Journal of Biochemistry and Molecular Biology, Vol. 37, No. 1, January 2004, pp. 93-106

JOURNAL OF Biochemistry and Molecular Biology

Review

Yeast as a Touchstone in Post-genomic Research: Strategies for Integrative Analysis in Functional Genomics

Juan I. Castrillo* and Stephen G. Oliver

School of Biological Sciences, University of Manchester, 2205 Stopford Building, Oxford Road, Manchester M13 9PT, UK

Received 10 December 2003

The new complexity arising from the genome sequencing projects requires new comprehensive post-genomic strategies: advanced studies in regulatory mechanisms, application of new high-throughput technologies at a genome-wide scale, at the different levels of cellular complexity (genome, transcriptome, proteome and metabolome), efficient analysis of the results, and application of new bioinformatic methods in an integrative or systems biology perspective. This can be accomplished in studies with model organisms under controlled conditions. In this review a perspective of the favourable characteristics of yeast as a touchstone model in post-genomic research is presented. The state-of-the art, latest advances in the field and bottlenecks, new strategies, new regulatory mechanisms, applications (patents) and high-throughput technologies, most of them being developed and validated in yeast, are presented. The optimal characteristics of yeast as a well-defined system for comprehensive studies under controlled conditions makes it a perfect model to be used in integrative, 'systems biology' studies to get new insights into the mechanisms of regulation (regulatory networks) responsible of specific phenotypes under particular environmental conditions, to be applied to more complex organisms (e.g. plants, human).

Keywords: Functional genomics, Integrative studies, Systems biology, *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*

Introduction

The completion of the first draft of the human genome sequence by the International Human Genome Mapping Consortium (2001a, 2001b) has revealed a complex picture and confirmed that it was presumptuous to assume that once the sequence had been obtained all problems would be solved (Shoemaker *et al.*, 2001; Levine and Tjian, 2003). In fact, more questions have arisen which show that the complexity of biological systems lies not only in the basic units of information (or genes) contained in the genome (Morange, 2002; Dillon, 2003; Pearson, 2003), but in the regulatory circuits which control the expression of these genes, synthesis of proteins and changes of internal metabolic pools which allow the cell to respond to environmental changes, mediated by specific sensing and regulatory mechanisms in an tightly-controlled way (Oliver, 1996; Oliver *et al.*, 1998; Kitano, 2002; Morange, 2002; Levine and Tjian, 2003; Tate and Cooper, 2003). In spite of this, however, many studies continue to work on the discovery of genes responsible for specific diseases, while ignoring the influence of epigenetic factors and the environment (Morange, 2002; Levy-Lahad and Plon, 2003).

The current 'genomic revolution' is generating large amounts of valuable information in the form of genome sequences. This new knowledge has to be accompanied by post-genomic studies based on new advanced methods, strategies and technologies which have to be continuously developed and improved. These methods and strategies should be mainly directed to the elucidation of the function of new genes, new mechanisms of regulation, and gene regulatory networks responsible for specific phenotypes at the different biological levels: genome, transcriptome, proteome and metabolome, in an integrative or systems biology perspective (Delneri *et al.*, 2001; Kitano, 2002; Oliver, 2002; Oliver *et al.*, 2002). In this context, the exploitation of model organisms (Table 1) (Cech, 2001; Bahls *et al.*, 2003) is central, not only to continued generation of basic knowledge on new regulatory mechanisms and cell biology, but also in the development of advanced techniques for the analysis of genomic information under controlled, reproducible conditions, which are difficult to apply to more complex organisms (e.g. plants and mammals) (Arabidopsis Genome Initiative, 2000; International Human Genome Mapping

*To whom correspondence should be addressed.
Tel: 44-161-606-7260; Fax: 44-161-606-7360
E-mail: Juan.I.Castrillo@man.ac.uk

Table 1. Eukaryotic model organisms (Cech, 2001; Bahls *et al.*, 2003)

Organism	References	Web resources
Human		
(*H. sapiens*)	(The International Human Genome Mapping Consortium, 2001a, 2001b, 2003)	http://www.sanger.ac.uk/HGP/ http://www.ncbi.nlm.nih.gov/genome/guide/human/
Mouse		
(*M. musculus*)	(Mouse Genome Sequencing Consortium, 2002)	http://www.ensembl.org/Mus_musculus/ http://www.informatics.jax.org
	(The FANTOM Consortium and the RIKEN Genome Exploration Research Group, 2002)	http://genome.gsc.riken.go.jp/ http://fantom2.gsc.riken.go.jp/
Fly		
(*D. melanogaster*)	(Adams *et al.*, 2000)	http://www.flybase.org
Plant		
(*A. thaliana*)	(*Arabidopsis* Genome Initiative, 2000)	http://www.arabidopsis.org
Fish		
(*D. rerio*)	(Bahls *et al.*, 2003)	http://zfin.org/ http://www.sanger.ac.uk/Projects/D_rerio/
Worm		
(*C. elegans*)	(The *Caenorhabditis elegans* Sequencing Consortium, 1998)	http://www.wormbase.org http://www.wormatlas.org
Yeast		
(*S. cerevisiae*)	(Goffeau *et al.*,1996, 1997)	http://www.yeastgenome.org/
(*Sz. pombe*)	(Wood *et al.*, 2002)	http://www.sanger.ac.uk/Projects/S_pombe/

Consortium, 2001; Mouse Genome Sequencing Consortium, 2002). Finally, the large volume of information generated (raw and processed data) has to be curated, stored and managed efficiently. For that purpose, robust bioinformatic tools, in the form of databases and methods for data mining, as well as the processing, analysis and dissemination of the information are necessary (Kell and King, 2000; Cornell *et al.*, 2003; Doniger *et al.*, 2003; Taylor *et al.*, 2003).

This review will present a perspective of the most recent developments in post-genomic studies. First, an overview of the last advances in basic knowledge and applied research at the genome, transcriptome, proteome and metabolome level is provided, with examples of studies in different organisms. Direct applications in the form of patents are also included. Based on this picture, a view of some overlooked features in post-genomic studies is presented. Particular emphasis is placed on the careful selection of the model organism, level of analysis, experimental design and controlled cultivation conditions in order to get reliable results, leading to valid conclusions. Finally, in the last part of the review attention is focused on the favourable characteristics of yeast as a model organism in post-genomic research. Examples of the most relevant scientific knowledge, methods and strategies developed in yeasts (*Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*) are presented. The optimal characteristics of yeasts make them valuable tools for comprehensive studies at the genome, transcriptome, proteome and metabolome levels, and global integrative strategies. These aspects are discussed, together with their implications and future prospects for application in systems biology research.

Post-genomic Studies. Perspective. New Challenges

The generation of huge volumes of genomic sequences is revealing new complexity. For example, the analysis of the number of genes in the human genome, lower than initially expected (Pearson, 2003) has resulted in a new perspective in which the complexity of an organism is progressively being analysed not as a direct consequence of the number of genes, but on the basis of the control mechanisms responsible for regulation or 'tuning' of gene action and interaction. The concept of the gene as a physical unit of information has been under debate since its origin, and is currently being revisited (Morange, 2002; Dillon, 2003). Moreover, the presence of a high proportion of non-coding sequences cannot be considered 'junk DNA' any more, but underlie important functions yet to be elucidated, as demonstrated by the fact that such sequences are significantly more conserved (80% to 90% the same on average across most of the species) than some protein-coding genes (Dermitzakis *et al.*, 2003). As a consequence of this, the identity of other functional elements of the DNA sequence (promoters and transcriptional regulatory sequences) along with determinants of chromosome structure and function has to be established, and new initiatives are being directed to this objective (e.g. ENCODE project: ENCyclopedia Of DNA Elements. http://www.genome.gov/10005107). The latest studies are providing new insights into the mechanisms of transcriptional regulation and gene silencing by RNA interference (RNAi) (Hannon, 2002; Schramke and Allshire, 2003; Volpe *et al.*, 2003) and the small ubiquitin-related modifier (SUMO) family of

proteins (Muratani and Tansey, 2003; Shiio and Eisenman, 2003; Verger *et al.*, 2003). This new complexity presents one of the main challenges in post-genomic studies.

The most advanced post-genomic strategies are directed to the elucidation of new genes, their function and mechanisms of regulation, using new techniques at different levels of study. Thus:

1) Strategies at the genome level. Having the completed sequence of an organism as a reference (e.g. Table 1 and references therein) new molecular techniques are being developed for identifying genes and functional elements. Thus, comparison of the sequences between closely and less related organisms is providing valuable information from the conserved and non-conserved sequences between species (Blaxter, 2003; Cliften *et al.*, 2003; Dermitzakis *et al.*, 2003; Guigo *et al.*, 2003; Kellis *et al.*, 2003). Also, strategies for genome-wide analysis of disruption phenotypes, or generation of knockout mutants are being applied to different model organisms (e.g. yeast and the mouse, Ross-MacDonald *et al.*, 1999; Winzeler *et al.*, 1999; Warringer and Blomberg, 2003; Mouse knockout and mutation database: http://research.bmn.com/mkmd). In addition, new transcomplementation strategies are being developed to provide valuable information on the function of human genes (Zhang *et al.*, 2003).

2) Strategies at the gene expression (transcriptome) level. In the last decades, several methods have been developed for the analysis of messenger RNA, with the objective of comparing patterns of gene expression between cells or under different conditions (for more information, see Liang and Pardee, 2003). However, the application of some of these techniques at a global, genome-wide, scale can be too onerous (Brown *et al.*, 2001). With the recent development of hybridisation-array technology (Schena *et al.*, 1995; Lashkari *et al.*, 1997; Wodicka *et al.*, 1997; Gasch *et al.*, 2000), gene expression studies are increasingly being used for functional discovery, analysis and experimental annotation of reference genomes (Hughes *et al.*, 2000; Shoemaker *et al.*, 2001; Hayes *et al.*, 2002; Bono *et al.*, 2003; Levine and Tjian, 2003; Yamada *et al.*, 2003). Microarrays are also being applied in analyses of differential expression in cells and tissues, gene discovery, drug screening and diagnostics (US Patent 6544742; 2003124581; 2003180774; WO Patent 0218646).

