Minireview

Four years of post-genomic life with 6000 yeast genes

André Goffeau*

Unité de Biochimie Physiologique, Université Catholique de Louvain, Place Croix du Sud 2/20, B-1348 Louvain-la-Neuve, Belgium

Received 19 May 2000

Edited by Gianni Cesareni

Abstract Four years after disclosure of the full yeast genome sequence, a series of resources including tens of thousands of mutant strains, plasmids bearing isolated genes and disruption cassettes are becoming publicly available. Deletions of each of the 6000 putative yeast genes are being screened systematically for dozens of phenotypic traits. In addition, new global approaches such as DNA hybridization arrays, quantitative proteomics and two-hybrid interactions are being steadily improved. They progressively build up an immense computation network of billions of data points which will, within the next decade, characterize all molecular interactions occurring in a simple eukaryotic cell. In this process of acquisition of new basic knowledge, an international community of over 1000 laboratories cooperates with a remarkable willingness to share projects and results.

Key words: Genomics; Yeast; DNA hybridization array; Proteomics; Two-hybrid interaction

1. Introduction

The first chromosome ever sequenced was yeast chromosome III, published in 1992 (Table 1). Between 1994 and 1996, all 15 other yeast chromosomes were completed and made public by 16 different teams of the international community of 'yeast scientists'. This series was only interrupted in mid-1995 by the publication of the bacterial chromosomes from *Haemophilus influenzae* and *Mycoplasma genitalia*, both sequenced by TIGR. The total yeast genome was put on the web by MIPS (Munich information center for protein sequences) in a fully organized and annotated form, on 24 April 1996. This is, the first, and possibly still the only case where fully organized and annotated genome data (and not just the raw sequences of the day) were made public well before the 'Lords of Publication' gave their official imprimatur. In this respect, the yeast genome has pioneered the new trendy policy of 'publication on the web'.

A remarkable feature of the yeast genome data is the use of a non-ambiguous and non-redundant nomenclature for open reading frames (ORFs), proteins and genes, well curated by three complementary databases: MIPS (http://www.mips.biochem.mpg.de), SGD (http://genome-www.stanford.edu) and YPD (http://www.proteome.com). However, the total numbers of ORFs encoded by the yeast genome, and the exact number of ORFs of unknown function are still uncertain.

For instance, on 15 May 2000, the yeast genome was estimated by YPD to encode 6149 putative proteins of which 4270 are characterized by genetics, biochemistry or sequence homology. On the same date and according to MIPS, the genome contains 6368 ORFs of which 178 correspond to small proteins of less than 100 amino acids, most of which were identified by the SAGE technique [26]. MIPS estimates that a total of 4344 ORFs are known by genetics, biochemistry or homology.

Stanislaw Cebrat et al. (Wroclaw, Poland) estimate on the basis of codon usage in the second position that up to 1200 'unknown' ORFs could be non-coding [27]. Most of them would correspond to overlapping antisense ORFs which were replicated independently of their sense partner during evolution. In contrast, MIPS lists only 447 'questionable' ORFs unlikely to be expressed.

In brief, because of different definitions of 'unknown' or 'hypothetical' and 'uncoding' or 'questionable' ORFs, the number of yeast proteins of which the function remains to be identified is estimated to be 300 (the Cebrat 'uncoding') or 1568 (the MIPS 'hypothetical') or 1879 (the YPD 'unknown').

An important complement of the yeast genome was the sequencing of the mitochondrial DNA from the strain S288c which had also been used for nuclear genome sequencing [25]. Yeast mitochondrial DNA is highly polymorphic and before its sequencing by Françoise Foury (Louvain-la-Neuve, Belgium) only a virtual and incomplete assembly of mitochondrial DNA from different strains was available [28]. The final version of mtDNA unraveled more than 10 kb of new mitochondrial DNA with seven small putative ORFs probably not coding. It also identified three hotspots of point mutations in the exons bordering group I intron-related sequences. The mtDNA sequence also allowed Bernard Dujon (Pasteur, France) to identify 34 mitochondrial DNA fragments (22–230 bp in length) inserted, probably by double strand break repair, in the non-coding sequence of the nuclear genome [29].

