Box 1. Structure of the EUROFAN network

Service consortia
- A1 Central coordination
- A2 EUROFAN informatics resource centre
- A3 Industrial liaison: yeast industry platform (YIP)
- A4 Genetic archive and stock centre

Research consortia
- B0 Generation of deletants, clones and plasmid tools
- B1 Qualitative phenotypic analysis
- B2 RNA-level expression analysis
- B4 Protein-level expression analysis II: gene fusions
- B5 Gene interactions: two-hybrid analyses
- B6 Metabolic control analyses
- B7 Subcellular structure and organelles
- B8 Relation to other genomes
- B9 Development and evaluation of novel methodologies for genome analysis

Specific functional analysis nodes
- N1 DNA synthesis and cell cycle
- N2 RNA synthesis and processing
- N3 Translation
- N4 Stress responses
- N5 Cell wall and morphogenesis
- N6 Transport
- N7 Energy and carbohydrate metabolism
- N8 Lipid metabolism
- N9 Special metabolism
- N10 Development
- N11 Mutagenesis, repair, recombination and meiosis
- N12 Chromosome structure
- N13 Cell architecture
- N14 Secretion and protein trafficking

made using a PCR-mediated gene-replacement approach developed by Achim Wach and Peter Philippsen[1,2]. In this technique, PCR is used to create a specific deletion cassette that contains a bacterial kanamycin resistance gene (called *kanMX*) flanked by promoter and terminator sequences derived from the filamentous fungus, *Ashbya gossypii*. This constitutes a selectable marker that confers geneticin resistance on yeast, but has no homology to the yeast genome. That homology is provided by either long or short tails, which are added to the selection cassette using PCR. The tails are designed such that transformation of the replacement cassette into yeast will permit the selection of geneticin-resistant transformants in which at least 80% of a specific open reading frame (ORF) has been replaced by *kanMX*. The accuracy of the deletion is checked and then a series of basic phenotypic analyses are carried out (e.g. for viability, growth rate, the ability to respire, mate and sporulate) before depositing the deletant in the EUROFAN stock centre for distribution to the other consortia and specific functional analysis nodes in the network. The consortia (B1–B8 in Box 1) carry out analyses that can be applied on a genome-wide scale and not all of which require the specific deletant strains. For example, one consortium (B9, coordinated by Bernard Dujon, Paris, France) has the responsibility of developing novel analytical methods for use by the other consortia. Among these is an alternative deletion strategy, known as 'mass murder' in which several ORFs are deleted at a time in an overlapping manner that allows assignment of phenotypes to specific genes. The nodes (N1–N14 in Box 1) use the deletion strains, and the phenotypic information provided by the consortia, to start to assign the functionality of particular ORFs to specific areas of the organism's biological activity.

This structure should permit us to carry out genetics in the reverse direction to the one traditionally pursued, that is, from gene (or DNA sequence) to biological function, rather than from heritable phenotype to gene. If we are successful, then we will have created an informatics and material resource that will significantly accelerate the assignment of functions to novel genes found in larger genomes. This will be done by identifying equivalence between yeast gene products and those of higher eukaryotes at either the structural (via sequence comparisons) or functional (via the complementation of specific deletion mutations in yeast with cDNA clones derived from a higher organism) level. In this way, the functionality of the genomes of crop plants, farm animals and humans can be mapped on to that of yeast, and yeast will become a major navigational aid in our journeys through these larger, and more complex, genomes.

References

1 Wach, A., Brachat, A., Poehlmann, R. and Philippsen, P. (1994) *Yeast* 10, 1783–1808

2 Wach, A. (1996) *Yeast* 12, 259–265

Towards a complete understanding of how a simple eukaryotic cell works

MARK JOHNSTON
mj@sequencer.wustl.edu

DEPARTMENT OF GENETICS, BOX 8232, WASHINGTON UNIVERSITY MEDICAL SCHOOL, 4566 SCOTT AVE, ST LOUIS, MO 63110, USA

Systematic sequencing of the bakers' yeast genome has yielded a rich vein of previously unidentified genes whose functions must be determined if yeast geneticists are to approach their ultimate goal: a comprehensive understanding of how a simple eukaryotic cell functions. This will, of course, be much more difficult than determining the sequence, and will require the efforts of the entire yeast community, along with help from our colleagues studying other organisms. Unlike the sequencing, it is a project that will never really be finished, but it is a goal that we can now imagine approaching.

What is the best way to reveal the functions of these novel genes? One possibility is to continue the systematic approach initiated by the sequencers. Each gene can be disrupted, and a standard set of tests carried out on the resultant mutants. The hope is that they will exhibit

Copyright © 1996 Elsevier Science Ltd. All rights reserved. 0168-9525/96/$15.00
PII: S0168-9525(96)30064-4

phenotypes that might reveal the function of the missing protein.

There are two related problems that will limit the success of this approach. First, most mutants (60–80%) do not display obvious phenotypes[1–5]. The fraction of mutants manifesting a phenotype will undoubtedly increase if systematically analyzed, because more tests will be carried out more carefully in such a program, but it seems likely that a significant number of genes will resist such examination. Second, and perhaps more problematic, the phenotypes that many mutants exhibit will not obviously inform us about the function of the encoded protein. What does the failure of a mutant to grow at high temperature (or on high salt-containing medium) tell us about the function of the missing protein? There is no doubt that in some cases this kind of information will complement other pieces of evidence and contribute to an understanding of protein function, but for a large number of genes the phenotypic information is likely to remain uninformative. Thus, a purely systematic approach to functional analysis is unlikely to bring us close to our goal.

