The role and aims of the FYSSION project

John Armstrong, Neil Bone, James Dodgson and Timothy Beck

Advance Access publication date 31 March 2007

Abstract
FYSSION is a resource for researchers working on the fission yeast Schizosaccharomyces pombe. It currently comprises libraries of temperature-sensitive mutants in essential genes, and insertional mutants in non-essential genes, available for screening by visiting workers. Here we outline methods for constructing and using the libraries, and describe future prospects for functional genomics of this organism, here and elsewhere.

Keywords: Schizosaccharomyces pombe; functional genomics; insertional mutagenesis; phenotype ontology

FUNCTIONAL GENOMICS OF SCHIZOSACCHAROMYCES POMBE

The fission yeast Schizosaccharomyces pombe has become a very popular model eukaryote for molecular analysis, being readily amenable to both classical and ‘reverse’ genetics [1]. In comparison with budding yeast Saccharomyces cerevisiae, studies of S. pombe have tended to concentrate on a few areas of biology, particularly control of the eukaryotic cell cycle, a general theme which has now expanded into control of growth and shape during the cell cycle. However, the completion of the genome sequence in 2002 [2] provided a stimulus to use the organism to study many more areas of biology. At the time the sequence, determined under the auspices of the Sanger Institute, was almost certainly the most accurate eukaryotic genome available. Continued analysis has led to a very comprehensive database of genes, known and predicted, and their likely functions [3]. However, again in comparison with S. cerevisiae, the development of systematic resources for functional genomics of this organism has been relatively slow. Our aim in establishing FYSSION (Fission Yeast Strains at Sussex; the remainder of the acronym has never been agreed) has been to widen the accessibility of S. pombe as an experimental organism, so that the barriers to studying any aspect of its biology are lowered, particularly for smaller groups.

FYSSION was initiated under the Wellcome Trust’s Functional Genomics Initiative. The first aim was very simple: to provide mutant libraries available for visitors to screen for any desired mutant.

TEMPERATURE-SENSITIVE MUTANTS OF ESSENTIAL GENES

It is currently estimated that 17.5% of the S. pombe genome, or about 900 genes, are essential under laboratory conditions [4]. Many mutants in these genes had previously been isolated by first screening for temperature-sensitive (ts) growth. Often this seems to have involved needless repetition, as unwanted mutants were discarded. A collection of ts mutants had been created and stored by the group of P. Nurse and extended in the laboratory of A. Carr to approximately 2000 strains. We carried out a further mutagenesis with ethyl methane sulphonate to ~50% lethality, and isolated ts mutants until the collection reached its present size of 4000 or approximately 4 per essential gene. Unlike in a screen for a particular phenotype, this required repeated replica plating to check that strains were stably ts.

How big should such a library be? Given the construction of the genetic code and the particular chemical action of the mutagen, any one protocol

Available at http://pombe.biols.susx.ac.uk/
can only create a small proportion of the 380 possible single amino-acid substitutions, and some gene products may be simply refractory to production of a ts form. Thus a collection of this size may include mutants in nearly all of the available genes.

**INSERTIONAL MUTANTS OF NON-ESSENTIAL GENES**

The remaining, non-essential, part of the *S. pombe* genome is clearly unlikely to be represented in the ts collection. This does not, of course, mean that the genes are without interest. In many organisms, such genes are conveniently studied by targeted or random insertion of a selectable marker. A simple technique for large-scale isolation of insertion mutants in *S. pombe* has been reported [5], using transformation with a marker gene to which a few random nucleotides have been added at each end. Although no such method is likely to produce truly random integrations, available evidence indicates that this technique can generate disruptants in many different genes [5].

A collection of such mutants were produced using the *ura4* gene and a strain deleted in this gene, following exactly the published protocol. As with the ts mutants, the aim was to produce a collection of stable mutants, therefore transformants were repeatedly plated on non-selective than selective media.

**LOGISTICS AND HANDLING OF MUTANT LIBRARIES**

The brave new world of functional genomics may appear to require large factory laboratories. In fact, for an organism of 4900 genes, this is not the case: a mutant in every gene can be stored in a stack of 96-well plates occupying less than a freezer shelf. Likewise, the practical technology for handling numbers of this magnitude is very simple.

Our libraries are propagated on conventional Petri dishes or in liquid in multiwell plates. We use a 96-syringe Robbins ‘Hydra’-filling device for filling plates, but it is equally feasible to use multichannel pipettes. For transfer between liquid and solid media we use homemade 48-pin metal replicators (Figure 1); disposable plastic equivalents are available commercially. Although we have had access to a completely automated laboratory robot, we have not found that it offers any advantages in speed or convenience for microbiological work of this type and scale.

