The Dissertation Committee for Wei Niu Certifies that this is the approved version of the following dissertation:

Development of imaging-based high-throughput genetic assays and genomic evaluation of yeast gene function in cell cycle progression

Committee:

Edward Marcotte, Supervisor

Andy Ellington

Makkuni Jayaram

Vishwanath Iyer

Scott Stevens

Development of imaging-based high-throughput genetic assays and genomic evaluation of yeast gene function in cell cycle progression

by

Wei Niu, B.S.; M.S.; M.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2007

Dedication

To my parents and my husband

Acknowledgements

First of all, I would like to thank my advisor, Dr. Edward Marcotte for his continuous encouragement, support and mentoring throughout my graduate studies. I would also like to thank my committee members, Dr. Andy Ellington, Dr. Vishy Iyer, Dr. Makkuni Jayaram, and Dr. Scott Stevens for their valuable comments and advices. I would like extend my gratitude to Dr. Andy Ellington and Dr. Vishy Iyer for their support and collaborations.

I am also grateful for working in a great research environment of the laboratory of Dr. Edward Marcotte. I thank every current and past member in the lab, especially Rammohan Narayanaswamy, Alex Scouras, G. Traver Hart, Emily Moradi, Zack Simpson, and Matt Davis for the collaboration on cell chip projects, Zhihua Li for the thoughtful scientific discussions and collaborations on cell cycle project, Kris McGary, Traver Hart, and Mark Carlson for proofreading my proposals, manuscripts and dissertation. I also thank members in Dr. Vishwanath Iyer's lab, Jonathan Davis for starting cell chip project with me, Kerri Jeffers for the help on array printing, Wenjing Zhan for the help with cDNA microarray experiment. I also would like to thank Hong Cui from Dr. Makkuni Jayaram's lab for the help of yeast genetic methods and for providing me the plasmids for mitotic stability assay.

I would like to thank my parents for their support and love. And last but not the least, I would specially thank my dear husband, Qian Chen, who has given me valuable advices, full support, and love. Without his support, it is impossible to finish my degree.

Development of imaging-based high-throughput genetic assays and genomic evaluation of yeast gene function in cell cycle progression

Publication No._____

Wei Niu, Ph.D

The University of Texas at Austin, 2007

Supervisor: Edward Marcotte

Systems biology studies the complex interactions between components of biological systems. One major goal of systems biology is to reconstruct the network of interactions between genes in response to normal and perturbed conditions. In order to accomplish this goal, large-scale data are needed. Accordingly, diverse powerful and high-throughput methods must be developed for this purpose. We have developed novel high-throughput technologies focusing on cellular phenotype profiling and now provide additional genome-scale analysis of gene and protein function.

Few high-throughput methods can perform large-scale and high-throughput cellular phenotype profiling. However, analyzing gene expression patterns and protein behaviors in their cellular context will provide insights into important aspects of gene function. To complement current genomic approaches, we developed two technologies, the spotted cell microarray (cell chip) and the yeast spheroplast microarray, which allow high-throughput and highly-parallel cellular phenotype profiling including cell morphology and protein localization. These methods are based on printing collections of

cells, combined with automated high-throughput microscopy, allowing systematic cellular phenotypic characterization. We used spotted cell microarrays to identify 15 new genes involved in the response of yeast to mating pheromone, 80 proteins associated with shmoo-tip 'localizome' upon pheromone stimulation and 5 genes involved in regulating the localization pattern of a group II intron encoded reverse transcriptase, LtrA, in *Escherichia coli*. Furthermore, in addition to morphology assays, yeast spheroplast microarrays were built for high-throughput immunofluorescence microscopy, allowing large-scale protein and RNA localization studies.

In order to identify additional cell cycle genes, especially those difficult to identify in loss-of-function studies, we performed a genome-scale screen to identify yeast genes with overexpression-induced defects in cell cycle progression. After measuring the fraction of cells in G1 and G2/M phases of the cell cycle *via* high-throughput flow cytometry for each of ~5,800 ORFs and performing the validation and secondary assays, we observed that overexpression of 108 genes leads to reproducible and significant delay in the G1 or G2/M phase. Of 108 genes, 82 are newly implicated in the cell cycle and are likely to affect cell cycle progression *via* a gain-of-function mechanism. The G2/M category consists of 87 genes that showed dramatic enrichment in the regulation of mitotic cell cycle and related biological processes. *YPR015C* and *SHE1* in the G2/M category were further characterized for their roles in cell cycle progression. We found that the G2/M delay caused by the overexpression of *YPR015C* and *SHE1* likely results from the malfunction of spindle and chromosome segregation, which was supported by the observations of highly elevated population of large-budded cells in the pre-M phase, super-sensitivity to nocodazole, and high chromosome loss rates in these two overexpression strains. While the genes in the G2/M category were strongly enriched for cell cycle associated functions, no pathway was significantly enriched in the G1 category

that is composed of 21 genes. However, the strongest enrichment for the G1 category consists of the genes involved in negative regulation of transcription. For instance, the overexpression of *SKO1*, a transcription repressor, resulted in strong cell cycle delay at G1 phase. Moreover, we found that the overexpression of *SKO1* results in cell morphology changes that resembles mating yeast cells (shmoos) and activates the mating pheromone response pathway, thus explaining the G1 cell cycle arrest phenotype of *SKO1* ORF strains.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Systems biology has become a new trend in molecular biology due to the revolutionary development of genomics, although systems analysis has been performed in many areas of biology such as ecology, immunology and developmental biology decades ago (Westerhoff and Palsson, 2004). Systems biology focuses on comprehensive characterization of components of a biological system, including DNA, RNA, proteins, metabolites and informational pathways, and emphasizes the analysis of the responses of these components to systematic perturbations as a whole system (Aderem, 2005; Griffin et al., 2002). Systems biologists have fundamentally different interests from traditional molecular biologists who focus on identifying and analyzing a few aspects of an organism at a time and have relatively limited insights about how the whole biological system responses to perturbations (Ussery and Jensen, 2005). Although systems biology is still in its infancy, the final objectives of systems biology are to reconstruct the network of functional interactions among the genes and proteins in response to perturbed environmental conditions, to identify the markers and mechanisms that are critical to the function of the perturbed system, and perhaps eventually to establish computational models to predict the response of a system to any given perturbation (Butcher et al., 2004; Griffin et al., 2002).

In order to accomplish such goals, high-throughput tools and large-scale studies that can systematically measure the effects of perturbations are needed (Aderem, 2005; Kirschner, 2005). High-throughput technologies have scaled up the scope of systems biology studies, enabling us to view the genome as a 'system' to study (Westerhoff and Palsson, 2004). Different types of high-throughput experiment data were produced by functional genomic technologies such as transcriptomics, proteomics, and metabolomics.

These data greatly facilitate the analysis of higher-order functions involving various cellular processes and eventually lead to the development of the integrated field of systems biology (Burja et al., 2003).

DNA microarray technology has become an effective tool for gene expression or transcriptional analysis (Lockhart and Winzeler, 2000; Schena et al., 1995). A DNA microarray is a glass slide deposited with arrayed DNA fragments, which often come from cDNA or oligonucleotide (Fodor et al., 1991; Hughes et al., 2001; Iyer et al., 1999). cDNA microarrays consist of double-stranded cDNA or PCR products spotted on a glass slide, which can easily allow incorporation of complete human transcriptome (30,000-40,000 genes) (Lander et al., 2001). Oligonucleotide arrays are synthesized or spotted on glass slides at densities that can hold 6 million 25-mer probes per array (Guo et al., 1994; Hughes et al., 2001; Liu, 2007). The measurement of transcribed mRNA has proven to be powerful to distinguish different cell types, differentiate between states in a particular cell type, and discover molecular markers for perturbations. However, in many cases mRNA abundance does not correlate with protein abundance, and the measurement of mRNA abundance can not accurately predict the attenuation of protein abundance or activity caused by post-transcriptional regulation or protein modification and the subcellular localization of expressed protein (Blagoev and Pandey, 2001; Griffin et al., 2002; Lockhart and Winzeler, 2000).

Different from transcriptomics, proteomics focuses on the characterization of many proteins, ideally the whole proteome, within a cell type simultaneously, involving analyses of protein identity and abundance, protein interactions, protein modifications, etc. (Aebersold and Mann, 2003; Ideker et al., 2001). It is an essential component of systems biology research because of enriched information of proteins. For organisms whose genome has been sequenced, mass spectrometry is one of the most sensitive and

powerful global approaches in proteomic research (Ideker et al., 2001; Pandey and Mann, 2000). Shotgun proteomics is a gel-free approach based on multidimensional liquid chromatography coupled to tandem mass spectrometry (LC/LC/MS/MS) (Washburn et al. 2001). In this method, mixed peptides fragments, which are digested from a protein complex or cell lysates, are separated by multi-dimensional liquid chromatography, followed by the fragmentation in the tandem mass spectrometer. The acquired fragmented peptide spectrums are matched to translated genomic databases using SEQUEST algorithm (Link et al., 1999). Proteomics has been considerably broadened in the past few years, yet it is need to improve with respect to sensitivity and dynamic range of mass spectrometry and capability of analyzing more than two cell states or more than one modification at a time (Aebersold and Mann, 2003; Steen and Pandey, 2002).

In addition to transcriptomics and proteomics, a third main branch in 'omics' is metabolomics. It represents an emerging technology in postgenomic era, which aims to identify and quantify the entire complement of small molecules and metabolites in the cell under various conditions (Borneman et al., 2007). The changes observed in the metabolome depend on the composition of both the intra- and extracellular environments, and directly suggest which genes affected which function of a particular pathway (Borneman et al., 2007; Burja et al., 2003). Mass spectrometry (MS) and nuclear magnetic resonance spectroscopy (NMR) have currently been used in metabolism profiling (Rochfort, 2005; Weckwerth, 2003). However, unlike transcriptomics and proteomics, metabolomics is currently a method-based science, with limited numbers of studies that are generally focused on subsets of metabolome. Once the technology has been caught up, the amounts of research data should accumulated quickly for better understanding the dynamics of metabolic networks in response to specific perturbations (Burja et al., 2003; Weckwerth, 2003).

Akin to transcriptomics, proteomics and metabolomics, “cellomics” or “cellular genomics” is another newborn ‘omics’ technology. It is the fusion of genomics and cell biology, seeking to understand cellular systems by analyzing single cells (Levsky et al., 2002; Perutka et al., 2004). Transcriptomics, proteomics and metabolomics rely on the sample extracted from cell lysates by destroying the cellular structural context, and it is almost impossible to monitor the gene and protein behaviors on spatial or temporal level in intact cellular compartments (Levsky et al., 2002). However, such cellular data provide important cues for understanding genetic circuitry (Levsky et al., 2002; Perlman et al., 2004). For example, DNA microarray data can tell us when and to what extent a gene is expressed, but it is unable to shed insight on the co-ordination between mRNA expression and specific sites where the corresponding transcription occurs (Levsky et al., 2002). Levsky *et al.* describe a single-cell gene expression profiling technology, “cellular transcription profiling”, which can monitor mRNA synthesis by visualizing specific sites of transcription using fluorescence *in situ* hybridization (Levsky et al., 2002). By probing mRNAs with 10 different fluorophore labeled oligomer DNA probes, they were able to simultaneously monitor 10 mRNAs’ syntheses site and expression level in a single human colon adenocarcinoma (DLD-1) cell. This work provided a way to analyze the coordinated transcription events and organization of gene expression at the cellular level.

Currently, chip-based microfluidic systems (Palkova et al., 2004; Takahashi et al., 2004) and transfected cell array technology (Conrad et al., 2004; Silva et al., 2004; Ziauddin and Sabatini, 2001) have been major tools in cellular genomics. Microfluidic technology is a complementation of flow cytometric cell assay in microfluidic devices, consisting of microfluidic glass chip and optical detection system (Palkova et al., 2004; Takahashi et al., 2004). Briefly, cells prestained with fluorescent markers and dyes are loaded into sample wells in the microfluidic glass chip, and then moved in microfluidic

channels by a pressure-driven flow system. Cellular fluorescence parameters are automatically detected by a fluorescence microscope, one microfluidic channel at a time. Data are normally displayed as electropherograms or as virtual gels. Compared to traditional microscopy analysis, this technology has the advantage of analyzing larger cell population (up to several million cells per sample). However, this microfluidic system can only contain several different samples at a time and has a limitation for cell size due to microfluidic channel dimensions (Palkova et al., 2004).

In contrast, transfected cell array technology holds a great promise for high-throughput characterization of cellular phenotypes. Transfected cell arrays are manufactured by depositing nanoliter amounts of plasmid-gelatin solution containing cDNA expression constructs or RNAi constructs onto glass slides, followed by the treatment of the slide with lipid-based transfection reagent and the attachment of cells to the slides for taking up the cDNA or RNAi constructs in a petri dish. After successful transfection, cells can be visualized by incubation with fluorescence-labeled antibodies against the proteins, direct fluorescent microscopy (GFP) or other detection procedures (Conrad et al., 2004; Ziauddin and Sabatini, 2001). No doubt, transfected cell array provides a new type of high-throughput platform for large-scale functional genomic studies. However, transfected cell arrays have major limitations: the difficulties of making the genomic cDNA expression constructs, achieving high transfection efficiency in some cell lines, and separating the processes of imaging and cell growth.

To exploit the availability of strain collections such as yeast haploid deletion strains (Giaever et al., 2002; Winzeler et al., 1999) and GFP-tagged yeast strains (Huh et al., 2003), we developed the spotted cell microarray (cell chips), a high-throughput technology for measuring cellular phenotypes including cell morphology, protein and RNA localization. As described in chapter 2, spotted cell microarrays, unlike transfected

cell microarray, are made by directly depositing suspensions of cells from an arrayed library onto coated glass slides using a printing arrayer and imaging the stained subcellular features, which avoids the laborious cDNA expression or RNAi construct procedure and tedious trouble shooting for high-efficient transfection. We first screened genes affecting normal cellular morphology and genes affecting the response of yeast to mating pheromone across the yeast haploid deletion collection (Narayanaswamy et al., 2006), as described in chapter 1. Cell chips have applications beyond yeast morphology screens and can in principle be extended to any organism or cell type for which defined libraries of cells can be arrayed, such as other easily manipulated organisms, banks of bacteria, and deletion libraries for other microorganisms (Narayanaswamy et al., 2006). Two major applications are described in chapter 3: utilizing pre-established yeast GFP-tagged strain collection to systematically survey the spatial dynamics of the yeast proteome in polarized growth and constructing *Escherichia coli* cell microarrays from ~5000 *E. coli* knockout strains transformed with plasmids (pACD2X-GFP/LtrA) for screening genes affecting the polar localization of LtrA, a group II intron-coded retrotranscriptase.

We eventually anticipate that the biggest advantage of cell chips will be analyzing the localization of proteins and RNAs by high-throughput immunostaining and *in situ* hybridization. For example, we can print multiple identical cell chips, and then probe each slide with a different set of dyes or antibodies. Each chip then becomes a unique assay for the dye or antibody target across the set of genetically distinct strains (Narayanaswamy et al., 2006). In chapter 4, yeast spheroplast chips were developed to analyze protein localization by high-throughput immunofluorescence microscopy. Cells from an arrayed yeast library are fixed, spheroplasted, and spotted onto poly-Lysine coated slides, and then stained on cell chips instead of in the sample wells, which

effectively separate the growth of the cells from the imaging process (a strategy difficult to achieve with plate assays or transfected cell arrays). The resulting images would indicate synthetic genetic interactions between the probe targets and the deleted genes. As a proof of principle, we constructed yeast spheroplast chips from yeast tetracycline-repressible promoter strain collection for examining microtubule and spindle dynamics (Mnaimneh et al., 2004), demonstrating that yeast spheroplast chip will eventually provide a useful tool to perform large-scale protein localization analysis by adapting traditional cell biology technology (immunofluorescence microscopy) to high-throughput genomic technology (cell chips).

As mentioned earlier, large-scale data are important for systems biology studies. In chapter 5, we performed a large-scale screen for identifying genes that affect cell cycle progression when they were overexpressed, in the interest of providing a systematic and comprehensive data set for gene function in cell cycle progression. Budding yeast shares a very similar process with other eukaryotic organisms except that the nuclear envelope does not break down during mitosis and the daughter cell comes from bud emergence (Hartwell, 1974). Because of its rich genomic information and versatile genetic manipulations, the budding yeast has become a fertile field to study cell cycle division, which would accelerate our understanding of cell cycle control in other organisms, and finally contribute to the understanding of the human diseases caused by the defects of cell cycle progression. Although there had been extensive effort on discovering new cell cycle genes, additional cell division cycle (CDC) genes remain to be uncovered, especially genes that are difficult to identify by recessive loss-of function approaches (Stevenson et al., 2001). In an effort to complement previous cell cycle studies and in the interest to discover new cell cycle genes, we took a comprehensive screen for genes

function in cell cycle progression by using the most complete yeast overexpression strain collection that covers 91% of the yeast genome (Gelperin et al., 2005).

In summary, the present study is focused on two important aspects of systems biology: developing new high-throughput technologies for cellular phenotype profiling and providing more comprehensive information on the well known but essential biology processes. The following chapters focus on the detailed procedures that were carried out and the analyses of the results that were obtained in the present study. Chapter 2 describes the construction of spotted cell microarrays (cell chips). Chapter 3 describes the applications of cell chips to different yeast libraries and species other than the budding yeast. Chapter 4 describes the development of yeast spheroplast microarray that focuses on large-scale protein and RNA localizations. Chapter 5 focuses on the study of identifying new genes that confer defects in cell cycle progression when overexpressed. Major conclusions and future directions for present study are summarized in chapter 6.

Chapter 2: Systematic profiling of cellular phenotypes with spotted cell microarrays reveals new mating pheromone response genes

2.1 INTRODUCTION

A major goal in systems biology is to define the biological functions of the genes encoded in each genome and to reconstruct the network of functional interactions that underlies normal and altered cellular and organismal biology (Collins et al., 2003). DNA microarrays, mass spectrometry, and protein interaction screens have been powerful tools in this regard (Lockhart and Winzeler, 2000; Pandey and Mann, 2000), but it is important to employ diverse technologies addressing independent aspects of gene function in order to generate complementary datasets (Hughes et al., 2004). In particular, spatial, temporal and phenotypic data provide important clues for understanding genetic circuitry.

In this chapter, we describe a technology for measuring cell morphology and subcellular localization phenotypes, applied to a model system in which yeast change morphology in response to mating pheromone (Elion, 2000; Fields, 1990). Wild-type haploid yeast cells, on detecting pheromone of the opposite mating type *via* a cell surface receptor, heterotrimeric G protein, and mitogen-activated protein (MAP) kinase-mediated signal transduction cascade, arrest their cell cycles in G1 phase and grow in a polarized fashion towards the mating partners, forming a characteristic cell shape termed a "shmoo" (Dohlman and Thorner, 2001). Cells of opposite mating type fuse, producing a diploid organism. Several hundred genes change expression during this process (Roberts et al., 2000). The pheromone response MAP kinase cascade is broadly conserved across eukaryotes, yet even this signal transduction pathway is incomplete. Here, we describe the development of spotted cell microarrays and their application in defining additional genes controlling the response of yeast cells to mating pheromone.

2.2 MATERIALS AND METHODS

2.2.1 Yeast deletion strains and growth conditions

Cell microarrays (cell chips) were manufactured containing all the strains from the *S. cerevisiae* haploid deletion collection. This is a set of strains in the BY4741 genetic background (MATa *his3*Δ *leu2*Δ *met15*Δ *ura3*Δ) generated by the international yeast deletion consortium, in which each strain contains a chromosomal gene replacement of a non-essential gene with a selectable KanMX4 marker that confers resistance to the antibiotic G418 (Giaever et al., 2002; Winzeler et al., 1999). We obtained the arrayed library of 4852 such haploid deletion strains from Invitrogen. Frozen cell cultures in 96-well plates were thawed and used to inoculate 96-well Costar tissue culture plates with 200 μl YPD medium containing the antibiotic G418 (200 mg/L) and 17% glycerol, using a Beckman Biomek FX 96-well pipetting robot to perform all pipetting operations. Copies of the strain collection were incubated for growth at 30 ^{o}C, monitoring cell growth by optical density measurement at 600 nm using a 96-well plate reader. Quality control for sterility and cross-contamination was performed by monitoring control wells empty of cells in the master plates. After growth at 30 ^{o}C for 2 days, copy plates were agitated with a plate shaker, sealed and frozen at -80 ^{o}C, with each copy thawed prior to use for printing microarrays of cells.

2.2.2 Alpha factor treatment

The yeast deletion collection was grown to saturation in 96 well plates. Each plate was sub-cultured into fresh YPD medium in 96 well Costar tissue culture plates and allowed to grow for 36 hours at 30 ^{o}C without shaking. The plates were spun and washed multiple times in YPD medium, pH 3.5 to inactivate the Bar1p protease. Alpha factor was added to each sample well at a concentration of 350 μg/ml, a concentration

sufficiently high to induce shmoo formation in >90% of wild type BY4741 cells grown at different densities. After 4 hours of treatment at 30 ^{o}C, the cells were fixed in 3.7% formaldehyde for 1 hour at room temperature and washed with YPD medium containing 17% (w/v) glycerol. At this stage, 20 mM $CaCl_2$ and 20 mM $MnCl_2$ were added to each well. The cells were now ready to be spotted onto pre-cleaned glass slides coated with concanavalin A (ConA).

2.2.3 Printing of cell microarrays

Cell microarrays were printed by contact deposition of suspensions of yeast cells from the alpha factor treated yeast strain collection onto ConA or poly-L-lysine coated glass slides using a custom-built DNA microarray printing robot. Printing was carried out using conically tapered 1/16-inch diameter stainless steel printing tips with 0.0015-inch slots (Majer Precision Engineering; MicroQuill 2000) that are sterilized between print runs. The resulting spots are ~200 μm in diameter, spaced 410 μm apart; spots were printed in 12 blocks, each 21 spots in width. In a standard print run, the tips were rinsed and vacuum dried 3 times after each loading and printing step, sufficient to prevent carryover of cells from one well to the next, as judged by microscopic inspection of putatively empty spots. During printing, the 96 well plates were kept under a clean acrylic cover at all times except during pick up of the cells. All surfaces, including the vacuum slide platter and the underside of the acrylic dust cover are sterilized by wiping with 70% ethanol. In our experience, these procedures ensure that there is no detectable contamination of wells in the plates or of spots on the microarray. After printing, the slides were centrifuged flat at 1500x g for 5 minutes in a swinging bucket centrifuge adaptor to promote the adherence of the cells to the slide surface. Cell microarrays can be imaged immediately at this point or stored at 4 ^{o}C or - 80 ^{o}C for extended periods of time. To prevent condensation when thawing slides stored at -80 ^{o}C, frozen slides were rapidly

thawed by dipping briefly in room temperature 95% ethanol, and then centrifuged dry in an empty 50 ml conical tube at 600 rpm for 5 minutes.

2.2.4 Scanning of cell microarrays

Cell microarrays were imaged in two steps: first, the lattice of cell spots was determined using a standard DNA microarray scanner, then each spot was imaged using an automated microscope. Prior to staining and imaging, each slide was marked with four reference marks using a diamond scribe (two of these marks are visible at the bottom of the image in Figure 2.1B), then the slides were scanned using an Axon GenePix 4000A/B microarray scanner. The spots of cells were detected as bright spots in the 545 nm detector channel because of light scattering by the cells and by the droplet of media or dried liquid at each spot (see Figure 2.1B). GenePix scanner software was then used to fit a two-dimensional grid over the spots to define the block, row, and column location of each spot, thus providing an x, y coordinate with each spot in the scanner's system of coordinates. These x, y coordinates were written out to a GenePix GPR-format file, as well as the associated strain identities (stored as a GenePix Gene Array List (GAL) file). At this point, each slide had an associated set of coordinates describing the relative locations of each cell spot, their identities, and the locations of the reference marks. Spot coordinates were then converted from the GenePix coordinate system to the optical microscope coordinate system through the use of the four reference points and an affine transformation. Slides were then stained or otherwise manipulated prior to microscopy. For typical bright field or DIC microscopy, slides were washed with water after scanning in order to remove glycerol, dried *via* 5 minutes of centrifugation, and one drop of mounting media are applied containing 100ng/ml DAPI nuclear stain. Slides were then covered with 24x60mm coverslips and sealed with nail polish.

Automated microscopy was carried out using a Nikon E800 with the CF160 optical system, and outfitted with a motorized X-Y stage with 0.1 micron resolution, a piezoelectric auto-focus device for 9.7 nm focusing resolution, a Photometrix Coolsnap camera with 1392x1040x12 bit pixel resolution, filters for Differential Interference Contrast (DIC), fluorescence, and visible wavelengths, and MetaMorph software. First, the reference marks on the slide were found and their positions recorded using the microscopes coordinate system. An affine transformation matrix was derived that converts coordinates in the GenePix coordinate system to that in the microscope coordinate system, then applied to all points in the GPR file output from the GenePix scanner, creating a MetaMorph format STG file containing the coordinates of all spots converted into the microscope's coordinate system. Images were collected at each spot by executing a MetaMorph 'journal' macro at each spot listed in the STG file that autofocused and captured DIC and fluorescent images, saving each image in TIFF and JPEG format. An entire slide with ~5000 spots can be imaged in ~10 hours, capturing both fluorescent and DIC/brightfield images.

2.2.5 Scoring cellular morphology phenotypes

Two graders independently evaluated the set of images from a yeast cell microarray for strains with atypical morphologies. Phenotypes were scored related to cell size (mutant phenotypes being *large* or *small* with respect to the wild-type control), cell shape (mutant phenotypes being *round*, *elongated*, *pointed* with respect to the wild-type ovoid shape) or either a *pseudohyphal*, *clumped* or *polarized bud growth* or other *budding* defects. While the scoring of phenotypes on these alpha-factor treated cell chips was in progress, we also examined shmoo phenotypes of several handpicked deletion mutants that had previously been identified as cell morphology mutants in our earlier cell microarray analyses. The shmoo phenotype of these mutants was compared to that of

wild-type cells as well as cells defective for genes known to have a role in the pheromone response-signaling pathway.

2.2.6 Yeast growth curves +/- alpha factor

To test if yeast strains arrested growth in the presence of alpha factor, selected strains were picked from the yeast deletion library and grown in YPD overnight until they attained log phase growth. The cultures were spun and washed with YPD pH3.5 to inactivate Bar1p protease. The cultures were subsequently split into replicate 96 well plates, with and without alpha factor at a final concentration of 25μg/ml, when OD600 reached at ~0.2-0.5. The plates were incubated at 30 ^{o}C for 10 hours without shaking and their absorbance was recorded at 600 nm each hour. The slope of each growth curve was calculated from a plot of log OD600 vs. time. The affect of alpha factor on the strains was obtained as the ratio of the slope from the untreated sample to that of the alpha factor treated sample. Average slope ratios were calculated from 2-3 independent assays. This analysis, in combination with the cell microarray alpha factor treatment analysis, allowed us to identify a number of genes that were affected in their ability to form shmoos after alpha-factor.

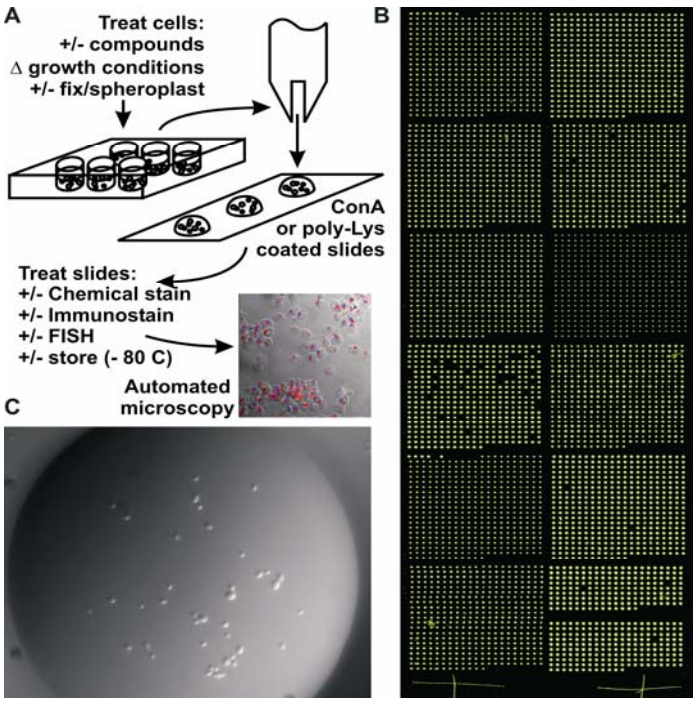


Figure 2.1 An overview of spotted cell microarrays**.** (A) "Cell chips" are constructed using slotted steel pins to robotically print cells from 96 well plates onto poly L-lysine or ConA coated glass slides. The sample image shows arrayed yeast cells immunostained for tubulin using FITC-conjugated-goat anti-rat IgG/rat anti-α-tubulin (red), overlaid on a bright field image and a DAPI-stained image (blue) of the cells' nuclei. (B) Wide-field light scattering image of a cell microarray (approx. 2 cm x 6 cm) containing ~4,800 viable, haploid yeast deletion strains. The bright dots arise from light scattered when scanning the array with a Genepix DNA microarray scanner. Spots are ~200 μm in diameter, separated by 410 μm. (C) Close-up of a typical spot from the microarray showing distinct cells at 40X magnification. This image was taken immediately after printing, so growth medium (YPD, 17% glycerol, 200 mg/L G418) is still visible.