During the last 4 years, a large scientific community, estimated to amount to at least 5000 scientists in more than 1000 laboratories, has published nearly 7000 papers on yeast genes and genomics (Jim Garrel and Peter Hodges, Proteome, personal communication). Approximately 60% of these laboratories are located in the USA and about 35% are in Europe. In those 4 years the full yeast community has deciphered experimentally the biochemical or physiological function of 1060 yeast proteins which were unknown in 1996 (see YPD). This community also set up a series of large-scale resources which all are publicly available.

*Fax: (32)-10-473872.
E-mail: goffeau@fysa.ucl.ac.be

0014-5793/00/$20.00
PII: S0014-5793(00)01775-0

2. Deletion mutant libraries

The efficient homologous DNA recombination in yeast cells has allowed the construction of several massive deletion mutant libraries. The EUROFAN project deleted and grossly characterized the mutant phenotypes from 758 ORFs of unknown function among the 4400 ORFs (71% of the total genome) initially sequenced by the European laboratories (see MIPS/EUROFAN). The same procedure was used by the German Yeast Network which deleted 322 genes in several strains [30]. The 'mass murder' project organized by Bernard Dujon has analyzed 129 overlapping large deletions from chromosome XI covering a total of 217 ORFs [31]. Systematic deletions of chromosome III by Piotr Slonimski (Gif sur Yvette, France) and VIII by Johannes Hegemann (Giessen, Germany) were also carried out [32,33]. The most ambitious project, called 'Transatlantic Consortium for Bar-Coded Deletion', which aims at deleting all the 6000 yeast genes, is near completion. Each deletant is marked by unique 20-mer nucleotide signatures which allows recognition under selective conditions [34]. Deletants, plasmids containing individual genes and disruptant cassettes are available at EUROSCARF (http://www.uni-frankfurt.de/FB/mikro/euroscarf/index-htlm) or RESEARCH GENETICS (http://www.resgen.com).

Among the 3662 deleted ORFs tested so far, only 884 are presently recorded as being essential for cell growth (see MIPS). They include 42 'hypothetical' proteins (with no similarity) and curiously, 15 'questionable' proteins (which are considered unlikely to be expressed).

The 758 EUROFAN deletants are presently being screened systematically by more than 100 European laboratories under the scientific general coordination of Steve Oliver (Manchester, UK) and originally that of Peter Philippsen (Basel, Switzerland). A series of phenotypes are examined systematically, such as growth, temperature, resistance to metal ions, to high osmotic pressure, to ethanol, to antibiotics and other drugs, meiosis, recombination repair, telomeric structure, transport, organelles, lipids, secretion, trafficking, cytoskeleton, cell wall, morphogenesis, sporulation, stationary phase recovery, mating, genetic redundancy. The systematic search of lethal codeletants of certain ORFs is proceeding. This work should be completed and made public before the end of the year 2000. It will produce tens of thousands of data points and is expected to identify several hundreds of new phenotypes. Preliminary data can be consulted at the MIPS web site.

3. The transposon insertion mutant library

Early projects obtained a limited number of mutants interrupted by yeast Ty1 [35] or bacterial mini Mu [36] transposons. More recently, Michael Snyder and collaborators [37] have generated randomly, by insertion of a bacterial transposon, nearly 7800 yeast mutants. Many of these mutants have been characterized for level of expression (using a transposon-encoded lacZ reporter), as well as for 27 different phenotypes such as growth, temperature sensitivity, sensitivity to chemicals (by large-scale macroarray analysis), and for 19 the subcellular localization has been determined (using FGP or HA epitope). A total of 250 000 data points were provided by preliminary screenings of aliquots of the mutant library. For example, this approach has unraveled the involvement in sporulation of seven genes previously not known and 15 pre-

