Minireview

Tackling an essential problem in functional proteomics of *Saccharomyces cerevisiae*
Oscar M Aparicio

Address: Molecular and Computational Biology Program, Department of Biological Sciences, University of Southern California, Los Angeles, California 90089-1340, USA. E-mail: oaparici@usc.edu

Abstract

Gene inactivation is the cornerstone of functional genetic analysis, but the analysis of essential genes requires conditional inactivation of the gene product. A new study has adapted a simple method for creating conditional alleles to allow large-scale analysis of essential genes in *Saccharomyces cerevisiae* and has identified a role in DNA replication for a newly identified protein complex.

Over thirty years ago, Hartwell and colleagues recognized that the identification and analysis of genes controlling the eukaryotic cell cycle represented a special challenge (and opportunity) because many, if not most, cell-cycle genes would be essential for viability [1]. The isolation of conditional, temperature-sensitive (*ts*) mutant strains of yeast and subsequent phenotypic screening of these strains led to the identification of many cell-division-cycle (*CDC*) genes, setting the stage for tremendous progress in understanding the eukaryotic cell cycle [2]. Today, the availability of complete genome sequences and the advent of proteomic methods promise similar advances in determining the functions of large numbers of uncharacterized genes. In one example, high-throughput gene deletion methods have been used to construct a collection of *Saccharomyces cerevisiae* mutant strains each containing a deletion in one non-essential gene [3]. The study found that 18% of the roughly 6,200 genes in the *S. cerevisiae* genome were essential (because haploid strains with deletions in these genes could not be isolated), and therefore could not be analyzed further. Because most genes with a known role in chromosomal DNA replication are essential and a significant number of essential genes in yeast remain uncharacterized, it seems likely that this 18% includes novel DNA-replication genes and is ripe for analysis. But, analysis of these genes requires a method to allow conditional inactivation of each of the gene products. A recently developed method facilitates analysis of essential genes in *S. cerevisiae* and has revealed a function in DNA replication for a newly identified protein complex [4].

Conditional mutations, and most notably *ts* mutants, have been used with great success to analyze essential genes: certain amino-acid changes in a protein destabilize its structure at an elevated (or lowered) temperature [1]. The potential for generating a useful *ts* mutation in any given gene is not equal, however, because some gene products are less prone to such destabilizing mutations than others; hence, some genes may elude identification in traditional (random) genetic screens for conditional alleles [2]. Although *in vitro* mutagenesis can be used to generate *ts* mutations directly in individual genes of interest [5], these methods are not suited to the rapid analysis of large numbers of genes.

To analyze gene function in yeast on a large scale, methods enabling relatively simple genome modification with universal DNA cassettes that contain repressible promoters have been developed [6-8]. Although these methods are often successful at greatly reducing the levels of gene transcription and eventually protein, rapid and sufficient elimination of protein function remains a problem for many proteins. This is particularly true of cell-cycle studies, where slow gene inactivation can activate one or more of the cell-cycle checkpoints, which block the cell cycle if earlier steps are not completed adequately (reviewed in [9]), and can thereby mask the actual arrest point for a mutant, limiting the direct functional insights.

A potential solution to this problem would be a method that modified individual genes such that the resulting gene

product is tagged for conditional elimination of its function. In principle, the most effective way of eliminating the function of any protein would be through complete proteolysis. Varshavsky and colleagues developed the thermolabile Degron, an amino-terminal protein-degradation signal that is targeted for proteolysis specifically at 37°C through the ubiquitin-mediated 'N-end rule' pathway, which selects aberrant proteins for degradation [10,11]. Fusion of the Degron coding sequence to the amino terminus of a heterologous gene of interest creates a temperature-inducible Degron (td) fusion protein that may be susceptible to temperature-dependent proteolysis [11]. Despite individual successes, degradation of many td fusion proteins is often not sufficiently efficient to eliminate their function [12].

Diffley and colleagues significantly ameliorated this problem by regulating the expression of *UBR1*, which encodes a recognition factor for proteolysis by the N-end pathway [12]. Overproduction of *UBR1* from the galactose-inducible (*GAL*) promoter significantly expands the range of td proteins that are rapidly and effectively degraded. This updated td approach has several advantages over the other methods discussed above. First, a time-consuming and possibly unsuccessful search for a *ts* allele can be avoided. Second, the function of td proteins is typically normal at the permissive condition of 23°C (and even at 37°C when *GAL-UBR1* is off), unlike ts proteins that are often partially defective under permissive conditions. And third, degradation of the td

proteins, and thus elimination of their function, can be very rapid and effective at 37°C when *GAL-UBR1* is induced. This method has been used to re-examine a number of essential replication factors individually, enabling detailed analysis of their functions [12,13].