3) Post-transcriptional gene silencing strategies. Nucleic-acid-based inhibitors of gene expression (antisense agents) have been used over the past 25 years. Among the more commons agents currently used are antisense oligonucleotides (ODNs), ribozymes, DNAzymes and RNA interference (RNAi) (Scherer and Rossi, 2003). In the last few years, interest has focused on the capacity of double-stranded RNAs (dsRNAs) to target degradation of complementary sequences, referred as RNA interference (RNAi) (Hannon, 2002; Scherer and Rossi, 2003). The exact mechanisms underlying these effects are the subject of current investigations (Schramke and Allshire, 2003; Volpe *et al.*, 2003). However, despite the existence of possible limitations in the use of antisense agents in different organisms (Cho *et al.*, 2001; Cho-Chung and Becker, 2003; Scherer *et al.*, 2003) the fact that RNAi may elicit targeted specific knockdown of gene expression makes it a valuable tool for studying gene function (Elbashir *et al.*, 2001; Kamath *et al.*, 2003). At this moment RNAi is an emerging field, and significant investments are being made in research and in the development of new applications (Howard, 2003; WO Patent 03070918; 03087371).

4) Strategies at the proteome level. The study of the global set of a cell's proteins (the proteome) constitutes one of the most daunting challenges in biology. The most advanced studies are focused on new techniques for purification and analysis of proteins and protein phosphorylation sites (Rigaut *et al.*, 1999; Brancia *et al.*, 2001; Ficarro *et al.*, 2002; Knight *et al.*, 2003), the study of protein-protein interactions and protein complexes (Uetz *et al.*, 2000; Ito *et al.*, 2001; Gavin *et al.*, 2002; Ho *et al.*, 2002; von Mering *et al.*, 2002) and their subcellular localization (Ghaemmaghami *et al.*, 2003; Huh *et al.*, 2003; Wohlschlegel and Yates, 2003). At the mechanistic level, the studies on how the ubiquitin-proteasome system, particularly the small ubiquitin-related modifier (SUMO) family of proteins controls transcription, and the mechanisms of protein turnover are also subject of increasing study (Pratt *et al.*, 2002; Muratani and Tansey, 2003; Shiio and Eisenman, 2003; Verger *et al.*, 2003). Together with this, the comparison of proteomes between organisms is increasingly being used (Chervitz *et al.*, 1998; Costanzo *et al.*, 2000; Peri *et al.*, 20003; Taylor *et al.*, 2003), with the objective of understanding the protein complexes and protein organization in human cells (Gavin and Superti-Furga, 2003). Finally, in an important breakthrough in the field, the construction of protein microarrays containing a whole proteome on a chip (the yeast proteome) has recently been achieved (Zhu *et al.*, 2001; Zhu and Snyder, 2003; Zhu *et al.*, 2003). This technology, although not exempt from limitations and still subject to improvement (Kodadek, 2001; Mitchell, 2002; Cutler, 2003) opens the possibility of proteome-wide studies, global analyses of protein interactions or activities at a proteomic scale, with direct applications. At this moment, while new strategies and technologies are being studied to overcome the limitations (Ouyang *et al.*, 2003) new applications, including analysis of antibody specificities, yeast protein activities and new antibody arrays, are being developed (Lal *et al.*, 2002; Michaud *et al.*, 2003; WO Patent 0194946; 02092118; 0239120; 03025213; Zhu *et al.*, 2003).

5) Strategies at the metabolome level. The complete pool of cellular metabolites (the metabolome) comprises the whole range of molecules and intermediary metabolites subject to biochemical conversion through the tightly-regulated network of metabolic pathways, in order to generate energy and building blocks for growth and the maintenance of cellular functions. The concentrations and variations in the levels of metabolites directly reflect the metabolic state of the cell, and the metabolome is considered the closest level of analysis to the cell's phenotype (Trethewey *et al.*, 1999; Raamsdonk *et*

al., 2001; Adams, 2003). Metabolic profiling of internal metabolites (metabolic fingerprinting) is currently being used in a wide variety of organisms (yeast, plants, mammalian cells) (Trethewey *et al.*, 1999; Raamsdonk *et al.*, 2001; Fiehn, 2002; Watkins and German, 2002; Castrillo *et al.,* 2003). The main current limitation at this point is the need for sensitive, high-throughput methods for extraction and global analysis of different families of metabolites, and for the screening of large numbers of samples (Oliver *et al.,* 2002; Castrillo *et al.,* 2003 and references therein). In this context, metabolic profiling of metabolites extruded to the medium (metabolic footprinting) may constitute a complementary approach, and new studies are showing the validity of these strategies for specific purposes (e.g. classification of yeast mutants; Kell and Mendes, 2000; Allen *et al.,* 2003). Among the most advanced metabolome studies reported are new methods and strategies for metabolic profiling in plants and mammalian cells (Fiehn *et al.*, 2000; Watkins and German, 2002) and metabolic control analysis and strategies for the elucidation of the function of new genes and metabolic pathways (Teusink *et al.*, 1998; Raamsdonk *et al.*, 2001; Trethewey, 2001; de la Fuente *et al.*, 2002b; Weckwerth and Fiehn, 2002). At the mechanistic level, the role of some amino acids and other specific metabolites in controlling gene expression has also being reported (Fafournoux *et al.*, 2000; Hansen and Johannesen, 2000; So and Crowe, 2000), which opens the field to new advanced studies. Finally, the analysis of the metabolome and use of metabolic profiling techniques has found direct applications in the investigation of molecules for evaluation of health and disease states (biomarkers) for application in diagnostics, and for the screening of new drugs (Griffin *et al.*, 2001; Watkins and German, 2002; US Patent 2003180800; WO Patent 9957306; 0178652; 02057989).

6) Bioinformatic methods. The high volume of information generated at the different functional genomic levels has to be processed efficiently. Advanced bioinformatic strategies, databases, and management methods are progressively being implemented to cope with the requirements of the post-genomic era. The main tools can be summarized as follows:

a) Methods and protocols for efficient extraction of genomic information (e.g. Minimum Information About a Microarray Experiment, MIAME, Brazma *et al.*, 2001; Proteomics Standards Initiative, PSI, Orchard *et al.,* 2003, and Proteomics Experiment Data Repository strategies, PEDRo, Taylor *et al.*, 2003; http://www.mged.org/miame; http://psidev.sourceforge.net; http://pedro.man.ac.uk/home.shtml).

b) Data repositories or data warehouses. Databases for storage and rapid access to raw data. Some examples are: *Saccharomyces* Genome Database (SGD) and other genome databases from other organisms (Table 1); Microarray databases (e.g. Stanford Microarray Database and ArrayExpress repositories; http://genome-www5.Stanford.edu; http://www.ebi.ac.uk/arrayexpress). Yeast, worm, human protein databases (YPD, WormPD, HPRD; Payne and Garrels, 1997; Costanzo *et al.*, 2000; Peri *et al.*, 2003) and biomolecular interaction databases (e.g. BIND, Bader *et al.*, 2003; http://www.blueprint.org/bind/bind.php); Metabolic pathways and metabolic databases, KEGG (http://www.genome.ad.jp/kegg/); MetaFluxNet (Lee *et al.,* 2003b) (see also later section, Table 2).

c) Methods for data annotation and curation under community standards (e.g. Microarray Gene Expression Data standards, MGED, and Proteomics Standards Initiative, PSI; Brazma *et al.,* 2001; Orchard *et al.,* 2003; Hermjakob *et al.,* in press; http://www.mged.org/index.html; http://psidev.sourceforge.net).

d) Strategies for processing of data, normalization and statistical analysis (Fang *et al.*, 2003; Tilstone, 2003).

e) Methods for advanced analysis of genomic data by applying multivariate analysis tools (chemometric methods; e.g. unsupervised or supervised methods, Kell and King, 2000; Bryant *et al.*, 2001) for comparative studies of patterns of expression of internal relationships, and clustering analysis (i. e. gene networks, transcription factor networks, metabolic networks; de la Fuente *et al.*, 2002a; Mendes, 2002; Wu *et al.*, 2002; Fiehn and Weckwerth, 2003).

f) Global management systems of genomic information at the different levels of study. GIMS (Cornell *et al.*, 2003); MAPPFinder and geneMAPP (Doniger *et al.*, 2003). Part of these bioinformatic tools, data management and methods of analysis can be accessed freely (e.g. web resources with free access to all or some specific areas), or restricted to specific applications (e.g. bionformatic methods for specific studies on microarrays and analysis of metabolic profiles; EPO Patent 1300778; WO Patent 0140896).