Table 1
The first yeast and bacterial chromosome sequences

	Chromosome	Length (in kb)	Coordinator	Public availability	Ref.
S. cerevisiae	III	315	S. Oliver (UK)	March 1992	[1]
S. cerevisiae	XI	666	B. Dujon (France)	June 1994	[2]
S. cerevisiae	VIII	589+*CUP1*	M. Johnston (USA)	September 1994	[3]
S. cerevisiae	II	813	H. Feldmann (Germany)	December 1994	[4]
S. cerevisiae	I	230	H. Bussey (Canada)	April 1995	[5]
S. cerevisiae	VI	271	Y. Murakami (Japan)	July 1995	[6]
H. influenzae		1850	R. Fleischmann, C. Venter (USA)	July 1995	[7]
M. genitalium		471	C. Fraser, C. Venter (USA)	October 1995	[8]
S. cerevisiae	Total genome (web)	13 478	W. Mewes, J. Sgouros, K. Kleine, J. Hani, A. Zollner (MIPS, Germany), M. Cherry (SGD, USA)	April 1996	[9]
S. cerevisiae	X	745	F. Galibert (France)	June 1996	[10]
M. jannaschii		1660	C. Bult, C. Venter (USA)	August 1996	[11]
S. cerevisiae	Total genome (paper)	13 478	A. Goffeau (Belgium)	October 1996	[12]
M. pneumoniae		816	R. Himmelreich (Germany)	November 1996	[13]
Synechocystis		3573	T. Kaneko, S. Tabata (Japan)	November 1996	[14]
S. cerevisiae	Genome directory		A. Goffeau and 641 others (Europe, USA, Japan, Canada)	May 1997	[15]
S. cerevisiae	IV	1532+*PMR2*	C. Jacq (France)	May 1997	[16]
S. cerevisiae	V	577	F. Dietrich (USA)	May 1997	[17]
S. cerevisiae	VII	1091	H. Tettelin (Belgium)	May 1997	[18]
S. cerevisiae	IX	440	B. Barrell (UK)	May 1997	[19]
S. cerevisiae	XII	1078+1 Mb rDNA	M. Johnston (USA), J. Hoheisel (Germany)	May 1997	[20]
S. cerevisiae	XIII	924	B. Barrell (UK)	May 1997	[21]
S. cerevisiae	XIV	784	P. Philippsen (Switzerland)	May 1997	[22]
S. cerevisiae	XV	1091	B. Dujon (France)	May 1997	[23]
S. cerevisiae	XVI	948	H. Bussey (Canada), B. Barrell (UK), K. Davis (USA), M. Johnston (USA), A. Goffeau (Belgium)	May 1997	[24]
S. cerevisiae	mtDNA	86	F. Foury (Belgium)	December 1998	[25]

viously uncharacterized genes. All constructions are freely available from on-line order (http://fondue.med.yale.edu/ygac/triples.htm).

4. DNA hybridization arrays

New global approaches aimed at characterization of a large number of genes by screening the full yeast genome have been developed these last 4 years. The most popular one is the DNA hybridization array on glass slides or membranes which probes simultaneously the mRNA level of all yeast genes under specific genetic and physiological conditions. The data output of such analyses is fabulous: 6000 mRNA levels per chip per week. Important biological parameters have already been tested, such as diauxy, cell cycle, meiosis, sporulation, ploidy, rich versus minimal medium, mitogen-activated protein kinases and other mutants, salt stress, osmotic stress, oxidative and chemical stresses, deletion, overexpression or activation of transcription factors (*TUP1*, *YAP1*, *PDR1*, *PDR3*). Very recently, up to 300 complete expression profiles were generated in which transcript levels of 287 mutants and 13 compound-treated cultures were analyzed [38]. An immense wealth of data of over 1 billion data points has already been obtained which is impossible to summarize here (see http://www.biologie.ens.fr/yeast-publi.html for references and links).

The following rules emerge. The present method seems very simple. A do-it-yourself protocol is on the web (http://emgm.stanford.edu/pbrown/mguide/index.html). In reality, considerable time and skills are required for mastering the numerous experimental parameters modulating the data. Even so, the method is of limited sensitivity; only ratios between sample and control above 2 are considered significant. Small-magnitude fluctuations in transcript abundance were analyzed in a series of 63 seemingly identical cultures [38]. These fluctuations are believed to represent a form of biological noise due to the regulation of several sets of genes by nutritional or stress factors. They account for virtually all the transcriptional changes when two conditions are compared which do not modify the growth rate [38]. Careful design of some of the DNA probes is necessary to avoid cross-hybridization by mRNAs from highly similar genes (e.g. the large HXT family of glucose transporters). The level of mRNA often varies considerably within minutes after setting up a given physiological condition. In these cases, a kinetic analysis is required. Only two arrays of the yeast genome are commercially available: the Research Genetics membranes and the Affymetrics chips. Neither covers the complete yeast genome unambiguously. Their current price is out of reach of most laboratories since measurements should in principle be reproduced in three distinct experiments with at least three time points each. These and other pitfalls are expected to be solved in the future [39]. Still, the DNA array approach already provides valid data because large differences in expression profiles are generally analyzed. As a safe rule, the major conclusions should be confirmed by other methods such as Northern, lacZ fusion, proteomic or biochemical measurements. In this respect, the near availability of transcript quantifications of about 1000 yeast ORFs by Northern analysis under eight physiological conditions (glucose regulation, nitrogen starvation, stationary phase and various stresses) organized by Rudi Planta (Amsterdam, The Netherlands) and Alistair Brown (Aberdeen, UK) within the EUROFAN project (see MIPS/EUROFAN) will provide useful reference points for the other DNA array data. Despite the present shortcomings of the technique, it is not difficult to predict that within a few years, thousands of yeast microarray analyses will become available. An immense network of data reporting the mRNA levels of all yeast genes in numerous physiological and genetic conditions is building up which will be manageable only by informatic network software presently under development. It is likely that most of the data will be published only on the web.