2.3 RESULTS

2.3.1 A high-throughput screen of yeast cellular morphology

Cells were printed from each of the 4,848 distinct haploid yeast deletion strains onto glass microscope slides coated with poly L-lysine or concanavalin A using a custom-built high-speed robotic arrayer that is normally used to manufacture DNA microarrays (Lashkari et al., 1997). Figure 2.1B shows an image of a cell microarray printed using this methodology. Each spot normally contains ~20-40 cells from a single deletion strain, as seen in Figure 2.1C using a standard microscope. Our preliminary data indicate that arrayed cells remain viable and physiologically normal after printing and washing, although cells are typically fixed for imaging purposes. A cell chip is analyzed using an automated fluorescence microscope to sequentially autofocus and image each spot.

As an initial proof-of-concept, we first performed genome-wide differential interference contrast (DIC) imaging to examine the effects of deleting each yeast nonessential gene on basic aspects of cellular morphology such as cell shape, size, budding pattern and clumping, from which we expected to find genes controlling fundamental cell growth processes. Systematic analysis of the haploid yeast deletion strain phenotypes on 2 slides (~10,000 images) reveals that ~2,000 of the 4,848 strains exhibited atypical morphologies of varying degree. Two independent graders manually assigned numerical scores to phenotypes by severity, penetrance in the population, and type (large, small, elongated, round, and clumped (Giaever et al., 2002), as well as polarized bud growth and pseudohyphal-like morphology). Control experiments were performed by constructing cell chips from known morphology mutants and wild-type strains, and grading these in the same grading scheme. 383 deletion strains (8%), were considered to have severe morphology defects (Figure 2.2A) to a degree considered

significant in the control experiments, with an estimated precision of 82% and recall of 26%.

Genes deleted from strains with an observed morphology defect were often functionally diverse. Nonetheless, certain general functions were enriched: elongated strains were enriched (p<0.01, as calculated by ref. (Robinson et al., 2002)) for genes operating in nucleic acid metabolism, cell cycle defects, transcription, and meiosis; large strains were enriched for transporter defects; round strains for cell wall, budding, cell polarity, and cell differentiation genes; small strains for mitochondrial, carbohydrate metabolism, and phosphate transport genes; and strains with polarized bud growth defects for budding, cell polarity, and filament formation genes. Large and elongated strains significantly (p<0.01) overlapped strains previously identified with these phenotypes during analysis of the homozygous diploid yeast deletion strains (Giaever et al., 2002).

2.3.2 Systematic identification of genes controlling mating pheromone response

Having established the typical morphology of each haploid deletion strain, we examined the primary morphological differentiation pathway in budding yeast, the response of the cells to the mating pheromone alpha factor during sexual conjugation. Although this pathway is well-studied (Dohlman and Thorner, 2001), it has yet to be analyzed to completion. We reasoned that additional genes affecting the pheromone response pathway, either directly or indirectly could be identified by examining shmoo phenotypes when the deletion collection was treated with alpha factor. We treated the entire mating type “a” haploid yeast deletion collection with alpha factor, then constructed and imaged spotted cell microarrays from the treated and fixed cells. Two graders manually examined the cell images for the absence of shmoos, grading the images on a numerical scale. Consistency between graders was high, and no systematic grading differences were apparent. 142 strains appeared to have defects in shmooing,

forming either no shmoos or barely detectable shmoos in the imaged fields of cells (Figure 2.2B). These 142 strains represent a mixture of genes participating in the pathway and false positive results in the large-scale screen. We filtered this set for reproducible shmoo defects by manually retesting the 142 strains twice *via* alpha factor addition and microscopic imaging; 54 of the 142 strains showed consistent shmoo defects. Of these strains, 10 were previously identified as diploid or MAT alpha strains in the MATa haploid strain collection (A. Tong & C. Boone, personal communication), which correctly appear insensitive to alpha factor in this screen. Removing these strains and 6 strains whose deletions could not be confirmed by PCR or whose phenotype failed to reproduce in a reconstructed strain leaves 38 MATa haploid strains reproducibly defective in shmoo formation.

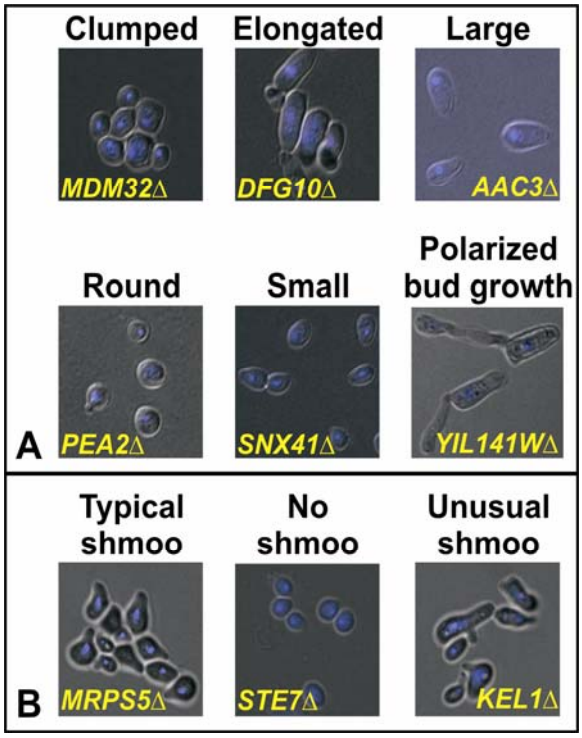


Figure 2.2 Characteristic yeast cell phenotypes observed on cell arrays**.** Images were collected automatically as DIC images at 60X magnification with DAPI-stained nuclei superimposed in blue pseudocolor. (A) Six phenotypic classes observed among the haploid yeast deletion strains. YIL141W overlaps the AXL2 gene, whose disruption in the deletion strain probably provides the observed morphology. (B) Changes in cell morphology observed after treating the deletion collection with mating pheromone. Many mutants, such as the MRPS5 deletion strain (left), form 'wild-type'-like mating projections upon adding alpha factor, while cells lacking STE7 (middle) fail to form mating projections, and cells lacking KEL1 (right) form mating projections of unusual morphology.

2.3.3 Independent validation of mating pheromone response genes

To validate the involvement of these genes in the pheromone response pathway, we followed up the high-throughput cell chip screen by conducting growth assays measuring the tendency of the strains to arrest growth upon pheromone exposure. We tested the complete set of 142 deletion strains (i.e., the 38 reproducibly defectively shmooing strains and the remaining strains whose defects failed to reproduce) plus 271 additional deletion strains as controls with either normal shmooing (wild-type like, as determined from the cell microarray screen) or enhanced shmooing (marked by increased frequency of shmoos in the cell population), as well as strains deleted for 28 of the 41 genes previously known to be involved in the pheromone response pathway. The positive controls are clearly differentiated from the normally shmooing strains in this assay (Figure 2.3), except for inhibitors of the pathway. In all, 5 of the positive control strains fail to arrest growth in this assay, including strains deleted for *BAR1*, the protease that

degrades mating pheromone (Ballensiefen and Schmitt, 1997), and *DIG2*, which inhibits pheromone-responsive transcription (Olson et al., 2000).

Figure 2.3 shows 30 of the 38 reproducible shmoo defective strains fail to arrest growth upon exposure to alpha factor to an extent comparable to the positive controls. Lack of growth arrest agreed well with reproducible shmoo defects. These strains were defective in both shmoo formation and growth arrest, implicating the deleted genes in the pathway. An additional 4 MATa haploid strains first identified as shmoo defective, but not among the 38 reproducibly shmoo defective strains, also fail to arrest growth upon alpha factor exposure, implicating the deleted genes in the pathway. Enhanced shmooing strains arrest even more strongly and appear systematically hypersensitive to the pheromone (Figure 2.3). Thus, the extent of growth arrest in this assay correlates well with the cell-chip measured penetrance of shmooing across the populations of cells.

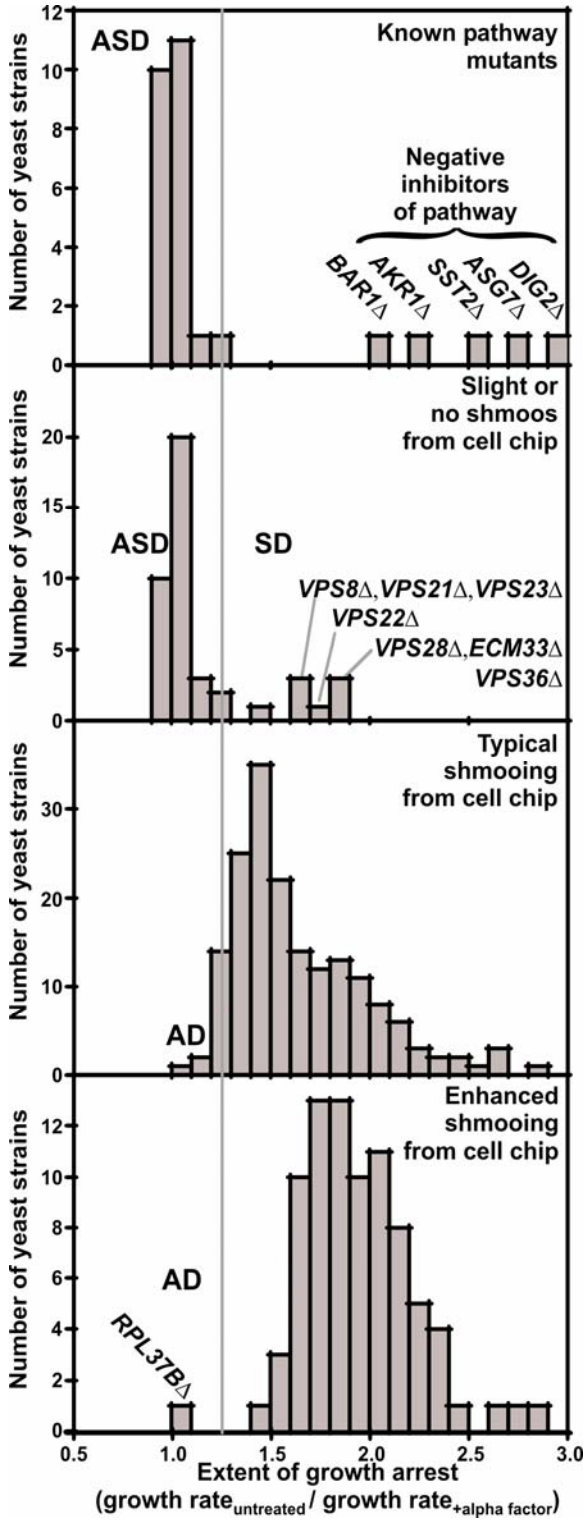


Figure 2.3 Results of a cell microarray-based genome-wide screen for genes participating in the mating pheromone response pathway**.** Strains defective in the pathway fail to arrest growth when treated with alpha factor, unlike wild-type cells. The histograms report the average results of 2-3 replicate growth assays for (from top to bottom) 28 strains containing deletions of genes known to participate in pheromone response, 38 strains identified from cell microarrays as failing to shmoo properly, 178 strains forming typical shmoos, and 91 strains forming shmoos with a notably enhanced frequency in the cell population. The true positive alpha factor response pathway mutants (ASD, arrest+shmoo defective) are well separated from non-pathway mutants. Additional mutant categories identified were those defective only in the shmoo pathway (SD, shmoo defective), and those defective only in the growth arrest pathway (AD, arrest defective).

2.3.4 Comparison with known pathway implicates new genes in pheromone response and shmoo formation

As two distinct phenotypes were assayed, growth arrest and shmoo formation, we expected to find genes defective in either or both pathways—a defect in both implicates the gene in the initial alpha factor response pathway or in both downstream pathways, while a defect in only one implicates the gene in the corresponding downstream pathway (Figure 2.4). We first investigated mutants exhibiting both defects (termed 'ASD', for 'Arrest and Shmoo Defective'), implicated in pheromone detection and signaling. Comparison with the known pathway (Dohlman and Thorner, 2001) (Figure 2.5) shows that of the 41 genes previously known to be in the pathway, 15 were recovered in the cell microarray experiment. Examination of the remaining genes is revealing: 10 genes are not represented in the deletion library (many are essential), 13 genes are inhibitors of the pathway and are thus not expected to be observed in either screen, as the deletion strains still shmoo, and the remaining 3 genes were missed for technical reasons related to image focus or low cell count. Thus, of the 31 genes expected to be found in this screen, 15 (48%) were correctly identified, including components of the receptor coupled heterotrimeric G protein (*STE4*, *GPA1*), the MAP kinase signal transduction cascade (*STE20*, *STE11*, *STE5*, *STE7*, *FUS3*, *FAR1*, *STE50*), and silencers of mating loci (*SIR1*, *SIR2*, *SIR3*). Recognizing that negative regulators may not be found in this screen raises the recovery rate to 15/18 genes, or 83%. Interestingly, strains with deletions of certain negative regulators such as *HSL7* and *DIG1* are shmoo defective and we correctly identify them in the screen (Figure 2.5).

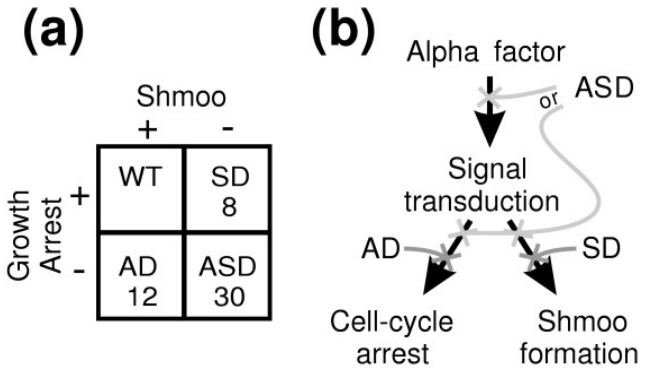


Figure 2.4 Summary of cell-chip/growth assay results**.** With two independent screens, we expected 3 classes of mutants: true positive alpha factor response pathway mutants (ASD), those defective only in the shmoo pathway (SD), and those defective only in the growth arrest (AD). (a) shows the number of genes identified in each category, (b) shows their interpretation. Only 413 strains were tested by growth assay, so the number of strains with wild-type phenotypes (WT) is omitted.

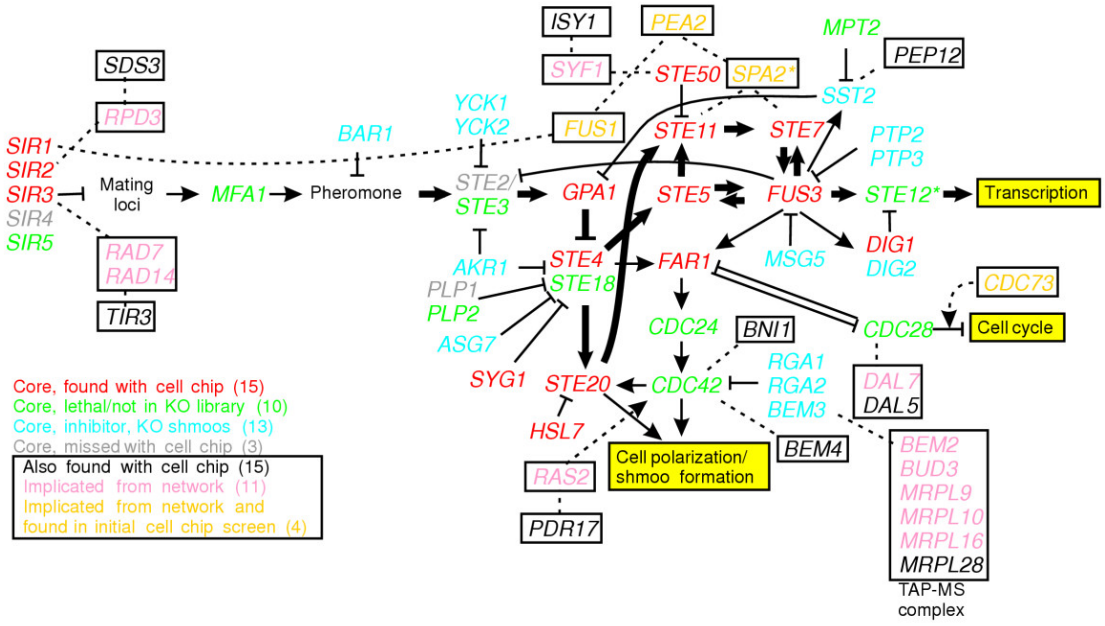


Figure 2.5 Comparison with the known response pathway revealed new genes in pheromone response pathway**.** The comparison reveals that of the 18 known genes expected to be found in this screen, 15 were recovered (red labels). 13 genes are pathway inhibitors (blue labels) whose corresponding deletion strains shmoo. 10 known pathway genes are absent from the deletion collection (green labels). Of the 15 additional genes found, 9 (black labels, boxed) could be associated with the core pathway via protein interactions or mRNA co-expression with intermediates (pink labels, boxed). 4 network-implicated intermediates (orange labels, boxed) were also found in the initial cell chip screen, though not reconfirmed. Bold arrows mark the canonical signal transduction cascade leading to transcriptional changes. Arrows indicate activation; flathead arrows, inhibition; dotted lines, functional genomics linkages. Genes with asterisks are also implicated in filamentous growth.

Beyond the known pathway, 15 genes were found that fail to shmoo and fail to arrest growth upon exposure to alpha factor. Examples include genes with clear functions in polarized growth (*BEM4* and *BNI1*), as well as *PEP7(VPS19)*, *PEP12(VPS6)*, and *VPS3(PEP6)*, three genes involved in vesicular trafficking and vacuole protein sorting (Katzmann et al., 2002), suggesting specific involvement of this system in pheromone response, possibly related to vesicular trafficking's roles in pheromone receptor localization, endocytosis, or recycling (Dulic and Riezman, 1990; Miller et al., 2000). There is a general implication of genes affecting membrane properties, including *PDR17*, controlling phospholipid synthesis/transport (van den Hazel et al., 1999) and *LAS21*, controlling glycosylphosphatidylinositol-linked protein transport/remodeling (Benachour et al., 1999). Several plasma membrane transporters are identified (*QDR2* and *DAL5*), as is the cell wall biosynthetic enzyme *YEA4* and mannoprotein *TIR3*. Loss of any of these genes disrupts pheromone response, possibly indicating membrane properties feeding back into control of mating response, consistent with the important role of plasma membrane reorganization in shmooing (Bagnat and Simons, 2002).

Finally, we identified strains defective in only one of the two assayed phenotypes, implicating the genes in downstream pathways. The set of strains that fail to arrest yet shmoo properly (termed 'AD' for 'Arrest Defective') was functionally diverse as well as small (in part because only ~8% of the deletion collection was tested for growth arrest—we expect more such mutants given a complete screen for growth arrest). These strains were deleted for *FMP35, RPL37B*, *YHL042W*, *YDR360W*, *YGL214W*, *PUB1*, *PMT2*, *TRX2*, *SFK1*, *MUP3*, *SPL2*, and *STM1*. Conversely, 8 genes were identified arresting normally yet failing to shmoo (termed 'SD' for 'Shmoo Defective'). Interestingly, 6 of these (*VPS8*, *VPS21*, *VPS22*, *VPS23*, *VPS28*, *VPS36*) are involved in vacuolar protein sorting, with all but *VPS8* and *VPS21* specific to class E sorting and resulting in

inefficient transport out of the endosome (Katzmann et al., 2002), suggesting a critical role of this system in shmoo formation (i.e., downstream of pheromone signaling), possibly related to plasma membrane reorganization (Bagnat and Simons, 2002; Dulic and Riezman, 1990). The remaining two proteins are involved in polarized growth (*ECM33*) and transcriptional regulation (the histone acetylase *EAF3*).

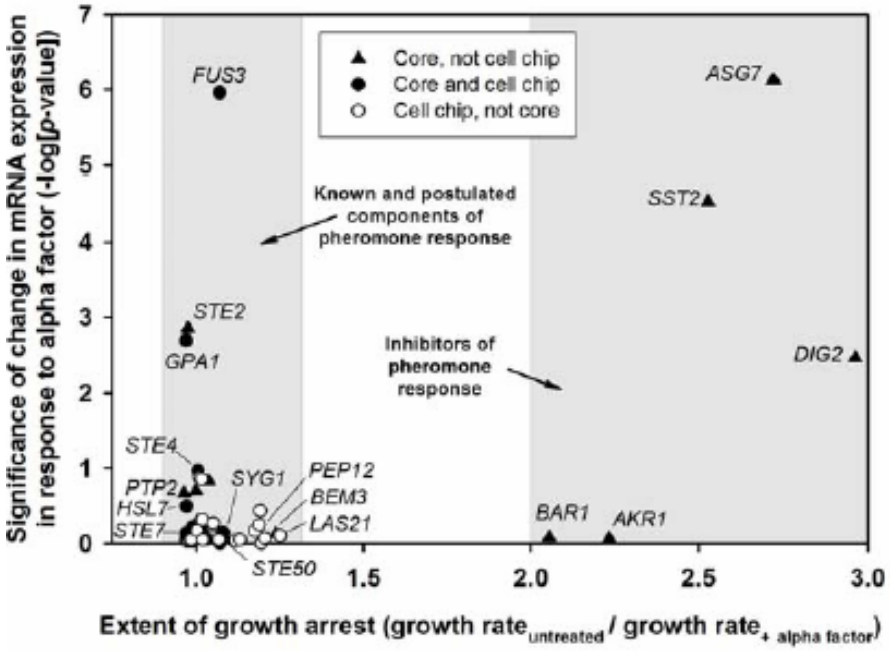


Figure 2.6 Neither known core pheromone response pathway genes (filled symbols) nor the cell-chip identified components (circles) show a systematic mRNA expression change during alpha factor stimulation. The changes of genes' mRNA levels in wild-type cells in response to addition of 50 nM alpha factor for 30 minutes (y-axis, expressed as the log [probability of a significant expression change]) are plotted opposite the extent of alpha-factor induced growth arrest of strains deleted for the corresponding genes (x-axis; data from this study). Participation in the pathway does not imply changes in mRNA expression of the pathway genes. Instead, core pathway genes fall into two categories

(gray regions)—primary pathway components and inhibitors of pathway components; only subsets of each change expression as a function of alpha factor. Other than lacking inhibitors, genes identified in the cell chip assay show the same trend.

2.4 DISCUSSION

We attempted to connect the 15 new pheromone-response implicated genes (the 'ASD' set) to the known pathway (the 'core set') using available functional genomics data by searching for the shortest pathways through protein interaction and mRNA co-expression networks (Lee et al., 2004) that connected the new genes to the core set. 9 of the new genes could be reasonably connected to the core set by two interactions or less (Figure 2.5), indicating that these genes may have direct, rather than indirect, roles in the pheromone response pathway. As connecting 9 of 15 genes to the core is no more than expected by random trials, these linkages serve only as hypotheses to provide a starting point for experiments validating the associations.

One gene connected in this manner is *SDS3*, a component of the Rpd3/Sin3 histone deacetylase complex implicated in gene silencing (Lechner et al., 2000), and it is likely that the implication of *SDS3* in the pheromone response pathway probably stems from the action of this complex on mating loci silencing. Likewise, another gene implicated in the screen, the ubiquitin protein ligase *UBR2*, is an interaction partner of *DOT1*, a participant in Sir-mediated gene silencing (San-Segundo and Roeder, 2000), and thus a reasonable inference is that deletion of *UBR2* may also influence silencing. Another gene from the screen, *ISY1*, is pleiotropic but connected to control of the cell cycle, participating in mRNA splicing and the spindle checkpoint (Dahan and Kupiec, 2002). *ISY1* exhibits some connection to polarized growth: homozygous diploid deletions of *ISY1* exhibit abnormal axial budding (Blagoev and Pandey, 2001). Although *MRPL28*

can be connected the core network in this manner, its shmoo defect might also arise by a disruption in the deletion strain of the proper functioning of the adjacent *MFA1* alpha factor mating pheromone gene.

Cell morphology phenotypes are rich in information, and although we have focused on strains exhibiting a failure to shmoo, additional strains were identified with morphological defects in the mating projections, such as shown for the *KEL1*Δ strain of Figure 2.2B. We flagged a total of 29 strains producing shmoos of aberrant morphology. These strains are deleted for genes involved in a statistically significant (p<0.01 (Robinson et al., 2002)) fashion in mating, especially for genes of polarized growth (*CDC10*, *KEL1*, and *BUD19*), but also genes of transcriptional and translational regulation, including components of transcription and chromatin remodeling (*SNF6*, *SPT3*, *SPT10*, *HTL1*, and *SIN4*), translational regulation (*CBP6*, *ASC1*, and *SRO9*), and rRNA processing/ribosome biogenesis (*NSR1*, *RPP1A*, *RPL31A*, *RPS16B*, and *RAI1*). There is also some interplay between cell morphology and pheromone response phenotypes—for example, the *MRPL28* Δ strain exhibits a large cell phenotype until alpha factor is added, whereupon the cell size defect is corrected, although the cells fail to shmoo.

Interestingly, we also find the extent of alpha factor-induced growth arrest appears largely uncorrelated with the change in expression of the corresponding genes following alpha factor treatment in wild-type cells (Roberts et al., 2000), even for known genes in the core pathway (Figure 2.6). Instead, the known pathway genes fall into two categories: those whose deletion strains show strong alpha factor-induced growth arrest or those that fail to arrest. The former category is exclusively composed of inhibitors of pheromone response components. The majority of known pathway genes do not change expression following alpha factor treatment (Roberts et al., 2000), nor do the majority of

new genes implicated in the pathway by the combined cell chip/growth inhibition assay. Therefore, the cell chip based screen complements the information available from microarrays.

2.5 CONCLUSIONS

In conclusion, we describe a new genomic-scale technology for microscopy on genetically distinct cells, applied here to measuring the cell morphologies of yeast in the haploid deletion strain collection and to the mapping of genes participating in the response of yeast cells to mating pheromone. Although this chapter focuses on cell morphology, cell chips have utility beyond this and can in principle be extended to any organism or cell-type for which defined libraries of cells can be arrayed, such as other easily manipulated organisms, banks of bacteria, and deletion libraries for other microorganisms. We expect diverse collections of strains can be arrayed, such as GFP-tagged yeast proteins. Just as it proved possible to identify pathways modulated by alpha factor, it should be possible to quickly identify mutants and pathways differentially affected by drugs. A major advantage of the cell chips is the minimal use of expensive reagents on the chips, afforded by limiting the use of antibodies and dyes to single microscope slides, as compared to the approx. fifty 96-well plates required to image the complete deletion collection. However, the key principle distinguishing cell chips from other approaches (such as immunoassays in 96-well plates) is the separation of cell growth from imaging. Thus, we anticipate the strongest advantage of cell chips will be their use for analyzing the localization of proteins or RNAs by high-throughput *in situ* hybridization and antibody-based immunoassays.

Authors' contributions

The author, Rammohan Narayanaswamy, and Alex Scouras contributed equally to this work that has been published in Genome Biology, 2006. The author and Rammohan

Narayanaswamy carried out the majority of experimental work and image grading. Alex Scouras designed and constructed the Cellma database and the algorithms for conversing GPR files to STG files for automatic microscopy. Traver Hart contributed to the mathematic and statistical analysis. Jonathan Davies participated in the construction of cell chips.

Chapter 3: Applications of spotted cell microarrays

3.1 INTRODUCTION

Although the previous study focuses on yeast morphology using *S. cerevisiae* deletion library, cell chips should have applications beyond morphology screens and could be applied to other defined collections of yeast, and to other organisms with arrayed strain collections. In this chapter, we applied cell chip technology to yeast GFP-tagged clone collection for studying yeast proteome dynamics in yeast polarization growth and to a *mariner* transposon insertion library of *E. coli* cells for identifying genes regulating the polar localization of LtrA, a reverse transcriptase encoded by *Lactococcus lactis* L1.LtrB group II intron. These two examples from yeast and bacteria indicate that cell chip could be a powerful tool for subcellular localization profiling either in yeast or in any other organisms with arrayed strain collections.

3.1.1 Cell polarity in Yeast

Cell polarity is a fundamental property of cells, which is an asymmetric distribution of regulatory molecules, cellular components, cell shape, and cell functions. Nearly every cell type, from simple single-cell organisms to multi-cell organisms, can polarize (Nelson, 2003; Pruyne and Bretscher, 2000). Cell polarity is essential to the partitioning of cell fate in embryonic development, the generation of axons during neuronal development, and etc. The polarization of yeast shares many features with those of more complex systems, including regulation by both intrinsic and extrinsic cues, conserved regulatory molecules such as Cdc42 GTPase, and asymmetry of the cytoskeleton as its center component (Chant, 1999). Because it is such a highly accessible

experimental system, *S. cerevisiae* serves as a fundamental model system for deciphering molecular mechanisms underlying the generation of cell polarity.