In a further advance, Labib and colleagues have now adapted the td approach for more rapid analysis of essential genes in *S. cerevisiae* [4]. They have designed a universal cassette that allows simple construction of individual td fusion proteins under the control of the repressible *CUP1* promoter (Figure 1). This cassette is PCR-amplified with a gene-specific pair of long oligonucleotides (approximately 70mers) so that the resulting targeted cassette is flanked by DNA sequences that allow its precise insertion, by homologous recombination, at the amino terminus of the gene of interest. Transformation of the cassette into yeast cells harboring *GAL-UBR1* allows specific degradation of the gene product of interest (Figure 1).

In a post-genomic era re-enactment of Hartwell's seminal screen for *CDC* genes [1], Labib's group has now [4] constructed td fusion proteins of over 60 essential genes of unknown function (about half of the total set of unknown essential genes). Figure 2 shows the scheme that was used to analyze this collection of *td* strains. Cultures were grown at 23°C in the presence of copper and raffinose (permissive conditions). Copper was removed to repress expression of
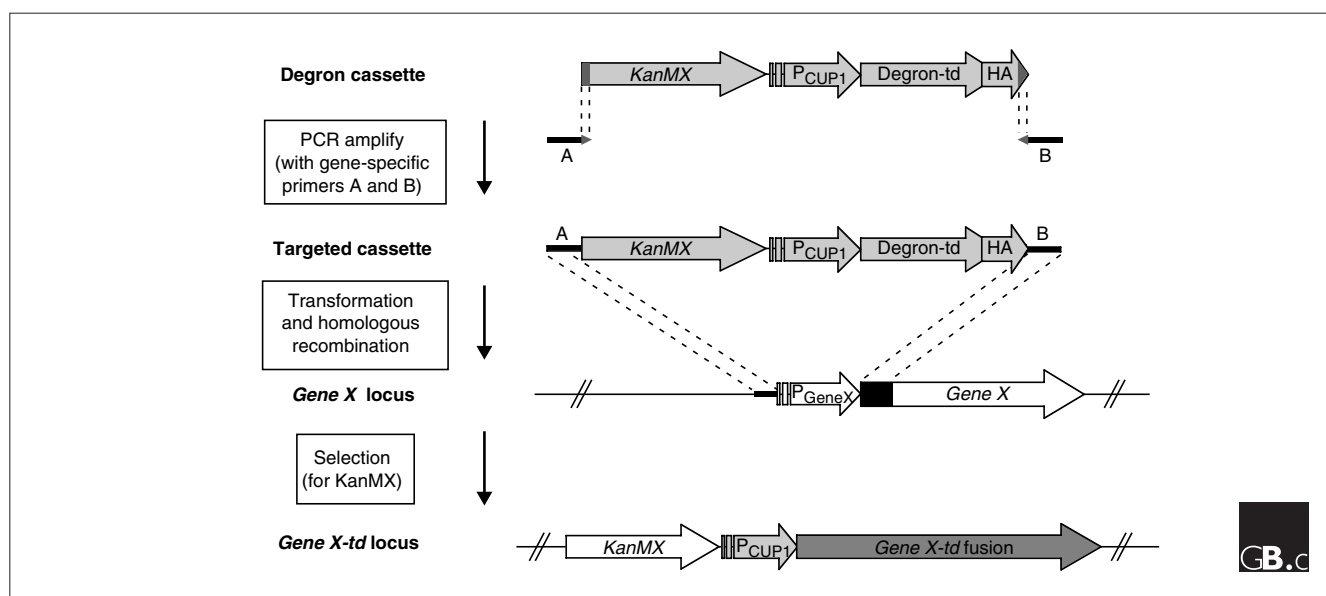

Figure 1
The Degron cassette enables the construction of potential *td* alleles for any gene. The cassette contains *KanMX* for selection (G418 resistance), the *CUP1* promoter (P$_{CUP1}$) that is expressed in the presence of copper and repressed in its absence, and the gene encoding the temperature-sensitive Degron (td) protein fused to the HA epitope. A gene-specific pair of oligonucleotides (A and B), with 3′ ends matching the cassette (shown by vertical dashed lines), is used to amplify the cassette by PCR. Because the 5′ end of each oligonucleotide matches sequences of *gene X* (shown by slanted dashed lines), transformation of the cassette into yeast cells targets the cassette for insertion upstream of the coding sequence by homologous recombination, replacing the native promoter and creating a *gene X-td* fusion. Materials and details of construction are available at [20].