5. Proteome

The mRNA level is not always related to the level of active protein. Therefore, the proteome analysis has to be developed. The proteome has been defined by the Australians Marc Wilkins and Keith Williams et al. [40] as the set of proteins from a given cell, tissue or species identified by two-dimensional electrophoresis. Unfortunately, the integral membrane or small proteins cannot be separated by current techniques and the method requires at least one femtomol of each analyzed protein. Often, tryptic digestion has to be carried out. Their mass can only be identified in proteomes from fully sequenced genomes. These limitations explain why only about 2000 yeast protein spots can be analyzed today in total cell extracts. Among those, less than 300 yeast proteins (http://www.ibgc.u-bordeaux2.fr/ypm) have been identified by mass spectroscopy or other methods [41]. A recent important improvement has been the commercial availability of immobilized pH gradient gels as support of the first dimension electrophoresis which allows data to be highly reproducible from one laboratory to another [42]. Here again, yeast, because of its simplicity and the exhaustive exploration of its genome and transcriptome, is the test-bed of choice for the development of the proteome technique.

Among others, the following parameters have been recently analyzed at the proteome level: oxidative stress [43], cyclic AMP mutants [44], and multidrug-resistant mutants [45]. The latter analysis is the first one where DNA microarray [46] and proteome data could be compared under exactly identical genetic and physiological conditions. The following conclusions could be drawn. The proteome analysis is more cumbersome and requires more heavy equipment and know-how than the DNA hybridization analysis. However, it seems to be more reproducible. Differences of the order of 30% between the level of given protein spots assessed in different conditions can be significant at the 99% probability level, when the average of three independent cultures is used for each measurement [45]. The metabolic and physiological consequences of a 30% increase or decrease in an enzyme level can be quite significant. The proteome data on multiple drug resistance determinants overlap only partly with the DNA microarray data. This results from the greater relative sensitivity of the proteome approach which makes it possible to pinpoint fine-tuning regulations of protein synthesis which escape the DNA microarray approach. Also, protein post-translation modifications (proteolysis, phosphorylation and others) are detected.

In brief, next to biochemical activity measurements which rarely can be carried out on a large scale, the proteome analysis is the method of choice for global investigations of fine-tuning of metabolic pathways induced by physiological or

genetic parameters. Watch, however, the first attempt of genomic-scale biochemical screening [47] based on construction of glutathione *S*-transferase fusion ORFs purified and tested in defined pools. Once more, yeast is the test-bed of choice for future development and very soon an immense network of proteome and biochemical data, probably physiologically more meaningful than those provided by DNA microarray, will be available on the web.

6. Two-hybrid analysis

This ingenious method aims at the identification of protein–protein physical interactions by coexpression in the same cell [48]. *Saccharomyces cerevisiae* has been found to be the best host for such an approach, which however remains inappropriate for membrane-bound proteins. Three large-scale analyses have been carried out on yeast proteins. None of them has yet analyzed all possible 36 million combinations among all proteins of *S. cerevisiae*. Ito et al. [49] have tested 4 million different combinations (11% of the total) which revealed 183 independent two-hybrid interactions. In particular, this set of data revealed new putative connections between distinct steps of vesicular transport. Uetz et al. [50] have tested 1.2 million interactions (3% of the total) and identified 48 possible partners. In particular, new interactions in arginine metabolism, in the regulation of the cell cycle and generation of the meiotic synaptonemal complex were unraveled. Fromont-Racine et al. [51] have further refined the method. A library of 3 million fragments with an average of nearly 250 amino acids was tested iteratively as prey for interaction with a subset of the 6000 full yeast gene products as bait. In a first screening against 150 full yeast gene products (2.5% of the genome), more than 1000 different ORFs were found to interact (P. Legrain, this issue). In particular, a role for a new complex involved on the one hand in the nuclear spliceosome and on the other hand in the cytoplasmic degradation of mRNA has been unraveled.

