In Focus

Functional genomics in yeast

Completion of the *Saccharomyces cerevisiae* genome sequencing project in 1996 led to an incredible explosion of research on basic cellular processes and has provided the opportunity to determine how genes and their products are regulated and function on a genome-wide scale. The technologies that were developed from this provided an incredible array of tools to study cellular processes in great detail and were a paradigm for developments from subsequent sequencing projects.

On completion of the *S. cerevisiae* sequencing project it was estimated that there was understanding of the role of only about 50% of the genes identified, and for many genes this knowledge was only partial. The completed sequence provided the catalyst for development of a wide range of technologies that allow analysis of how the entire complement of genes in the genome participates in particular processes. Over the intervening decade the proportion of genes for which there is some understanding of function has risen considerably with a light-hearted estimate of the completion of the annotation process by 01 April 2007 in time for the 23rd International Conference on Yeast Genetics and Molecular Biology (Melbourne, 01 – 06 July 2007).

This understanding is now leading to interesting developments in the systems biology approach to modelling how an organism functions, as exemplified by the analysis of how yeast cells respond to osmotic stress. These technologies have also been valuable in studying cellular functions in eukaryotic organisms other than *S. cerevisiae*, including humans. About 60% of yeast genes have a human homologue, and in many cases the human gene can function in yeast. Moreover, nearly 400 proteins in *S. cerevisiae* are orthologues of proteins associated with human diseases. Therefore, *S. cerevisiae* can provide a platform to study molecular processes and structure-function relationships, and to analyse the toxicity and effectiveness of drugs in a more rapid and convenient way than in the particular organism under study.

The technologies developed include microarray for gene expression studies, with extension to ChIP-chip analysis for each transcription factor to determine where it binds in the entire yeast genome. By labelling every gene individually with GFP the location of each gene product in the cell can be determined and a database is available on the internet. Several protein-protein interaction screens have provided valuable databases that inform the fundamental subassemblies that make up the cellular machinery. However, this article discusses those approaches that have led to the generation of mutations affecting every gene in the yeast genome since these have been very fruitful, not only in studying yeast genes, but also in studying structure-function relationships for other, higher eukaryotes.

Genome-wide mutation analysis

The use of mutants as a specific way to study gene function has been very widely used for a long time. With the availability of the genomic sequence came the opportunity to generate mutants on a genome-wide scale to try and find a phenotype for the mutations that would provide clues to the function of unknown genes and to identify the role of all genes in specific cellular processes.

With *S. cerevisiae* there were several such ambitious programs. An initial program (Eurofan), inspired by the sequencing partners in Europe, set out to delete 1000 genes (of the ~6000 identified) in a cooperative way. This was superseded by a joint US/European program that led to the generation of mutants in every non-essential yeast gene, and conditional mutants in essential genes. This set of yeast mutants has every mutant specifically and individually ‘tagged’ with unique 20 bp ‘barcodes’ at the start and end of the kanamycin resistance cassette used to generate the deletion. These mutants have been used individually to identify phenotypes. For example, in our laboratory, initial screening of the genome-wide deletion collection on five different oxidants identified about 456 mutations affecting cellular responses to reactive oxygen species, and increased the number of known mutants by more than ten-fold. Many of the unknown mutants identified were as sensitive as other well-characterised mutants obtained by random mutagenesis and tedious screening procedures that were obviously far from saturating.

Since the tags (molecular bar codes) in the deletion strains can be identified by hybridisation of DNA from the population to

Ian W Dawes, Geoffrey D Kornfeld & Gabriel G Perrone
Ramaciotti Centre for Gene Function Analysis
School of Biotechnology and Biomolecular Sciences
University of New South Wales, Sydney NSW 2052
Tel: (02) 9385 2089
Fax: (02) 9385 1050
Email: i.dawes@unsw.edu.au
In Focus

custom microarrays, these strains can also be used to determine the contribution that a gene makes to the ‘fitness’ of the organism. By growing a mixed population of all mutant strains under a particular condition (e.g. in the presence of a toxic metabolite) one can detect strains that have been lost by the absence of the bar code hybridisation signal. With the large number of screens that have been done of how cells respond to various antibiotics and toxic compounds, it is possible in some cases to identify the target of an unknown compound by comparison to the spectrum of responses to known drugs. Each mutation is available in haploid $\alpha$ and alpha strains, in the heterozygous diploid and in the homozygous diploid state. Heterozygous diploids have been exploited in haploinsufficiency studies to identify gene functions and these are reported to give a clearer understanding of gene involvement in metabolic processes, since they identify genes that need to be fully expressed for a function to occur normally.