**DESIGN AND OPERATION OF MUTANT SCREENS**

Once the libraries are plated on Petri dishes, the principles and pitfalls of screening are not particularly different from those for a conventional mutant hunt. The ts library has the additional constraint that at ‘permissive’ temperature, strains show variable growth rates and some may disappear on repeated replication.

The most straightforward screens are, naturally, those looking for failed or greatly reduced growth in particular conditions. We have conducted a variety of screens for sensitivity to toxins or stresses. Generally, it is advisable to have a simple first screen, if necessary allowing a significant rate of false positives. The reduced numbers can then be checked more carefully. However, we have also carried out more sophisticated screens in which the first step is to change the genotype. With the insertional mutants, a new mutation can be incorporated if it carries a selectable trait and is present in a strain of opposite mating type to the library and lacking the *ura4* gene.
The tester strain is first inoculated as a lawn onto mating media, and the library strains are added using the 48-pin replicator. Once mating and sporulation have occurred, the plate is replicated to rich media for germination, and then to an appropriate medium to select for recombinants. A simple application of this procedure is as a method to introduce a plasmid. For example, we have transformed the library with a plasmid expressing a GFP fusion protein, and then searched microscopically for strains in which the protein is mislocalized.

One feature of the ts library is the proportion of strains with abnormal morphology, even at permissive temperature. This has previously been the basis of some very effective mutant screens in \textit{S. pombe} [6]. We have screened this library for morphological mutants, at the same time staining with vital dyes for internal membranes [7]. Thus, several screens can effectively be carried out at once.

**IDENTIFYING THE GENE**

The first step in analysing the mutants, as with any genetic screen, is backcrossing to show that the phenotype is stable and is the result of a single mutation. From both libraries this leads to a certain attrition rate as described below.

Since the ts library could not be backcrossed at the time it was made, for practical reasons, some mutants inevitably are found to be multigenic. The morphological screen noted above gave quite a rigorous test of this. From 16 mutants with a variety of morphological phenotypes, four were sterile, two had no linkage between ts and morphological phenotypes, two carried two ts mutations unlinked to the morphological phenotype, and the remaining eight had single linked mutations for both phenotypes. Thus the majority of isolated mutants could be pursued. Since the screen was carried out at permissive (ambient) temperature, this is probably a worst case; simpler screens for viability might give a closer correlation to the ts phenotype.

Identifying the mutant gene in ts strains can be a considerable challenge. The general approaches are rescue by transformation with a plasmid genomic library, or genetic mapping. The former can be quite quick, but suffers from the fact that plasmids are not under strict copy-number control in \textit{S. pombe}, and therefore genes which are sensitive to expression level may be refractory to this approach. In addition, it relies on stability of the original ts phenotype; even a small reversion rate can cause confusion. The second approach, mapping, can be very laborious, involving multiple crosses and tetrad analysis. Simply finding enough suitable markers can be a challenge. Given the progress in mapping genes from other species, including man, by molecular methods, a rapid method for mapping \textit{S. pombe} mutations would be of great value (see below).

The insertional mutants have a different problem. The library was made by repeated replications on selective then non-selective medium, to check for stability of the marker in mitotic growth. However, up to 10\% can then lose the marker during meiosis. This probably reflects the considerable anecdotal evidence that DNA can be propagated in \textit{S. pombe} in a state which is neither in a recognizable plasmid nor integrated into the genome; in this state, like plasmids, it is readily lost during meiosis. For the remaining clones, the attraction of insertional mutagenesis is that the mutated gene is now marked with a known DNA sequence. It can therefore be identified by a variety of methods. We use a variation of adaptor-primed PCR [8]. In brief, genomic DNA is digested with restriction enzymes whose recognition sites are absent from the marker gene, and the products made blunt-ended if they are already not. A short double-stranded oligonucleotide adaptor is then ligated, and PCR is carried out using a primer from within the marker gene and one from the adaptor. The design of the adaptor allows amplification only from the restriction fragment containing the marker gene, not the remainder of the genome. The product is then sequenced, identifying the insertion point. We have had good results with this method, but it can fail for various reasons. The distance between the insertion and the restriction sites is unknown, and very large products are less likely to be amplified; so generally several different enzymes are used. Also the method can be confounded by multiple head-to-head or head-to-tail insertions of the marker gene. In different screens, 10–20\% of clones have failed to give a useful product, for these or unknown reasons. Perhaps surprisingly, insertion of the marker at multiple points in the genome seems to be very rare.