the gene of interest and galactose was added to induce *UBR1*; at this point the cells were shifted to 37°C (non-permissive conditions) to allow *UBR1*-dependent degradation of the td fusion protein. Following inactivation of the td fusion protein, Labib and colleagues analyzed the cellular and nuclear morphology of the cultures to determine whether a characteristic *CDC* phenotype [1] (accumulation of cells with a uniform cellular morphology) was observed. Although most strains did not exhibit a *CDC* phenotype, a handful of strains did, suggesting a defect in cell-cycle progression. Further examination of these strains by DNA-content analysis identified three strains with apparent defects in chromosomal DNA replication, together with additional *CDC* genes that appeared to be involved in other cell-cycle processes but were not described in the current report. Each of these three genes involved in DNA replication, which were named *CDC101* (which was partially viable as a td fusion), *CDC102*, and *CDC105*, has an apparent homolog of unknown function in other eukaryotes, including humans. Interestingly, yeast genome-wide two-hybrid screens for protein interactions had suggested that Cdc101, Cdc102 and Cdc105 interact with each other, possibly forming a complex of unknown function [14]. Biochemical analysis by Labib's group [4] demonstrates that the three proteins do indeed exist in a complex with one other essential protein of unknown function (YOL146W, independently identified as PSF3, see below), which was also viable as a td fusion. Quite remarkably, this four-protein complex is identical to the GINS complex, which was very recently described as having an essential role in chromosome replication in yeast and *Xenopus* [15,16]. GINS (Go, Ichi, Nii, San; five, one, two, and three in Japanese), which consists of Sld5 (Cdc105), Psf1 (Cdc101), Psf2 (Cdc102), and Psf3 (Cdc103), interacts with multiple proteins involved in origin loading of DNA polymerase epsilon and appears to associate with replication forks following initiation, suggesting possible function(s) in initiation and/or elongation. But *ts* mutations of *SLD5* and *PSF1* did not inactivate the function of these proteins sufficiently to discern a specific elongation defect [15].

Focusing on Cdc102 and Cdc105, Labib's group [4] examined the effect on DNA synthesis of depleting each protein at different points during the cell cycle. The efficient degradation of Cdc102-td and Cdc105-td allow clear replication defects to be detected in cells arrested before and after the establishment of replication forks, strongly suggesting a role for these proteins in the establishment and progression of replication forks. These 'execution point' experiments are commonly used in cell-cycle studies to determine the requirement for an individual factor in different cell-cycle processes and to determine the order of action of different factors within a process [2,17]. The success of experiments of this type relies strongly on efficient elimination of gene function, emphasizing the value of the td approach.
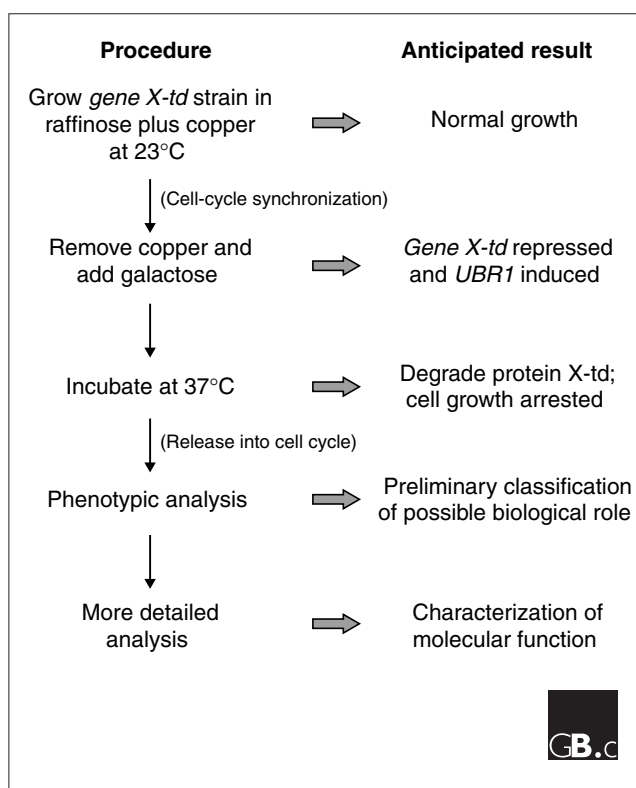

Figure 2
Schematic outline of the procedure for *td*-gene inactivation. Horizontal arrows indicate the anticipated results at each step of the process. Procedures in parentheses are optional steps that may be used to characterize the td protein's function within the cell cycle. Alternative treatments may be used before and/or after td-protein inactivation to determine its effect on different processes. See text for additional description of the scheme.