Comprehensive Methods of Analysis Require Rigor in Experimental Design

The comprehensive scope and high-throughput nature of functional genomic analyses have caused understandable excitement in the research community. However the global character and speed of these analyses require investigators to take particular care in the design of experiments in order to remove confounding variables and obtain only data that are relevant to the biological questions being posed. Not all post-genomic analyses have shown sufficient rigor in experimental design and data analysis. Moreover, some of the methods applied in post-genomic studies are still in their first stages of development, and may need refinement and proper validation in touchstone models under controlled conditions (see next section). All of us need to be careful not to overlook these problems in our post-genomic studies. The most relevant problems include:

a) Conceptual. Genes are defined by function and there are other functional elements in the DNA sequence besides the reading frames (Dillon, 2003; Pearson, 2003). Some studies have shown discrepancies between mRNA abundance and protein content, or enzyme activity (Gygi *et al.,* 1999; ter

Table 2. Yeast resources, databases and tools for global analysis of genomic data

Resources	Links. References.
Genome databases	
S. cerevisiae genome database	http://www.yeastgenome.org
Munich Information Centre for Protein Sequences (MIPS)	http://mips.gsf.de/genre/proj/yeast/index.jsp
S. cerevisiae mutant collection (EUROSCARF)	http://www.uni-frankfurt.de/fb15/mikro/euroscarf/complete.html
S. cerevisiae resource center	http://www.depts.washington.edu/~yeastrc
S. pombe project	http://www.sanger.ac.uk/Projects/S_pombe
Gene expression resources (transcriptome)	
Stanford Microarray Database	http://genome-www5.stanford.edu
Yeast transcription factors and related components (YTF)	http://biochemie.web.med.uni-muenchen.de/YTFD/
S. cerevisiae promoter database (SCPD)	http://cgsigma.cshl.org/jian/
ArrayExpress	http://www.ebi.ac.uk/arrayexpress
Microarray standards (MIAME)	http://www.mged.org/miame
Proteome resources	
Yeast Proteome Database YPD (Incyte)	http://www.incyte.com/control/researchproducts/insilico/proteome
Protein interactions	(von Mering *et al.,* 2002)
Yeast proteins localization	http://yeastgfp.ucsf.edu/
Yeast protein microarrays	(Zhu and Snyder, 2003)
Human protein reference database (HPRD)	(Peri *et al.,* 2003) http://www.hprd.org/
Proteomics Standards Initiative	http://psidev.sourceforge.net
Metabolic pathways databases	
KEGG	http://www.genome.ad.jp/kegg/
BioCyC	http://biocyc.org
PathDB	http://www.ncgr.org/pathdb/
RIKEN	http://genome.gsc.riken.go.jp/DNA-Book/metabolome.shtml
Biomolecular interactions. BIND	http://www.blueprint.org/bind/bind.php
Bioinformatics. Tools for analysis and management of global genomic information	
GeneOntology; FatiGO	http://www.geneontology.org http://fatigo.bioinfo.cnio.es/
GenMAPP; MAPPFinder	http://www.genmapp.org/MAPPFinder.html
PEDRo	http://pedro.man.ac.uk/home.shtml
MetaFluxNet™	(Lee *et al.,* 2003b) http://mbel.kaist.ac.kr/mfn/
KEGG	http://www.genome.ad.jp/kegg/
Genome Information Management System (GIMS)	(Cornell *et al.,* 2003) http://www.cs.man.ac.uk/img/gims/

Kuile, 1999; Glanemann *et al.,* 2003). Thus, the existence of mechanisms or regulation at the post-transcriptional, post-translational, and metabolic levels should not be ignored (Fafournoux *et al.,* 2000; Hansen and Johannesen, 2000; So and Crowe, 2000; Muratani and Tansey, 2003; Schramke and Allshire, 2003; Verger *et al.,* 2003; Volpe *et al.,* 2003).

b) Methodological. Culture systems. Different cultivation systems (Petri dishes, flasks, fermenters, tissue cultures) and culture media (complex, defined) have different advantages and limitations, offering a range of possibilities for their use in post-genomic studies. Thus, as an example, rich complete or complex media formulations can be used in classical microbiological and molecular biological studies at a small scale (Petri dishes, flasks). However, in careful physiological studies at a higher scale, a knowledge of the identity and concentration of all components in the medium is necessary, and synthetic defined media (e.g., Baganz *et al.,* 1997) are recommended. Moreover, when using complex media at this scale, simple factors such as the autoclaving procedure, manufacturer, and even batch reproducibility from the same manufacturer can result in different medium composition (Chatterjee *et al.,* 2001) and may seriously compromise the reproducibility of results and the conclusions derived from them, making a proper comparison of data between different laboratories impossible.

In liquid cultures, there are also intrinsic factors that have to be taken into account. One of them is the existence of an intrinsic variability within a population of microbial cells, with cells growing at different growth rates, or being at different phases of their division cycles. The existence of this heterogeneity has led some authors to recommend the implementation of new strategies to undertake analyses at the 'single cell' level (Lidstrom and Meldrum, 2003). However, multiple small-scale batch cultures (Lasko *et al.,* 2000;

Warringer and Blomberg, 2003), and other systems strategies, such as continuous culture systems (Castrillo and Ugalde, 1994; Weusthuis *et al.*, 1994; Lane *et al.*, 1999; ter Linde *et al.*, 1999), competition experiments (Baganz *et al.*, 1998; Giaever *et al.*, 2002) and synchronization of cultures for cell cycle studies (Spellman *et al.*, 1998; Walker, 1999; Muller *et al.*, 2003) are extensively used.

The new high-throughput methods of post-genomic analysis at the levels of transcriptome, proteome and metabolome present the possibility of generating comprehensive sets of data. However, the considerable cost of these systems (e.g. microarrays) may limit the number of experiments that can be performed. A careful selection of the objectives, including a decision about if statistical analysis needs to be included, has to be the highest priority. In some cases, the high cost can make it tempting to reduce the number of replicate experiments or analyses. However, this can compromise the interpretation of the results, if good reproducibility, accuracy, and a proper statistical treatment are necessary (Hayes *et al.*, 2002; Tilstone, 2003).

c) In the technological context, finally, some sources of variability and limitations are still present in studies of gene-expression data (microarrays for transcriptome studies) (Grunenfelder and Winzeler, 2002; Spruill *et al.*, 2002), and appropriate strategies, controlled conditions and quality control methods are strongly recommended (Hayes *et al.*, 2002; US Patent 2003170672). Existence of possible nonsequence specific effects in antisense RNAs strategies has also been described (Cho *et al.*, 2001; Cho-Chung and Becker, 2003; Kamath *et al.*, 2003). Some limitations in protein-protein interaction studies (von Mering *et al.*, 2002) and protein microarrays (Kodadek, 2001; Mitchell, 2002) have also been reported, and new methods for analysis of a wider range of families of metabolites are under study (Castrillo *et al.*, 2003).

Finally, in the context of bioinformatic studies, there is a general consensus on the necessity for application of uniform methods under international standards (e.g. MIAME and Proteomics Standards Initiative, PSI; PEDRo, Brazma *et al.*, 2001, Orchard *et al.*, 2003 Taylor *et al.*, 2003; http://www.mged.org/miame; http://psidev.sourceforge.net; http://pedro.man.ac.uk/home.shtml) and the adoption of new strategies for proper filtering and curation of the data. All of this suggests that it is prudent to employ a number of independent analytical techniques in functional genomic studies (e.g. in transcriptome studies, microarray analysis combined with Northern, quantitative real-time PCR or differential display; Brown *et al.*, 2001; Levine and Tjian, 2003). The congruence of results between independent analyses or between different strategic approaches (e.g. transcriptome and metabolome; Urbanczyk-Wochniak *et al.*, 2003) will permit a truly integrative perspective (see last section of the review) and may constitute the most useful approach to obtain reliable conclusions in post-genomic studies.

Yeast as a 'Touchstone' Model in Post-genomic Research

Saccharomyces cerevisiae exhibits a number of favourable characteristics, which have made it an optimal system for genetic, molecular and physiological studies. The most relevant are: simple methods of cultivation under controlled conditions, well characterized genetics and facile techniques of genetic manipulation. Many basic cellular mechanisms and biochemical processes were first elucidated in yeast, and a wide knowledge of the genetics, biochemistry and physiology of this yeast is presently available (Rose and Harrison, 1987-1995).

Together with this, *S. cerevisiae* was the first eukaryotic organism for which the whole genome sequence was completed (Goffeau *et al.*, 1996). The knowledge of the sequence, combined with the existence of a comprehensive collection of gene deletion mutants (http://www.uni-frankfurt.de/fb15/mikro/euroscarf/complete.html), and new high-throughput technologies for global analyses at a genome-wide scale are rapidly extending the range of applications. Yeast is being increasingly used as a model organism in post-genomic studies. As an example of this, the performance of the majority of new genome-wide methods are first validated in yeast, using well-controlled reproducible conditions. To exemplify this, the most relevant methods and post-genomic strategies investigated using yeast as a model system are summarized here: a) At the level of the genome, the utilization of genetic methods, gene disruption and construction of a collection of deletion mutants for functional profiling and characterization (Ross-MacDonald *et al.*, 1999; Winzeler *et al.*, 1999; Giaever *et al.*, 2002), together with new methods of gene transcomplementation between human cells and yeasts, for the elucidation of the function of human genes (Zhang *et al.*, 2003). b) At the gene expression (transcriptome) level, yeast microarrays are used extensively (Lashkari *et al.*, 1997; Wodicka *et al.*, 1997; Spellman *et al.*, 1998; Hayes *et al.*, 2002). In this field, transcriptional responses and patterns of expression of yeast under carefully controlled conditions (e.g. chemostat culture) are progressively being studied (ter Linde *et al.*, 1999; Hayes *et al.*, 2002; Boer *et al*, 2003; Lim *et al.*, 2003). c) At the proteome level. The first whole-proteome microarray has been developed in yeast (Mitchell, 2002; Michaud *et al.*, 2003; Zhu and Snyder, 2003; Zhu *et al.*, 2003) as well as new strategies for the preparation of proteins arrays (Washburn, 2003). The first studies, reported in yeast, on subcellular localization of proteins on a proteome-wide scale (Ghaemmaghami *et al.*, 2003; Huh *et al.*, 2003), phosphoproteome analysis (Ficarro *et al.*, 2002; Knight *et al.*, 2003) and protein-protein interaction maps (Uetz *et al.*, 2000; Ito *et al.*, 2001; Gavin *et al.*, 2002; Ho *et al.*, 2002). Also, a new strategy for the investigation of enzymatic activities associated to specific metabolic pathways has been recently reported (Chen *et al.*, 2003), as well as the role of post-translational effects (protein turnover) as an overlooked

dimension in proteomics (Pratt *et al.,* 2002). d) At the metabolome level. New methods for analysis of yeast metabolites (Castrillo *et al.,* 2003), strategies to ascribe function to unknown genes (Raamsdonk *et al.,* 2001) and classification of yeast mutants using metabolic footprinting (Allen *et al.,* 2003). e) Some of the best examples may be found at the metabolic control and bioinformatic level, such as the development of new machine learning methods for the analysis of transcriptome, proteome and metabolome data and for the study of regulatory networks (Kell and Mendes, 2000; de la Fuente *et al.,* 2002a; Mendes, 2002; Fiehn and Weckwerth, 2003). Many of these advanced resources were first conceived for use in yeast (Payne and Garrels, 1997; Cornell *et al.,* 2003). A summary of the most relevant resources, databases and methods for global analysis of yeast genomic data is presented in Table 2.