An alternative view is that we can reach our goal by continuing to follow the path we have taken in the past: individual laboratories uncovering the function of a few proteins in the course of pursuing their interests in particular well-defined biological problems. While this directed '*ad hoc*' approach has clearly stood us in good stead, causing bakers' yeast to be arguably the best-understood eukaryote, it too seems unlikely to bring us close to our goal in the foreseeable future. I say this because of the realization that geneticists and biochemists employing this approach have discovered less than half of the genes of bakers' yeast (see review by B. Dujon, this issue), despite intensive investigation with highly sophisticated tools over the past several decades. Of course, as analysis becomes more intensive, and as our tools become even more sophisticated, *ad hoc* analysis will reveal the function of more of these genes, but it seems likely that the function of a significant fraction of proteins will remain mysterious for a long time if we proceed with business as usual.

I am optimistic that we can come close to our goal of a comprehensive understanding of how a simple eukaryotic cell works if we move in two directions. First, we should arrange a marriage of the systematic and *ad hoc* approaches. As with most marriages, it will work best if the strengths of each partner are complementary. The expertise and organizational skills of the scientists involved in the genome sequencing project should be employed to rapidly produce a complete set of gene disruptants. Knowledge of the complete DNA sequence of the yeast genome, along with facile gene disruption techniques that have recently been developed[6], make this goal realistic.

The expertise and enthusiasm of the entire international yeast community should be tapped for the *ad hoc* analysis of this collection of mutants. If the complete set of approximately 6000 mutants were sent to all yeast laboratories (possibly 1000 worldwide), with the request that they perform their favorite test on each mutant, a wealth of information would be generated quickly, cheaply and reliably. If only 10% of the laboratories complied, we would have about 100 pieces of information on each gene! This information would, in many cases, inform ongoing *ad hoc* analysis, and would probably direct individual investigators to many genes they might otherwise have overlooked. In addition, having a complete set of mutants in their freezer might stimulate investigators to pursue new approaches to genetic analysis (certainly many mutant hunts would be performed differently than they are now!). At the very least, the complete set of disruptants would save individual workers the effort and expense of disrupting genes (much like the sequencing project spared them the task of determining DNA sequence). It seems certain that the wide and timely distribution of a uniform set of yeast mutants will accelerate the pace of discovery of protein function in bakers' yeast.

There are lines of investigation that, perhaps, can best be pursued systematically, such as protein identification by two-dimensional gel electrophoresis[7,8], determination of the localization and levels of proteins[4], and mapping of protein–protein interactions[9]. The timely production of a complete set of disruptants is likely to enhance and complement such efforts.

The second direction we should move in is to take advantage of the tools provided by other organisms. Yeasts have long been touted as model organisms that can inform analysis of larger eukaryotes, but the converse is also true. As sequencing of more complex organisms proceeds, an increasing number of yeast proteins are being recognized as homologs of proteins in other eukaryotes. This is good for scientists working on the higher eukaryote, because yeast provides them with a tractable experimental system for studying their proteins, but it is also good for yeast geneticists, because the higher eukaryote can provide them with a rich source of phenotypes and cytology that often gives clues about the function of the yeast protein[10]. This might provide one solution to the problem of the relative dearth of phenotypes exhibited by yeast mutants.

Mammalian geneticists are increasingly taking advantage of what yeasts have to offer for the analysis of mammalian proteins; yeast geneticists need to exploit the mammalian experimental systems more aggressively to aid their search for yeast protein function. Already the distinction between mammalian and yeast geneticists is blurring, and this is sure to accelerate as yeast becomes ever more accessible to mammalian geneticists. The availability of the yeast genome sequence has greatly accelerated this cross-fertilization between researchers[11]; having a uniform set of disruptants will make yeast even more accessible to those not familiar with the organism. This would be a welcome development, because we will need the contributions of our colleagues studying other organisms if we are to reach our goal of a complete understanding of how a eukaryotic cell functions.

References

1 Goebl, M.G. and Petes, T.D. (1986) *Cell* 46, 983–992
2 Riles, L. and Olson, M.V. (1988) *Genetics* 118, 601–607
3 Oliver, S.G. *et al.* (1992) *Nature* 357, 38–46
4 Burns, N. *et al.* (1994) *Genes Dev.* 8, 1087–1105
5 Bussey, H. *et al.* (1995) *Proc. Natl Acad. Sci. USA* 92, 3809–3813
6 Baudin, A. *et al.* (1993) *Nucleic Acids Res.* 21, 3329–3330
7 Garrels, J.I. *et al.* (1994) *Electrophoresis* 15, 1466–1486
8 Boucherie, H. *et al.* (1996) *Yeast* 11, 601–613
9 Bartel, P.L. *et al.* (1996) *Nat. Genet.* 12, 72–77
10 Tugendreich, S. *et al.* (1995) *Cell* 81, 261–268
11 Bassett, D.E., Jr *et al.* (1995) *Trends Genet.* 11, 372–373