Yeast cell has two polarization processes: budding and mating. The overall cellular organization in these two processes is similar, except that the budding cells have a neck constriction between mother and bud cells (Chant, 1999). In brief, when cells sense internal or external cues for choosing a direction for polarization, Cdc42p and other activated Rho GTPases are localized to the cell surface, establishing the potential polarized site, coordinating and reorienting the actin cytoskeleton along the axis of polarization (Chant, 1999; Pruyne and Bretscher, 2000). Reoriented actin cytoskeleton, consisting of actin patch and actin cable, directs the secretion vesicles (exocyst complex) to the bud or mating projection for delivering biosynthetic enzymes and membranes thereby causing this region to grow selectively (Finger et al., 1998), as well as contributing to the partitioning of certain organelles such as mitochondrion to daughter cells (Simon et al., 1995). Astral microtubules that are stabilized at mother-bud neck or the tip of mating projection pull the nucleus to the neck or to the tip of projection in preparation for nuclear fusion (Read et al., 1992; Shaw et al., 1997).

Despite these similarities, there are major differences between budding and mating. During budding (haploid cells), the new bud emergence is immediately adjacent to the previous site of cell separation. This internal cue for bud initiation is established during previous budding events by the BUD gene products, allowing a Ras-related protein, Bud1p, to bind to Cdc24p and Bem1p at a discrete region of the plasma membrane during early G1 (Chant, 1999; Park et al., 1999; Pruyne and Bretscher, 2000; Zheng et al., 1995). The binding, in turn, triggers the recruitment of Cdc42p, which defines the nascent bud site and allows bud emergence to begin. In contrast, during mating, external gradients of pheromone guide the polarized growth (Segall, 1993). The

mating pheromone stimulation activates G-protein-coupled receptors that free Gβγ from Gα, which in turn recruits a polarity determinant Far1p to the plasma membrane (Butty et al., 1998). Together, Gβγ and Far1p recruit Bem1p, Cdc24p and Ste20p to assemble a Cdc42p-dependent signaling complex (Butty et al., 1998; Leeuw et al., 1998), which, in turn, activates the downstream mating signaling pathway (MAP kinase cascade) and the pathway for orienting cells towards a partner. In the orienting process, Cdc42p complexes are clustered into a patch, directed toward the pheromone source (Ayscough and Drubin, 1998). This patch then orients the actin cytoskeleton and directs shmoo growth towards a mating partner. Mutation of three classes genes affect polarization during mating: genes required for orienting axes chemotropically such as *STE4*, *FAR1*, and *CDC24* (Butty et al., 1998; Nern and Arkowitz, 1998; Valtz et al., 1995), genes originally required for budding polarization such as *CDC42* and associated factors, and genes in polarisome complexes (e.g. *SPA2*, *BNI1*, and *PEA2*) (Chenevert et al., 1992; Gehrung and Snyder, 1990; Valtz and Herskowitz, 1996). However, genes normally required for orienting axes of budding polarity (*BUD* and *AXL* gene) are largely dispensable for polarization in mating (Chant, 1999).

Tremendous progress has been made in deciphering the molecular mechanisms used by yeast cells. Some of the machinery are yeast specific, such as the bud site landmark proteins, whereas some machinery is highly conserved in function, such as *CDC42* controlling axis formation or G-protein coupled receptors guiding chemotropism or chemotaxis (Chant, 1999). The problem of deciphering the links between exogenous pheromones and ultimately cytoskeletal polarization in yeast is highly analogous to the problem of chemotaxis in higher eukaryotic systems (Chant, 1999). Therefore, our understanding of the polarization growth in response to the mating pheromone by yeast cells will accelerate our understanding of chemotaxis in other eukaryotic organisms. With

the completion of yeast genome and the development of new genomic technologies, more accurate and broader view of the molecules and pathways involved in this process will be provided. Here, we are particularly interested in reconstituting the spatial dynamics of yeast whole proteome in the polarized growth during mating, and revealing the molecular machinery that might be shared between budding and mating. In chapter 2, we demonstrated that cell chip is a powerful tool to perform a large-scale and high-throughput screen for identifying additional genes affecting yeast morphology when they were deleted. Therefore, By utilizing a similar approach and a Bayesian classifier that could potentially reduce the false negative rate in large-scale screens, we expected to provide a significantly improved picture on the spatial dynamics of yeast proteome when yeast cells initiate shmoo-associated polarized growth *via* systematically examining the localization of each of the 4200 GFP-tagged proteins upon alpha factor treatment.

3.1.2 LtrA, a reverse transcriptase encoded by Lactococcus lactis L1.LtrB group II intron

In collaboration with Junhua Zhao from Dr. Alan Lambowitz's lab, *E. coli* cell microarrays were constructed for screening genes affecting LtrA pole-localization in *E. coli.*

LtrA is a reverse transcriptase encoded by *Lactococcus lactis* L1.LtrB group II intron, which binds to intron RNA and promotes RNA splicing and intron mobility. Zhao showed that LtrA localizes to cellular poles of *E. coli* (Zhao and Lambowitz, 2005). The preferential insertion sites of randomized Ll.LtrB intron library locate in the origin (Ori) region in *E. coli*. Similarly, the wild-type Ll.LtrB intron retrotransposes into *E. coli* genome mainly in the Ori and Ter region, which are also pole-localized in most cell growth stages (Coros et al., 2005). Furthermore, either the polar localization of LtrA or the Ori domain preference of intron target site distribution is independent of functioning

of replication origin (Zhao, 2005; Zhao, unpublished data). These results indicated that the polar localization of LtrA could be responsible for the clustering of L1.LtrB insertion sites at Ori and Ter regions of *E. coli* chromosome. Moreover, LtrA protein interferes with polar localization of the *Shigella spp*. outer-membrane protein IcsA that regulates polarized actin tail assembly and has a similar polar localization pattern (Zhao and Lambowitz, 2005). Therefore, the authors suggested that LtrA and IcsA may have similar localization mechanisms, and LtrA's localization may be regulated by those that localize other proteins to cellular poles and potential cell division site, not the proteins associated with active replication origin. Thus, a genome-wide screen by examining LtrA localization changes caused by gene function disruption is necessary for seeking the potential proteins that could regulate LtrA pole-localization. However, it has been difficult to find the mutations that affect bipolar localization of proteins in *E. coli*. For example, only one mutant was identified by Nilsen *et al*., after manually screening ~7,000 *E. coli* mutants for altered localization of the IcsA protein (Nilsen et al., 2005).

Here, we attempted to use spotted cell microarray to automate the screening process and increase the size of the screen. After screening ~9600 mutants using spotted cell microarrays, we found that the disruption of five *E. coli* genes leads to both a more diffused intracellular distribution of LtrA and a more uniform genomic distribution of Ll.LtrB insertion sites, in agreement with the hypothesis of LtrA localization being a major determinant of insertion-site preference.

3.2 MATERIALS AND METHODS

3.2.1 *S. cerevisiae* GFP cell chips

3.2.1.1 Yeast strains and growth conditions

Cell chips were manufactured from the *S. cerevisiae* GFP-tagged clone collection (Invitrogen), in which each of ~4200 individual strains (MATa *his3Δ leu2Δ met15Δ ura3Δ*) was C-terminally tagged with GFP tag *via* homologous recombination (Huh et al., 2003). Frozen cell cultures in 96-well plates were thawed and inoculated to new 96-well plates with 200 µl YPD medium containing 17% glycerol. After growth at 30 ^{o}C for 2 days, copied plates were agitated with a plate shaker, sealed and frozen at -80 ^{o}C.

3.2.1.2 Alpha factor treatment

GFP-tagged strains were treated with alpha factor in the following manner: A copy of the GFP-tagged clone collection was thawed from -80 ^{o}C and 5 µl cells were used to seed a fresh copy in YPD media. After growth for ~ 36 hrs at 30 ^{o}C, 5 µl cells from this intermediate copy were inoculated to a new set of 96-well plates. This copy was grown to desired OD (~0.5) at 30 ^{o}C and was washed three times to inactivate secreted extracellular Bar1p protease that degrades alpha factor. Subsequently, alpha factor was added to each sample well at a final concentration of 75 µg/ml and grown for ~ 3hrs at 30 ^{o}C. The cells were fixed in YPD medium with freshly prepared 2% formaldehyde for 1 hr at 30 ^{o}C and then washed three times with YPD medium. Cells were then resuspended in l00 µl of YPD medium containing 17% glycerol and were ready for print or storage at -80 ^{o}C.

3.2.1.3 Manual follow-up of shmoo-tip localized proteins

All shmoo-tip localized proteins from the initial cell chip screen were manually picked up from the GFP-tagged library. Cells were grown in SD medium supplemented

with 4 amino acids (histinine, leucine, methionine, uracil), and treated with 75 μg/ml alpha factor when cells reached OD at 0.3~0.5. Images were manually captured through GFP and DIC filters *via* a Nikon E800 fluorescence microscope.

3.2.1.4 Construction of strains for co-localization experiments

GFP-RFP doubly tagged MATa strains for co-localization were constructed by mating and tetrad dissection. 9 RFP-tagged marker strains (BY4742: MATα *his3Δ leu2Δ lys2Δ ura3Δ*) represent for several organelles and cellular structures: actin (Sac6p), endosome (Snf7p), ER to Golgi vesicle (Sec13p), Golgi apparatus (Anp1p), late Golgi (Chc1p), spindle pole body (Spc42p) lipid particle (Erg6p), nucleolus (Sik1p), and peroxisome (Pex3p). They were individually mated with *FUS1*/GFP strains (MATa *his3Δ leu2Δ met15Δ ura3Δ*), and diploids were selected on SD, -Lys, -Met agar plates. After selection, diploids were grown for 6-8 days in sporulation medium, and tetrads were dissected on YPD plates. Doubly-tagged MATa type strains were confirmed by examining the presence of the shmoo morphology and the shmoo-tip localization of Fus1p upon alpha factor treatment.

3.2.1.5 Classifier construction

The genes identified in the high-throughput screen were used to train a naïve bayesian classifier (in Weka). Six mitochondrial genes were manually removed from the gene set to prevent training on them. Additional features were aggregated from data provided by the UCSF GFP screen (Huh et al., 2003) and the functional network by Lee et al. (Lee et al., 2007). The features collected for each gene were: the sum of log likelihood scores (LLS) to the set of shmoo genes, the ratio of the LLS sum linking genes to the shmoo set divided by the LLS sum of that genes linking to all genes in the Lee network (Lee et al., 2007), protein abundance (Ghaemmaghami et al., 2003), and cell

location (Huh et al., 2003). The test set of 5,804 genes, labeled as shmoo or not-shmoo was also used as the training set. Ten-fold cross validation had very similar results. After training, the classifier recovered 20 of the 37 shmoo genes (cross-validation: 19). An additional 151 (cross-validation: 153) genes not identified in the initial screen were also classified as shmoo genes using a 0.5 probability cutoff. Of the 151 genes, 118 were present in the GFP library.

3.2.2 *E. coli* cell chips

3.2.2.1 Bacterial strains and growth conditions

Mariner transposon gene disruption strains was grown in LB medium supplemented with kanamycin, electroporated with the intron-expression plasmid pACD2X-GFP/LtrA, which carries a cap^R marker, and plated on LB containing chloramphenicol and kanamycin. ~9,600 colonies were picked, grown in 96-well plates, and then stored at -80°C.

3.2.2.2 E. coli cell chip construction and imaging

E. coli mutants carrying pACD2X-GFP/LtrA were inoculated in 96-well plates containing LB medium with chloramphenicol and incubated overnight at 37°C. The cell cultures were then inoculated 1:10 into fresh LB medium with 17% glycerol in new 96-well plates and grown for 5 h at 37°C. In one experiment, cells in fifty-one 96-well plates were induced with 100 μM IPTG at 37°C. In the second experiment, cells in forty-nine 96-well plates were induced with 500 μM IPTG at 30°C. Culture transfer and media additions were performed by a Biomek FX laboratory automation workstation (Beckman Coulter, Fullerton, CA). Cell microarrays (cell chips) were constructed as described before (Narayanaswamy et al., 2006). In brief, ~5,000 knockouts plus a wild-type HMS174 (DE3) control were printed onto poly-L-lysine coated microscope slides using a

custom-built DNA microarray printing robot. In each experiment, ~30 cell chips were manufactured, two of them were used for imaging, and the remaining cell chips were stored at -80°C. Before imaging, cell chips were briefly washed with ddH_2O to remove printing buffer and debris, and then mounted with VECTASHIELD hard-set mounting medium containing 1.5μg/ml DAPI (4'-6-diamidino-2-phenylindole, Vector Laboratories, Inc). Cell images were collected as described in Chapter 2. Images were stored in cell microarray image database (Cellma) and manually examined to identify strains with altered GFP/LtrA localization patterns.

3.3 RESULTS

3.3.1 A systematic survey of proteome spatial dynamics in yeast polarized growth

3.3.1.1 Identification of shmoo tip localized proteins using yeast GFP chip

Having established the typical morphology screen using cell chips, we examined protein localization changes in response to the mating pheromone, which is one of the primary polarization pathways in budding yeast. Although this pathway is well-studied, it has yet to be completed for examining dynamics of yeast proteome during mating. Additionally, we reasoned that additional genes regulating yeast polarization, either directly or indirectly could be identified by examining their localization changes when the GFP-tagged clone collection was treated with alpha factor. Two sets of cell chips were manufactured from ~4,200 GFP-tagged strains treated with alpha factor and the replicated GFP-tagged strains without alpha factor treatment. We manually examined the cell images for the presence of protein localization to shmoo tip when cells were treated with alpha factor (Figure 3.1). 188 proteins localized to the shmoo-tip when cells were treated with alpha factor. We validated this set for reproducible shmoo-tip localization by manually retesting the 188 strains *via* alpha factor addition and microscopic imaging. 47

of the 188 strains showed reproducible shmoo-tip localization. Of these strains, 13 were previously identified as shmoo-tip localized proteins. Of the 47 genes annotated as shmoo-tip localization in SGD, 40 genes were available in the GFP-tagged strain collection. Therefore, we only recovered 33% of genes known to be shmoo localized. However, many proteins, especially those involved in signaling pathway are low abundant, and are difficult to be detected even through GFP microscopy. If proteins with expression level more than 2500 molecules/cell are counted, we captured 63% of the known ones.

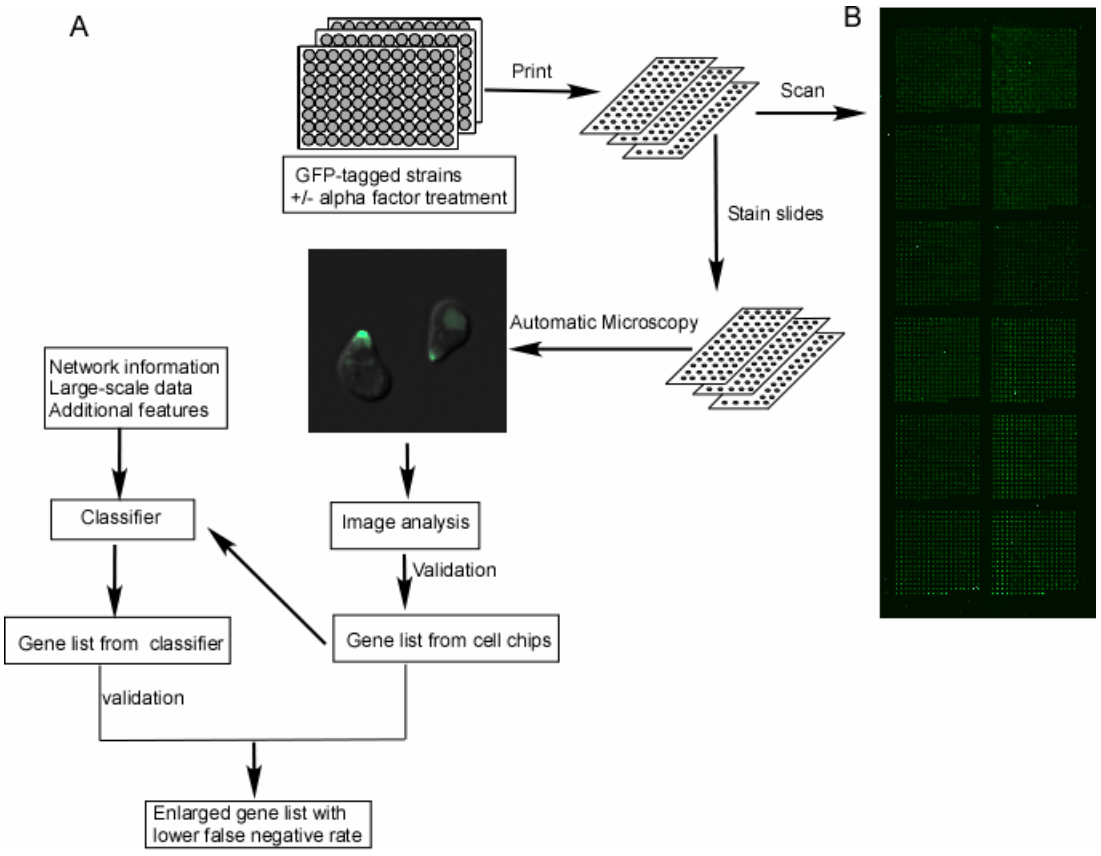


Figure 3.1 The experimental schema of yeast GFP chips**.** (A) The diagram of cell chip experiment, including cell chip construction, classifier, follow-up, an image analysis. Yeast GFP chips were constructed from ~4,200 GFP-tagged strains under two conditions:

alpha factor treated and control sets. We examined the localizations of GFP-tagged proteins and identified 43 proteins localized to shmoo tip, and then we trained Bayesian classifier with the various related features. 37 additional shmoo-tip localized proteins were identified through classifier methods. (B) The wide-field light scattering image of a cell microarray containing ~4,200 yeast GFP strains.

3.3.1.2 Identification of additional shmoo-tip localized proteins using classifier

To counter the high false-negative rate that is often seen in high-throughput screens, we used a naïve Bayesian classifier to identify additional shmoo-tip localized proteins. The classifier was trained with the localization information from UCSF (Huh et al., 2003) and the functional network by Lee *et al.* (Lee et al., 2004). Removing 20 genes identified in cell chip screen, 151 additional genes were predicted from classifier and were not identified in the initial screen. Of the 151 genes, 118 were present in the GFP library. 37 genes were validated to be shmoo-tip localized *via* manual alpha factor treatment and microscope examination. In the end, we identified 80 shmoo-tip localized proteins, which almost doubled the number of genes from the original cell chip screen (Table 3.1).

3.3.1.3 Functional enrichment of 80 shmoo tip localized proteins

We then examined the functional enrichments in these shmoo-tip localized proteins using a web-based yeast cluster database, FunSpec (p-value < 0.01) (Robinson et al., 2002). 80 shmoo-tip localized genes showed dramatic enrichment for biological processes involved in budding and mating such as actin cytoskeleton organization and biogenesis (GO:0030036, p-value<$1x10^{-14}$), establishment and/or maintenance of cell polarity (GO:0030012, p-value<$1x10^{-14}$), polar budding (GO:0007121, p-value<$1x10^{-14}$),

and related pathways. These 80 genes are also highly enriched in cellular components involved in polarized growth such as exocyst (GO:0000145, p-value<$1x10^{-14}$), cytoskeleton (GO:0005856, p-value<$1x10^{-14}$), bud tip (GO:0005934, p-value<$1x10^{-14}$), bud (GO:0005933, p-value<$1x10^{-14}$), shmoo tip (GO:0005937 p-value<$2x10^{-6}$), polarisome (GO:0000133, p-value<$2x10^{-5}$) and related systems.

We identified all 8 exocyst proteins known for mediating polarized targeting of secretory vesicles to the active sites of exocytosis (*EXO84*, *SEC5*, *SEC3*, *SEC15*, *SEC6*, *EXO70*, *SEC10*, and *SEC8*). Of the 27 proteins localized to bud tip, 11 were captured in this screen as shmoo-tip localized proteins (*EDE1*, *BEM1*, *SYP1*, *PEA2*, *SMI1*, *KEL2*, *KEL1*, *SMY1*, *BUD6*, *BNI1*, and *SEC8*). Of 71 proteins annotated as bud-localized proteins, 17 were also identified as shmoo-localized proteins. Most interestingly, 6 proteins (*EDE1*, *BOI1*, *CHS3*, *RGD1*, *CDC10*, and *SMI1*) that are only annotated as bud, bud neck, or bud-tip localized proteins in SGD were identified to be shmoo-tip localized, which indicates that these proteins required in budding process might be reused in shmoo formation. Additionally, 9 uncharacterized genes (*YCR043C*, *YDR061W*, *YDR348C*, *YER071C*, *YIR003W*, *YMR295C*, *YNL100W*, *YOR304C-A*, *YPR171W*) are associated with shmoo formation (Figure 3.2A).

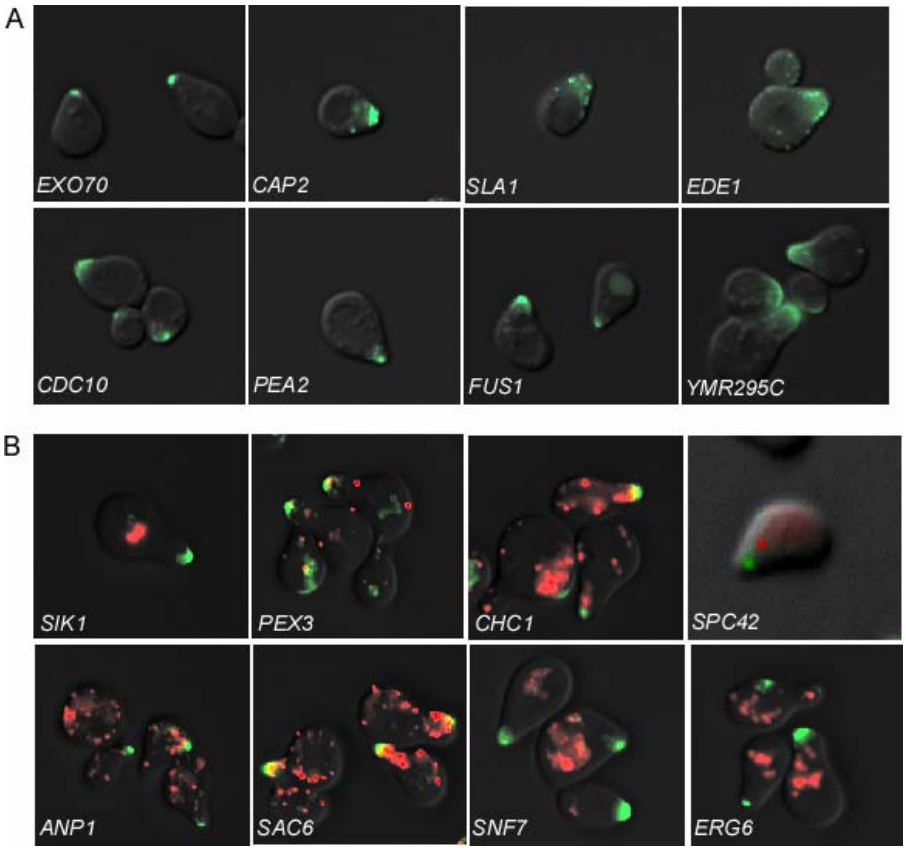


Figure 3.2 Example images of proteins localized to shmoo tip**.** (A) Each image represents a cellular component: *EXO70* for exocyst; *CAP2* for actin cortical patch; *SLA1* for actin cytoskeleton; *EDE1* for bud tip; *CDC10* for bud and bud neck, *PEA2* for polarisome; *FUS1* for shmoo tip. *YMR295C* is uncharacterized gene. (B) Example images of GFP-RFP doubly tagged strains. RFP: actin (Sac6p), endosome (Snf7p), ER to Golgi vesicle (Sec13p), Golgi apparatus (Anp1p), late Golgi (Chc1p), spindle pole body (Spc42p), lipid particle (Erg6p), nucleolus (Sik1p), and peroxisome (Pex3p); GFP: Fus1p.

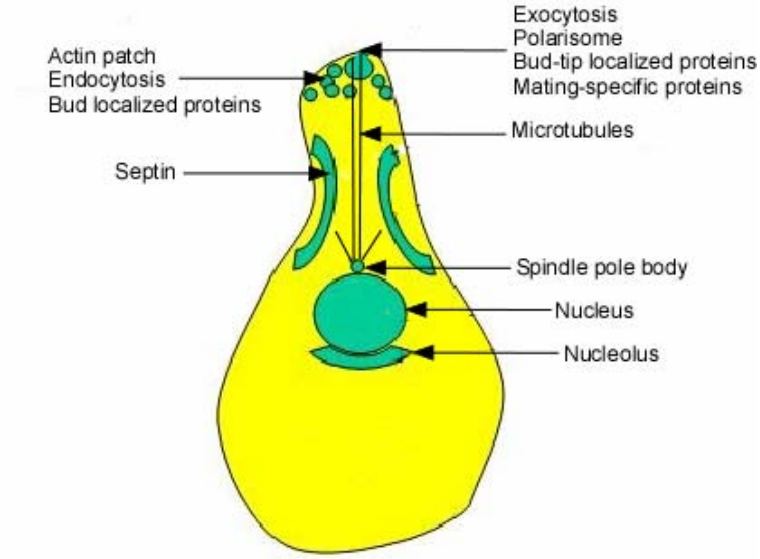


Figure 3.3 The shmooing cell with proteins and organelles labeled by position. Secreted vesicles are accumulated at the site where the projection formed, followed by actin patches and endocytosis sites. Other intracellular organelles such as nucleus, spindle pole body, rough endoplasmic reticulum and mitochondria were also rearranged, showing a polar organization of the cytoplasm during projection formation.

3.3.1.4 The reorganization of organelles in polarized growth

To demonstrate the movement patterns of yeast organelles in response to alpha factor, 9 MATa strains in which each stain was doubly tagged with GFP and RFP: GFP for shmoo-tip marker (*FUS1*) and RFP for an organelle marker protein (actin, endosome, ER to Golgi vesicle, Golgi apparatus, late Golgi, lipid particle, nucleolus, peroxisome, or spindle pole body) (Figure 3.2B). Quantitative image analysis was performed to measure the distance between organelle and shmoo tip, showing that intracellular organelles were

in ordered organization during shmoo formation, which approximately agreed with the previous finding (Baba et al., 1989). Actin patches, exocyst, and endocytosis machinery were accumulated in the shmoo tip; mitochondria was rearranged in the direction of shmoo tip; nucleus was also moved towards shmoo tip with spindle pole body attaching to the side towards to projection tip and with the nucleolus locating at the opposite side of shmoo tip (Figure 3.3).

3.3.1.5 Conclusions

There are very limited studies showing that how many proteins are involved in cellular organization when shmoo is formed and which proteins might be involved in two polarization pathways, mating and budding. Our systematic survey of the yeast localizome provides the first intact and detailed picture of the proteome spatial dynamics in polarization growth during mating, and reveals the possibly adaptive reuse of the polarized growth machinery between budding and mating. Beyond the biological findings, this study may provide a useful experimental pipeline to reduce the false negative rate in large-scale screens.

Authors' contributions

The author, Rammohan Narayanaswamy, and Traver Hart constructed Cell chip. Initial cell chips imaging screen was carried by Dr. Edward Marcotte and Rammohan Narayanaswamy; The author and Traver Hart manually tested 188 strains from the initial cell chip screen. 9 MATa strains in which each stain contains both Fus1-GFP and an RFP-organelle protein were constructed by the author and Rammohan Narayanaswamy, followed by common yeast mating and dissection protocols. The time-lapse videos were collected by Traver Hart and Matt Davis. Kris McGary and Matt Davis built classifier. Rammohan Narayanaswamy and Kris McGary manually screened 118 shmoo genes identified by classifier. Emily Moradi performed quantitative image analysis.

3.3.2 Identification of genes regulating polar localization of LtrA in *E. coli* using cell chips

3.3.2.1 Cell chip screen for E. coli mutants with altered LtrA localization patterns

To screen for *E. coli* mutants with altered LtrA localization patterns, we constructed a *mariner* transposon insertion library in *E. coli* HMS174 (DE3) transformed with plasmid pACD2X-GFP/LtrA, which employs a T7 lac promoter to express the Ll.LtrB-ΔORF intron with short flanking exons, followed by a GFP/LtrA fusion protein. After induction with IPTG, cells from above *E. coli* mutant collection were robotically deposited from 96-well plates onto poly-L-lysine coated slides and examined for altered GFP/LtrA localization patterns (Figure 3.4). ~9,600 mutants were screened, using two sets of Ll.LtrB intron-induction conditions (4,700 with 500 μM IPTG at 30^{o}C and 4,900 induced with 100 μM IPTG at 37 ^{o}C), potentially enabling detection of a wide range of mutants. Of 277 initial candidates, 36 mutants showed consistently altered GFP/LtrA localization patterns in duplicate arrays, of which 5 mutants showed the most reproducibly altered GFP/LtrA localization patterns after manual validation in liquid culture (Figure 3.5).