At about the same time, the laboratories of Mike Snyder and Shirleen Roeder set out to generate transposon insertion mutations in every yeast gene. By clever design of the transposable cassette (Figure 1) the mutants can be used not only to determine the phenotype of the insertion mutant, but also the regulation of the gene due to the presence in the transposon of a lacZ reporter gene lacking an ATG start codon. The location of the gene product in the cell could be determined by using 3X haemagglutinin epitope tags fused to the yeast gene as part of the transposon. An interesting refinement was to engineer loxP sites, arranged so that in any mutant expression of cre recombinase would remove all but 39 bp of the transposon up to the haemagglutinin tags. This could generate ‘hypomorphic’ alleles (i.e. with reduced or conditional activity) instead of null alleles providing further tools for analysis of gene function. Basically this procedure provided three complementary hits to functional analysis in the one system: analysis of the loss of function of the gene; analysis of its expression; and identification of the location of the gene product.

Within five years of them becoming available, over 100 experimental conditions were analysed and more than 5000 novel phenotypic traits were assigned to yeast genes using these various mutant collections, and these numbers have increased substantially since then. This has led to the identification of numerous previously unknown components of well-known processes in yeast.
pathways and processes, such as cell division, protein trafficking, meiosis and spore formation and germination.

The deletion mutant collection has also been used to analyse functional aspects of genes from other organisms. A clear example is the study by Lindquist’s group of the effects of over-expressing the human alpha-synuclein gene that is implicated in several neurodegenerative disorders, such as Parkinson’s disease, in every individual member of the mutant set. When expressed in yeast, alpha-synuclein selectively associated with the plasma membrane, before forming cytoplasmic inclusions. This system provided an opportunity to dissect the molecular pathways underlying normal alpha-synuclein biology and the pathogenic consequences of its misfolding. A similar systematic genome-wide study has been done to identify host functions affecting replication of the positive-strand Brome mosaic virus.

In collaboration with Phil Hogg’s group in the Centre for Vascular Biology we screened the anticancer compound GSAO (glutathionarsenoxide, which functions to inhibit angiogenesis) using the genome-wide knockout collection to rapidly provide clues about what is happening in higher eukaryotes.

**Synthetic lethality screening**

Synthetic lethality is the genetic interaction of two non-lethal mutations that lead to inviability in combination. The related process, synthetic fitness, describes the interaction of two non-lethal mutations that lead to a non-lethal growth defect in combination. Although synthetic lethality and fitness have been used commonly for many years, first in *Drosophila* and later extensively in yeast, the technique really achieved its potential for the elucidation of pathway topology with the work of Tong et al. using the genome-wide knockout collection.

In an ordered approach, Tong crossed the viable haploid strains of the knockout collection with a strain carrying a ‘query’ mutation. The resultant diploids where sporulated and the