In the post-genomic era, apart from the development of new methods, yeasts can be used in fundamental research for the elucidation of new mechanisms of regulation. The present perspective is that, although a few processes may work differently in yeast and mammalian cells (Zeng *et al.,* 2001; Conlon and Raff, 2003), the evidence points to a high degree of conservation of the fundamental processes in all eukaryotic cells (Rose and Harrison, 1987-1995; Gould and Nurse, 1989; Hartwell and Weinert, 1989; Boube *et al.,* 2002), and yeast is regularly used as a model system for the study of new regulatory mechanisms. In this context, examples of the most relevant studies are: a) new insights into the mechanisms of transcriptional regulation of gene expression (Gancedo, 1998; Levine and Tjian, 2003; Mosley *et al.,* 2003), mechanisms of nutrient sensing and signal transduction pathways (Rohde and Cardenas, 2003; Tate and Cooper, 2003) and, in fission yeast, mechanisms of pre-mRNA splicing (Kuhn and Käufer, 2003), DNA repair and genomic stability (Ahmad *et al.,* 2002), and RNA interference (Raponi and Arndt, 2003; Volpe *et al.,* 2003). b) In order to explain the lack of correlation between transcriptional patterns and protein levels, or *in vivo* fluxes (Gygi *et al.,* 1999; Glanemann *et al.,* 2003; Daran-Lapujade *et al.,* in press), new mechanisms are being studied, such as protein turnover (Pratt *et al.,* 2002), the role of the ubiquitin-proteasome system and histone deubiquitination in transcriptional regulation (Lipford and Deshaies, 2003; Muratani and Tansey, 2003; Verger *et al.,* 2003; Daniel *et al.,* in press; Dong and Xu, in press) and the role of specific metabolites in the regulation of gene expression (Hansen and Johannesen, 2000; So and Crowe, 2000; Muller *et al.,* 2003).

In summary, the simplest eukaryote, *S. cerevisiae*, has been used extensively as a reference model in the past, in the elucidation of the fundamental molecular biology and biochemistry of the eukaryotic cell (Rose and Harrison, 1987-1995). The optimal characteristics of this organism, its potential and validity in genome-wide studies at the different cellular levels (genome, transcriptome, proteome and metabolome) under controlled conditions, makes it an optimum model, both for the implementation of new advanced strategies and for the study of new mechanisms of regulation at the different levels of cellular complexity. Yeast is being used presently as a model organism to study the cell cycle and checkpoints (Gould and Nurse, 1989; Hartwell and Weinert, 1989; Spellman *et al.,* 1998); cell polarity (Chang and Peter, 2003); mechanisms of evolution and speciation (Delneri *et al.,* 2003); ageing and extension of lifespan (Howitz *et al.,* 2003); mechanisms of infection and propagation of prions (Fernandez-Bellot and Cullin, 2001; Bach *et al.,* 2003; Kryndushkin *et al.,* 2003; Sherman and Muchowski, 2003), and as a model to gain insight into the molecular pathology of neurodegenerative diseases (Outeiro and Lindsquit, 2003; Willingham *et al.,* 2003).

With this perspective, yeast is expected to continue providing new knowledge and insights in cell biology, as a relevant touchstone in the post-genomic era. In order to reach this goal, integrative studies are necessary, combining relevant information obtained from a single organism at the different functional genomic levels. In this context, yeast researchers are in a privileged position, offering them an excellent opportunity to employ new comprehensive integrative studies, whose knowledge can be related to information from other organisms, towards the objective of a better understanding of the cell biology of more complex systems (e.g. fly, plants, mouse, and human).

New Integrative Studies: Towards Systems Biology

The high volumes of complex data being generated in functional genomics studies may make the analysis and interpretation of results difficult and recommends the use of more comprehensive integrative studies in the post-genomic era (Delneri *et al.,* 2001; Oliver *et al.,* 2002; Phelps *et al.,* 2002; Urbanczyk-Wochniak *et al.,* 2003; WO Patent 03058238).

Systems biology focuses on the importance of a global integrative view of biological processes, including new holistic approaches to elucidate cell complexity by combining analysis of data sets and results obtained from different parallel approaches under controlled conditions (Kitano, 2002; Peri *et al.,* 2003). For these purposes the development of new tools and methods to link information from different parallel analyses is of central importance (Yao, 2002; Weckwerth, 2003). The majority of these approaches can be sub-divided into three levels of analysis: 1) Advanced methods for extraction and direct comparison of basic genomic information. Thus, for example, comparative genomics for the direct comparison of conserved and non-conserved sequences (Kellis *et al.,* 2003; Salzberg, 2003), or proteome organization between species (Costanzo *et al.,* 2000; Gavin and Superti-Furga, 2003) using databases and tools for the global management of data (Table 2). 2) New methods and strategies for elucidating complex regulatory networks at each specific level of analysis (genome,

transcriptome, proteome and metabolome), and exploring the intricate interrelations between them. For this purpose, new algorithms, knowledge discovery methods, and advanced tools for *in silico* analysis of specific patterns (of gene expression, proteome interactions, or internal metabolite patterns) can be used (de la Fuente *et al.,* 2002a; Mendes, 2002; Wu *et al.,* 2002; Fiehn and Weckwerth, 2003). These studies are mainly focused on the linking and clustering of specific patterns, based on different criteria (unsupervised or supervised methods). These can be used not only for the elucidation of possible roles of ORFs of unknown function, but also to identify the main regulatory network or regulatory level (e.g. regulation at the transcriptional, translational, or post-translational level) responsible for the global physiological response of the cell (phenotype) under specific environmental conditions (Bryant *et al.,* 2001; Fiehn, 2001; Wu *et al.,* 2002; Fiehn and Weckwerth, 2003; Sandelin *et al.,* 2003) and the hierarchical organization of such networks in the cell (ter Kuile and Westerhoff, 2001; Ihmles *et al.,* in press). These studies on transcriptome, proteome and metabolic networks can provide crucial information and clues for the development of new drug targets and therapeutic strategies, but are critically dependent on the accuracy and reliability of the experiments and the raw data they generate. 3) Methodical, organized strategies combining two or more levels of functional genomic analysis, or new advanced techniques, can provide insights into mechanisms or role of unknown genes. Good examples of this are: The parallel study and global analysis of transcriptome and proteome (Lee *et al.,* 2003a) and the identification of genes associated with specific protein activities from defined metabolic pathways (Chen *et al.,* 2003).

In the case of yeast, its favourable characteristics make it an optimum platform for integrative studies. Relevant examples are the integrated approaches to identify previously overlooked genes (Kumar *et al.,* 2002) and the studies of the hierarchical organization of transcriptome and metabolome networks in *S. cerevisiae* (ter Kuile and Westerhoff, 2001; Ihmels *et al.,* in press). These, and other studies already referred in this review, illustrate the huge potential of yeast as a touchstone model for integrative studies in the post genomics era (Oliver, 1997, 1998; Oliver *et al.,* 1998; Delneri and Oliver, 2000; Delneri *et al.,* 2001).

Conclusions. Future Perspectives

The initial excitement that the completion of the human genome sequence generated, not only among scientists but also among the general public, was based mainly on the new information and the huge possibilities that this knowledge could provide. Despite, or perhaps because, the human genome contained fewer genes than expected, the task of determining how these genes act and interact to build a human being remains a daunting one.

Even the complete solution of gene action and interaction to determine the workings of a yeast cell is a major challenge. A comprehensive and integrative approach is required, first studying the yeast cell at the different levels of complexity (genome, transcriptome, proteome and metabolome) and the regulatory networks of each level and their interrelation in the global system, using advanced bioinformatic methods and strategies towards the understanding of the cellular system as a global entity. Fortunately, the yeast *Saccharomyces cerevisiae,* as the first eukaryote to have its genome sequenced, is a superb 'touchstone' for post-genomic studies, and constitutes the best system in which to develop the required integrative or systems biology approach. These techniques and conceptual approaches developed with yeast, and the information obtained, will be invaluable in the analysis of progressively more complex organisms.

Acknowledgments This work was supported by the BBSRC (COGEME Consortium) and by an EC contract to SGO within the frame of the Garnish network of FP5.

References

Adams, A. (2003) Metabolomics: Small-molecule omics. *The Scientist* **17**, 38-40.