The *mariner* transposon insertion sites in the five disruptants were amplified and sequenced *via* thermal-asymmetric-interlaced (TAIL) PCR and found to be in the *gppA, uhpT, wcaK, ynbC, zntR* genes. *gppA* encodes guanosine pentaphosphatase A, required for the synthesis of the stringent response regulator, ppGpp ("magic spot") (Keasling et al., 1993); *uhpT* encodes a component of the hexose phosphate transport system (Hall and Maloney, 2001); *wcaK* is predicted to encode a colanic acid kinase, pyruvyl-transferases (Stevenson et al., 1996); *ynbC* encodes a 585-aa ORF of unknown function (Blattner et al., 1997); and *zntR* encodes the zinc-responsive transcriptional regulator (Newberry and Brennan, 2004). All of the disrupted genes are in single-gene transcription units, except

uhpT, which is the last gene of the *uhp* operon. Thus, the observed alteration in GFP/LtrA localization patterns is likely due to disruptions of the gene where the *mariner* transposon is inserted.

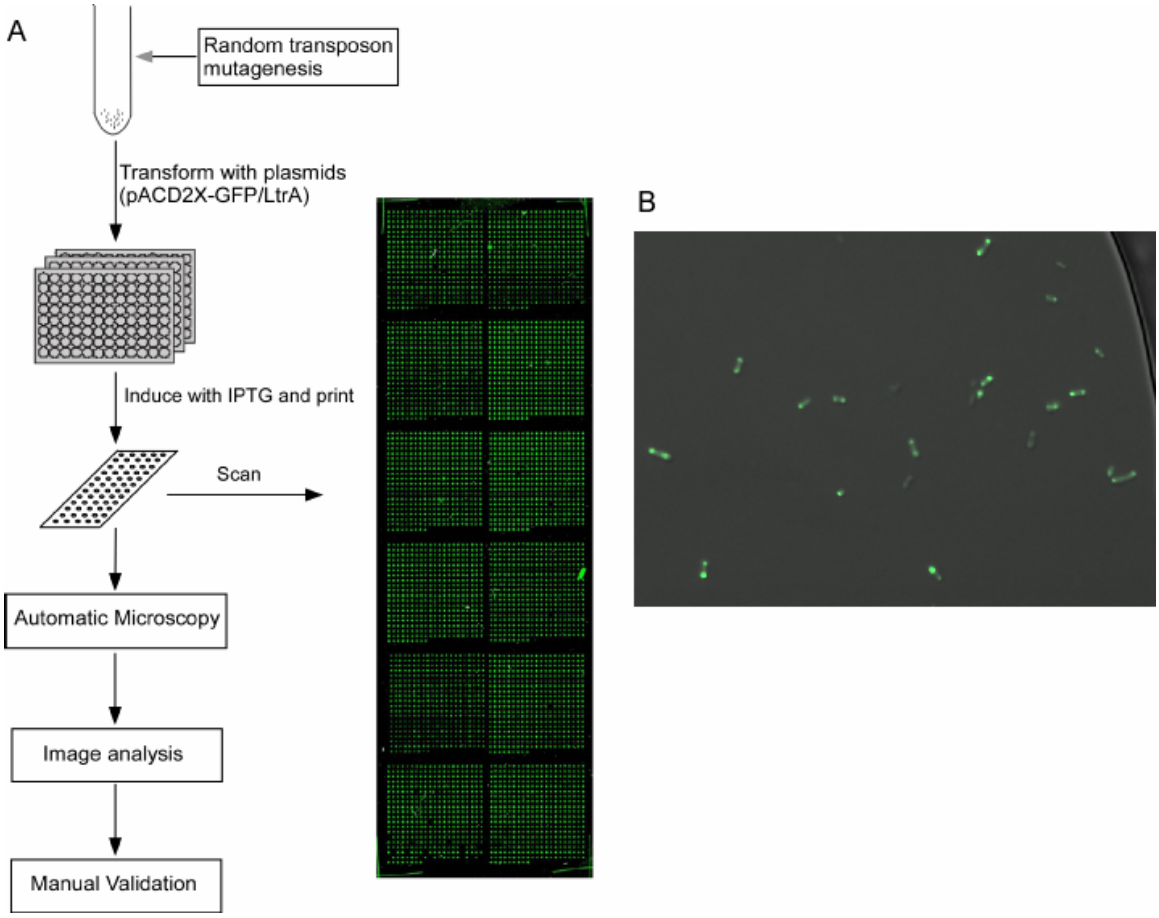


Figure 3.4 The construction of *E. coli* cell chips. (A) shows the cell chip procedure: A mariner transposon insertion library of *E. coli* HMS174 (DE3) cells carrying pACDX-GFP/LtrA was arrayed onto microscope slides and screened by automated fluorescence microscopy to identify mutants with altered GFP/LtrA localization patterns. (B) shows the wild-type pattern of bipolar localization of GFP/LtrA.

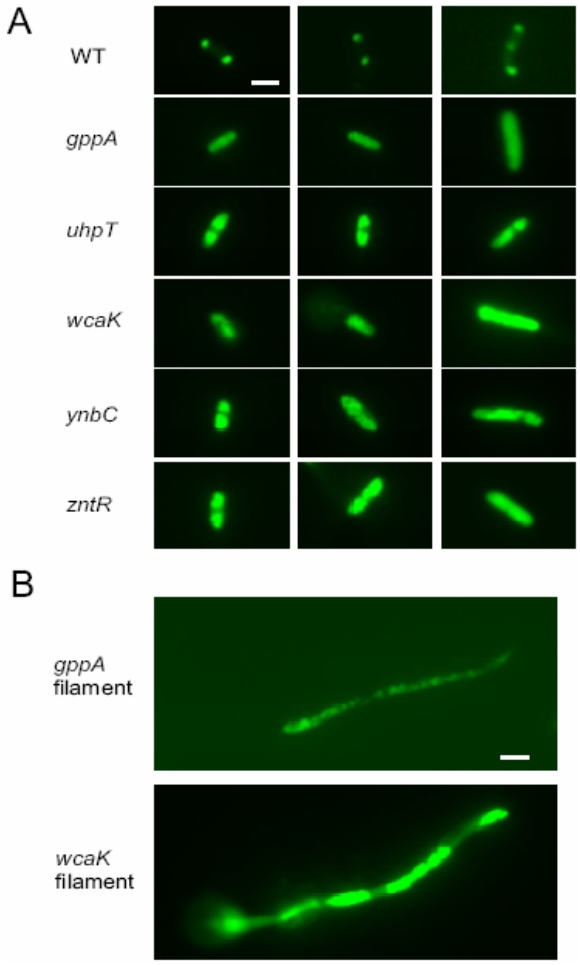


Figure 3.5 Example images of wild-type strain and mutants with altered GFP/LtrA localization patterns identified by cell array screen**.** Wild-type HMS174 (DE3) (WT) and mutants containing pACD2X-GFP/LtrA were induced with 500 μM IPTG overnight at 30^{o}C. (A) Wild-type cells showing bipolar GFP/LtrA localization and mutants showing altered GFP/LtrA localization patterns. (B) Examples of filamentous *gppA* and *wcaK* cells showing GFP/LtrA localization as multiple foci, diffuse patches, or a combination of the two. Bar = 2 μm.

3.3.2.2 Functional analysis of five genes identified from cell chip screens

Further studies showed that LtrA localization is a major determinant of insertion-site preference. In the wild-type strain, the proportion of Ll.LtrB insertion sites found in the Ori region was 79%. In contrast, only 35-51% of the Ll.LtrB insertions sites in the disruptants were located in this Ori region, with the lowest proportions being found in the *zntR* and *ynbC* disruptants (35% and 38%, respectively). Therefore, a more diffuse intracellular distribution of LtrA correlates with a more uniform distribution of Ll.LtrB insertion sites throughout the *E. coli* genome. Surprisingly, we found that the diffused intracellular localization of LtrA is due to the accumulation of intracellular polyphosphate (poly(P)) and that these effects on protein localization are not limited to LtrA but also for other pole-localized proteins (data not shown). Our results suggest that poly(P) accumulation may be part of a mechanism that globally relocalizes proteins in response to cell stress or entry into stationary phase, with potentially wide physiological consequences.

3.3.2.3 Conclusion

We used spotted cell microarray technology to screen ~9,600 transposon gene insertion clones in *E. coli* for identifying genes whose mutation affects the polar localization of LtrA. Our result reveals genes regulating localization pattern of LtrA. Most importantly, it demonstrates the potential of using cell microarrays for high-throughput screens in identifying mutations affecting cell morphology or protein localization in organisms beyond the budding yeast.

Authors' contributions

The author and Junhua Zhao constructed cell chips. Junhua Zhao constructed *E. coli* transposon insertion library and follow up experiments.

Table 3.1 lists of shmoo genes from cell chip screen and classifier.

Gene list from cell chip screen

43 genes were identified from yeast GFP cell chip screen.

Gene	Common name	Gene	Common name
YCR088W	ABP1	YLR332W	MID2
YMR092C	AIP1	YER149C	PEA2
YPL115C	BEM3	YNR052C	POP2
YKL007W	CAP1	YGL107C	RMD9
YIL034C	CAP2	YLR166C	SEC10
YPL111W	CAR1	YBR080C	SEC18
YNL161W	CBK1	YNL272C	SEC2
YCR002C	CDC10	YER008C	SEC3
YJR076C	CDC11	YDR166C	SEC5
YDL126C	CDC48	YIL068C	SEC6
YGR255C	COQ6	YPR055W	SEC8
YER141W	COX15	YLR058C	SHM2
YIL111W	COX5B	YDL212W	SHR3
YBL047C	EDE1	YBL007C	SLA1
YNL084C	END3	YOR008C	SLG1
YDL161W	ENT1	YKL079W	SMY1
YJL085W	EXO70	YCR043C	YCR043C
YBR102C	EXO84	YDR061W	YDR061W
YCL027W	FUS1	YMR295C	YMR295C
YNL106C	INP52	YNL100W	YNL100W
YHR158C	KEL1	YOR304C-A	YOR304C-A
YGL099W	LSG1		

Gene list from classifier

37 additional genes were identified from classifier

Gene	Common name	Gene	Common name
YOR239W	ABP140	YCR009C	RVS161
YNL020C	ARK1	YDR388W	RVS167
YJL095W	BCK1	YDR129C	SAC6
YBR200W	BEM1	YGL233W	SEC15
YNL271C	BNI1	YDL195W	SEC31
YBL085W	BOI1	YHR098C	SFB3
YPR171W	BSP1	YDL225W	SHS1
YLR319C	BUD6	YNL243W	SLA2
YHR114W	BZZ1	YGR229C	SMI1
YBR023C	CHS3	YNL138W	SRV2

YLR330W	CHS5	YCR030C	SYP1
YLR206W	ENT2	YGR080W	TWF1
YGR238C	KEL2	YLR337C	VRP1
YOR181W	LAS17	YNL283C	WSC2
YOR326W	MYO2	YOL105C	WSC3
YMR109W	MYO5	YDR348C	YDR348C
YIR006C	PAN1	YER071C	YER071C
YIL095W	PRK1	YIR003W	YIR003W
YBR260C	RGD1		

Table 3.2 Five mutants screened from the cell chip assay.

Data of fluorescence pattern, mobility, and target site distribution were from experiments carried out at 30 °C.

* At least 100 cells were counted in each strain.

**10% genome is 5% region flanking either side of replication origin (00). ≥50 target sites were studied in each strain.

*** In *zntR* knockout, several target sites were repeatedly detected, so that the percentage of target sites within the 10% genome is less than 10%.

(Provided by Junhua Zhao)

K/Os	Genes	Protein functions	% of diffused & randomly distributed fluorescence*	Intron mobility	Target site distribution in 10% genome**
wild type			2.8%	100.0%	60%
25D12	*wcaK*	putative galactokinase	33%	18.1%	30%
43F3	*gppA*	guanosine pentaphosphate phosphohydrolase	34%	173.0%	23%
49A3	*uhpT*	hexose phosphate transport protein	23%	85.1%	31%
D2	*ynbC*	predicted hydrolase, unknown function	20%	112.8%	22%
D3	*zntR*	zinc-responsive transcriptional regulator	20%	121.9%	6%***

Chapter 4: Development of Yeast Spheroplast Microarrays

4.1 INTRODUCTION

DNA Microarrays and mass spectrometry have been powerful tools for direct assay of gene and protein function in a genome-wide scale. However, these approaches can not monitor the gene and protein behavior in an intact cellular compartment. We wished to measure single cell phenotypes by high-throughput microscopy, taking advantage of the availability of libraries of genetically distinct cells. Here we described a new technology based on the spotted cell microarray platform (Narayanaswamy et al., 2006), yeast spheroplast microarrays. This technology integrates technologies in two fields, cell biology and genomics, and will have applications beyond morphology screening and can be extended to the analysis of localization phenotypes of thousands of genetically distinct cells in parallel.

As a proof of concept, we constructed yeast spheroplast chips containing ~700 yeast tetracycline-repressible promoter strain collection, in which the expression of yeast essential gene is repressed by addition of doxycycline into culture medium, for demonstrating the potential of spheroplast chips in large-scale protein localization profiling *via* recapturing the known genes interfering with spindle formation. Moreover, we could identify new essential genes affecting microtubules and spindle formation, considering the precise molecular and genetic function of many essential yeast proteins have not been systematically studied due to the difficulty to make mutant strains for essential genes in a systematic manner (Mnaimneh et al., 2004; Yu et al., 2006).

In *S. cerevisiae*, the microtubule system consists of microtubules, microtubule-associated proteins (MAPs) and the microtubule organizing centre (MTOC), known as the spindle pole body (SPB). Microtubules are composed of tubulin, a heterodimer of α-

and β-tubulin. α-tubulin is encoded by two highly related isogenes *TUB1* and *TUB3*; *TUB2* encodes the essential β-tubulin. The SPB of budding yeast is a multi-layered structure that comprises outer, central, and inner plaque, the half bridge, the bridge, and the satellite and is embedded in the nuclear envelope throughout the cell cycle. The SPB organizes two functionally and spatially distinct microtubule arrays: cytoplasmic and nuclear microtubules. The cytoplasmic microtubules are directed toward the cytoplasm and needed for nuclear positioning and the migration of the nucleus from mother cell into bud cell. The nuclear microtubules are directed toward nucleoplasm and essential for SPB separation, mitotic and meiotic spindle formation, and chromosome segregation (Jacobs et al., 1988; Winsor and Schiebel, 1997). The function and behavior of the yeast microtubule system is in cell cycle dependent way (Kilmartin and Adams, 1984; Winsor and Schiebel, 1997).

In the G1 phase, each yeast cell contains one SPB that is associated with cytoplasmic and nuclear microtubules. Cytoplasmic arrays originate from half bridge of the SPB, push against cell cortex, and propel the nucleus in the opposite direction (Kilmartin and Adams, 1984; Knop et al., 1999; Shaw et al., 1998). SPBs are duplicated by the end of the G1 phase. The formation of a short spindle is completed at the end of S phase, and SPBs are separated at the earlier G2 phase. Meanwhile, cytoplasmic astral microtubules are reoriented towards bud tip and penetrate the bud in order to assist nucleus migrating to bud neck (Winsor and Schiebel, 1997). When cells enter mitosis, microtubules are rearranged into bipolar spindles and positioned through bud neck (Hanna et al., 1995). The bipolar spindle is one of key components to ensure fidelity of sister chromosome segregation. In anaphase, sister chromosomes move towards opposite poles of spindle by the shortening of the cytoplasmic microtubules and the elongation of spindles (Winey et al., 1995), and sister chromosomes are segregated by the end of

anaphase. Upon cytokinesis is completed and spindle disassembles, mother and bud cell enter a new cell cycle.

Since yeast microtubule and spindle have distinct organization patterns during cell cycle and many proteins involved in spindle formation and nuclear segregation are essential for cell growth (Kilmartin and Adams, 1984; Winsor and Schiebel, 1997), we constructed yeast spheroplast chips covering 700 *TetO*$_7$ promoter strains and probed cell chips with anti-alpha tubulin antibody to visualize microtubule structure, expecting to demonstrate that yeast spheroplast chip could be a useful tool for performing large-scale protein localization analysis.

4.2 MATERIALS AND METHODS

4.2.1 Yeast strains and growth conditions

The *TetO*$_7$ promoter collection used in this study contains ~700 essential yeast genes for which the endogenous promoter has been replaced with a Tet-titratable promoter in the genome (Mnaimneh et al., 2004). Thus, the expression of the gene can be switched off by the addition of doxycycline to the yeast's growth medium. Cells in 96-well plates were grown in YPD medium supplemented with doxycycline. After 10-hour growth, cells were immediately fixed in growth medium with 1/10 volume 37% formaldehyde for 1 hour at 30^{o}C.

4.2.2 Zymolyase treatment and Preparation of cell chips

The fixed cell cultures were pelleted and washed twice with spheroplast buffer (1.2 M sorbitol, 0.1 M KH_2PO_4, pH 7.5). Cells were then resuspended in 200µl spheroplast buffer with 0.025mg/ml zymolyase 20T (Seikagaku corporation) and 0.4µl β-mercaptoethanol and incubated for 2 hour at 30 ^{o}C. After spheroplasted, cells should be handled very carefully and centrifuged at low speeds (<3000 rpm). Cells were then

washed twice with spheroplast buffer and resuspended in 200µl spheroplast buffer. The spheroplasted cells were spotted onto poly-L-lysine coated slides by using the same printing format as described in chapter 2.

4.2.3 Immunofluorescence microscopy

The immunofluorescence (IF) staining protocol is revised from Burke et al (Burke et al., 2000). Cells on the chip were permeabilized in cold methanol for 5 minutes, followed by 30 seconds in cold acetone (both solutions were pre-chilled at -20 ^{o}C). Slides were then air-dried completely. Before incubating with antibodies, cell chips were incubated in blocking buffer (3%BSA in 1XPBS) for 30 minutes at 30 ^{o}C in a humidity chamber. Cell chips were stained with 60µl 4µg/ml mouse anti alpha-tubulin monoclonal antibodies (Molecular probes) for 1 hour, followed by 3 rinses with 1XPBS, and then incubated with 60µl 4µg/ml Texas Red conjugated goat anti–mouse secondary antibody (Molecular probes) for 1-2 hours. After washed three times with 1XPBS, slides were mounted with 60µl VECTASHIELD hard set mounting medium with 1.5µg/ml DAPI (Vector Laboratories, Inc) and covered with coverslips. Images were acquired as described in Narayanaswamy *et al* (2006).

4.3 RESULTS

4.3.1 Construction of yeast spheroplast microarrays

Cells from each of the ~700 distinct yeast tet-promoter strains were grown in YPD medium with doxycycline and fixed in culture medium with formaldehyde at final concentration of 3.7%. After zymolyase treatment, yeast spheroplasts from arrayed 96-well plates were printed onto poly L-lysine coated glass microscope slide using a high-speed robotic printing machine. As seen in Figure 4.1B using microarray scanner, a light scatter image of spheroplast chip is still visible after cells were spheroplasted. Each spot is from a single tet-promoter strain. Spheroplast chips can be immunostained immediately

or stored at -80 ^{o}C, and then cell chips can be analyzed using an automated fluorescence microscope to sequentially auto-focus and image each spot. Figure 4.2 shows yeast cells immunostained for alpha-tubulin (a) and acetyl-histone H4 (b).

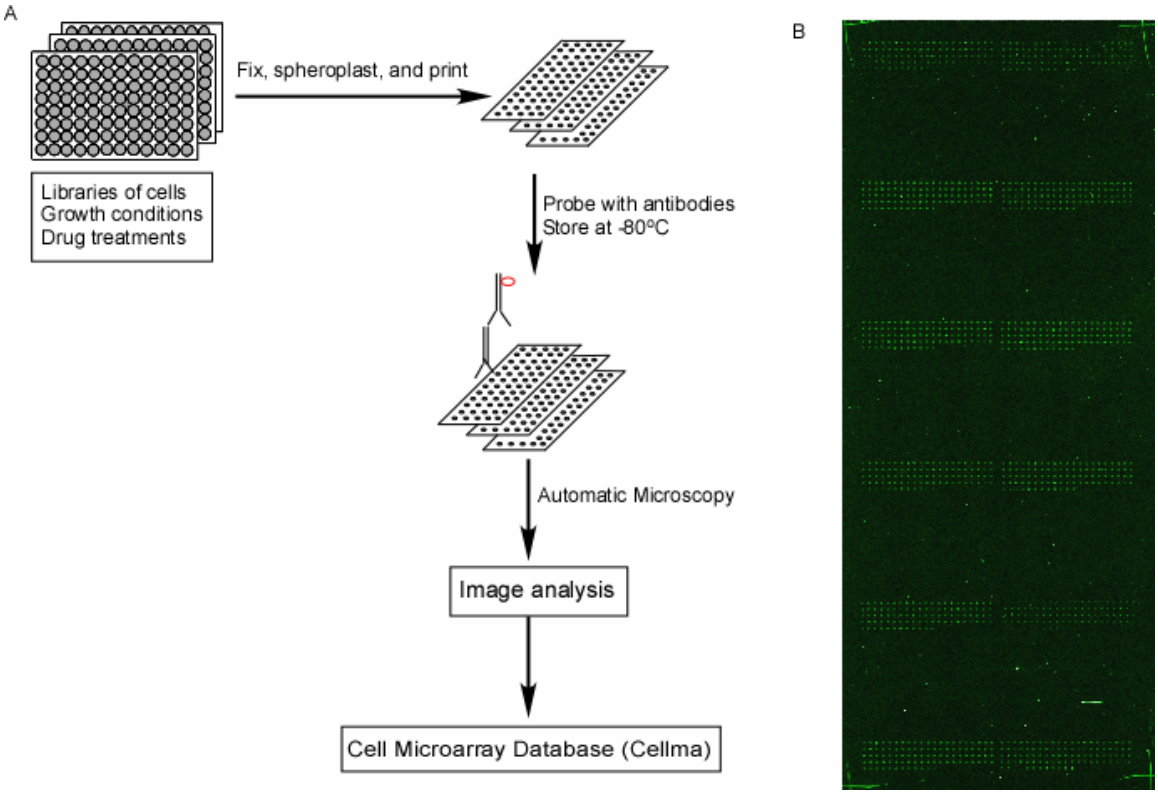


Figure 4.1 An overview of yeast spheroplast microarrays. (A) Cells are cultured, fixed, and spheroplasted in 96-well plates. Subsequently, they are robotically printed onto poly-L-lysine microscope slides using slotted steel pins, containing up to 5,000 spots. Arrays are scanned with a Genepix DNA microarray scanner to obtain spot coordinates, and then each array is probed with 70 µl of a specific antibody, followed by automatic high-throughput imaging. (B) The wide-field light scattering image of yeast spheroplast chips with ~700 yeast *TetO7* promoter strains.

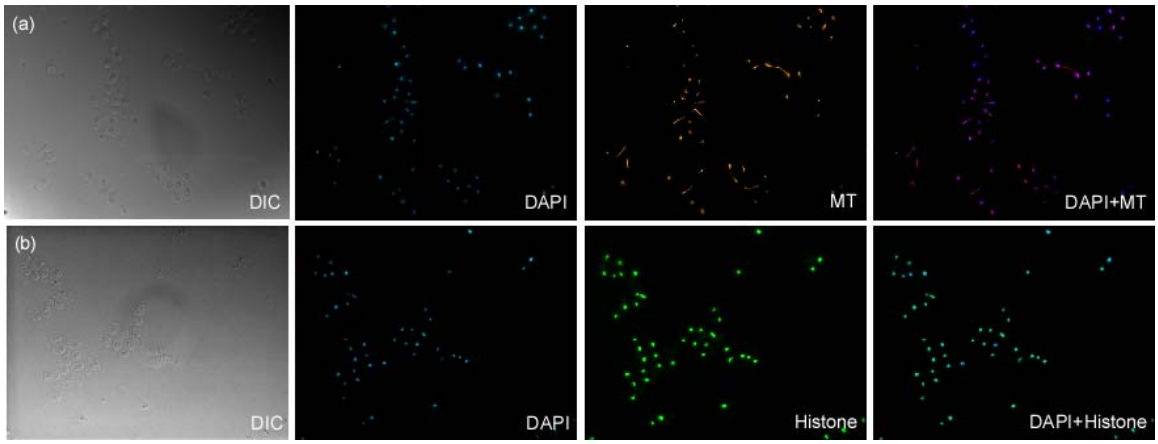


Figure 4.2 Example images from yeast spheroplast chips immunostained with different antibodies. (a) Arrayed cells were immunostained with anti-alpha tubulin primary antibody and Texas-red conjugated secondary antibody. (b) Arrayed cells were probed with anti-acetyl-histone H4 primary antibody and FITC conjugated secondary antibody.

4.3.2 The optimization of the spheroplasting procedure

One of the most crucial steps in making yeast spheroplast chips is the spheroplasting treatment with zymolyase. Yeast cells have rigid cell walls, the most significant component of the yeast cell wall is polysaccharides, including polymers of mannose, chitin and β-1,3-glucan (Aguilar-Uscanga and Francois, 2003). In order to allow antibodies accessing to the cells efficiently, the first step of making spheroplast chips is to remove the yeast cell wall. The essential activities of Zymolyase are β-1, 3-glucan laminaripentaohydrolase and β-1, 3-glucanse, which hydrolyze glucose polymers at the β-1, 3-glucan linkage and lyse the yeast cell wall. The cell should be a dark, translucent gray after appropriate digestion (Figure 4.3 D). Bright cells are insufficiently digested (Figure 4.3 A-C). Ghost cells (pale gray with little any internal structure) have been over digested (Figure 4.3 E-F). Most protocols use a relative high concentration of zymolyase and short digestion time (less than 30min) (Hanna et al., 1995; Schwartz et al.,

1997). Since the yeast spheroplast microarray is designed to deal with large-scale experiments that usually consist of hundreds or thousands of strains at a time, it can easily take longer time to spheroplast all the strains simultaneously, compared to small-scale experiments that generally deal with several samples at a time. Thus, we tried to titrate the concentrations of zymolyase in order to be able to get an appropriate digestion, but using longer digestion time. The optimized concentration used in our assay was 0.025mg/ml; the digestion time was 2 hours (Figure 4.3 D).

4.3.3 The storage conditions of cell chips

The storage of cell chips is another important issue for this technology. DNA microarrays can be simply stored at room temperature for several months because DNA molecules are relatively stable. However, since spheroplasted chips aim for protein localization and mRNA in situ hybridization, they cannot be simply stored at room temperature. For mRNA in situ hybridization, fixed and spheroplasted cells can be well preserved for weeks at -20 ^{o}C after cells are incubated at least overnight in 70% ethanol at -20 ^{o}C (Long et al., 1995). There is no report on how long and how to store the spheroplast cells used in immunostaining. We tried to store the chips with different treatments and at different temperatures. We found that we still detected intact microtubule structures (Figure 4.4B) after chips treated with methanol and acetone were stored at -80 ^{o}C, or after they were simply stored at -80 ^{o}C without any methanol and acetone treatment. Currently, yeast spheroplast chips can be stored for at least one month at -80 ^{o}C, which makes it convenient to perform replicate experiments later.

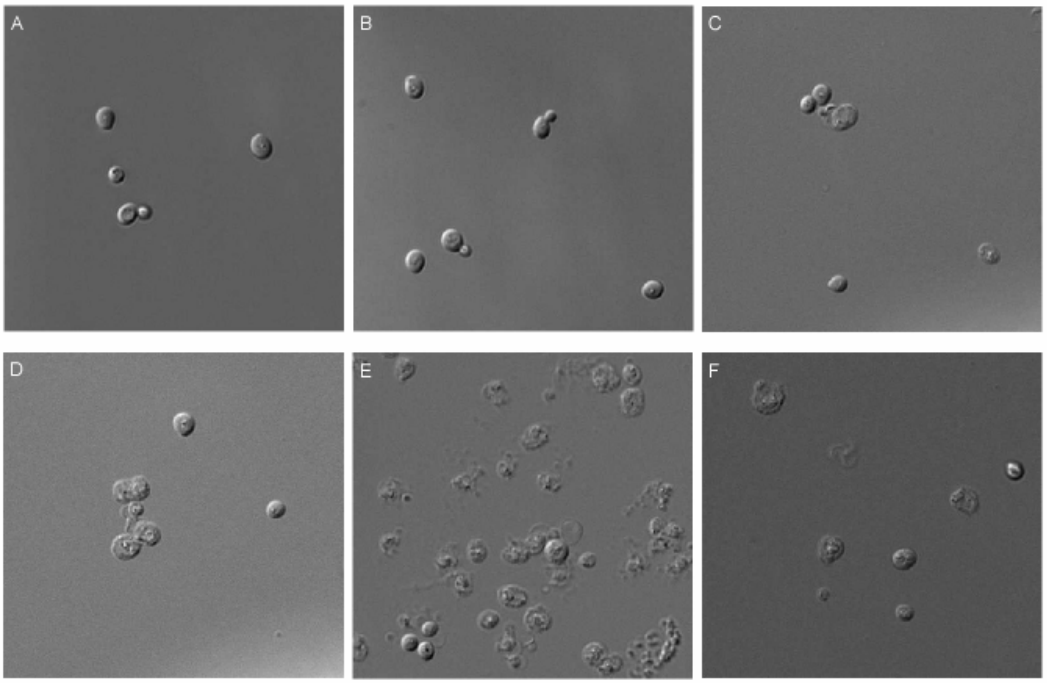


Figure 4.3 The morphology changes of yeast cells treated with Zymolyase at different concentrations. A: no treatment, B: 0.001mg/ml, C: 0.01mg/ml, D: 0.025mg/ml; E: 0.05mg/ml; F: 0.1mg/ml. These images were taken manually at 60X magnification.