### Different deletion mutants and their functions

<table>
<thead>
<tr>
<th>Deleted Gene</th>
<th>IC50, mM</th>
<th>Function of gene product</th>
</tr>
</thead>
<tbody>
<tr>
<td>CYS4</td>
<td>0.032</td>
<td>First enzyme in cysteine biosynthesis</td>
</tr>
<tr>
<td>CYS3</td>
<td></td>
<td>Catalyses cysteine for cystathionine</td>
</tr>
<tr>
<td>MET14</td>
<td>0.073</td>
<td>Sulfate assimilation and methionine biosynthesis</td>
</tr>
<tr>
<td>MET16</td>
<td>0.080</td>
<td>Sulfate assimilation and methionine metabolism</td>
</tr>
<tr>
<td>MET10</td>
<td>0.083</td>
<td>Conversion of sulfite into sulfate</td>
</tr>
<tr>
<td>MET22</td>
<td>0.108</td>
<td>Phosphate and glycogen metabolism and cell cycle control</td>
</tr>
<tr>
<td>MET28</td>
<td>0.109</td>
<td>Transcription factor involved in sulfur metabolism</td>
</tr>
<tr>
<td>PHO85</td>
<td>0.110</td>
<td>Phosphorylate and phosphatase and cell cycle control</td>
</tr>
<tr>
<td>PRM7</td>
<td>0.110</td>
<td>Unknown function</td>
</tr>
<tr>
<td>MET73</td>
<td>0.120</td>
<td>Sulfur balance, methionine synthesis and sulfate assimilation</td>
</tr>
<tr>
<td>MET77</td>
<td>0.122</td>
<td>Heat shock protein involved in mitochondrial protein import</td>
</tr>
<tr>
<td>MTH1</td>
<td>0.129</td>
<td>Carbon catabolite repression, oxidative stress</td>
</tr>
<tr>
<td>BRE1</td>
<td>0.135</td>
<td>Degradation of histone H2B</td>
</tr>
<tr>
<td>RRS32</td>
<td>0.135</td>
<td>Unknown function</td>
</tr>
<tr>
<td>RPP20</td>
<td>0.140</td>
<td>Glyceraldehyde 3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>RML30</td>
<td></td>
<td>Mitochondrial ribosome small subunit component</td>
</tr>
<tr>
<td>HOM2</td>
<td>0.186</td>
<td>First step in common methionine and threonine biosynthesis</td>
</tr>
<tr>
<td>HOM1</td>
<td>0.188</td>
<td>Vascular import and degradation</td>
</tr>
<tr>
<td>HOM4</td>
<td>0.188</td>
<td>Methionine and threonine biosynthesis</td>
</tr>
<tr>
<td>MTH17</td>
<td>0.209</td>
<td>Required for sulfur amino acid synthesis</td>
</tr>
</tbody>
</table>

Figure 2.
MAPa haploids that carried both the original library mutation and the query mutation were isolated. Genes that conferred synthetic lethality resulted in loss of the strain from the plate. Ooi simplified the technique by integratively transforming pools of deletion strains, growing the cells en masse, and isolating the total genomic DNA to screen the pool using microarrays that can identify the barcodes that are built into the knockout collection 23. The technique was termed SLAM for 'synthetic lethality analysis by microarray'. Recently described is an improvement on this technique, termed dSLAM for 'heterozygote-based diploid synthetic lethality analysis by microarray', in which the knockout construct is integrated into a mixed population of heterozygous diploid knockout strains, which are then sporulated and the progeny selected for both the single and double knockouts for use in microarray analysis 23. Use of the more genetically stable diploids and the single mutant controls in the analysis results in an increase in the reliability of the data. This technique can also be adapted to examine the less common phenomenon of synthetic haploinsufficiency, further adding to our understanding of pathway dynamics.

In a much more adventurous approach one group has developed the concept of the synthetic genetic array analysis in which every pathway can be examined. This technique can also be adapted to examine the less common phenomenon of synthetic haploinsufficiency, further adding to our understanding of pathway dynamics.

Whilst synthetically lethal point mutations may be the result of either loss of redundant pathways or the cumulative effect of two debilitating mutations in the same pathway, synthetic lethality analysis of knockouts unambiguously pinpoints parallel pathways. Although Saccharomyces duplicated its genome after the divergence of Saccharomyces from Kluyveromyces approximately 107 years ago leading to approximately 400 duplicated gene pairs 21, results indicated that the majority of genes conferring synthetic lethality (and therefore having redundant function) examined so far, arise from non-homologous genes rather than paralogs 21 indicating that parallel pathways are often nonparalogous, possibly to increase their robustness.

Concluding comment

The genome-wide analysis of single gene knockouts and the use of synthetic lethality screens can be seen as orthogonal dimensions to the data that can be obtained from genome-wide expression analysis (DNA microarray) and protein-protein interactions (PPI). While the latter are useful and powerful techniques, in many ways the mutational systems are far more powerful because they can discern the function of genes that are not regulated at the level of transcription (or only weakly so) or that require significant physical interactions. In contrast to the physical interaction map of the proteome provided by PPI techniques, synthetic lethality analysis generates a genetic interaction map that can discern gene products in parallel pathways that are not physically coupled or even in the same subcellular compartment. Despite these comments, it has been clear from our experience and that of other groups studying cellular responses to the environment at the genomic techniques of microarray, that gene over-expression and mutant analysis are very complementary. Each technique identifies a very different subset of genes, which provides further clues to function. For example, in the response to oxidants, deletion screening identifies functions that are required to maintain resistance when an oxidant is applied, whereas microarray analysis identifies those genes that are induced as a result of oxidant damage and these are very different sets of genes 7.

References