Adams, M. D., Celniker, S. E., Holt, R. A., Evans, C. A., Gocayne, J. D., Amanatides, P. G., Scherer, S. E., Li, P. W., Hoskins, R. A., Galle, R. F. *et al.* (2000) The genome sequence of *Drosophila melanogaster. Science* **287**, 2185-2195.

Ahmad, F., Kaplan, C. D. and Stewart, E. (2002). Helicase activity is only partially required for *Schizosaccharomyces pombe* Rqh1p function. *Yeast* **19**, 1381-1398.

Allen, J., Davey, H. M., Broadhurst, D., Heald, J. K., Rowland, J. J., Oliver, S. G. and Kell, D. B. (2003) High-throughput classification of yeast mutants using metabolic footprinting. *Nat. Biotechnol.* **21**, 692-696.

Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* **408**, 796-815.

Bach, S., Talarek, N., Andrieu, T., Vierfond, J. M., Mettey, Y., Galons, H., Dormont, D., Meijer, L., Cullin, C. and Blondel, M. (2003) Isolation of drugs active against mammalian prions using a yeast-based screening assay. *Nat. Biotechnol.* **21**, 1075-1081.

Bader, G. D., Betel, D. and Hogue, C. W. V. (2003) BIND: the biomolecular interaction network database. *Nucleic Acids Res.* **31**, 248-250.

Baganz, F., Hayes, A., Marren, D., Gardner, D. C. J. and Oliver, S. G. (1997) Suitability of replacement markers for functional analysis studies in *Saccharomyces cerevisiae*. *Yeast* **13**, 1563-1573.

Baganz, F., Hayes, A., Farquhar, R., Butler, P. R., Gardner, D. C. J. and Oliver, S. G. (1998) Quantitative analysis of yeast gene function using competition experiments in continuous culture. *Yeast* **14**, 1417-1427.

Bahls, C., Weitzma, J. and Gallaguer, R. (2003). Model organisms. *The Scientist* **17**, Suppl. 1.

Blaxter, M. (2003) Comparative genomics: Two worms are better than one. *Nature* **426**, 395-396.

Boer, V. M., de Winde, J. H., Pronk, J. T. and Piper, M. D. W. (2003). The genome-wide transcriptional responses of *Saccharomyces cerevisiae* grown on glucose in aerobic chemostat cultures limited for carbon, nitrogen, phosphorus, or sulfur. *J. Biol. Chem.* **278**, 3265-3274.

Bono, H., Yagi, K., Kasukawa, T., Nikaido, I., Tominaga, N., Miki, R., Mizuno, Y., Tomaru, Y., Goto, H., Nitanda, H. *et al.* (2003) Systematic expression profiling of the mouse transcriptome using RIKEN cDNA microarrays. *Genome Res.* **13**, 1318-1323.

Boube, M., Joulia, L., Cribbs, D. L., Bourbon, H. M. (2002) Evidence for a mediator of RNA polymerase II transcriptional regulation conserved from yeast to man. *Cell* **110**, 143-151.

Brancia, F. L., Butt, A., Beynon, R. J., Hubbard, S. J., Gaskell, S. J. and Oliver, S. G. (2001) A combination of chemical derivatisation and improved bioinformatic tools optimises protein identification for proteomics. *Electrophoresis* **22**, 552-559.

Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C. A., Causton, H. C., *et al.* (2001) Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. *Nature Genetics* **29**, 365-371.

Brown, A. J. P., Planta, R. J., Restuhadi, F., Bailey, D. A., Butler, P. R., Cadahia, J. L., Cerdan, M. E., De Jonge, M., Gardner, D. C. J., Gent, M. E. *et al.* (2001) Transcript analysis of 1003 novel yeast genes using high-throughput northern hybridisations. *EMBO J.* **20**, 3177-3186**.**

Bryant, C. H., Muggleton, S. H., Oliver, S. G., Kell, D. B., Reiser, P. and King, R. D. (2001) Combining active inductive programming, active learning and robotics to discover the function of genes. *Electron. Trans. Artif. Intell.* **5**, 1-36.

Castrillo, J. I. and Ugalde, U. O. (1994) A general model of yeast energy metabolism in aerobic chemostat culture. *Yeast* **10**, 185-197.

Castrillo, J. I., Hayes, A., Mohammed, S., Gaskell, S. J. and Oliver, S. G. (2003) An optimised protocol for metabolome analysis in yeast using direct infusion electrospray mass spectrometry. *Phytochemistry* **62**, 929-937.

Cech, T. R. (2001) The genes we share with yeast, flies, worms and mice. New clues to human health and disease. *8th Biochemical Science Report Howard Hughes Medical Institute*. On-line publication (http://www.hhmi.org).

Chang, F. and Peter, M. (2003) Yeasts make their mark. *Nature Cell Biol.* **5**, 294-299.

Chatterjee, M. T., Khalawan, S. A. and Curran, B. P. (2001). Subtle alterations in growth medium composition can dramatically alter the percentage of unsaturated fatty acids in the yeast *Saccharomyces cerevisiae*. *Yeast* **18,** 81-88.

Chen, C. N., Porubleva, L., Shearer, G., Svrakic, M., Holden, L. G., Dover, J. L., Johnston, M., Chitnis, P. R. and Kohl, D. H. (2003) Associating protein activities with their genes: rapid identification of a gene encoding a methylglyoxal reductase in the yeast *Saccharomyces cerevisiae*. *Yeast* **20**, 545-554.

Chervitz, S. A., Aravind, L., Sherlock, G., Ball, C. A., Koonin, E. V., Dwight, S. S., Harris, M. A., Dolinski, K., Mohr, S., Smith, T., Weng, S., Cherry, J. M. and Botstein, D. (1998) Comparison of the complete protein sets of worm and yeast: orthology and divergence. *Science* **282**, 2022-2028.

Cho, Y. S., Kim, M. K., Cheadle, C., Neary, C., Becker, K. G. and Cho-Chung, Y. S. (2001) Antisense DNAs as multisite genomic modulators identified by DNA microarray. *Proc. Natl. Acad. Sci. USA* **98**, 9819-9823.

Cho-Chung, Y. S. and Becker, K. G. (2003) A genome-wide view of antisense. *Nat. Biotechnol.* **21**, 492.

Cliften, P., Sudarsanam, P., Desikan, A., Fulton, L., Fulton, B., Majors, J., Waterston, R., Cohen, B. A. and Johnston, M. (2003) Finding functional features in *Saccharomyces* genomes by phylogenetic footprinting. *Science* **301**, 71-76.

Conlon, I. and Raff, M. (2003) Differences in the way a mammalian cell and yeast cells coordinate cell growth and cell-cycle progression. *J. Biol.* **2**, 7. On-line publication (http://jbiol.com/content/2/1/7).

Cornell, M., Paton, N. W., Hedeler, C., Kirby, P., Delneri, D., Hayes, A. and Oliver, S. G. (2003) GIMS: an integrated data storage and analysis environment for genomic and functional data. *Yeast* **20**, 1291-1306.

Costanzo, M. C., Hogan, J. D., Cusick, M. E., Davis, B. P., Fancher, A. M., Hodges, P. E., Kondu, P., Lengieza, C., Lew-Smith, J. E., Lingner, C., Roberg-Perez, K. J., Tillberg, M., Brooks, J. E. and Garrels, J. I. (2000) The yeast proteome database (YPD) and *Caenorhabditis elegans* proteome database (WormPD): comprehensive resources for the organization and comparison of model organism protein information. *Nucleic Acids Res.* **28**, 73-76.

Cutler, P. (2003) Protein arrays: the current state-of-the-art. *Proteomics* **3**, 3-18.

Daniel, J. A., Torok, M. S., Sun, Z. W., Schieltz, D., Allis, C. D., Yates, J. R. and Grant, P. A. (2004) Deubiquitination of histone H2B by a yeast acetyltransferase complex regulates transcription. *J. Biol. Chem.*, in press. doi:10.1074/jbc.C300494200.

Daran-Lapujade, P., Jansen, M. L. A., Daran, J. M., van Gulik, W., de Winde, J. H. and Pronk, J. T. (2004) Role of transcriptional regulation in controlling fluxes in central carbon metabolism of *Saccharomyces cerevisiae*, a chemostat culture study. *J. Biol. Chem.*, in press. doi:10.1074/jbc.M309578200.

Delneri, D. and Oliver, S. G. (2000) Functional genomics with yeast; in *The Yeast Nucleus: Frontiers in Molecular Biology*, Fantes, P. A. and Beggs, J. D. (eds.), pp. 1-18, Oxford University Press. Oxford, UK.

Delneri, D., Brancia, F. L. and Oliver, S. G. (2001) Towards a truly integrative biology through the functional genomics of yeast. *Curr. Opin. Biotechnol.* **12**, 87-91.

Delneri, D., Colson, I., Grammenoudi, S., Roberts, I. N., Louis, E. J. and Oliver, S. G. (2003) Engineering evolution to study speciation in yeasts. *Nature* **422**, 68-72.

Dermitzakis, E. T., Reymond, A., Scamuffa, N., Ucla, C., Kirkness, E., Rossier, C. and Antonarakis, S. E. (2003) Evolutionary discrimination of mammalian conserved non-genic sequences (CNGs) *Science* **302**, 1033-1035.

Dillon, N. (2003) Positions, please. *Nature* **425**, 457-458.

Dong, L. and Xu, C. W. (2004) Carbohydrates induce mono-ubiquitination of H2B in yeast. *J. Biol. Chem.*, in press. doi 10.1074/jbc.C300505200.

Doniger, S. W., Salomonis, N., Dahlquist, K. D., Vranizan, K., Lawlor, S. C. and Conklin, B. R. (2003) MAPPFinder: using gene ontology and genMAPP to create a global gene-

expression profile from microarray data. *Genome Biol.* **4**, R7. On-line publication (http://genomebiology.com/2003/4/1/R7).

Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K. and Tuschl, T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature* **411**, 494-498.

European Patent Office (2003) EP Patent 1300778. Microarray data warehouse.

Fafournoux, P., Bruhat, A. and Jousse, C. (2000) Amino acid regulation of gene expression. *Biochem. J.* **351**, 1-12.

Fang, Y., Brass, A., Hoyle, D. C., Hayes, A., Bashein, A., Oliver, S. G., Waddington, D. and Rattray, M. (2003) A model-based analysis of microarray experimental error and normalisation. *Nucleic Acids Res*. **31**, e96.

Fernandez-Bellot, E. and Cullin, C. (2001) The protein-only theory and the yeast *Saccharomyces cerevisiae*: the prions and the propagons. *Cell. Mol. Life Sci.* **58**, 1857-1878.

Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F. and White, F. M. (2002) Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat. Biotechnol.* **20**, 301-305.

Fiehn, O. (2001) Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comp. Funct. Genomics* **2**, 155-168.

Fiehn, O. (2002) Metabolomics - the link between genotypes and phenotypes. *Plant Mol. Biol.* **48**, 155-171.

Fiehn, O. and Weckwerth, W. (2003) Deciphering metabolic networks. *Eur. J. Biochem.* **270**, 579-588.

Fiehn, O., Kopka, J., Dormann, P., Altmann, T., Trethewey, R. N. and Willmitzer, L. (2000) Metabolite profiling for plant functional genomics. *Nat. Biotechnol.* **18**, 1157-1161.

de la Fuente, A., Brazhnik, P. and Mendes, P. (2002a) Linking the genes: inferring quantitative gene networks from microarray data. *Trends Genet.* **18**, 395-398.

de la Fuente, A., Snoep J. L., Westerhoff, H. V. and Mendes, P. (2002b) Metabolic control in integrated biochemical systems. *Eur. J. Biochem.* **269**, 4399-4408.

Gancedo, J. M. (1998) Yeast carbon catabolite repression. *Microbiol. Mol. Biol. Rev.* **62**, 334-361.

Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., Botstein, D. and Brown, P. O. (2000) Genomic expression programs in the response of yeast cells to environmental changes. *Mol. Biol. Cell*. **11**, 4241-4257.

Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J. M., Michon, A. M., Cruciat, C. M. *et al.* (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* **415**, 141-147.

Gavin, A. C. and Superti-Furga, G. (2003) Protein complexes and proteome organization from yeast to man. *Curr. Opin. Chem. Biol.* **7**, 21-27.

Ghaemmaghami, S., Huh, W. K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O'Shea, E. K. and Weissman, J. S. (2003) Global analysis of protein expression in yeast. *Nature* **425**, 737-741.

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B. *et al.* (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**, 387-391.

Glanemann, C., Loos, A., Gorret, N., Willis, L. B., O'Brien, X. M., Lessard, P. A. and Sinskey, A. J. (2003). Disparity between changes in mRNA abundance and enzyme activity in *Corynebacterium glutamicum* and implications for DNA microarray analysis. *Appl. Microbiol. Biotechnol.* **61**, 61-68.

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M. *et al.* (1996) Life with 6000 genes. *Science* **274**, 546-567.

Goffeau, A., Aert, R., Agostini-Carbone, M. L., Ahmed, A., Aigle, M., Alberghina, L., Albermann, K., Albers, M., Aldea, M., Alexandraki, D., *et al.* (1997) The yeast genome directory. *Nature* **387**. Suppl. 5. 5-105.

Gould, K. L. and Nurse, P. (1989). Tyrosine phosphorylation of the fission yeast cdc2+ protein kinase regulates entry into mitosis. *Nature* **342**, 39-45.

Griffin, J. L., Williams, H. J., Sang, E., Clarke, K., Rae, C. and Nicholson, J. K. (2001) Metabolic profiling of genetic disorders: a multitissue (1)H nuclear magnetic resonance spectroscopic and pattern recognition study into dystrophic tissue. *Anal. Biochem.* **293**, 16-21.

Grunenfelder, B. and Winzeler, E. A. (2002) Treasures and traps in genome-wide data sets: case examples from yeast. *Nature Rev. Genet.* **3**, 653-661.

Guigo, R., Dermitzakis, E. T., Agarwal, P., Ponting, C. P., Parra, G., Reymond, A., Abril, J. F., Keibler, E., Lyle, R., Ucla, C., Antonarakis, S. E. and Brent, M. R. (2003) Comparison of mouse and human genomes followed by experimental verification yields an estimated 1,019 additional genes. *Proc. Natl. Acad. Sci. USA* **100**, 1140-1145.

Gygi, S. P., Rochon, Y., Franza, B. R. and Aebersold, R. (1999) Correlation between protein and mRNA abundance in yeast. *Mol. Cell. Biol.* **19**, 1720-1730.

Hannon, G. (2002) RNA interference. *Nature* **418**, 244-251.

Hansen, J. and Johannesen, P. F. (2000) Cysteine is essential for transcriptional regulation of the sulfur assimilation genes in *Saccharomyces cerevisiae*. *Mol. Gen. Genet.* **263**, 535-542.

Hartwell, L. H. and Weinert, T. A. (1989). Checkpoints: controls that ensure the order of cell cycle events. *Science* **246**, 629-634.

Hayes, A., Zhang, N., Wu, J., Butler, P. R., Hauser, N. C., Hoheisel, J. D., Lim, F., Sharrocks, A. D. and Oliver, S. G. (2002) Hybridization array technology coupled with chemostat culture: tools to interrogate gene expression in *Saccharomyces cerevisiae. Methods* **26**, 281-290.

Hermjakob, H., Montecchi-Palazzi, L., Bader, G., Wojcik, J., Salwinski, L., Moore, S., Orchard, S., Sarkans, U., von Mering, C., Roechert, B. *et al.* (2004) The HUPO PSI Molecular Interaction Format - A community standard for the representation of protein interaction data. *Nature Biotechnology*, in press.

Ho, Y., Gruhler, A., Heilbut, A., Bader, G. D., Moore, L., Adams, S. L., Millar, A., Taylor, P., Bennett, K., Boutilier, K. *et al.* (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* **415**, 180-183.

Howard, K. (2003) Unlocking the money-making potential of RNAi. *Nat. Biotechnol.* **12**, 1441-1446.

Howitz, K. T., Bitterman, K. J., Cohen, H. Y., Lamming, D. W., Lavu, S., Wood, J. G., Zipkin, R. E., Chung, P., Kisielewski, A., Zhang, L. L., Scherer, B. and Sinclair, D. A. (2003) Small molecule activators of sirtuins extend *Saccharomyces cerevisiae*

lifespan. *Nature* **425**, 191-196.

Hughes, T. R., Marton, M. J., Jones, A. R., Roberts, C. J., Stoughton, R., Armour, C. D., Bennett, H. A., Coffey, E., Dai, H., He, Y. D. *et al.* (2000) Functional discovery via a compendium of expression profiles. *Cell* **102**, 109-126.

Huh, W. K., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S. and O'Shea, E. K. (2003) Global analysis of protein localization in budding yeast. *Nature* **425**, 686-691.

Ihmels, J., Levy, R. and Barkai, N. (2004) Principles of transcriptional control in the metabolic network of *Saccharomyces cerevisiae*. *Nat. Biotechnol*., in press. doi:10.1038/nbt918.

Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M. and Sakaki, Y. (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. USA.* **98**, 4569-4574.

Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., Welchman, D. P., Zipperlen, P. and Ahringer, J. (2003) Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. *Nature* **421**, 231-237.

Kell, D. B. and King, R. D. (2000) On the optimization of classes for the assignment of unidentified reading frames in functional genomics programmes: the need for machine learning. *Trends Biotechnol.* **18**, 93-98.

Kell, D. B. and Mendes, P. (2000) Snapshots of systems: metabolic control analysis and biotechnology in the post-genomic era; in *Technological and Medical Implications of Metabolic Control Analysis*, Cornish-Bowden, A. and Cardenas, M. L. (eds.), pp. 3-25, Kluwer Academic Publishers, Dordrecht, Netherlands.

Kellis, M., Patterson, N., Endrizzi, M., Birren, B. and Lander, E. S. (2003) Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423**, 241-254.

Kitano, H. (2002) Systems biology: a brief overview. *Science* **295**, 1662-1664.

Knight, Z. A., Schilling, B., Row, R. H., Kenski, D. M., Gibson, B. W. and Shokat, K. M. (2003) Phosphospecific proteolysis for mapping sites of protein phosphorylation. *Nat. Biotechnol.* **21**, 1047-1054.

Kodadek, T. (2001) Protein microarrays: prospects and problems. *Chem. Biol.* **8**, 105-115.

Kryndushkin, D. S., Alexandrov, I. M., Ter-Avanesyan, M. D. and Kushnirov, V. V. (2003) Yeast [PSI+] prion aggregates are formed by small Sup35 polymers fragmented by Hsp104. *J. Biol. Chem.* **278**, 49636-49643.

Kuhn, A. N. and Käufer, N. F. (2003) Pre-mRNA splicing in *Schizosaccharomyces pombe.* Regulatory role of a kinase conserved from fission yeast to mammals. *Curr. Genet.* **42**, 241-251.

Kumar, A., Harrison, P. M., Cheung, K-H., Lan, N., Echols, N., Bertone, P., Miller, P., Gerstein, M. B. and Snyder, M. (2002) An integrated approach for finding overlooked genes in yeast. *Nat. Biotechnol.* **20**, 58-63.

Lal, S. P., Christopherson, R. I. and dos Remedios, C. G. (2002) Antibody arrays: an embryonic but rapidly growing technology. *Drug Discov. Today* **7**, S143-S149.

