeLuna, A., Church, G. M. & Kishony, R. 2005 Modular epistasis in yeast metabolism. *Nat. Genet.* **37**, 77–83.
- Sherman, F. 1998 An introduction to the genetics and molecular biology of the yeast *Saccharomyces cerevisiae* http://dbb.urmc.rochester.edu/labs/sherman_f/yeast/.
- Sherman, F. 2002 Getting started with yeast. *Meth. Enzymol.* **350**, 3–41. (Updated in http://dbb.urmc.rochester.edu/labs/sherman_f/startedyeast.pdf)
- Spellman, P. T. *et al.* 2002 Design and implementation of microarray gene expression markup language (MAGE-ML). *Genome Biol.* **3**. (doi:10.1186/gb-2002-3-9-research0046) research00461-0046.9.
- Taylor, C. F. *et al.* 2003 A systematic approach to modelling, capturing and disseminating proteomics experimental data. *Biotechnology* **21**, 247–254.
- Teusink, B., Baganz, F., Westerhoff, H. V. & Oliver, S. G. 1998 Metabolic control analysis as a tool in the elucidation of the function of novel genes. *Meth. Microbiol.* **26**, 297–336.
- Tong, A. H. Y. *et al.* 2004 Global mapping of the yeast genetic interaction network. *Science* **303**, 808–813. (doi:10.1126/science.1091317)
- Venter, J. C. *et al.* 2001 The sequence of the human genome. *Science* **291**, 1304–1351. (doi:10.1126/science.1058040)
- Winzler, E. A. *et al.* 1999 Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901–906. (doi:10.1126/science.285.5429.901)
- Wu, J., Zhang, N., Hayes, A., Panoutsopoulou, K. & Oliver, S. G. 2004 Global analysis of nutrient control of gene expression in *Saccharomyces cerevisiae* during growth and starvation. *Proc. Natl Acad. Sci. USA.* **101**, 3148–3153. (doi:10.1073/pnas.0308321100)