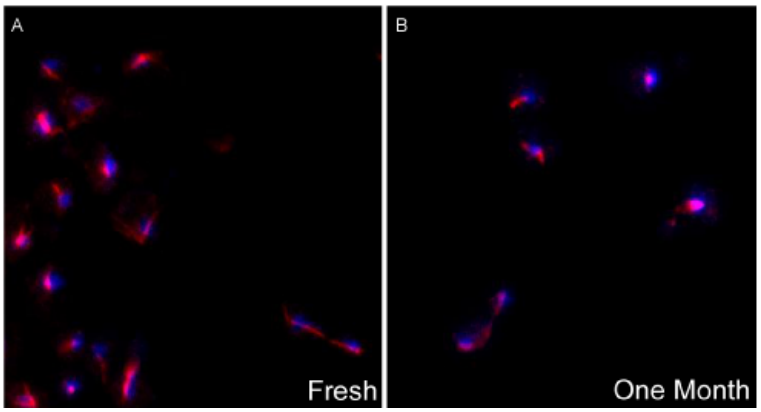


Figure 4.4 Comparison of the storage conditions by checking the stability of microtubule structure. Compared to fresh chips (A), chips stored at -80^{o}C still maintained intact

microtubule structures after one month (B). The images were captured through DAPI and Texas-red filter at 60X. Nucleus is shown in blue; microtubule is shown in red.

4.3.4 Evaluation of yeast essential gene function in microtubule organization using yeast spheroplast microarrays

As a proof of principle, we immunostained cell chips constructed from yeast tet promoter collection with anti-alpha-tubulin antibody, and we successfully showed the microtubule dynamics in wild type cells and recaptured genes whose mutations are known to cause spindle formation defects. In wild type cells, microtubules are star-like cytoplasmic arrays radiating from the microtubule organization center in the G1 phase (Figure 4.5A). During S/G2 phase, cytoplasmic astral microtubules are reoriented towards the bud tip and penetrate the bud (Figure 4.5B). When cells enter mitosis, microtubules are rearranged into bipolar spindle and positioned through the bud neck. In anaphase, cytoplasmic microtubule shortens, the spindle elongates, and sister chromosomes are segregated (Figure 4.5C). When cytokinesis is completed, the spindle disassembles (Figure 4.5D).

Spc97p and Spc98p are two major components of microtubule nucleating Tub4p complex, involved in microtubule nucleation and mitotic spindle organization (Knop et al., 1999). Down regulation of *SPC97* and *SPC98* resulted in short spindles, elongated cytoplasmic microtubules and defective chromosome segregation, which showed similar defect to that of the corresponding temperature sensitive mutants (Figure 4.6 A-B). Tid3p is a component of the kinetochore associated Ndc80 complex, involved in chromosome segregation and microtubule nucleation (Wigge and Kilmartin, 2001; Zheng et al., 1999). Down regulation of *TID3* caused defects in spindle elongation and nuclear division (Figure 4.6C). Irr1p is a subunit of cohesin complex, down regulation of which resulted

in large-budded cells with short spindles (Figure 4.6D). Besides recapturing the above known genes involved in microtubule organization, we also found the abnormal microtubule organization when *DNA2* and *POP1* were down regulated (Figure 4.6E-F), while *DNA2* and *POP1* are known to be involved in DNA repair and rRNA processing respectively.

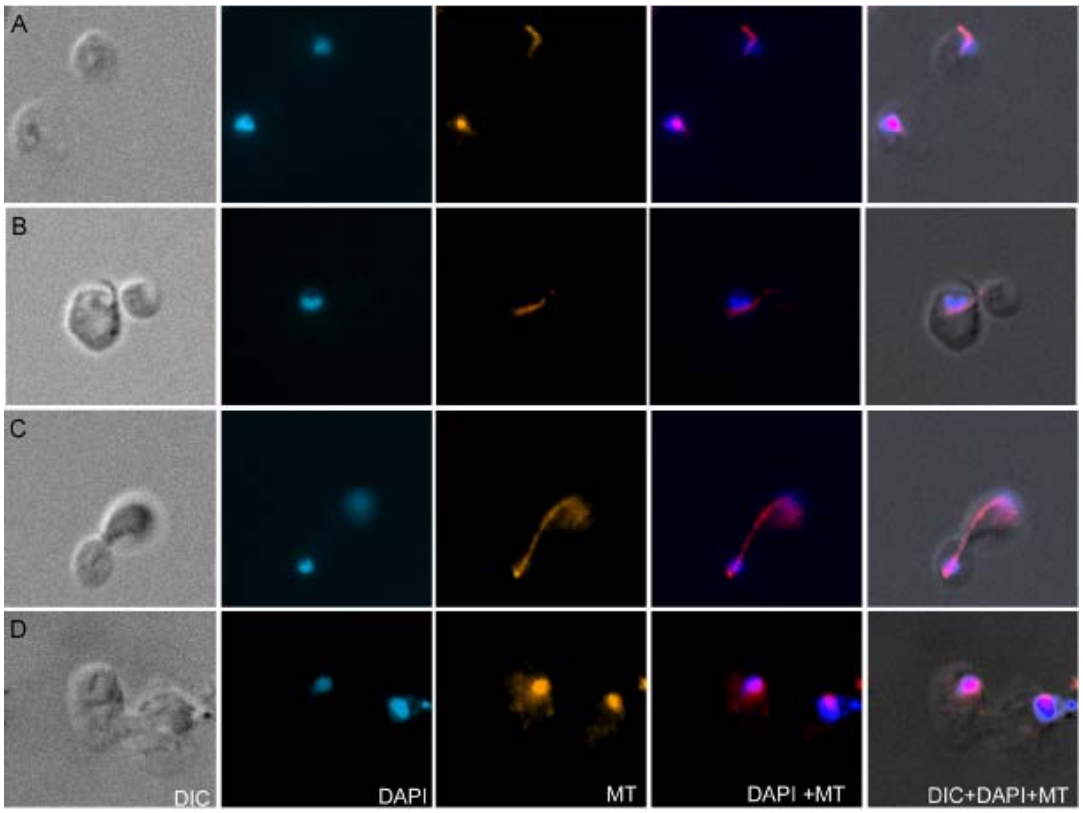


Figure 4.5 Example images of microtubule and spindle morphology in different cell cycle stages. (A) G1 phase, (B) G2 phase, and (C-D) anaphase. Nucleus is shown in blue; microtubule is shown in red. The images were captured through DAPI and Texas-red filter at 60X.

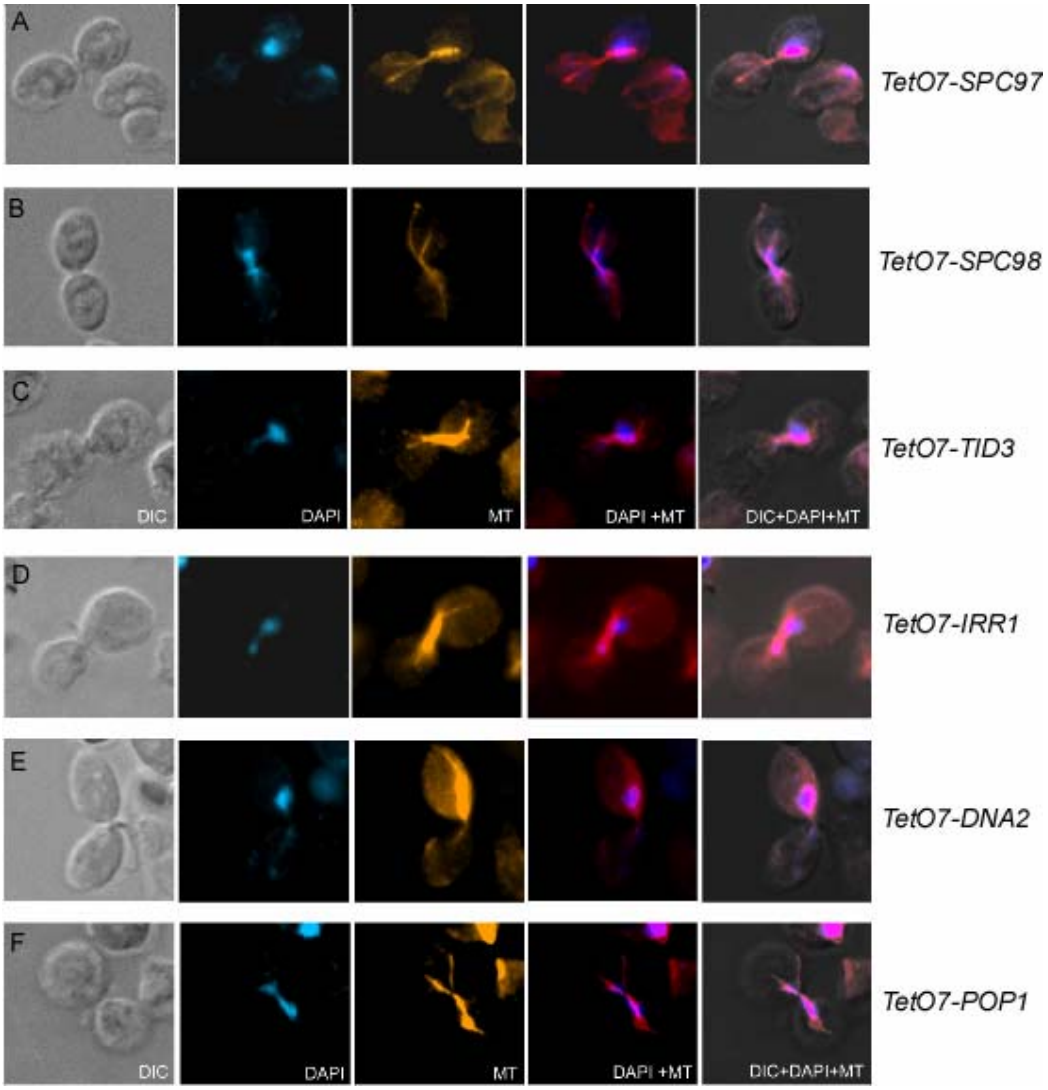


Figure 4.6 Example images of TetO7 promoter strains having defective microtubule and spindle structure. A and B: Down regulation of SPC97 and SPC98 resulted in short spindles and elongated cytoplasmic microtubules. C: Down regulation of TID3 caused defective spindle elongation and uneven nuclear division. D: Down regulation of IRR1 resulted in large budded cells with short spindles and abnormal chromosome segregation. E and F: Abnormal spindle morphology was observed in TetO7-DNA2 and TetO7-POP1

strains. Nucleus is shown in blue; microtubule is shown in red. The images were captured through DAPI and Texas-red filter at 60X.

4.4 DISCUSSION

We developed yeast spheroplast chips in order to analyze protein localization across thousands of yeast strains with distinct genetic backgrounds *via* low-cost and high-throughput immunofluorescence microscopy. As a proof of concept, yeast spheroplast chips were built from the *TetO7* promoter collection covering ~700 essential genes, and probed with anti-alpha tubulin antibody to examine the effects on the organization of microtubule and spindle when essential genes were down regulated. We recaptured known mutants having defective spindles and discovered new mutants with abnormal spindle structure. The proof of principle experiment showed the potential of spheroplast chips for high-throughput and high-parallel protein localization profiling by analyzing the interactions between probed targets and down-regulated essential genes.

Technically, we minimized the use of expensive reagents and time to screen a whole collection, afforded by limiting the use of antibodies and dyes to a single microscope slide and automatically capturing 2,100 hundred images (DIC, DAPI, and immunostained microtubules) in 10 hours as compared to use almost 700 times more antibodies and time to image the same collection. Moreover, we can print hundreds of chips at one time, store them at -80 ^{o}C, and probe chips manufactured at the same condition with different antibodies or the same antibodies as that used in the initial screen, which saves us time on preparing samples and allows us to reproduce the initial experiment under the exact same conditions.

Considering all the features mentioned above, the major advantages of yeast spheroplast chips are the minimal usage of reagents, high reproducibility, and highly

parallel analysis of protein localization using high-throughput immunofluorescence microscopy. With this method, we can potentially conduct synthetic phenotype analysis across a deletion library or perform protein co-localization profiling across 4,200 GFP-tagged strains with any available probe on a single chip, which will avoid the strain construction, mating and dissecting processes. Furthermore, this technology can be potentially adapted for any systems with arrayed strain collections, such as siRNA or microRNA tranfected human cell line collections. These single cell level and gene specific data from such experiments will provide useful functional connections between genes, and ultimately lead to the completion of functional network that underlines the normal and perturbed conditions**.**

Chapter 5: A systematic investigation of *Saccharomyces cerevisiae* genome for characterizing gene function in cell cycle progression

5.1 INTRODUCTION

The cell cycle is a ubiquitous, but complex process, involving numerous regulatory proteins that direct cells through a specific sequence of growth stages and the production of two daughter cells (Mendenhall and Hodge, 1998; Schafer, 1998). It begins from G1 phase wherein cells increase their size and volume, followed by DNA synthesis (S phase). After the second growth phase (G2 phase), cells enter mitosis (M phase) and two copies of DNA are equally separated to two daughter cells. Finally two daughter cells are separated during the cytokinesis phase and prepare for the next cell cycle. The budding yeast *Saccharomyces cerevisiae* has these processes very similar to other eukaryotic organisms except for the nuclear envelope not breaking down during mitosis and the daughter cell coming from bud emergence. Due to its rapid division, the availability of genetic tools, and homology to higher eukaryotic cell cycle processes, the budding yeast has become a fundamental model system to study eukaryotic cell cycle progression (Hartwell, 1974).

Since the classical screens for cell division cycle (CDC) genes performed by Hartwell and colleagues in the early 1970s, cell cycle division has been studied extensively for decades. By identifying conditional temperature-sensitive mutants with specific arrest points, Hartwell *et al.,* discovered more than 50 CDC genes required for specific stages in cell cycle division (Hartwell, 1971a; Hartwell, 1971b; Hartwell, 1973; Hartwell et al., 1970). However, there are additional genes with cell cycle functions remained to be identified (Hartwell, 1973; Hartwell, 1974). Indeed, the recent genetic screens in yeast and other organisms identified additional cell cycle genes (Bjorklund et

al., 2006; Stevenson et al., 2001; Yu et al., 2006). However, most of these screens employed loss-of-function approaches, such as gene deletion, RNAi, and promoter shutoff, which are generally limited by less informative phenotypes. In contrast, overexpression often leads to more detectable effect on cellular function, which helps us to identify genes whose functions are normally compensated by redundant genes or genes acting as negative regulators in the biological pathways (Gelperin et al., 2005; Ouspenski et al., 1999). Stevenson *et al.,* performed a large-scale overexpression screen for cell cycle genes *via* expressing moderated GAL promoter-driven cDNA library and sheared genomic DNA library in ARS-CEN vectors (Stevenson et al., 2001). They identified 113 genes, including genes whose elevated expression caused very slight effects on the cell cycle. Their screen was clearly not saturated due to the coverage of their cDNA library and incomplete gene annotation. Therefore, Stevenson proposed that the completion of *S. cerevisiae* genome and the ability to systematically clone all genes into an overexpression vector might allow comprehensive analysis of the entire yeast genome (Stevenson et al., 2001).

In the present work, we performed a near-saturating screen for yeast genes having overexpression-induced defects in cell cycle progression, taking advantage of the availability of yeast open reading frame (ORF) collection, covering 91% of the yeast complete ORF set including dubious ORFs (Gelperin et al., 2005). After measuring the fraction of cells in different phases of the cell cycle *via* high-throughput flow cytometry for each of ~5,800 ORFs, we observed that overexpression of 108 genes leads to significant changes in G1 or G2/M populations, 82 of these genes are newly implicated in the cell cycle, and most the newly indicated genes are likely to affect cell cycle progression *via* a gain-of-function mechanism.

5.2 MATERIALS AND METHODS

5.2.1 Yeast strains

The yeast ORF collection was obtained from Open Biosystems, in which each of 5,854 yeast ORFs was cloned into a 2μ plasmid under control of the GAL1 promoter in order to provide highly elevated expression. Wild type control strains were constructed by transforming empty precursor vector BG1766 to the ORF host strain Y258 (MATa *pep4-3, his4-580, ura3-53, leu2-3,112*). A Fus1-GFP strain (MATa *his3*Δ *leu2*Δ *met15*Δ *ura3*) carrying P_{GAL1}-*SKO1* plasmid was also constructed *via* a standard yeast plasmid transformation protocol.

5.2.2 Induction of expression

Cells were initially grown in 96-well plates (Corning 3595) with 170μl SD-URA medium for 1-2 days at 30 ^{o}C, and then 5μl cells were re-inoculated into a new set of 96-well plates with 170μl SC-URA, 2% raffinose medium. After 12 hour growth in raffinose medium, cells were re-inoculated to a new set of plates with 100μl SC-URA, 2% raffinose medium at the final OD at 0.15 and grown for 1 hour, and then 70μl SC-URA medium with 5% galactose were added into this cell culture (the final concentration is 2%). Cells were then induced for 8-10 hours at 30 ^{o}C. Wild type control strains and positive control strains were treated in parallel.

5.2.3 High-throughput flow cytometry

Flow cytometry analyses were performed as adapted from Haase and Reed (Haase and Reed, 2002). Briefly, ~0.2×10^{7} cells were harvested and fixed in 200μl 70% ethanol, then treated with 1mg/ml RNAse A (Sigma) for 4 hours at 37 ^{o}C, followed by incubation with 1mg/ml Proteinase K (Sigma) for 1 hour at 50 ^{o}C. ~8×10^{5} Cells were then resuspended in 200μl 50mM sodium citrate with Sytox green (Invitrogen) at a final concentration of 1.5μM. All the above liquid transferring processes were performed using

a Biomek FX robot system (Beckman Culter). Samples were analyzed by flow cytometry, using a Becton Dickinson FACSCalibur with BD HTS auto sampler, controlled by plate manager software and Cellquest pro software (BD Biosciences). Well-to-well contamination was prevented by flushing with ddH_2O between each pair of samples. In order to get as many events as possible and shorten the time in acquiring data for ~5800 strains, 20,000 events/strain or 30-second acquiring time/strain was the instrument setting for data collection. The HTS auto sampler stops acquiring data once either of the setting has been achieved. Thus, for the extreme slow growth strains, we did not get the average of 20,000 events in 30 seconds.

5.2.4 Analysis of flow cytometry profiles

Analysis of DNA profiles was automated using ModFit 3.0 software (Verify Software house, Inc), fitting the distributions of G1 and G2/M cells with Gaussian distributions (Figure 5.1C) and calculating the goodness-of-fit *via* the Reduced Chi Square (RCS) method. For quality control, DNA profiles with RCS>5 and event number<5000 were discarded. The percentage of cells under each DNA peak (1C peak or 2C peak) was calculated by dividing the number of events under each peak by the total number of events under all peaks, and the ratio (1C/2C) of the percentage of cells under the 1C peak and that under the 2C peak was calculated for each strain. The ratios of Log2 (1C/2C) for all strains were then calculated, well-fitting with Gaussian distribution (R^2 =0.97) (Figure 5.2A).

5.2.5 Nuclear staining and bud size measurements

108 ORF strains showing reproducible cell cycle arrest were grown and induced as described above. After induction, cells were fixed in 70% ethanol and treated with 1mg/ml RNAse A, and then stained with 1μM Sytox green (Invitrogen). Cells were examined *via* phase contrast microscopy and fluorescence microscopy using a Nikon

eclipse fluorescence microscope (Nikon). We measured bud size using the differential interference contrast (DIC) images. We used ImageJ to measure the length of bud and mother cell for an average of 100 cells for each of 108 strains. The bud size was classified by dividing the bud length by the length of mother cell. If the ratio was 0, cells were classified as 'no bud'; cells were categorized into 'small bud' when the ratio was between 0 and 0.4; 'large bud' when the ratio was higher than 0.4 (Hartwell, 1971a). We further examined the large-budded cells and counted three types of nuclear morphology: an undivided nucleus in one cell body (class I), an undivided nucleus in the bud neck (class II), and divided nuclei in two cell bodies (class III) (Huffaker et al., 1988; Schwartz et al., 1997; Sobel and Snyder, 1995). An average of 50 cells was counted for each of 87 G2/M strains.

5.2.6 Growth assays

82 newly indicated genes causing cell cycle defects and 3 positive controls (*TUB2*, *PAC2*, and *CST9*) were included in the growth assays, and three conditions were tested: SC-URA, 2% galactose, SC-URA, 2% galactose plus nocodazole, and SC-URA, 2% galactose plus hydroxyurea. Cells were grown overnight in SD-URA medium, and then washed with SC-URA, 2% raffinose medium and grown in SC-URA, 2% raffinose medium for one hour at 30 ^{o}C before being spotted onto agar plates. Six 10-fold serial dilutions were made for each strain, with the OD of the first series at 0.2. 10µl of each series was spotted onto SC-URA, 2% galactose plates and SC-URA 2% galactose plates containing the appropriate drugs, and grown at 30 ^{o}C. Plates were photographed after 2-3 days growth in SC-URA, 2% galactose plates, 5-8 days in the plates supplemented with drugs. Drug concentrations were as follows: nocodazole, 15µg/ml, hydroxyurea, 50µM (Hoyt et al., 1997; Ouspenski et al., 1999).

5.3 RESULTS

5.3.1 High-throughput flow cytometry and automated analysis of DNA profiles

To analyze the effect of overexpression of yeast genes in cell cycle progression, we applied high-throughput flow cytometry to screen ~5,800 strains of yeast ORF collection for genes that cause delay or arrest at particular cell cycle stages when overexpressed (Gelperin et al., 2005). Figure 5.1 shows the approach and representative flow cytometry profiles of arrested stains. Accumulation of cells with either one copy (1C) or two copies (2C) DNA content indicates defect in progression through a particular cell cycle stage. Thus, in order to search for such defects induced by overexpression of a particular yeast gene, we analyzed asynchronous cell cultures and determined the distributions of DNA content, assaying if cells from each given ORF strain exhibit a skewed distribution from control cells. We included replicated analyses of ~140 wild type control cultures in parallel with the ORF strains. A total of ~6,000 DNA histograms were acquired and quantitatively analyzed, measuring the ratio of 1C/2C cells for each strain, i.e., the ratio of cells in the G1 phase and the cells in the G2/M phase. The Log2 (1C/2C) ratios of 140 control strains and ORF strains fitted with Gaussian distribution (R^2=0.97) (Figure 5.2A). We observed the Log2 (1C/2C) ratios of the 5,800 ORF strains and 140 control strains to be normally distributed ($p<0.01$). Therefore, for each strain, we calculated a Z score for its DNA content and could identify the ORF strains with significantly different accumulation of cells in the G1 or G2/M growth phases. Based on this Z score (Figure 5.2B), 2 categories were assigned: G1 and G2/M category. ORF strains with Log2 (1C/2C) ratios in the left tail of Gaussian distribution were considered as G2/M category in which cells were accumulated at two copies of DNA content, and ORF strains with Log2 (1C/2C) ratios in the right tail were likely to accumulate at one copy of DNA content. We could assign genes in G1 and G2/M categories using different

confidence levels (Figure 5.2C). Here, we used 95% confidence level and 198 genes were in initial set of genes whose overexpression caused cell cycle detects, excluding any wild type control strain. Of 198 strains, 108 were validated at least twice by manual flow cytometry analysis, as shown in Supplemental Figure 5.1. Of the 108 genes, 21 genes were identified in the G1 category, 87 genes were in the G2/M category. These genes are listed in full in Table 1 and the DNA profiles are shown in full in Supplemental Figure 1. Additionally, we manually assigned genes to diploid and 3C categories: 202 strains were in the diploid category and 64 strains were in 3C category. However, diploid and 3C strains are likely to be false positives because the copy numbers of DNA content did not change upon induction, which could be a result of the mixture of haploid and diploid strains (data not shown).

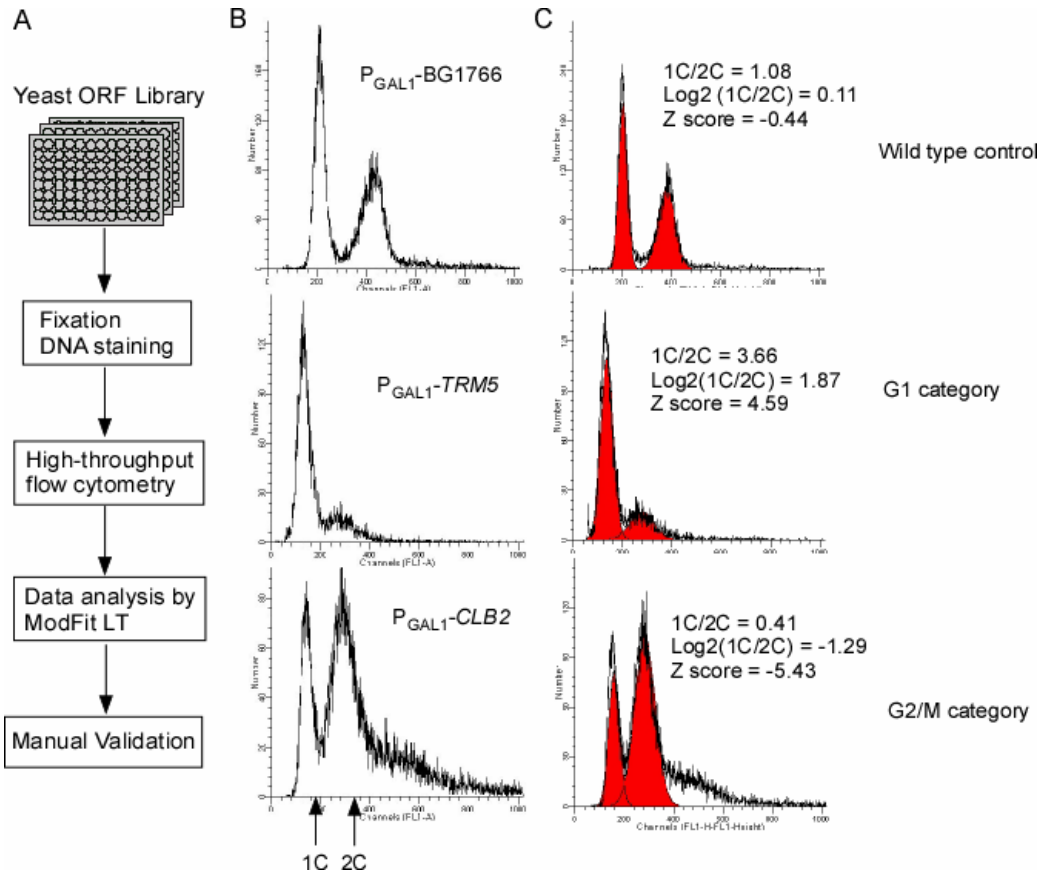


Figure 5.1 An overview for cell cycle screen**.** (A) Flow chart of cell cycle screen. Yeast ORF strains and control strains were induced in 96-well plates with SC-URA, 2% galactose medium, and analyzed *via* high-throughput flow cytometry. All flow cytometry histograms were analyzed by ModFit LT software. The ORF strains that showed cell cycle defects in the initial large-scale screen were validated twice in manual mode of flow cytometry. (B) Flow cytometry histograms of wild type strains and ORF strains are shown. The fluorescence intensity (DNA content) is indicated in x-axis, and the number of cells in a given intensity is plotted in y-axis. (C) The report files of DNA histograms analyzed by ModFit software. Each peak is fitted with a Gaussian distribution as shown in red.