Lane, P. G., Oliver, S. G. and Butler, P. R. (1999) Analysis of a continuous-culture technique for the selection of mutants tolerant to extreme environmental stress. *Biotechnol. Bioeng.* **65**, 397-406.

Lashkari, D. A., DeRisi, J. L., McCusker, J. H., Namath, A. F., Gentile, C., Hwang, S. Y., Brown, P. O. and Davis, R. W. (1997) Yeast microarrays for genome wide parallel genetic and gene expression analysis. *Proc. Natl. Acad. Sci. USA* **94**, 13057-13062.

Lasko, D. R., Zamboni, N. and Sauer, U. (2000) Bacterial response to acetate challenge: a comparison of tolerance amomg species. *Appl. Microbiol. Biotechnol.* **54**, 243-247.

Lee, J. H., Lee, D. E., Lee, B. U. and Kim, H. S. (2003a) Global analyses of transcriptomes and proteomes of a parent strain and an L-threonine-overproducing mutant strain**.** *J. Bacteriol.* **185**, 5442-5451.

Lee, D. Y., Yun, H., Park, S. and Lee, S. Y. (2003b) MetaFluxNet: the management of metabolic reaction information and quantitative metabolic flux analysis. *Bioinformatics* **19**, 2144-2146.

Levine, M. and Tjian, R. (2003) Transcriptional regulation and animal diversity. *Nature* **424**, 147-151.

Levy-Lahad, E. and Plon, S. E. (2003) A risky business. Assessing breast cancer risk. *Science* **302**, 574-575.

Liang, P. and Pardee, A. B. (2003) Analysing differential gene expression in cancer. *Nat. Rev. Cancer* **3**, 869-876.

Lidstrom, M. and Meldrum, D. R. (2003) Life on a chip. *Nat. Rev. Microbiol*. **1**, 158-164.

Lim, F. L., Hayes, A., West, A. G., Pic-Taylor, A., Morgan, B. A., Oliver, S. G. and Sharrocks, A. D. (2003) Mcm1p-induced DNA bending regulates the formation of ternary transcription factor complexes. *Mol. Cell. Biol*. **23**, 450-461.

Lipford, J. R. and Deshaies, R. J. (2003) Diverse roles for ubiquitin-dependent proteolysis in transcriptional activation. *Nat. Cell Biol.* **10**, 845 - 850.

Mendes, P. (2002) Emerging bioinformatics for the metabolome. *Brief. Bioinformatics* **3**, 134-145.

Michaud, G. A., Salcius, M., Zhou, F., Bangham, R., Bonin, J., Guo, H., Snyder, M., Predki, P. F. and Schweitzer, B. I. (2003) Analyzing antibody specificity with whole proteome microarrays. *Nat. Biotechnol*. **21**, 1509-1512.

Mitchell, P. (2002) A perspective on protein microarrays. *Nat. Biotechnol.* **20**, 225-229.

Morange, M. (2002) *The misunderstood gene*. Harvard University Press. Cambridge, USA.

Mosley, A. L., Lakshmanan, J., Aryal, B. K. and Ozcan, S. (2003) Glucose-mediated phosphorylation converts the transcription factor Rgt1 from a repressor to an activator. *J. Biol. Chem.* **278**, 10322-10327.

Mouse Genome Sequencing Consortium (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520-562.

Muller, D., Exler, S., Aguilera-Vazquez, L., Guerrero-Martin, E. and Reuss, M. (2003) Cyclic AMP mediates the cell cycle dynamics of energy metabolism in *Saccharomyces cerevisiae*. *Yeast* **20**, 351-367.

Muratani, M. and Tansey, W. P. (2003) How the ubiquitin-proteasome system controls transcription. *Nat. Rev. Mol. Cell. Biol*. **4**, 192-201.

Oliver, D. J., Nikolau, B. J. and Wurtele, E. S. (2002) Functional Genomics: High-throughput mRNA, protein, and metabolite analyses. *Metabolic engineering* **4**, 98-108.

Oliver, S. G. (1996) From DNA sequence to biological function.

Nature **379**, 597-600.

Oliver, S. G. (1997) Yeast as a navigational aid in genome analysis. *Microbiology* **143**, 1483-1487.

Oliver, S. G. (1998) Introduction to functional analysis of the yeast genome. *Methods in Microbiology* **26**, 1-13.

Oliver, S. G. (2002) Functional genomics: lessons from yeast. *Philos. Trans. Roy. Soc. B.*, **357**, 17-23.

Oliver, S. G., Winson, M. K., Kell, D. B. and Baganz, F. (1998) Systematic functional analysis of the yeast genome. *Trends Biotechnol.* **16**, 373-378.

Orchard, S., Hermjakob, H. and Apweiler, R. (2003) The proteomics standards initiative. *Proteomics* **3**, 1374-1376.

Outeiro, T. F. and Lindquist S. (2003) Yeast cells provide insight into alpha-synuclein biology and pathobiology. *Science* **302**, 1772-1775.

Ouyang, Z., Takats, Z., Blake, T. A., Gologan, B., Guymon, A. J., Wiseman, J. M., Oliver, J. C., Davisson, V. J. and Cooks, R. G. (2003) Preparing protein microarrays by soft-landing of mass-selected ions. *Science* **301**, 1351-1354.

Payne, W. E. and Garrels, J. I. (1997). Yeast protein database (YPD): a database for the complete proteome of *Saccharomyces cerevisiae*. *Nucleic Acids Res*. **25**, 57-62. (http://www.incyte.com/control/researchproducts/insilico/proteome).

Pearson, H. (2003) Geneticists play the numbers game in vain. *Nature* **423**, 576.

Peri, S., Navarro, J. D., Amanchy, R., Kristiansen, T. Z, Jonnalagadda, C. K., Surendranath, V., Niranjan, V., Muthusamy, B., Gandhi, T. K., Gronborg, M. *et al.* (2003) Development of human protein reference database as an initial platform for approaching systems biology in humans. *Genome Res.* **13**, 2363-2371.

Phelps, T. J., Palumbo, A. V. and Beliaev, A. S. (2002) Metabolomics and microarrays for improved understanding of phenotypic characteristics controlled by both genomics and environmental constraints. *Curr. Opin. Biotechnol.* **13**, 20-24.

Pratt, J. M., Petty, J., Riba-Garcia, I., Robertson, D. H. L., Gaskell, S. J., Oliver, S. G. and Beynon, R. J. (2002) Dynamics of protein turnover, a missing dimension in proteomics. *Mol. Cell Proteomics* **1**, 579-591.

Raamsdonk, L. M., Teusink, B., Broadhurst, D., Zhang, N., Hayes, A., Walsh, M. C., Berden, J. A., Brindle, K. M., Kell, D. B., Rowland, J. J., Westerhoff, H. V., van Dam, K. and Oliver, S. G. (2001) A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat Biotechnol.* **19**, 45-50.

Raponi, M. and Arndt, G. M. (2003) Double-stranded RNA-mediated gene silencing in fission yeast. *Nucleic Acids Res.* **31**, 4481-4489.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M. and Seraphin, B. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol.* **17**, 1030-1032.

Rohde, J. R. and Cardenas, M. E. (2003) The tor pathway regulates gene expression by linking nutrient sensing to histone acetylation. *Mol. Cell. Biol.* **23**, 629-635.

Rose, A. H. and Harrison, J. S. (1987-1995) *The yeasts*. Vol. 1-6. Academic Press, London, UK.

Ross-Macdonald, P., Coelho, P. S., Roemer, T., Agarwal, S., Kumar, A., Jansen, R., Cheung, K. H., Sheehan, A., Symoniatis, D., Umansky, L. *et al.* (1999) Large-scale analysis of the yeast genome by transposon tagging and gene disruption. *Nature* **402**, 413-418.

Salzberg, S. L. (2003) Genomics: Yeast rises again. *Nature* **423**, 233-234.

Sandelin, A., Hoglund, A., Lenhard, B. and Wasserman, W. W. (2003) Integrated analysis of yeast regulatory sequences for biologically linked clusters of genes. *Funct. Integr. Genomics* **3**, 125-134.

Schena, M., Shalon, D., Davis, R.W. and Brown, P. O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**, 467-470.

Scherer, L. J. and Rossi, J. J. (2003) Approaches for the sequence-specific knockdown of mRNA. *Nat. Biotechnol.* **21**, 1457-1465.

Schramke, V. and Allshire, R. (2003) Hairpin RNAs and retrotransposon LTRs effect RNAi and chromatin-based gene silencing. *Science* **301**, 1069-1074.

Sherman, M. Y. and Muchowski, P. J. (2003) Making yeast tremble: yeast models as tools to study neurodegenerative disorders. *Neuromolecular Med.* **4**, 133-146.

Shiio, Y. and Eisenman, R. N. (2003) Histone sumoylation is associated with transcriptional repression. *Proc. Natl. Acad. Sci. USA* **100**, 13225-13230.

Shoemaker, D. D., Schadt, E. E., Armour, C. D., He, Y. D., Garrett-Engele, P., McDonagh, P. D., Loerch, P. M., Leonardson, A., Lum, P. Y., Cavet, G. *et al.* (2001) Experimental annotation of the human genome using microarray technology. *Nature* **409**, 922-927.

So, E. N. and Crowe, D. L. (2000) Characterization of a retinoic acid responsive element in the human ets-1 promoter. *IUBMB Life* **50**, 365-370.

Spellman, P. T., Sherlock, G., Zhang, M. Q., Iyer, V. R., Anders, K., Eisen, M. B., Brown, P. O., Botstein, D. and Futcher, B. (1998) Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol. Biol. Cell* **9**, 3273-3297.