In brief, even though two-hybrid analysis of the full yeast genome remains a herculean task, there is no doubt that it is well under way and that the data output is accelerating. Here again the identified interactions must be regarded as putative and be confirmed by other methods such as cross-linking, coimmunoprecipitation, affinity column, copurification, etc. Using these methods, a total of more than 2000 protein–protein interactions are listed by YPD.

7. Other yeast genomes

Before the end of the year 2000, it is expected that the genome of the fission yeast *Schizosaccharomyces pombe* (http://www.sanger.ac.uk/Projects/S-pombe) will be completed. Interestingly, it seems that as much as 30% of the two genomes differ. Also, the genome of the pathogenic yeast *Candida albicans* is already largely covered by shotgun sequencing (http://www.stanford.edu/group/candida). This will circumvent the locking of the pharmaceutical companies which have privileged access to private genome data and hope to benefit from a small time lead in the development of new antifungals. The 8.8 Mb genome sequence of the cotton pathogen *Achbya* is near completion by Fred Dietrich and Peter Philippsen (Basel, Switzerland).

Finally, a French consortium coordinated by Jean-Luc Souciet (Strasbourg, France) has shot-gunned 14 Hemiascomycete genomes from species closely related to *S. cerevisiae*. A low coverage sequence of each genome has been achieved which provides information on a total of 22 000 genes from these newly analyzed yeast species. A wealth of data on the recent evolution of yeast species and on possible pathogenic or industrial gene products has been accumulated and will be described in a forthcoming issue of FEBS Letters.

8. Conclusion

Less than a decade ago, the full molecular description of any eukaryotic cell seemed unreachable. With the complete genome sequence of *S. cerevisiae*, a finite number of about 6000 genes was shown to be sufficient for encoding all the proteins from an eukaryotic cell. The large scientific community studying yeast, combined with the relative simplicity of the cell, its unique physiological and genetic assets, the emergence of powerful global tools such as DNA hybridization arrays, proteomics and two-hybrid interaction analyses should rapidly make available millions of new biological data points. Within the next decade, the yeast cell, this humble servant of mankind, will be upgraded to the status of the first eukaryotic cell from which informatic boolean networks will compile all molecular interactions between genes and gene products and all metabolic fluxes in hundreds of different physiological conditions.

Acknowledgements: We acknowledge the Belgian funding agencies, 'Service de la Politique Scientifique, Technique et Culturelle, Pôle d'Attraction Interuniversitaire' and 'Fonds National de la Recherche Scientifique'.

References

[1] Oliver, S. et al. (1992) Nature 357, 38–46.
[2] Dujon, B. et al. (1994) Nature 369, 371–378.
[3] Johnston, M. et al. (1994) Science 265, 2077–2082.
[4] Feldmann, H. et al. (1994) EMBO J. 13, 5795–5809.
[5] Bussey, H. et al. (1995) Proc. Natl. Acad. Sci. USA 92, 3809–3813.
[6] Murakami, Y. et al. (1995) Nat. Genet. 10, 261–268.
[7] Fleischmann, R. et al. (1995) Science 269, 496–512.
[8] Fraser, C. et al. (1995) Science 270, 397–403.
[9] http://www.mips.biochem.mpg.de/mips/yeast and http://www-genome.stanford.edu/Saccharomyces/.
[10] Galibert, F. et al. (1996) EMBO J. 15, 2031–2049.
[11] Bult, C. et al. (1996) Science 273, 1058–1073.
[12] Goffeau, A. et al. (1996) Science 274, 563–567.
[13] Himmelreich, R. et al. (1996) Nucleic Acids Res. 24, 4420–4449.
[14] Kaneko, T. et al. (1996) DNA Res. 3, 109–136.
[15] Goffeau, A. and 641 al. (1997) Nature 387 (Suppl.) 1–105.
[16] Jacq, C. et al. (1997) Nature 387 (Suppl.), 75–78.
[17] Dietrich, F.S. et al. (1997) Nature 387 (Suppl.), 78–81.
[18] Tettelin, H. et al. (1997) Nature 387 (Suppl.), 81–84.
[19] Churcher, C. et al. (1997) Nature 387 (Suppl.), 84–87.
[20] Johnston, M. et al. (1997) Nature 387 (Suppl.), 87–90.
[21] Bowman, S. et al. (1997) Nature 387 (Suppl.), 90–93.
[22] Philippsen, P. et al. (1997) Nature 387 (Suppl.), 93–98.
[23] Dujon, B. et al. (1997) Nature 387 (Suppl.), 98–102.
[24] Bussey, H. et al. (1997) Nature 387 (Suppl.), 103–105.
[25] Foury, F., Roganti, T., Lecrenier, N. and Purnelle, B. (1998) FEBS Lett. 440, 325–331.
[26] Velculescu, V.E. et al. (1997) Cell 88, 243–251.
[27] Kawalczuk, M., Mackiewicz, P., Gierlik, A., Dudek, M.R. and Cebrat, S. (1999) Yeast 15, 1031–1034.
[28] De Zamaroczy, M. and Bernardi, G. (1986) Gene 47, 155–177.