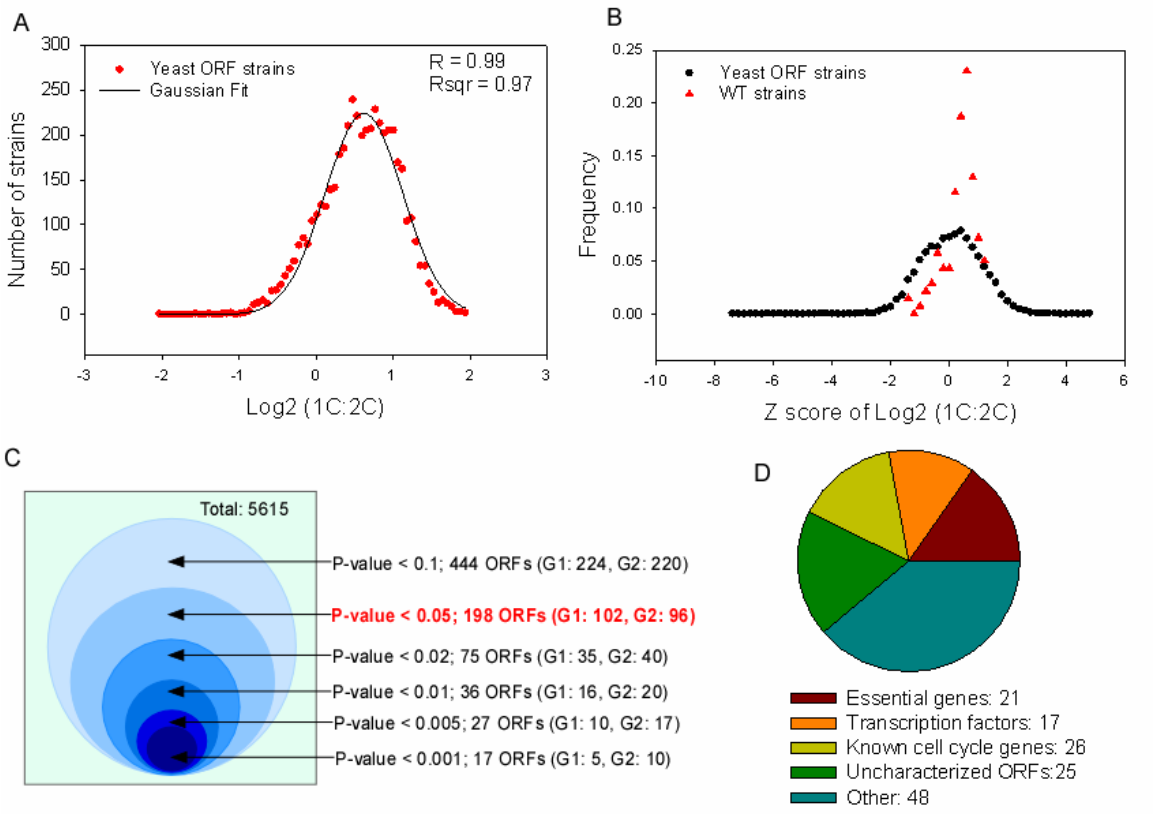


Figure 5.2 Summary of cell cycle screen**.** (A) A dot plot for Log2 (1C/2C) ratios of yeast ORF strains in the initial large-scale screen. These values were normally distributed and fitted with Gaussian distribution ($R^2 \approx 0.97$). We then calculated a Z score for its DNA content to identify the ORF strains with significantly different accumulation of cells in the G1 or G2/M growth phases. (B) A dot plot of frequency of Z scores calculated with mean and standard deviation of Log2 (1C/2C) ratios of yeast ORF strains in the initial screen. Control strains had a narrower distribution than ORF strains. (C) The numbers of ORF strains showing atypical cell cycle histograms in the initial screen when different confidence levels were used. The P-value used in the present study is highlighted in red bold font. There were 198 ORF strains having abnormal profiles when P-value was less

than 0.05. Of 198 genes, 108 were validated at least twice. (D) The classification of 108 validated genes.

5.3.2 Independent validation by bud size measurement

The size of the bud relative to the size of the mother cell is the most notable morphological landmark of the cell cycle stages in budding yeast. The bud size was the basis of classical cell cycle screens (Culotti and Hartwell, 1971; Hartwell, 1971a; Hartwell, 1971b; Hartwell, 1973; Hartwell et al., 1970; Moir et al., 1982). Changes in bud size allow the identification of mutants blocked at specific stages of the cell cycle. DNA replications occurs when bud size is small, nuclear division occurs when the bud is about three-fourths the size of the mother cell, and cell separation when the bud is approximately equal size to that of the mother cell. To validate genes in the G1 and G2/M categories using bud size, we measured the ratio of bud size and mother size for the 108 ORF strains identified by flow cytometry as having cell cycle defects. Genes in the G1 category caused clearly elevated populations of unbudded cells when overexpressed and all the genes in the G1 category tested for bud sizes (one is not included in the experiment) exhibited a higher percentage of unbudded cells than control strains. For example, 92% of cells were unbudded and only 2% of cells were large-budded when *TRM5* and *BUL1* were overexpressed. In contrast, only 57% of wild type control cells were unbudded, and 28% were large-budded (Figure 5.3BC and Table 5.1). Of 89 strains in the G2/M category, 85 exhibited a higher percentage of large-budded cells than control strains (Figure 5.4BC). For instance, more than 60% of cells had large buds when *TUB2* and *SPC97* were overproduced (Figure 5.3E and Table 5.1). Consistent with previous observations, *TRM5*, *TUB2* and *SPC97* are known to cause cell cycle delays when their normal function is perturbed (Fitch et al., 1992; Huffaker et al., 1988; Knop et al., 1997;

Yu et al., 2006). The bud size analysis thus provided a useful independent validation of the DNA content observations, with genes validated by both flow cytometry analysis and bud size distributions being the most likely to affect cell cycle progression. 13 strains in the G1 category are with significantly (95% confidence levels) high populations of unbudded cells (*BUL1*, *TRM5*, *SKO1*, *YOR131C*, *ARC1*, *NCB2*, *FMP36*, *CYT1*, *IES3*, *TMA64*, *GOS1*, *RPA14*, *ENO2*). 47 strains in the G2/M category are with significantly (95% confidence levels) elevated populations of large-budded cells (*RFA1*, *TUB2*, *NIP100*, *PPZ1*, *TEA1*, *CLB3*, *HOS3*, *ACT1*, *SPC97*, *SET3*, *CLB5*, *SGN1*, *CLB2*, *CDC39*, *YIL158W*, *GEA2*, *ENT3*, *RLI1*, *ARF1*, *PDR17*, *YIR016W*, *CST6*, *CBF1*, *ATG26*, *SAN1*, *MTH1*, *TEC1*, *YPL247C*, *PRP31*, *SHE1*, *PAC2*, *AVO2*, *WSC2*, *YAP1*, *YPR015C*, *YDR266C*, *YGR109W-A*, *SLK19*, *IME2*, *PMT5*, *MSN5*, *NTH1*, *SGF73*, *MNN10*, *ALG6*, *VTC4*, *PBS2*).

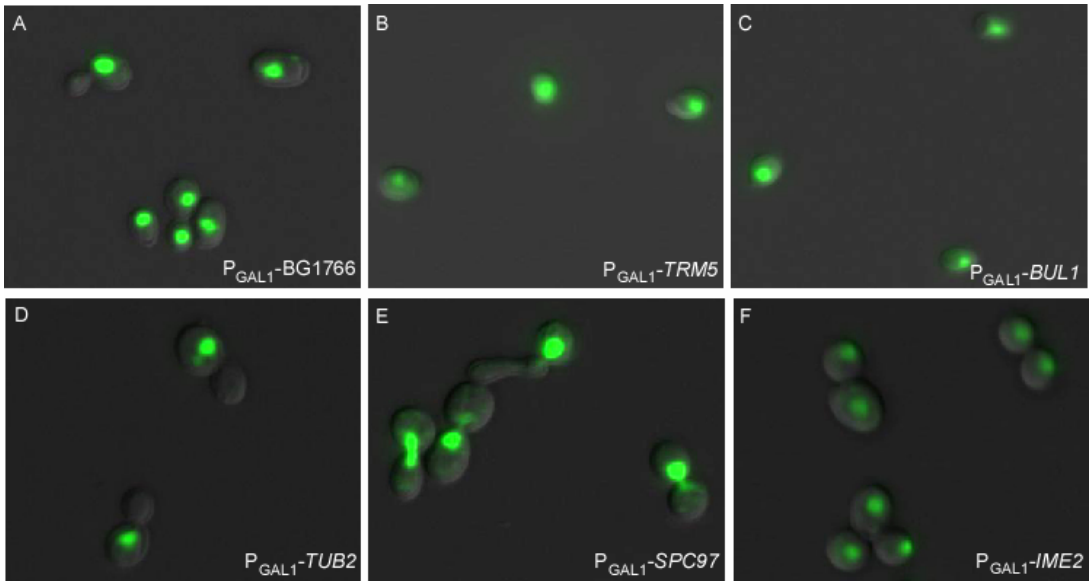
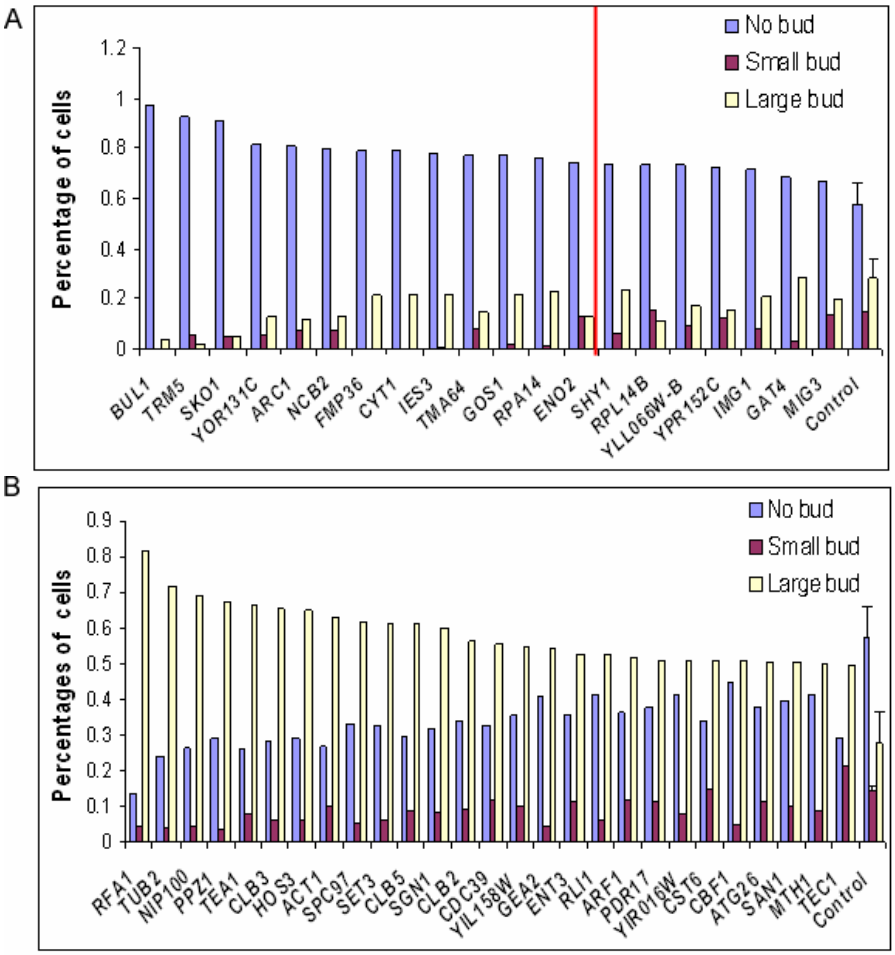


Figure 5.3 Example images of budding count and nuclear staining**.** Cells were stained with Sytox green, which stained nuclei, and cells were visualized through FITC and DIC filters. Overlaid images are shown. (A) Wild type cells with empty vector. (B-C) Selected

images for genes in the G1 category: C: *TRM5*; D: *BUL1*. (D-F) Selected images for genes in the G2/M category representing three classes of nuclear morphology: (D) Class I (pre-M): overexpression of *TUB2* caused elevation in large-budded mononucleate cells (E) Class II (early-M): overexpression of *SPC97* resulted in large budded cells with undivided nucleus in bud neck, (F) Class III (late-M): the elevated populations of large-budded cells that had completed nuclear DNA segregation when *IME2* was overexpressed.



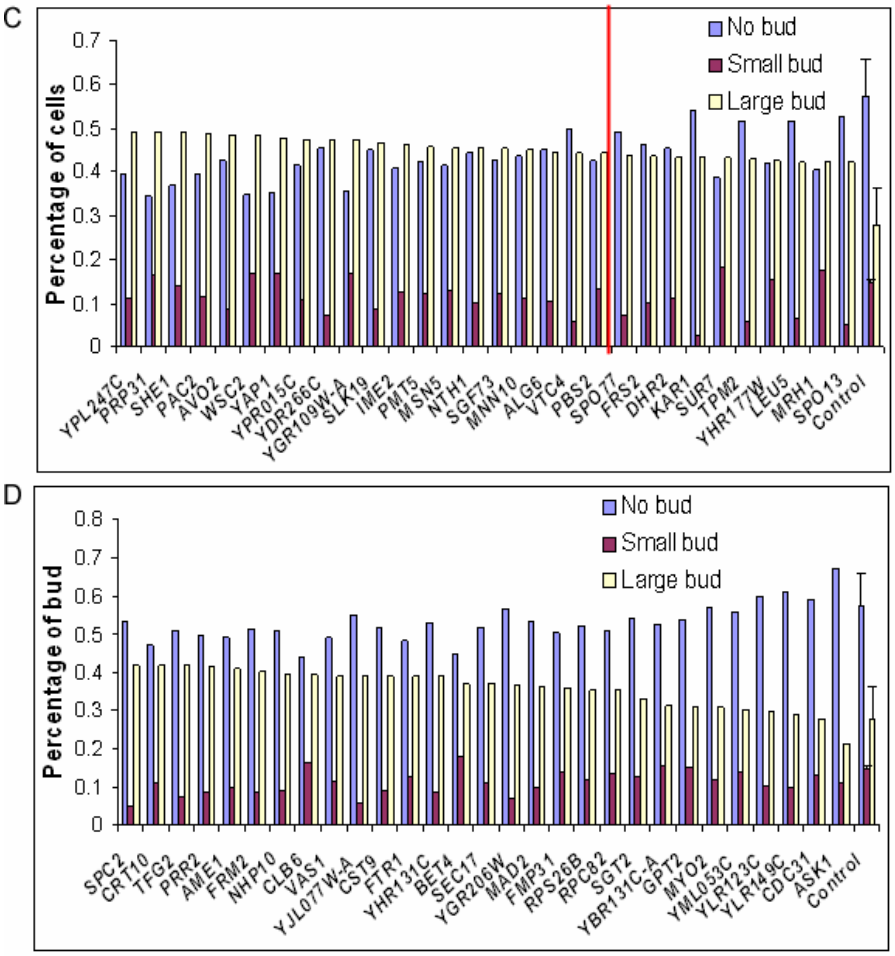


Figure 5.4 Summary of bud size measurements**.** (A) shows the Bud size measurements of strains in the G1 category. For better visualization, data were sorted by percentages of cells without buds. (B-D) shows the bud size measurements of strains in the G2/M category. Data were sorted by percentages of cells with large buds. Red lines mark the 95% confidence thresholds. Gene names are indicated in x-axis, percentages are indicated in y-axis.

5.3.3 Overexpression phenotypes occur primarily through a gain-of-function mechanism

Overexpression of a normal gene product affects cellular functions by different mechanisms. One of them is a dominant-negative effect that is similar to loss-of-function studies. Alternatively, elevated amounts of a protein can increase the activity of a normal cellular pathway that leads to a detectable phenotype (gain-of-function). In order to identify genes that affected cell cycle progression *via* a gain-of-function mechanism, we compared the bud size phenotype caused by overexpression with the same phenotype that caused by gene deletion provided, as measured in the *Saccharomyces cerevisiae* Morphology Database **(**SCMD) (Saito et al., 2004). Of 108 genes from our cell cycle screen, 77 also appear in SCMD (21 essential genes and 10 additional genes are not included in SCMD) (Figure 5.5C). We selected genes from our screen with significantly elevated populations (95% confidence level) of unbudded cells or large-budded cells. In the G1 category, there were 12 strains from our screen whose percentages of cells without buds were significantly higher than that of wild type. Of 12 genes, 2 also showed a significantly elevated population of unbudded cells upon deletion of the gene. Therefore, our rough estimate is that 10 genes in the G1 category may exhibit overexpression phenotype due to a gain-of-function mechanism versus to 2 by a dominant negative effect. Similarly, 44 genes in the G2/M category caused a significantly elevated proportion of large-budded cells when overexpressed but not deleted, versus 3 that could be explained by a dominant negative effect (Figure 5.5C, table 5.3).

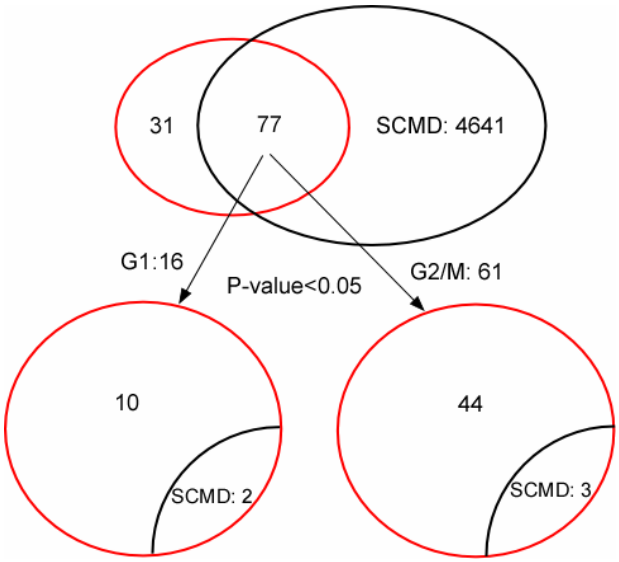


Figure 5.5 Overexpression phenotypes occur primarily through a gain-of-gain mechanism**.** Of 108 genes, 77 genes were included in SCMD: 16 genes were in the G1 category, 61 were in the G2/M category. When P-value was less than 0.05, 12 genes caused significantly higher percentages of cells without buds than control strains when overexpressed. Of 12 genes, only 2 genes lead to significantly high populations of unbudded cells when deleted. For G2/M genes, 47 genes caused significantly elevated percentages of cells with large buds upon overexpression; only 3 of them lead to significantly high populations of large-budded cells when deleted.

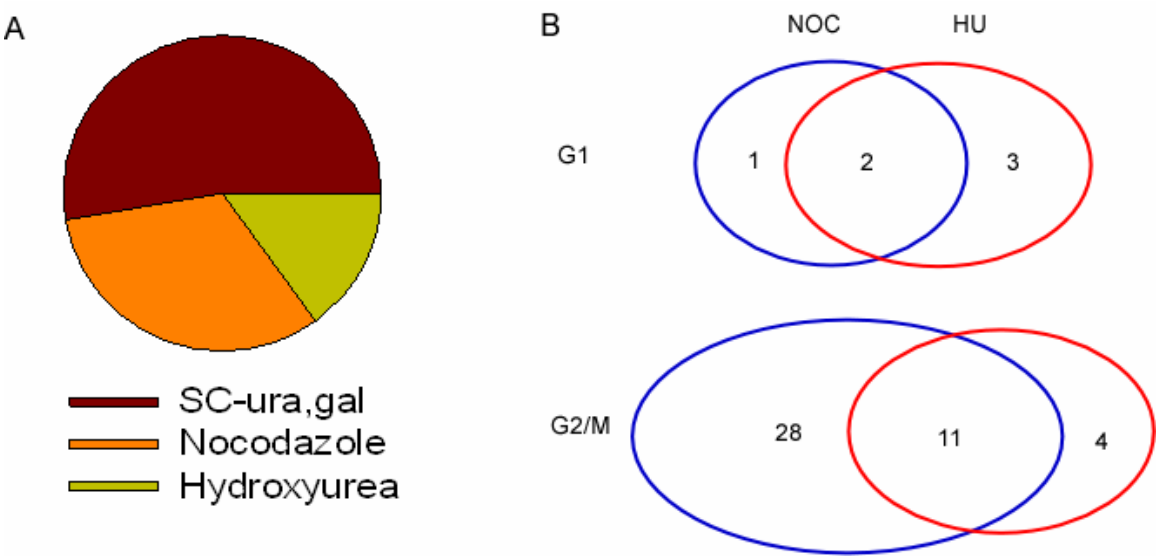


Figure 5.6 Summary plots for genes causing slow growth phenotypes upon overexpression**.** Positive control strains and 82 ORF strains that have not been previously reported to have cell cycle defects were grown on SC-URA, gal plates, and SC-URA, gal plates supplemented with an appropriate drug: nocodazole (NOC) or hydroxyurea (HU). (A) The numbers of ORF strains having slow growth phenotypes on each condition are shown. Gene set are listed full in Supplemental Table3. 68 were slow growth upon induction, 42 were sensitive to NOC, and 20 were sensitive to HU. (B) The overlapped gene set between genes sensitive to NOC and the genes sensitive to HU. 13 ORF strains were sensitive to both drugs (2 G1 genes; 11 G2/M strains); 29 were NOC-sensitive only (1 G1 gene; 28 G2/M genes), 7 were HU-sensitive only (3 G1 genes; 4 G2/M genes).

5.3.4 Cellular functions affected by overexpression of the genes newly implicated in the cell cycle

One major expected cause of defective cell cycle progression is chromosome instability, specially chromosome loss and non-disjunction. Chromosome loss is

characteristic of defects in DNA metabolism, while non-disjunction typically reflects defects in mitotic segregation (Ouspenski et al., 1999). To better speculate whether chromosomal functions were primarily affected by the overproduction of the identified ORFs, we examined the strains' sensitivity to hydroxyurea and nocodazole. Hydroxyurea (HU) is an inhibitor of ribonucleotide reductase, an enzyme necessary for DNA synthesis. Nocodazole (NOC) is a microtubule depolymerizing drug that prevents formation of the mitotic spindle. Genes involved in DNA metabolism and DNA replication checkpoint are often sensitive to HU, whereas genes sensitive to microtubule drugs are likely to interact with the mitotic checkpoint and mitotic spindle formation. Due to the presence of the spindle checkpoint control, yeast mutants affecting spindle structure normally show cell-cycle arrest in mitosis (Winsor and Schiebel, 1997). We observed 29 strains to be specifically sensitive to NOC, 7 to be specifically sensitive to HU, and 13 strains to show sensitivity to both, probably affecting multiple nuclear functions (Figure 5.6B, Table 5.4). The growth assay data are shown in full in Supplemental Figure 2. As expected, *TUB2* and *PAC2* exhibited a 'non-disjunction'-type phenotype, sensitivity to NOC but not HU because *TUB2* and *PAC2* are required for normal microtubule function and mitotic sister chromatid segregation (Hoyt et al., 1997; Huffaker et al., 1988). We might expect that genes in the same category as *TUB2* and *PAC2* might be directly or indirectly involved in microtubule function or functions related to chromosome segregation, explaining why most genes (28 of 29 genes) causing specific sensitivity to nocodazole arrested at the G2/M phase when overexpressed (Figure 5.6B).

5.3.5 Functional analysis of genes affecting cell cycle progression when overexpressed

We next examined the functions for the 108 genes that caused cell cycle defects when overexpressed. Among these genes, 21 are essential ORFs, 17 transcription factors

and 25 ORFs are uncharacterized or dubious (Figure 5.2D). Importantly, of the 26 genes known for having cell cycle defects identified in the screen, 24 were consistent with the previously observed phenotypes. Of 8 Cdc28p cyclins included in ORF collection, we were able to recapture 5 cyclins (*CLN1*, *CLB2*, *CLB3*, *CLB5*, and *CLB6*). Of 21 essential genes, 12 are available as $TetO_7$ alleles. We were able to recover 67% of essential genes whose overexpression caused the similar cell cycle arrests as $TetO_7$ alleles did (Yu et al., 2006). Additionally, the enrichment of biological processes was consistent with previously published cell cycle screens (Bjorklund et al., 2006; Sopko et al., 2006; Stevenson et al., 2001; Yu et al., 2006), validating the general quality of this screen. In the next two sections, we described the results for G1 and G2/M genes in more details.

5.3.5.1 G1 category

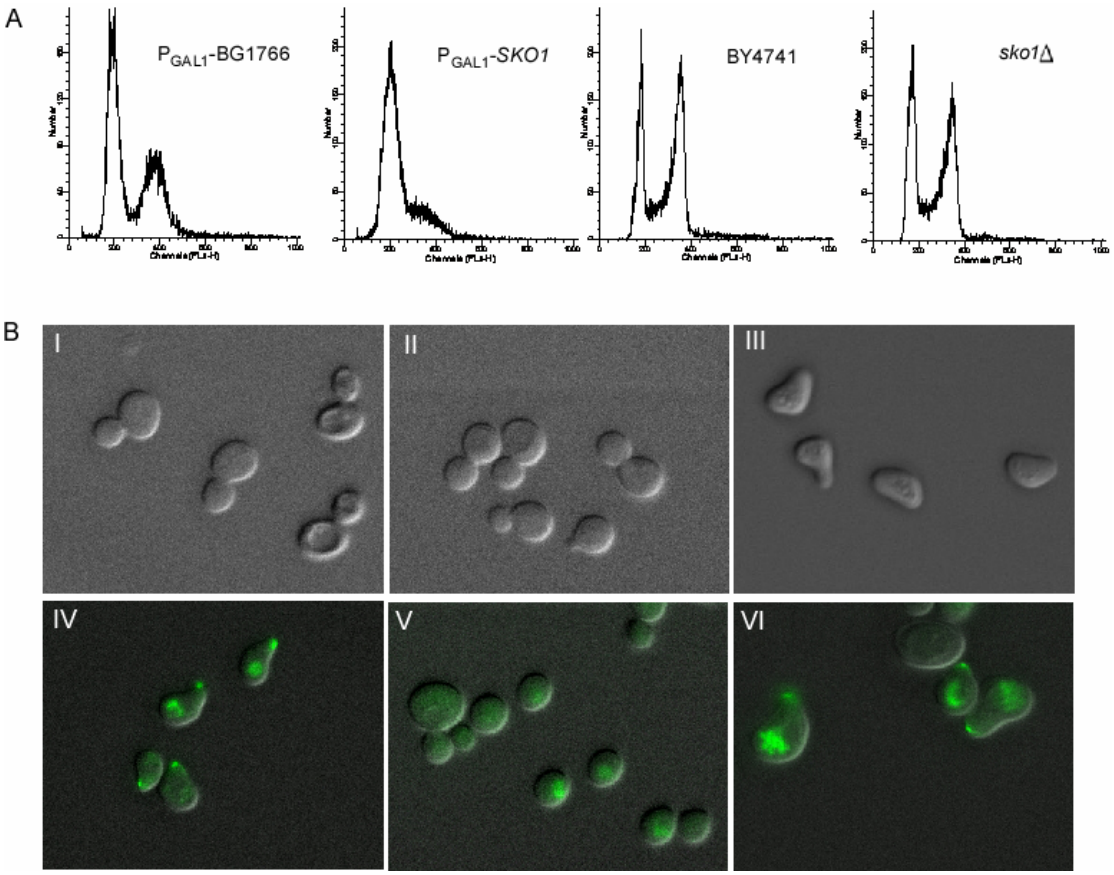
21 genes in the G1 category, only 2 of them has been previously reported to cause cell cycle arrest at the G1 phase. While the mutants arresting at the G2/M phase were strongly enriched for cell cycle associated functions, diverse mechanisms are known to induce G1 arrests (Sopko et al., 2006; Stevenson et al., 2001). This diversity was reflected in the enrichment of GO terms among the G1 arresting ORFs: no strong pathway was enriched at $P<0.001$ calculated by Robinson et al (Robinson et al., 2002) , which is consistent with the previous overexpression studies (Sopko et al., 2006; Stevenson et al., 2001). The strongest enrichment for the G1 category consists of the genes involved in negative regulation of transcription from RNA polymerase II promoter (GO:0000122; p-value$<4x10^{-4}$) when calculated by Hu et al (Hu et al., 2007).

This lower enrichment for GO terms reflects the fact that 29% (6) of genes in G1 category are uncharacterized ORFs. *YOR131C* and *YDR493W* were discovered in high-throughput studies. The only functional information available for them is localization: *YOR131C* is localized in nucleus and cytoplasm, and *YDR493W* is localized in

mitochondria (Giaever et al., 2002; Huh et al., 2003; Reinders et al., 2006). Our data further implicate these two barely studied genes in cell cycle progression. *TMA64* is another protein of unknown function, which was identified in mass spectrometry-based proteomic screen of yeast ribosomal complexes (Fleischer et al., 2006). *TMA64* associates with ribosomes, has a RNA binding domain and interacts with *RPS4B*, a component of the small (40S) ribosomal subunit (Krogan et al., 2004). Moreover, it has been suggested that there might be a strong connection between ribosomal biogenesis and G1 transit (Bjorklund et al., 2006; Yu et al., 2006). Therefore, G1 arrest phenotype caused by overexpression of *TMA64* suggested its role in ribosomal biogenesis and its potential role in ribosomal biogenesis might interfere with G1 transition.