Spruill, S., Hardy, S., Weir, B. S, and Lu, J. (2002) Assessing sources of variability in microarray gene expression data. *BioTechniques* **33**, 916-923.

Tate, J. J. and Cooper, T. G. (2003) Tor1/2 regulation of retrograde gene expression in *Saccharomyces cerevisiae* derives indirectly as a consequence of alterations in ammonia metabolism**.** *J. Biol. Chem.* **278**, 36924-36933.

Taylor, C. F., Paton, N. W., Garwood, K. L., Kirby, P. D., Stead, D. A., Yin, Z., Deutsch, E. W., Selway, L., Walker, J., Riba-Garcia, I. *et al.* (2003) A systematic approach to modelling capturing and disseminating proteomics experimental data. *Nat. Biotechnol.* **21**, 247-254.

ter Kuile, B. H. (1999) Regulation and adaptation of glucose metabolism of the parasitic protist *Leishmania donovani* at the enzyme and mRNA levels. *J. Bacteriol.* **181**, 4863-4872.

ter Kuile, B. H. and Westerhoff, H. V. (2001) Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway. *FEBS Lett.* **500**, 169-171.

ter Linde, J. J., Liang, H., Davis, R. W., Steensma, H. Y., van Dijken, J. P. and Pronk, J. T. (1999) Genome-wide transcriptional analysis of aerobic and anaerobic chemostat cultures of *Saccharomyces cerevisiae*. *J. Bacteriol.* **181**, 7409-7413.

Teusink, B., Baganz, F., Westerhoff, H. V. and Oliver, S. G. (1998) Metabolic control analysis as a tool in the elucidation of the

function of novel genes. *Methods Microbiol.* **26**, 297-336.

The *Caenorhabditis elegans* Sequencing Consortium (1998) Genome sequence of the nematode *C. elegans*: A platform for investigating biology. *Science* **282**, 2012-2018.

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team (2002) Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* **420**, 563-573.

The International Human Genome Mapping Consortium (2001a) Initial sequencing and analysis of the human genome. *Nature* **409**, 860-921.

The International Human Genome Mapping Consortium (2001b) A physical map of the human genome. *Nature* **409**, 934-941.

The International Human Genome Mapping Consortium (2003) Completion of the Human Genome Project. (Press release 14th April 2003) (http://www.genome.gov/11006929; http://www.sanger.ac.uk/Info/Press/2003/030414.shtml).

Tilstone, C. (2003) Vital statistics. *Nature* **424**, 610-613.

Trethewey, R. N. (2001) Gene discovery via metabolic profiling. *Curr. Opin. Biotechnol*. **12**, 135-138.

Trethewey, R. N., Krotzky, A. J. and Willmitzer, L. (1999) Metabolic profiling: a Rosetta stone for genomics? *Curr. Opin. Plant Biol.* **2**, 83-85.

Uetz, P., Giot, L., Cagney, G., Mansfield, T. A., Judson, R. S., Knight, J. R., Lockshon, D., Narayan, V., Srinivasan, M., Pochart, P. *et al.* (2000) A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* **403**, 623-627.

Urbanczyk-Wochniak, E., Luedemann, A., Kopka, J., Selbig, J., Roessner-Tunali, U., Willmitzer, L. and Fernie, A. R. (2003) Parallel analysis of transcript and metabolic profiles: a new approach in systems biology. *EMBO Rep.* **4**, 989-993.

US Patent Office (2003) US Patent 6544742. Detection of genes regulated by EGF in breast cancer.

US Patent Office (2003) US Patent 2003124581. Newborn screening for hemoglobinopathy by DNA microarray analysis.

US Patent Office (2003) US Patent 2003170672. Quality control method of DNA microarray.

US Patent Office (2003) US Patent 2003180774. Exploiting genomes in the search for new drugs.

US Patent Office (2003) US Patent 2003180800. Stable isotope based dynamic metabolic profiling of living organisms for characterization of metabolic diseases, drug testing and drug development.

Verger, A., Perdomo, J. and Crossley, M. (2003) Modification with SUMO. A role in transcriptional regulation. *EMBO Rep.* **4**, 137-142.

Volpe, T., Schramke, V., Hamilton, G. L., White, S. A., Teng, G., Martienssen, R. A. and Allshire, R. C. (2003) RNA interference is required for normal centromere function in fission yeast. *Chromosome Res.* **11**, 137-146.

von Mering, C., Krause, R., Snel, B., Cornell, M., Oliver, S. G., Fields, S. and Bork, P. (2002) Comparative assessment of large-scale data sets of proteinprotein interactions. *Nature* **417**, 399-403.

Walker, G. M. (1999) Synchronization of yeast cell populations. *Methods Cell Sci.* **21**, 87-93.

Warringer, J. and Blomberg, A. (2003) Automated screening in environmental arrays allows analysis of quantitative phenotypic profiles in *Saccharomyces cerevisiae*. *Yeast* **20**, 53-67.

Washburn, M. P. (2003) Soft landing for protein chips. *Nat. Biotechnol.* **21**, 1156-1157.

Watkins, S. M. and German, J. B. (2002) Metabolomics and biochemical profiling in drug discovery and development. *Curr. Opin. Mol. Ther.* **4**, 224-228.

Weckwerth, W. (2003) Metabolomics in systems biology. *Annu. Rev. Plant Biol.* **54**, 669-689.

Weckwerth, W. and Fiehn, O. (2002) Can we discover novel pathways using metabolomic analysis? *Curr. Opin. Biotechnol.* **13**, 156-160.

Weusthuis, R. A., Pronk, J. T., van den Broek, P. J. and van Dijken, J. P. (1994) Chemostat cultivation as a tool for studies on sugar transport in yeasts. *Microbiol. Rev.* **58**, 616-630.

Willingham, S., Outeiro, T. F., DeVit, M. J., Lindquist, S. L. and Muchowski, P. J. (2003) Yeast genes that enhance the toxicity of a mutant huntingtin fragment or α-synuclein. *Science* **302**, 1769-1772.

Winzeler, E. A., Shoemaker, D. D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J. D., Bussey, H. *et al.* (1999) Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901-906.

Wodicka, L., Dong, H., Mittmann, M., Ho, M. H. and Lockhart, D. J. (1997) Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nature Biotechnol.* **15**, 1359-1367.

Wohlschlegel, J. A. and Yates, J. R. (2003) Proteomics: Wheres Waldo in yeast? *Nat.* **425**, 671-672.

Wood, V., Gwilliam, R., Rajandream, M. A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S. *et al.* (2002) The genome sequence of *Schizosaccharomyces pombe*. *Nature* **415**, 871-880.

World Intellectual Property Organisation (1999) WO Patent 9957306. Method for diagnosing a vascular condition.

World Intellectual Property Organisation (2001) WO Patent 0140896. System and method for metabolic profiling.

World Intellectual Property Organisation (2001) WO Patent 0178652. Methods for drug discovery, disease treatment and diagnosis using metabolomics.

World Intellectual Property Organisation (2001) WO Patent 0194946. Microarrays for performing proteomic analyses.

World Intellectual Property Organisation (2002) WO Patent 02057989. Metabolome profiling methods using chromatographic and spectroscopic data in pattern recognition analysis.

World Intellectual Property Organisation (2002) WO Patent 02092118. Global analysis of protein activities using proteome chips.

World Intellectual Property Organisation (2002) WO Patent 0218646. Gene discovery using microarrays.

World Intellectual Property Organisation (2002) WO Patent 0239120. Method for identifying the proteome of cells using an antibody library microarray.

World Intellectual Property Organisation (2003) WO Patent 03025213. Yeast proteome analysis.

World Intellectual Property Organisation (2003) WO Patent 03058238. Combined metabolomic, proteomic and transcriptomic analysis from one, single sample and suitable statistical evaluation data.

World Intellectual Property Organisation (2003) WO Patent 03070918. RNA interference mediated inhibition of gene

expression using chemically modified short interfering nucleic acid.

World Intellectual Property Organisation (2003) WO Patent 03087371. Antiviral therapy on the basis of RNA interference.

Wu, L. F., Hughes, T. R., Davierwala, A. P., Robinson, M. D., Stoughton, R. and Altschuler S. J. (2002) Large-scale prediction of *Saccharomyces cerevisiae* gene function using overlapping transcriptional clusters. *Nat. Genet*. **31**, 255-265.

Yamada, K., Lim, J., Dale, J. M., Chen, H., Shinn, P., Palm, C. J., Southwick, A. M., Wu, H. C., Kim, C., Nguyen, M. *et al.* (2003) Empirical analysis of transcriptional activity in the *Arabidopsis* genome. *Science* **302**, 842-846.

Yao, T. (2002) Bioinformatics for the genomic sciences and towards systems biology. Japanese activities in the post-genome era. *Prog. Biophys. Mol. Biol.* **80**, 23-42.

Zeng, Q., Morales, A. J. and Cottarel, G. (2001) Fungi and humans: closer than you think. *Trends Genet.* **17**, 682-684.

Zhang, N., Osborn, M., Gitsham, P., Yen, K., Miller, J. R. and Oliver, S. G. (2003) Using yeast to place human genes in functional categories. *Gene* **303**, 121-129.

Zhu, H. and Snyder, M. (2003) Protein chip technology. *Curr. Opin. Chem. Biol.* **7**, 55-63.

Zhu, H., Bilgin, M. and Snyder, M. (2003) Proteomics. *Annu. Rev. Biochem.* **72**, 783-812.

Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lan, N., Jansen, R., Bidlingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R. A., Gerstein, M. and Snyder, M. (2001) Global analysis of protein activities using proteome chips**.** *Science* **293**, 2101-2105.