The main problem encountered with the use of *td* alleles remains the insufficient or ineffective proteolysis of a significant number of td fusion proteins. About 40% of the *td* strains in the Labib lab study [4] were viable, clearly indicating incomplete protein inactivation. Nevertheless, some of these, such as *CDC101-td*, remain informative, especially given the increasing numbers of potential interacting proteins identified by various large-scale experiments that may suggest additional candidates for analysis. In addition, the Degron cassette also introduces the hemagglutinin (HA)-epitope tag into the td protein, facilitating immediate characterization of its abundance and stability by western blot (Figure 1). The majority of *td* strains are therefore likely to be informative and their utility is limited mainly by the availability of appropriate assays for the many diverse biological functions represented in the mutant collection. Other potential problems, such as lethality (under any condition) due to replacement of the native gene promoter with the *CUP1* promoter or to inactivity of the gene fusion itself, were relatively infrequent: only 6% of the *td* strains were inviable [4]. The overall ease of strain construction and analysis makes the *td*

allele the current method of choice for initial phenotypic analysis of essential gene function in *S. cerevisiae*.

In principle, the td approach should be applicable to all eukaryotes, as the N-end pathway is conserved, but efficient proteolysis in other eukaryotes has not yet been achieved [18,19]. In addition, significant improvements in gene targeting will be required before this approach can be applied on a genome-wide scale in multicellular organisms. Until then, the functional characterization of the yeast proteome using *td* alleles will continue to reveal essential functions shared by all eukaryotes.

References

1. Hartwell LH, Culotti J, Reid B: **Genetic control of the cell-division cycle in yeast. I. Detection of mutants.** *Proc Natl Acad Sci USA* 1970, **66:**352-359.
2. Hartwell LH: **Twenty-five years of cell cycle genetics.** *Genetics* 1991, **129:**975-980.
3. Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, *et al*: **Functional profiling of the *Saccharomyces cerevisiae* genome.** *Nature* 2002, **418:**387-391.
4. Kanemaki M, Sanchez-Diaz A, Gambus A, Labib K: **Functional proteomic identification of DNA replication proteins by induced proteolysis *in vivo*.** *Nature* 2003, **423:**720-724.
5. Sikorski RS, Boeke FD: **In vitro mutagenesis and plasmid shuffling: from cloned gene to mutant yeast.** In: *Meth Enzymol.* 1991, **194:**302-318.
6. Belli G, Gari E, Aldea M, Herrero E: **Functional analysis of yeast essential genes using a promoter-substitution cassette and the tetracycline-regulatable dual expression system.** *Yeast* 1998, **14:**1127-1138.
7. Dragon F, Gallagher JE, Compagnone-Post PA, Mitchell BM, Porwancher KA, Wehner KA, Wormsley S, Settlage RE, Shabanowitz J, Osheim Y, *et al*: **A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis.** *Nature* 2002, **417:**967-970.
8. Wu LF, Hughes TR, Davierwala AP, Robinson MD, Stoughton R, Altschuler SJ: **Large-scale prediction of *Saccharomyces cerevisiae* gene function using overlapping transcriptional clusters.** *Nat Genet* 2002, **31:**255-265.
9. Elledge SJ: **Cell cycle checkpoints: preventing an identity crisis.** *Science* 1996, **274:**1664-1672.
10. Varshavsky A: **The N-end rule.** *Cell* 1992, **69:**725-735.
11. Dohmen RJ, Wu P, Varshavsky A: **Heat-inducible degron: a method for constructing temperature-sensitive mutants.** *Science* 1994, **263:**1273-1276.
12. Labib K, Tercero JA, Diffley JF: **Uninterrupted MCM2-7 function required for DNA replication fork progression**. *Science* 2000, **288:**1643-1647.
13. Tercero JA, Labib K, Diffley JF: **DNA synthesis at individual replication forks requires the essential initiation factor Cdc45p.** *EMBO J* 2000, **19:**2082-2093.
14. Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y: **A comprehensive two-hybrid analysis to explore the yeast protein interactome.** *Proc Natl Acad Sci USA* 2001, **98:**4569-4574.
15. Takayama Y, Kamimura Y, Okawa M, Muramatsu S, Sugino A, Araki H: **GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast.** *Genes Dev* 2003, **17:**1153-1165.
16. Kubota Y, Takase Y, Komori Y, Hashimoto Y, Arata T, Kamimura Y, Araki H, Takisawa H: **A novel ring-like complex of *Xenopus* proteins essential for the initiation of DNA replication.** *Genes Dev* 2003, **17:**1141-1152.
17. Hereford LM, Hartwell LH: **Sequential gene function in the initiation of *Saccharomyces cerevisiae* DNA synthesis.** *J Mol Biol* 1974, **84:**445-461.
18. Lindner K, Gregan J, Montgomery S, Kearsey SE: **Essential role of MCM proteins in premeiotic DNA replication.** *Mol Biol Cell* 2002, **13:**435-444.
19. Levy F, Johnston JA, Varshavsky A: **Analysis of a conditional degradation signal in yeast and mammalian cells.** *Eur J Biochem* 1999, **259:**244-252.
20. **European *Saccharomyces cerevisiae* archive for functional analysis - Degron kit** [http://www.uni-frankfurt.de/fb15/mikro/euroscarf/data/degron.html]