Nonetheless, 4 transcription factors involved in responding to the environmental stress were observed in the G1 category. Three are transcriptional repressors (*MIG3*, *NCB2*, and *SKO1*). *SKO1* is a basic leucine zipper (bZIP) transcription factor of the ATF/CREB family, involved in osmotic and oxidative stress responses. It forms a complex with Tup1p and Ssn6p to both activate and repress transcription (Nehlin et al., 1992; Pascual-Ahuir et al., 2001a; Pascual-Ahuir et al., 2001b). We observed overproduction of *SKO1* to strongly inhibit cell growth and arrest cells at the G1 phase (Figure 5.7A). Bud size analysis showed that 90% of cells had no bud when *SKO1* was overexpressed (Figure 5.7C). Surprisingly, overproduction of *SKO1* resulted in cell morphology changes that assembled mating yeast cells (shmoos) (Figure 5.7B-III). We reasoned that the elevated expression of *SKO1* might activate the pheromone response pathway either directly or indirectly, causing shmoo formation and a mating-associated G1 arrest. Since Fus1p is a marker protein induced during shmoo formation that localizes to the shmoo tip when the pheromone response pathway is activated (Trueheart et al., 1987), we tested *SKO1* activation of the pheromone response pathway by examining the

localization of Fus1p when *SKO1* was overexpressed. We transformed P_{GAL1}-*SKO1* plasmids into a MATa strain where *FUS1* was C-terminally tagged with GFP. Upon *SKO1* overexpression, Fus1p localized to the shmoo tip (Figure 5.7B-VI), resembling the localization pattern upon alpha factor treatment (Figure 5.7B-IV), demonstrating that the morphological changes are accompanied by general activation of the mating pathway, thus explaining the G1 cell cycle arrest phenotype of *SKO1* ORF strains.



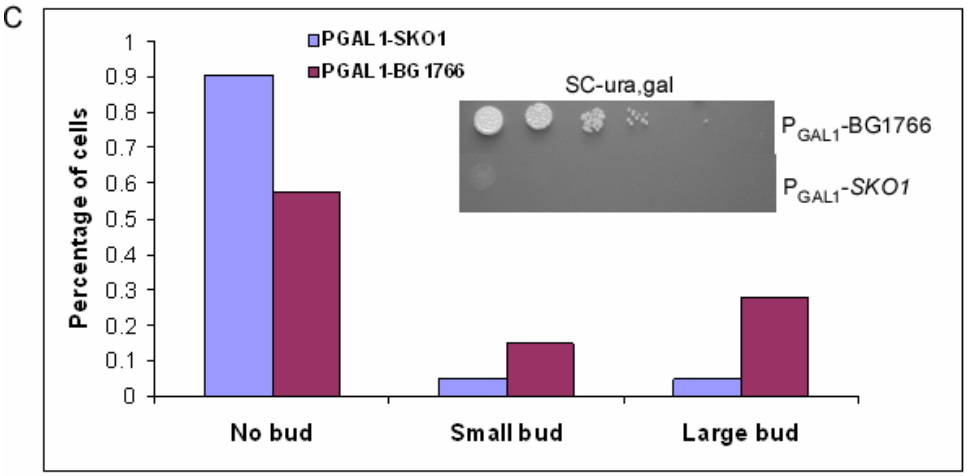


Figure 5.7 Overexpression of *SKO1* activates the pheromone response pathway**.** (A) Flow cytometry analysis of *sko1Δ* strain and pGAL1-*SKO1* strain. (A) Flow cytometry analysis of *sko1Δ* strain and P_{GAL1}-*SKO1* strain. Overproduction of *SKO1* caused a strong arrest at the G1 phase; 78% of P_{GAL1}-*SKO1* cells accumulated at the G1 phase *vs.* 58% of wild type cells at the G1 phase. In contrast, there was no obvious G1 arrest in *sko1Δ* strain. (B) Example images of P_{GAL1}-*SKO1* strain, *sko1Δ* strain, and P_{GAL1}-*SKO1*-Fus1-GFP strain. (I) *sko1Δ* cells: no shmoo; (II) Cells before *SKO1* was overexpressed: no shmoo; (III) Cells after overexpression of *SKO1*: shmoo; (IV) P_{GAL1}-*SKO1*- Fus1-GFP strains treated with alpha factor: shmoo; (V) P_{GAL1}-*SKO1*- Fus1-GFP cells before galactose induction: no shmoo; (VI) P_{GAL1}-*SKO1*- Fus1-GFP cells after galactose induction: shmoo and shmoo-tip localized Fus1p. Nuclei shown in green were overlaid with DIC images. (C) The bud size analysis is shown. 90% of cells had no buds when *SKO1* was overexpressed, while ~50% of wild type cells had no buds. The inserted images show that the overexpression of *SKO1* led to severe growth defect.

5.3.5.2 G2/M category

The G2/M category consisting of 87 genes showed dramatic enrichment in regulation of CDK activity [GO:0000079] (p-value<$9x10^{-7}$), microtubule-based process [GO:0007017] (p-value<$2x10^{-6}$), cell cycle [GO:0007049] (p-value<$4x10^{-6}$), cytoskeleton organization and biogenesis [GO:0007010] (p-value<$8x10^{-6}$), microtubule cytoskeleton organization and biogenesis [GO:0000226] (p-value<$8x10^{-6}$), G2/M transition of mitotic cell cycle [GO:0000086] (p-value<$5x10^{-5}$), DNA replication and chromosome cycle [GO:0000067] (p-value<$5x10^{-5}$), and related systems. It includes *CLB2*, *CLB3*, *CLB5*, *CDC31*, *KAR1*, *SPC97*, *PAC2*, *TUB2*, *NIP100*, *SLK19*, *ASK1*, *AME1*, *MAD2*, and *ACT1*, which have direct roles in regulating G2/M transition and related processes such as microtubule nucleation, chromosome segregation, and mitotic spindle checkpoint control. Additionally, 7 genes identified in the previous large-scale studies (*SPO13*, *SEC17*, *MYO2*, *PRP31*, *ARF1*, *TFG2*, and *SHE1*), although not directly involved in mitotic cell cycle control, have also been observed in the present study. Of 66 newly identified in the screen, 60 were observed to impair cell growth upon induction and the overexpression of 28 genes led to specific sensitivity to nocodazole.

Sub-categorization of genes in the G2/M category by the position of the nuclear DNA

In order to better classify the genes by the nature of their overexpression defects, i.e., as to whether the cells exhibited M phase arrest or chromosome segregation defects lead to G2/M arrest. 3 classes of nuclear morphology were assigned based on the patterns of DNA staining**,** as shown in Figure 5.3 D-F: D, an undivided nucleus in one cell body (class I, pre-M), E, an undivided nucleus in the bud neck (class II, early-M), and F, divided nuclei in two cell bodies (class III, late-M) (Stevenson et al., 2001). In wild type control strains, 60% of the cells exhibited class III nuclear morphology, with chromosomes in these cells successfully segregated, while only 11% of cells showed

class I morphology, and 26% of cells class II morphology. We observed 20 ORF strains to have significantly elevated percentages (95% confidence level) of cells with class I morphology, 13 ORF strains with class II, and 17 ORF strains with class III (Figure 5.8). Among 33 genes in the Class I and II, 10 are known to arrest at the G2/M phase, 52% of the newly indicated Class I and II genes were sensitive to nocodazole (Supplemental Figure 2). For example, Spc97p is a component of the microtubule-nucleating Tub4p (gamma-tubulin) complex and overproduction of *SPC97* causes microtubule defects, which in turn gives rise to a failure of chromosome segregation (Knop et al., 1997). *SPC97* was classified as a class II gene in our screen (Figure 5.3E). We therefore genes causing a similar phenotype to that of *SPC97* and sensitive to nocodazole might play direct or indirect role in chromosome segregation, as shown in Figure 5.8 for genes such as *YPR015C* and *SHE1*.

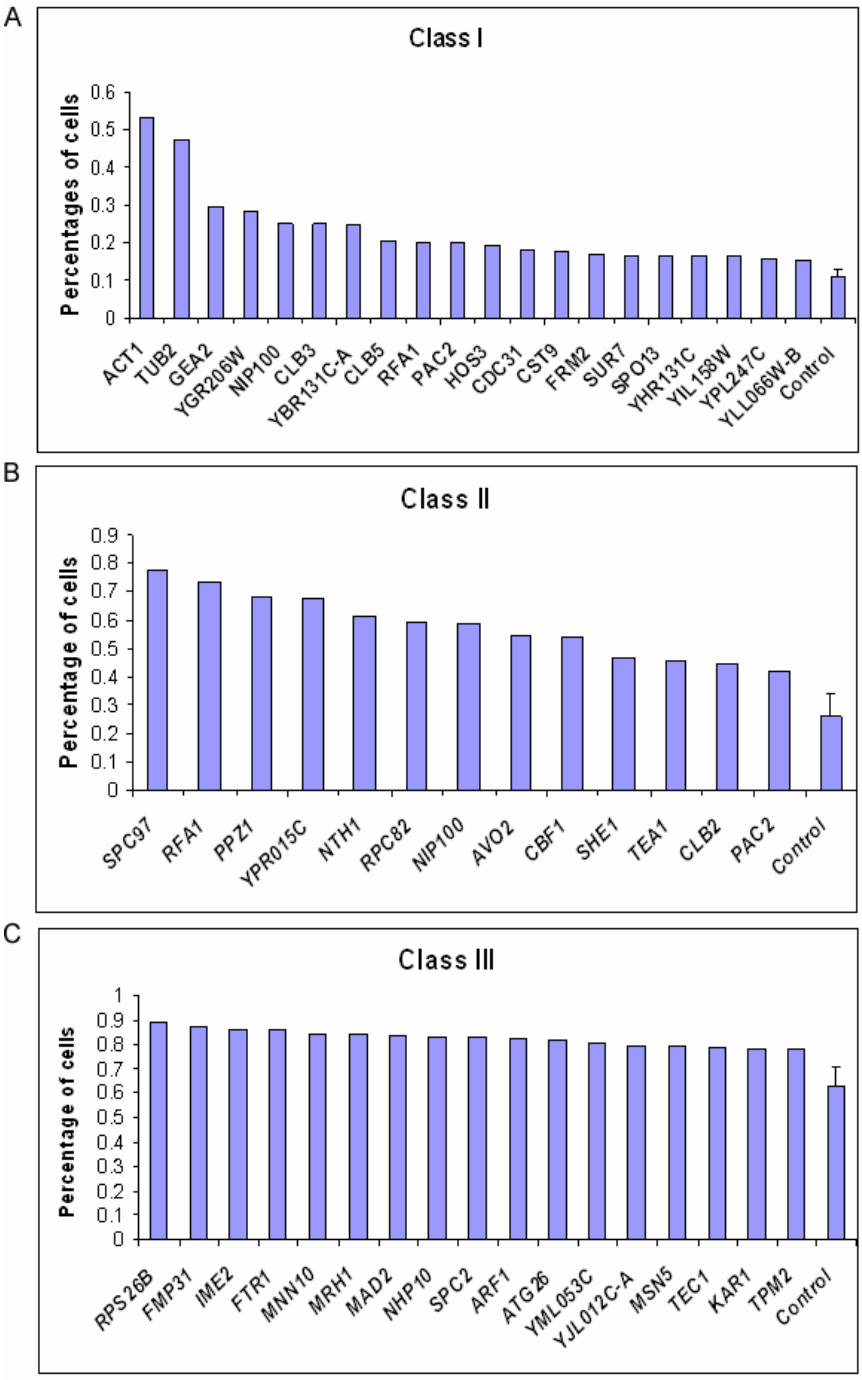


Figure 5.8 Sub-categories of genes in the G2/M category based on nuclear DNA staining in large-budded cells**.** (A) Category I (20 genes): an undivided nucleus in one cell body.

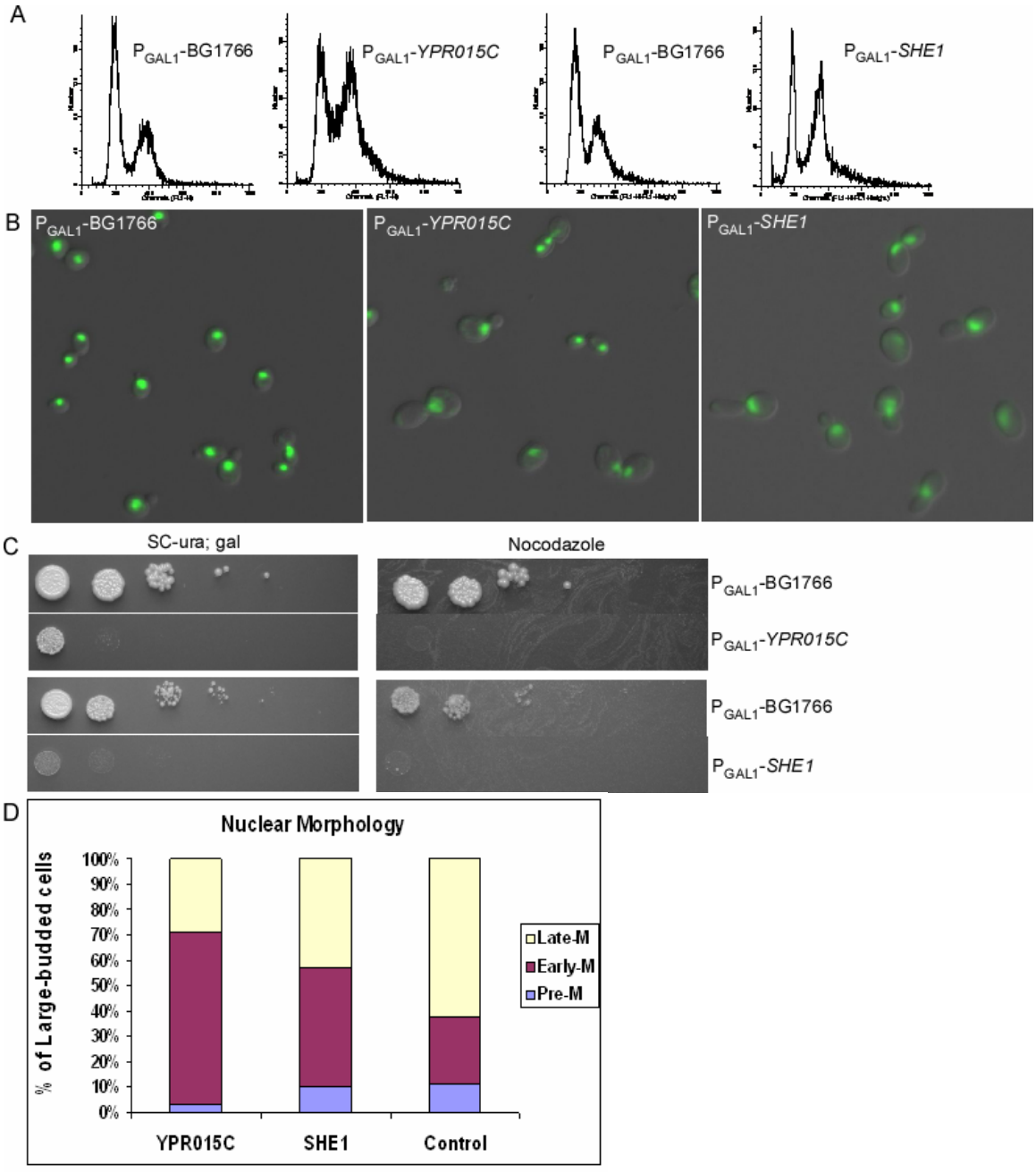
(B) Category II (13 genes): undivided nuclei in bud neck. (C) Category III (17 genes): two divided nuclei separated to two cell bodies.

Functional linkages between new G2/M genes and genes involved in known pathways

To further explore the possible mechanisms underlying the potential cell cycle genes whose overproduction caused defects in cell cycle progression, we analyzed the functional linkages of these genes (Lee et al., 2007) using tools for analyzing local sub-network of genes around each query. From this analysis, two genes, *YPR015C* and *SHE1*, could be further implicated in different processes of cell cycle division (Figure 5.9E). The uncharacterized gene, *YPR015C* exhibited synthetic lethality with *CTF4* (Tong et al., 2004). *CTF4* is a chromatin-associated protein, required for high-fidelity chromosome segregation and sister chromatid cohesion. The ctf4 null mutants increase mitotic recombination, cause G2/M arrest, and have elevated populations of large-budded cells with the nucleus in the bud neck (class II) (Hanna et al., 2001; Miles and Formosa, 1992; Spencer et al., 1990). Since overexpression of *YPR015C* gives rise to a very similar phenotype to that of *CTF4* mutants (Figure 5.9B and D) and *YPR015C* genetically interacts with *CTF4* (Tong et al., 2004), we reasoned that *YPR015C* might function similarly to that of *CTF4*. In support of this, YPR015C overexpression leads to high sensitivity to nocodazole (Figure 5.9C), and a reduction in large-budded cells with completed nuclear DNA segregation (Figure 5.9D). Therefore, overexpression of *YPR015C* apparently induces defects in mitotic chromosome segregation, which in turn arrested cell cycle at the G2/M phase.

SHE1 is another example in this case. *SHE1* is a cytoskeletal protein with unknown functions, except for slow growth and microtubule localization (Espinet et al., 1995; Huh et al., 2003; Sopko et al., 2006). In our study, we found that overexpression of

SHE1 caused specific sensitivity to nocodazole and an elevated percentage of cells with an undivided nucleus (Figure 5.9B and D). Based on yeast functional network, *SHE1* appears linked to proteins involved in microtubule positioning and orientation, actin filament organization, and mRNA localization. Moreover, *SHE1* has been observed to genetically interact with *KIP3*, a kinesin-related motor protein involved in mitotic spindle positioning (Tong et al., 2004). Considering the evidence, the *SHE1* overexpression-induced cell cycle defects is likely due to a malfunction of microtubule organization that in turn leads to defective chromosome segregation.


A
P GAL1 -BG1766
P GAL1 -YPR015C
P GAL1 -BG1766
P GAL1 -SHE1
B
P GAL1 -BG1766
P GAL1 -YPR015C
P GAL1 -SHE1
C
SC-ura; gal
Nocodazole
P GAL1 -BG1766
P GAL1 -YPR015C
P GAL1 -BG1766
P GAL1 -SHE1
D
Nuclear Morphology
% of Large-budded cells
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
YPR015C
SHE1
Control
Late-M
Early-M
Pre-M

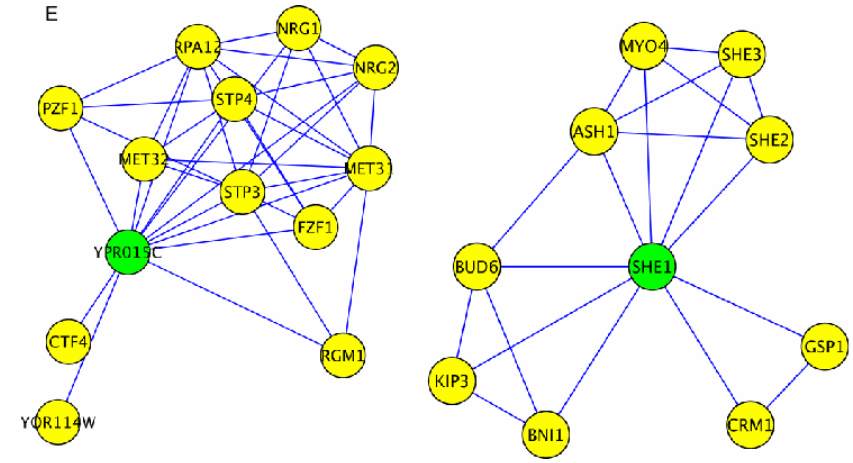


Figure 5.9 Functional analyses of *YPR015C* and *SHE1*. (A) Flow cytometry histograms of P_{GAL1}-*YPR015C* and P_{GAL1}-*SHE1* strains. (B) The overlaid images of P_{GAL1}-*YPR015C* and pGAL1-*SHE1* strains. Nuclei are shown in green, which are superimposed onto DIC images. Compared to wild type cells, these two strains are larger, and both of them have large buds. (C) Growth phenotypes of these two strains are shown. Both strains showed sensitivity to NOC, in addition to the slow growth phenotypes upon galactose induction. (D) The results from nuclear DNA staining are shown in histogram plot. P_{GAL1}-*YPR015C* and P_{GAL1}-*SHE1* strains had higher percentages of cells with undivided nuclei than wild type cells. (E) Genes functionally linked to *YPR015C* and *SHE1* are shown. As a possible transcription factor, *YPR015C* is connected to known transcription factors such as *PZF1*. More importantly, *YPR015C* is linked to *CTF4* whose null mutants show the phenotypes similar to the overexpression of *YPR015C*. *SHE1* is connected to genes involved in actin and microtubule organization, such as *KIP3*.

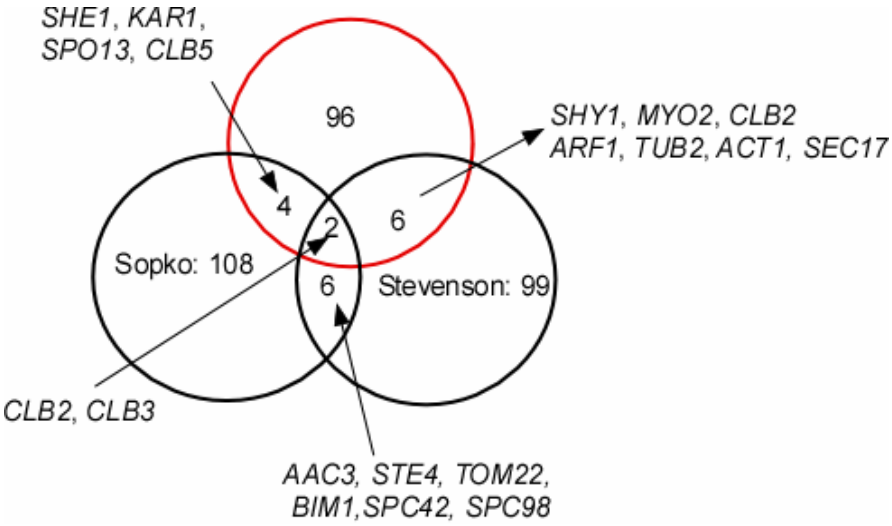


Figure 5.10 Few overlapped genes among large-scale cell cycle screens. The cell cycle gene sets from three large-scale screens are included in this figure: Sopko et al., Stevenson et al., and the gene set from the present study. The gene set from the present study is indicated in red.

5.4 DISCUSSION

In this paper, we describe the comprehensive genome-wide screen for yeast gene functioning in cell cycle progression and report the discovery of new candidates regulating cell cycle progression. The major difference between our screen and the other two large-scale overexpression screens (Sopko et al., 2006; Stevenson et al., 2001) is that we individually examined the effect of overexpression on cell cycle progression for each of ~5,800 yeast ORF, instead of just focusing on the genes whose overexpression causes severe growth defects. Here, we only reported the genes with most significant and reproducible cell cycle defects. The overexpression of 108 genes was observed to lead to cell cycle defects, most of which have not been reported in the previous large-scale screens (Figure 5.10) possibly due to different overexpression conditions and strain

backgrounds, and high false positive rates occurring in the large-scale cell cycle screens (Bjorklund et al., 2006). Among 108 genes, 18 genes have been fully characterized for the roles related to cell cycle progression, 8 were identified in the large-scale cell cycle screens and have not been studied in details (Sopko et al., 2006; Stevenson et al., 2001; Yu et al., 2006), and 82 genes are newly implicated in the cell cycle. In addition to the enriched pathways as described in result section, these genes are involved in a wide range of biological processes, for instance, mitochondria transport, ribosomal biogenesis and tRNA binding (G1 genes), and transcriptional control, cell wall organization, meiosis, vesicle-mediated transport, or DNA binding (G2/M genes).

Several major conclusions can be drawn from the results presented here. First, our screen is the most complete cell cycle screen so far, taking advantage of the availability of yeast ORF collection with recent gene annotation, covering 91% of the complete yeast ORF set (including dubious ORFs) (Gelperin et al., 2005). Our screen is different from the previous overexpression screen that utilized GAL-driven cDNA library and genomic library (Stevenson et al., 2001). The previous screen was not saturated due to the coverage of cDNA library and genomic library. Additionally, some genes from cDNA library or genomic library were not in full length or contiguous with each other, such as *PDS2* is N-terminally truncated (Stevenson et al., 2001). In contrast, Yeast ORF collection is well-constructed using high-efficiency and high-fidelity cloning process, 55% of clones were sequence-confirmed and 95% ORF fusion proteins were examined after transformation into yeast and galactose induction (Gelperin et al., 2005). Moreover, the affinity tags are fused to the C termini of cloned genes, which minimizes the possibility to interfere with proteins normal function, for example, transmembrane and secreted proteins (Gelperin et al., 2005).

A second conclusion is that we utilized an objective approach to determine the significance of DNA content changes caused by gene overexpression. A DNA histogram changes with dye concentration, carbohydrate source, cell size, growth condition, and instrumental settings (Haase and Reed, 2002), which leads to significant sample-to-sample variation even in wild type control cells. Determining which sample strains arrest during the cell cycle by simply comparing its histogram to a single DNA histogram of the control, especially for strains having slight DNA content changes, could introduce high false positives into data analysis. There is very little overlap between gene sets identified in many large-scale cell cycle screens (Figure 5.10), even when the same phenotype was assayed (Bjorklund et al., 2006; Sopko et al., 2006; Stevenson et al., 2001; Yu et al., 2006). This is apparently due to a large number of false positives in large-scale screens (Bjorklund et al., 2006). Therefore, in order to reduce false positives, it is important to know the variance in wild type cells before we compare overexpression strains with wild type strains. Therefore, we estimated sample-to-sample variation in wild type controls by measuring 140 control strains prepared in the same condition as ORF strains. Then, we used 95% confidence level to determine the significant gene sets, which exclude all 140 control strains, but include ORF strains having cell cycle defects as much as possible. The same strategy was also used in the manual follow-up flow cytometry experiments. We were able to recapture 5 of 8 Cdc28p cyclins included in ORF collection and recover 67% of essential genes whose overexpression caused the cell cycle defects similar to those of $TetO_7$ alleles (Yu et al., 2006). Additionally, functional enrichments were consistent with published studies (Bjorklund et al., 2006; Sopko et al., 2006; Stevenson et al., 2001; Yu et al., 2006). All these evidences suggested that the false positive rate was low without introducing a high false negative rate.

The third, we established potential connections between new cell cycle genes and the genes that are involved in different aspects of cell cycle division based on cellular phenotypic assays such as nuclear morphology and sensitivity to specific drugs. Chromosome segregation is one of critical events in cell cycle division. Many proteins play important roles in this process, such as condensins, cohensins, separins, securins, kinetichore proteins, tubulin, motor proteins, and regulatory proteins. When chromosome fails to segregate properly, non-disjunction is occurred. Cells having defects in chromosome segregation are normally large-budded, with a single nucleus in one cell body or in the bud neck (Euskirchen, 2002). Additionally, the sensitivity to microtubule depolymerizing drugs is another indication of non-disjunction (Ouspenski et al., 1999). *SPC97*, *NIP100*, *PAC2*, and *TUB2* are proteins involved in spindle formation and positioning. Disruption of their normal function causes defects in chromosome segregation (Hoyt et al., 1997; Huffaker et al., 1988; Knop et al., 1997). In addition to these well-known proteins that are involved in chromosome segregation, *YPR015C, AVO2, CBF1, TEA1,* and *SHE1* are likely to regulate chromosome segregation either directly or directly because the overproduction of these proteins not only gave rise to elevated populations of large-budded and mononuleate cells, but also caused sensitivity to nocodazole. Moreover, *YPR015C* and *SHE1* are functionally connected to known proteins involved in regulating chromosome segregation and spindle positioning, respectively (Lee et al., 2007).

Finally, we were able to approximately estimate genes that cause cell cycle defects *via* a gain-of-function mechanism based on bud size measurements due to its morphological landmark for the cell cycle stages. Of 12 genes in the G1 category having bud size measurements under overexpression and deletion conditions, 10 genes caused significantly elevated populations of unbudded cells when overexpressed but not deleted.

Of 47 genes in the G2/M category having both bud size measurements either upon overexpression or deletion, 44 genes in the G2/M category caused significantly elevated populations of large-budded cells upon overexpression but not deletion. Therefore, our approximate estimate is that the cell cycle defects caused by these 54 genes could be explained by a gain-of-function mechanism. For example, overexpression of *SKO1* caused a strong cell cycle delay at the G1 phase, but the deletion of *SKO1* had no detectable phenotype. Surprisingly, we found that overexpression of *SKO1* gave rise to a shmoo morphology and resulted in localization of the mating pathway marker *FUS1* to the shmoo tip. However, genes involved in the pheromone response pathway do not typically appear to be regulated by *SKO1* at least as measured by chromatin-immunoprecipitation of *SKO1* under normal culture conditions (Lee et al., 2002). Therefore, our results suggest that *SKO1* may regulate genes in the pheromone response pathway only when *SKO1* is overexpressed.

In summary, our high-throughput overexpression analysis of yeast cell cycle progression complements previous cell cycle experiments and has resulted with identification of 108 genes whose overexpression results in cell cycle progression defects, which is largely a novel up-regulated effect. A large number of new genes identified in the screen have not been previously linked to the cell cycle, and finally these results lay the foundation for future experiments to elucidate the precise roles of these genes in cell cycle progression. Furthermore, since overexpression has been most efficiently used in human cell culture and regulation of cell proliferation is an important aspect of studying human diseases, we anticipate that a similar effort in human cell lines will accelerate our understanding of cell cycle control in mammalian systems, which perhaps ultimately help to understand the connections between the cell cycle control and cancers.