[29] Ricchetti, M., Fairhead, C. and Dujon, B. (1999) Nature 402, 96–100.
[30] Entian, E.D. et al. (1999) Mol. Gen. Genet. 262, 683–702.
[31] Fairhead, C., Thierry, A., Denis, F., Eck, M. and Dujon, B. (1998) Gene 223, 33–46.
[32] Ricger, K., El-Alamada, M., Stein, G., Bradshow, C., Slonimski, P.P. and Moundrell, K. (1999) Yeast 15, 513–526.
[33] Niedenthal, R., Riles, L., Guldener, U., Klein, S. and Johnston, M. (1999) Yeast 15, 1775–1796.
[34] Winzeler, E.A. et al. (1999) Science 285, 901–906.
[35] Smith, V. et al. (1996) Science 274, 2069–2074.
[36] Ross-Macdonald, P. et al. (1999) Nature 402, 413–417.
[37] Daignan-Fornier, B. and Bolotin-Fukuhara, M. (1989) Yeast 5, 259–269.
[38] Hughes, T.R. et al. (2000) Cell, in press.
[39] Hoheisel, J.D. and Vingron, M. (2000) Res. Microb. 151, 113–119.
[40] Williams, K.L., Gooley, A.A. and Packer, N.H. (1996) Today's Life Sci. 8, 16–21.
[41] Perrot, M., Sagliocco, F., Mini, T., Monribot, C., Schneider, U., Shevchenko, A., Mann, M., Jeno, P. and Boucherie, H. (1999) Electrophoresis 20, 2280–2298.
[42] Norbeck, J., Fey, S.J., Larsen, P.M., Boepstorff, P., Degand, H., Boutry, M., Posch, A. and Gorg, A. (1995) Electrophoresis 10, 1935–1945.
[43] Godon, C., Lagniel, G., Lee, J., Buhler, J.M., Kieffer, S., Perrot, M., Boucherie, H., Toledano, M.B. and Labarre, J. (1998) J. Biol. Chem. 273, 22480–22489.
[44] Norbeck, J. and Blomberg, A. (2000) Yeast 16, 121–128.
[45] Navrocki, A., Fey, S.J., Goffeau, A., Roepstorff, P. and Larsen, P.M., submitted.
[46] DeRisi, J., van den Hazel, B., Marc, P., Balzi, E., Brown, P., Jacq, C. and Goffeau, A. (2000) FEBS Lett. 470, 156–160.
[47] Martzan, M.R., McCraith, S.M., Spinelli, S.L., Torres, F.M., Fields, S., Grayhack, E.J. and Phizicky, E.M. (1999) Science 286, 1153–1155.
[48] Fields, S. and Song, O. (1989) Nature 340, 245–246.
[49] Ito, T., Tashiro, K., Muta, S., Ozawa, B., Chiba, T., Nishizawa, M., Yamamoto, K., Kubara, S. and Sakaki, Y. (2000) Proc. Natl. Acad. Sci. USA 97, 1143–1147.
[50] Uetz, P. et al. (2000) Nature 403, 623–627.
[51] Fromont-Racine, M., Mayes, A.E., Brunet-Simon, A., Rain, J.C., Colley, A., Dix, I., Joly, N., Beggs, J.D. and Legrain, P. (2000) Comp. Funct. Genomics 176, in press.