**DESCRIBING THE DATA: A PHENOTYPE ONTOLOGY**

To record the data from all screens, a database was constructed using open source software: MySQL.
running on a Linux platform, with a web browser interface. However, it soon became apparent that we faced an unusual problem: how to record phenotype? For example a simple growth screen, such as for drug sensitivity, might be recorded in a database field as ‘hits’ for sensitive strains. What if subsequent analysis showed that different mutants were sensitive to different concentrations of drug? How would this be recorded? More difficult are morphological screens, in which the phenotypes are variable, e.g. long, short, ‘orb’ or ‘banana’ shapes [6]. The database could contain a text field with a simple narrative of results, but this would not capture the information in a systematic way. For this reason, we are designing a ‘phenotype ontology’ to record the results of screens within a controlled vocabulary, and ultimately produce a data structure much more amenable to automated analysis.

In contrast to most functional genomics work with *S. cerevisiae*, FYSSION is in the unusual position, as explained above, that phenotype information precedes genotype. Thus Gene Ontology (GO) [9] is not a solution to the problem. However, we are endeavouring to include some of the characteristics of GO in the phenotype ontology (Figure 2). Our aim is to produce a structure which might be extendable to many other sorts of biological data which would benefit from being made accessible by systematic recording in a controlled vocabulary.

**FUTURE DEVELOPMENTS IN FYSSION**

A question which regularly arises from potential users is whether the libraries can be used to identify mutants in a particular gene of interest. This is the reciprocal problem to efficient identification of mutant genes in the library. We have carried out pilot experiments in which the starting point is a strain deleted for a particular essential gene but carrying a plasmid copy. The aim is to cross this with the ts library, evict the plasmid and select for strains in which no recombination occurs between the marker for the deleted gene and temperature sensitivity. The technical challenges for this and similar high-throughput operations with the ts library revolve around the different growth rates and mating efficiencies of the mutants. However, it is possible that this approach will succeed with further development.

Another general method which would be of great value for FYSSION and for *S. pombe* work in general would be a rapid method for mapping mutations. It may seem surprising that for an organism with a small and sequenced genome, and well-developed genetics, no such method is currently available. In contrast to mapping human genes, and to *S. cerevisiae*, the limitation of *S. pombe* (which can also be a great advantage) is that nearly all laboratories work with strains derived from a single isolate. We have carried out sequence analysis of genes from a different isolate and found that it carries ~0.5% sequence differences from the standard strain, but the two can still mate, albeit at low efficiency. The new generation of sequencing methodologies should make it quite realistic to sequence this strain, revealing polymorphisms throughout the genome, and much more useful information besides. This would form the basis of a mapping technique in which molecular methods are used to look for linkage between the ts mutation and the polymorphisms in the second strain.

**PARALLEL AND FUTURE DEVELOPMENTS IN FUNCTIONAL GENOMICS OF FISSION YEAST**

Recently several other groups have developed resources for *S. pombe* which complement those of
FYSSION. Matsuyama et al. [10] have constructed GFP fusions for nearly every gene. Shimoda and colleagues [11] have formed a reference collection of characterized S. pombe mutants and plasmids. A collaboration between the Korea Research Institute of Bioscience and Biotechnology and the biotechnology company Bioneer has produced a collection of heterozygous diploid deletion mutants of most of the genome [12]. Strains from this collection can be purchased individually and converted into haploids, but the considerable time and resources expended on creating the library mean that at present, routine purchase of the entire collection is likely to beyond most research laboratories. However, in an example of enlightened cooperation between the commercial and academic sectors, several collaborations are in progress which may ultimately lead to the library becoming available to research laboratories at affordable cost. Clearly this will be a major contribution to work on S. pombe.

Currently under development in our laboratory is an efficient system for targeted gene replacement using expression of a bacteriophage recombination protein. The aim is that, once a gene has been ‘tagged’ once with, for example, GFP, the tag can be easily substituted with another, such as a domain for affinity purification. This is carried out by mating in high throughput. Eventually, a collection of tagged strains will be built up so that any new ‘tag’ can be introduced to all of them in a single operation. In combination with improved protocols for live cell microscopy and mass spectrometry of protein complexes, it is hoped to produce a resource which, in line with the original aim of FYSSION, will make the S. pombe genome available to researchers in any aspect of its biology.

Acknowledgements

FYSSION was originally supported by the Wellcome Trust Functional Genomics Initiative. Current work is supported by the Medical Research Council Functional Proteomics Initiative. We thank Thalia Antoniadi and John Walshaw for major contributions to its creation, Evelyn Amoah-Buahin, Antony Carr, Ronak Lakhani and Adam Watson for contributions to its development, numerous students and visiting workers for their involvement, and Paul Nurse and Antony Carr for providing ts strains.

References