Authors' contribution

The author carried out the initial cell cycle screen, all the follow-up experiments. Zhihua Li also participated in the initial cell cycle screen.

Table 5.1 Yeast ORF strains having cell cycle defects.

21 ORF strains are in the G1 category with the measurements of budding size. The measurements are represented in the percentages of cells with no budded, small bud, and large bud.

Systematic name	Standard name	% unbudded cells	% small budded cells	% large budded cells
YLR052W	IES3	0.78	0.01	0.21
YOR131C	YOR131C	0.81	0.06	0.13
YHL001W	RPL14B	0.74	0.15	0.11
YER028C	MIG3	0.67	0.14	0.2
YHR174W	ENO2	0.74	0.13	0.13
YDR117C	TMA64	0.77	0.08	0.15
YGR112W	SHY1	0.74	0.06	0.23
YDR156W	RPA14	0.76	0.01	0.23
YCR046C	IMG1	0.72	0.08	0.21
YDR493W	FMP36	0.79	0	0.21
YOR065W	CYT1	0.79	0	0.21
YPR152C	YPR152C	0.72	0.12	0.16
YDR397C	NCB2	0.8	0.07	0.13
YIR013C	GAT4	0.69	0.03	0.28
YHL031C	GOS1	0.77	0.02	0.21
YPL127C	HHO1	Not included	Not included	Not included
YHR070W	TRM5	0.92	0.06	0.02
YNL167C	SKO1	0.91	0.05	0.05
YMR275C	BUL1	0.97	0	0.03
YGL105W	ARC1	0.81	0.08	0.11
YLL066W-B	YLL066W-B	0.73	0.09	0.17
WT_control	WT_control	0.57	0.15	0.28

97 ORF strains are in the G2/M category with the measurements of bud size and nuclear morphology. The measurements are represented in the percentages of cells with no budded, small bud, and large bud, as well as the percentages of cells in three categories of nuclear morphology. Three categories of nuclear morphology are:
Class I: an undivided nucleus in one cell body
Class II: undivided nuclei in bud neck
Class III: divided nuclei separated into two cell bodies

Systematic name	Standard name	% unbudded cells	% small budded cells	% large budded cells	% class I	% class II	% class III
YKL052C	ASK1	0.67	0.11	0.21	0.11	0.14	0.75
YBR131C-A	YBR131C-A	0.53	0.16	0.31	0.24	0.22	0.53
YOR257W	CDC31	0.59	0.13	0.28	0.18	0.13	0.69

YCR093W	CDC39	0.33	0.12	0.55	0.12	0.17	0.71
YGR206W	YGR206W	0.57	0.07	0.37	0.29	0.09	0.63
YML055W	SPC2	0.53	0.05	0.42	0.1	0.07	0.83
YHR172W	SPC97	0.33	0.05	0.62	0.14	0.77	0.09
YIL138C	TPM2	0.52	0.05	0.43	0.04	0.18	0.78
YBL050W	SEC17	0.52	0.11	0.37	0.13	0.26	0.61
YOR326W	MYO2	0.57	0.12	0.31	0.15	0.22	0.63
YNL264C	PDR17	0.38	0.11	0.51	0.13	0.27	0.6
YDR277C	MTH1	0.41	0.09	0.5	0.04	0.38	0.58
YLR123C	YLR123C	0.6	0.1	0.3	0.13	0.2	0.67
YML052W	SUR7	0.39	0.18	0.43	0.17	0.26	0.57
YHR014W	SPO13	0.53	0.05	0.42	0.17	0.27	0.56
YHR002W	LEU5	0.52	0.06	0.42	0.07	0.24	0.69
YLR394W	CST9	0.52	0.09	0.39	0.18	0.28	0.54
YMR199W	CLN1	NA	NA	NA			
YJR060W	CBF1	0.44	0.05	0.51	0.13	0.54	0.33
YCL026C-A	FRM2	0.51	0.09	0.4	0.17	0.4	0.43
YPR015C	YPR015C	0.41	0.11	0.48	0.03	0.68	0.29
YOR286W	FMP31	0.5	0.14	0.36	0.03	0.09	0.88
YGR091W	PRP31	0.34	0.16	0.49	0.12	0.16	0.72
YDL002C	NHP10	0.51	0.09	0.4	0.03	0.13	0.83
YJL077W-A	YJL077W-A	0.55	0.06	0.39	0.03	0.21	0.76
YML007W	YAP1	0.35	0.17	0.48	0.11	0.36	0.53
YER145C	FTR1	0.48	0.13	0.39	0.05	0.09	0.86
YLR149C	YLR149C	0.61	0.1	0.29	0.04	0.25	0.71
YJL012C	VTC4	0.5	0.06	0.44	0.09	0.11	0.8
YLR341W	SPO77	0.49	0.07	0.44	0.07	0.27	0.66
YNL188W	KAR1	0.54	0.02	0.43	0.1	0.12	0.79
YOR195W	SLK19	0.45	0.08	0.47	0.03	0.39	0.58
YGR109C	CLB6	0.44	0.17	0.39	0.08	0.35	0.58
YBR211C	AME1	0.49	0.1	0.41	0.05	0.3	0.65
YDR245W	MNN10	0.44	0.11	0.45	0	0.16	0.84
YDR033W	MRH1	0.4	0.18	0.42	0.03	0.13	0.84
YJL030W	MAD2	0.54	0.1	0.36	0.02	0.14	0.84
YDR091C	RLI1	0.41	0.06	0.53	0.04	0.34	0.62
YIR001C	SGN1	0.32	0.08	0.6	0.06	0.29	0.64
YKL078W	DHR2	0.45	0.11	0.43	0.02	0.22	0.76
YPR190C	RPC82	0.51	0.14	0.35	0.02	0.59	0.39
YDR266C	YDR266C	0.45	0.07	0.48	0.04	0.25	0.71
YDL214C	PRR2	0.5	0.09	0.42	0.08	0.15	0.78
YDR001C	NTH1	0.44	0.1	0.45	0.05	0.61	0.33
YIR016W	YIR016W	0.41	0.08	0.51	0	0.23	0.77
YBR083W	TEC1	0.29	0.22	0.49	0.02	0.19	0.79

YKR067W	GPT2	0.54	0.15	0.31	0.07	0.18	0.75
YHR131C	YHR131C	0.53	0.08	0.39	0.17	0.13	0.71
YFL022C	FRS2	0.46	0.1	0.44	0.07	0.23	0.7
YDR143C	SAN1	0.4	0.1	0.5	0.12	0.18	0.71
YPL174C	NIP100	0.26	0.05	0.69	0.25	0.59	0.16
YOR002W	ALG6	0.45	0.11	0.44	0.11	0.21	0.68
YNL283C	WSC2	0.35	0.17	0.49	0.13	0.13	0.73
YPL247C	YPL247C	0.39	0.11	0.49	0.16	0.27	0.58
YJL031C	BET4	0.45	0.18	0.37	0.11	0.36	0.52
YPR119W	CLB2	0.34	0.09	0.56	0.08	0.45	0.47
YML053C	YML053C	0.56	0.14	0.3	0.09	0.11	0.81
YJL106W	IME2	0.41	0.13	0.46	0.1	0.04	0.86
YPR120C	CLB5	0.3	0.09	0.62	0.21	0.31	0.49
YML016C	PPZ1	0.29	0.03	0.67	0.11	0.68	0.21
YLR189C	ATG26	0.38	0.11	0.51	0.04	0.14	0.82
YER007W	PAC2	0.39	0.12	0.49	0.2	0.42	0.38
YEL022W	GEA2	0.41	0.05	0.54	0.29	0.35	0.36
YDL192W	ARF1	0.36	0.12	0.52	0.05	0.12	0.83
YDR335W	MSN5	0.41	0.13	0.45	0.02	0.18	0.8
YKR029C	SET3	0.33	0.06	0.62	0.07	0.22	0.72
YGR094W	VAS1	0.49	0.12	0.39	0.07	0.17	0.76
YBL031W	SHE1	0.37	0.14	0.49	0.1	0.47	0.43
YGR005C	TFG2	0.51	0.07	0.42	0.12	0.18	0.71
YGR109W-A	YGR109W-A	0.36	0.17	0.48	0.11	0.18	0.71
YPL116W	HOS3	0.29	0.06	0.65	0.19	0.19	0.61
YAR007C	RFA1	0.13	0.04	0.82	0.2	0.73	0.06
YDL093W	PMT5	0.42	0.12	0.46	0.1	0.21	0.69
YOL063C	CRT10	0.47	0.11	0.42	0.11	0.17	0.71
YJR125C	ENT3	0.36	0.12	0.53	0.05	0.36	0.59
YIL036W	CST6	0.34	0.15	0.51	0.07	0.18	0.76
YOR337W	TEA1	0.26	0.08	0.66	0.12	0.46	0.42
YOR007C	SGT2	0.54	0.13	0.33	0	0.22	0.78
YJL128C	PBS2	0.42	0.13	0.44	0.09	0.18	0.73
YIL158W	YIL158W	0.36	0.1	0.54	0.16	0.25	0.59
YHR177W	YHR177W	0.42	0.16	0.43	0.02	0.23	0.74
YGL066W	SGF73	0.42	0.12	0.45	0.11	0.22	0.67
YMR068W	AVO2	0.43	0.09	0.49	0.13	0.55	0.32
YER131W	RPS26B	0.52	0.12	0.36	0.04	0.06	0.89
YFL037W	TUB2	0.24	0.04	0.72	0.47	0.34	0.19
YDL155W	CLB3	0.28	0.06	0.66	0.25	0.29	0.46
YPR120C	CLB5	0.3	0.09	0.62	0.21	0.31	0.49
YFL039C	ACT1	0.27	0.1	0.63	0.53	0.27	0.2
WT_control	WT_control	0.57	0.15	0.28	0.11	0.26	0.6

Table 5.2 Summary of 108 strains with cell cycle defects: * genes in the G1 category

21 essential genes	26 known cdc genes	17 transcription factors	25 uncharacterized ORFs	48 others
*NCB2	*SHY1	*SKO1	*YLL066W-B	*HHO1
*TRM5	*TRM5	*GAT4	*YOR131C	*BUL1
ASK1	ASK1	*NCB2	*YPR152C	*GOS1
CDC31	CDC31	*MIG3	*TMA64	*RPL14B
CDC39	CDC39	TEA1	*YMR082C	*RPA14
SPC97	SPC97	TFG2	*FMP36	*ENO2
SEC17	SEC17	CDC39	YGR206W	*ARC1
MYO2	MYO2	CST6	YML053C	*CYT1
PRP31	SPO13	YPR015C	YPR015C	*IES3
KAR1	CLN1	CBF1	YHR131C	*IMG1
AME1	PRP31	TEC1	YLR123C	RFA1
RLI1	KAR1	CRT10	FRM2	HOS3
DHR2	SLK19	YDR266C	YIL158W	MRH1
RPC82	CLB6	YGR109W-A	YHR177W	SGN1
FRS2	AME1	MTH1	YDR266C	IME2
BET4	MAD2	YAP1	YGR109W-A	MNN10
VAS1	NIP100	SGF73	YIR016W	BET4
TFG2	CLB2		YLR149C	NTH1
RFA1	CLB5		FMP31	GEA2
TUB2	PAC2		YPL247C	GPT2
ACT1	ARF1		YBR131C-A	RPS26B
	SHE1		SHE1	SPC2
	TFG2		YJL077W-A	SPO77
	CLB3		FMP31	LEU5
	TUB2		YDL159W-A	TPM2
	ACT1			FRS2
	ASK1			RPC82
	CDC31			PMT5
				VTC4
				ALG6
				SGT2
				SET3
				VAS1
				SAN1
				AVO2
				ATG26
				ENT3
				SEC17
				PRR2
				FTR1
				NHP10

				PBS2
				WSC2
				MSN5
				PDR17
				CST9
				SUR7
				DHR2

Table 5.3 Comparison between gain-of-function and loss-of-function.

Genes in the G1 category

Overexpression only	SCMD only	Overexpression ∩ SCMD	Genes not included in SCMD
BUL1	SHY1	FMP36	TRM5
SKO1			NCB2
YOR131C			ENO2
ARC1			
CYT1			
IES3			
TMA64			
GOS1			
RPA14			

Genes in the G2/M category

Overexpression only	SCMD only	Overexpression ∩ SCMD	Genes not included in SCMD
NIP100		PDR17	TUB2
PPZ1		MNN10	ACT1
TEA1		ARF1	SPC97
CLB3			CDC39
HOS3			GEA2
SET3			RLI1
CLB5			PRP31
SGN1			YGR109W-A
CLB2			VTC4
YIL158W			
ENT3			
ARF1			
YIR016W			
CST6			
CBF1			
ATG26			
SAN1			
MTH1			

TEC1			
YPL247C			
SHE1			
PAC2			
AVO2			
WSC2			
YAP1			
YPR015C			
YDR266C			
SLK19			
IME2			
PMT5			
MSN5			
NTH1			
SGF73			
ALG6			
PBS2			

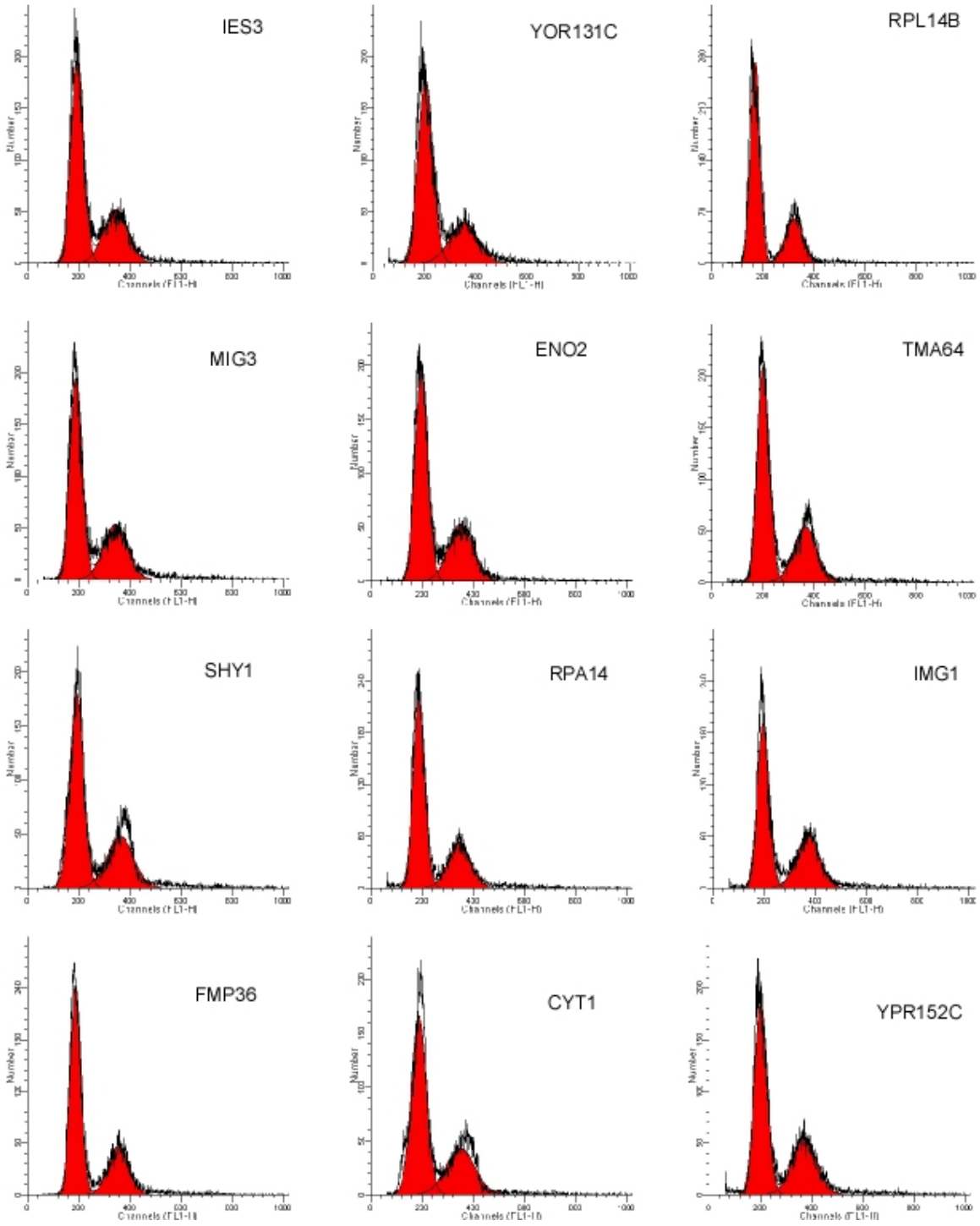
Table 5.4 Genes that caused slow growth and sensitivity to drugs upon induction.
Strains having cell cycle defects were tested for the sensitivity to nocodazole (NOC: 15μg/ml) and hydroxyurea (HU: 50μM).
Genes in the G1 category:

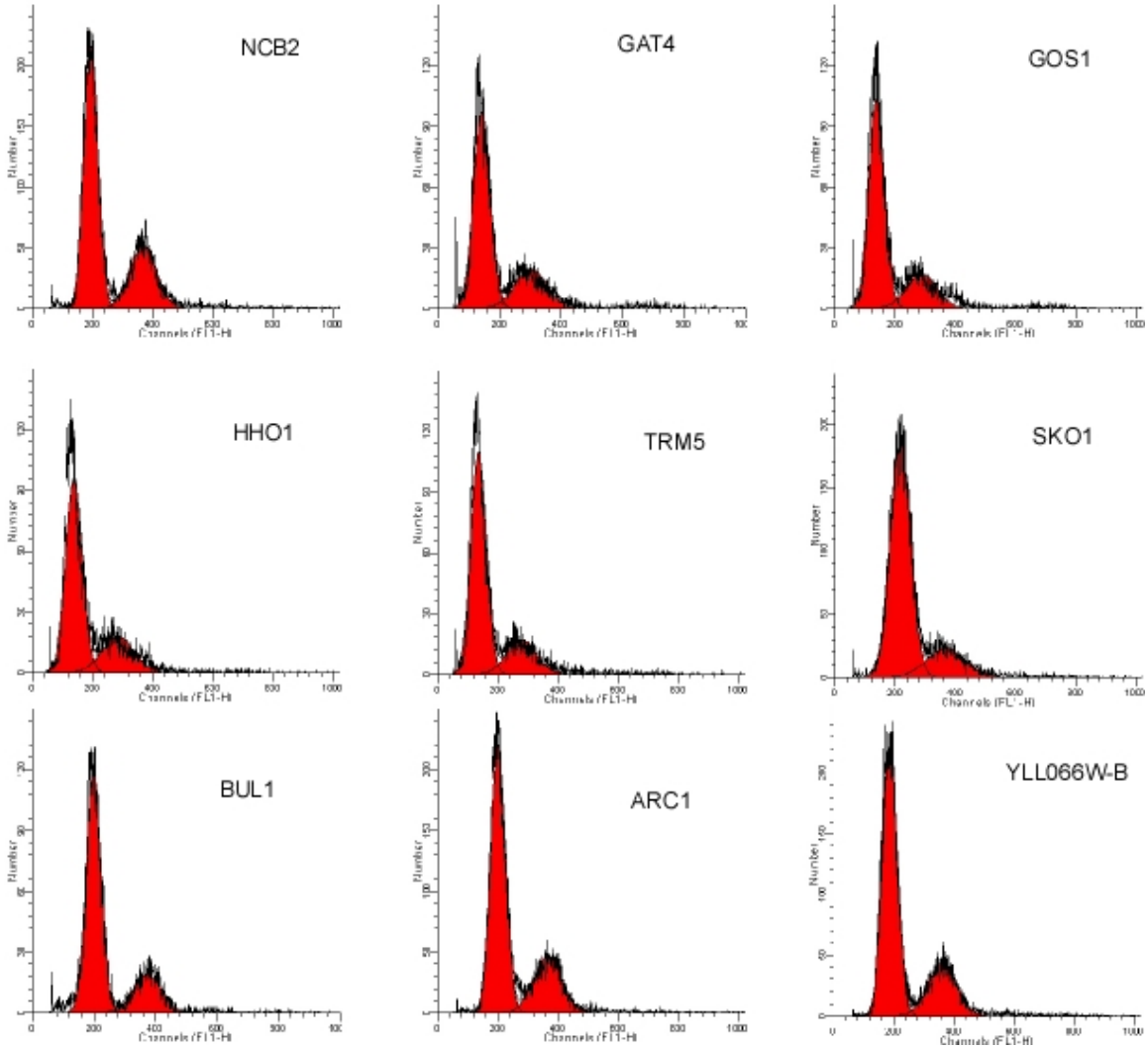
SC_GAL	SC_GAL_NOC	SC_GAL_HU	NOC and HU
IES3	FMP36	YOR131C	YHR131C
YOR131C		YPR152C	MIG3
RPL14B		GOS1	GAT4
MIG3			
ENO2			
CYT1			
YPR152C			
NCB2			
GAT4			
GOS1			
HHO1			
TRM5			
SKO1			
BUL1			
ARC1			
TMA64			
IES3			
YOR131C			
RPL14B			

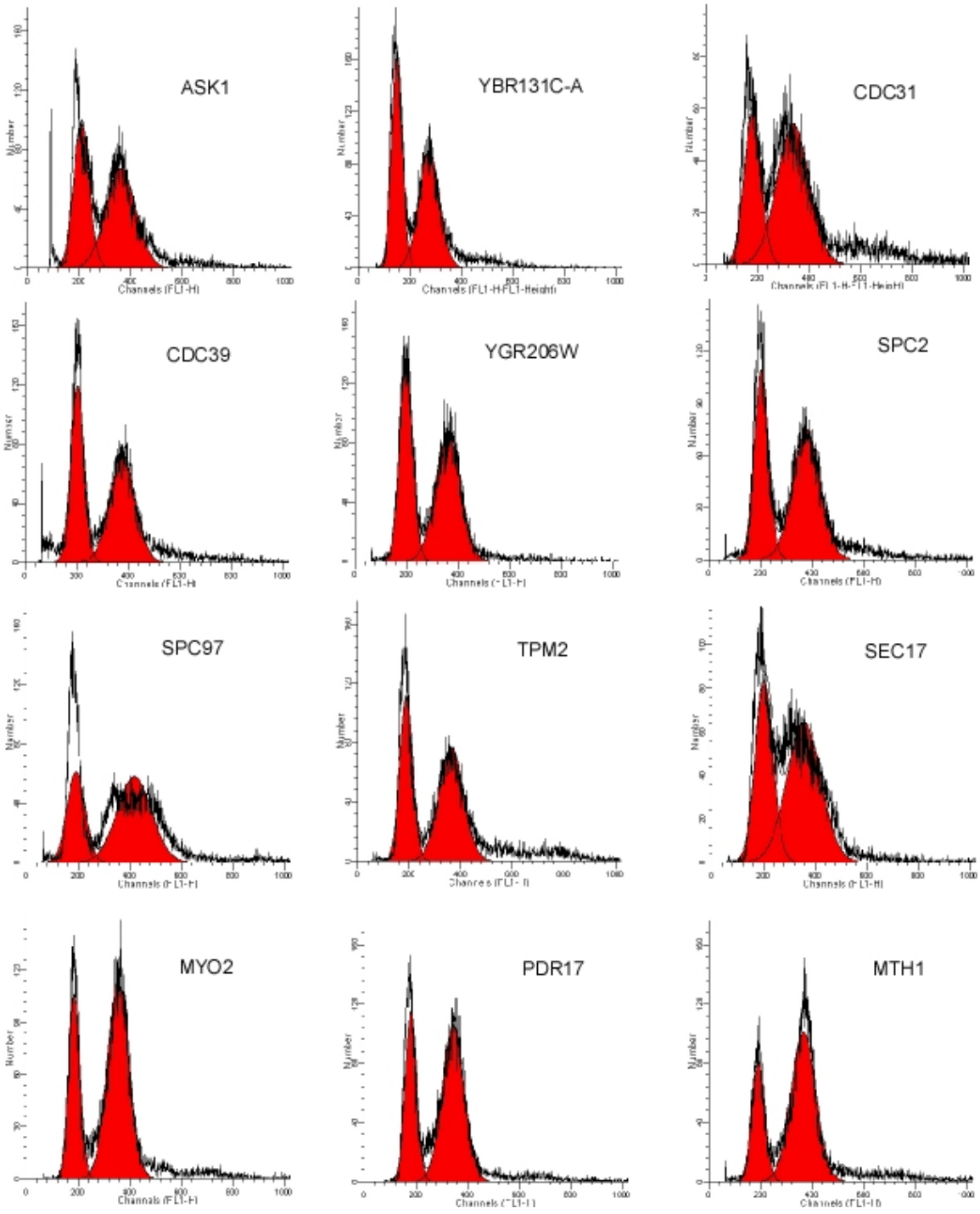
Gene in the G2/M category:

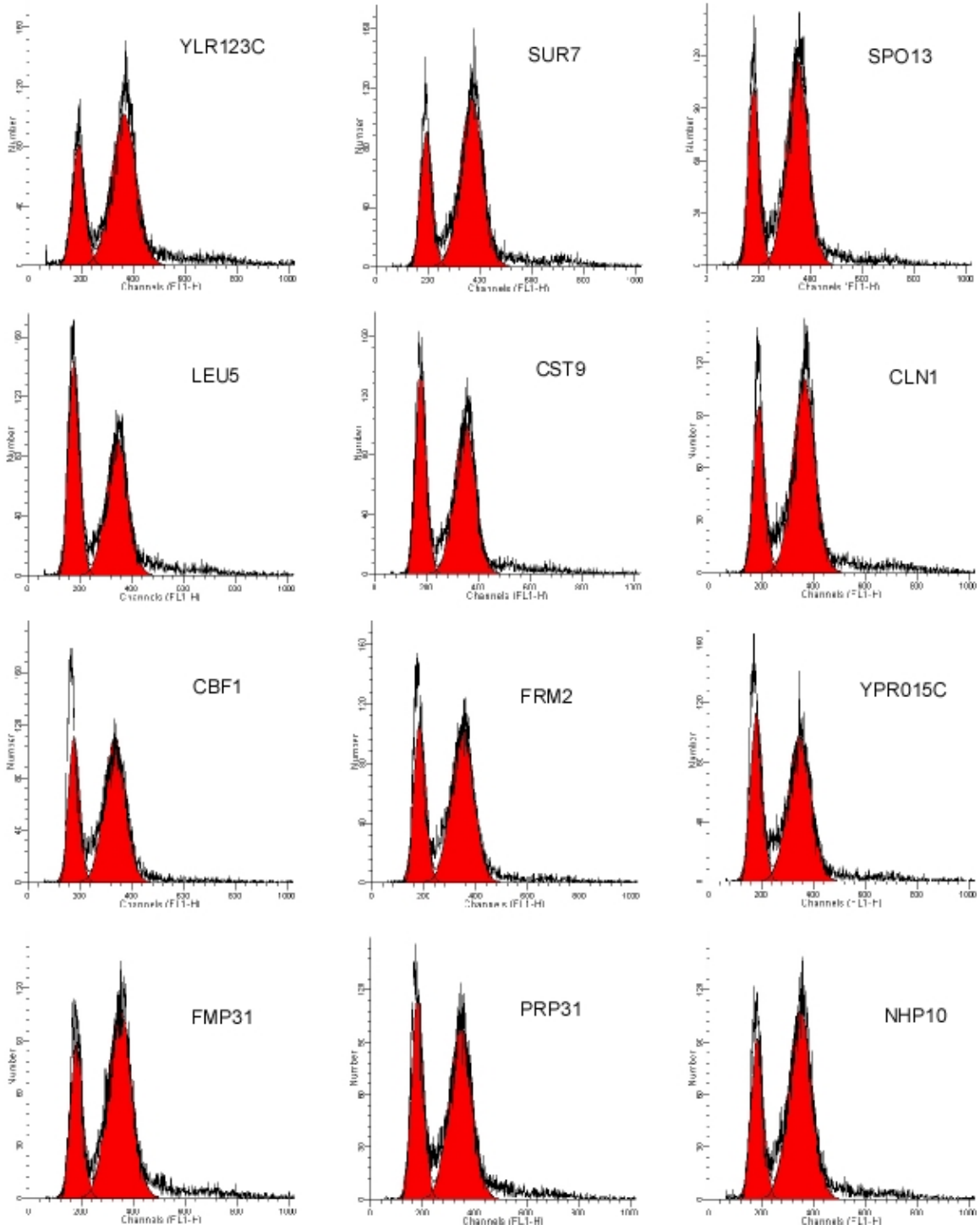
SC_GAL	SC_GAL_NOC	SC_GAL_HU	Resistant to HU	NOC and HU
SEC17	YML053C	LEU5	FRM2	YPL247C
PDR17	YPR015C	FRS2	YIR016W	RFA1
MTH1	YLR149C	WSC2	PMT5	MNN10
SUR7	MRH1	SGF73	FTR1	BET4
CST9	SHE1			VTC4
CBF1	HOS3			SET3
YPR015C	SGN1			SGT2
FMP31	IME2			ALG6
NHP10	CST6			CST9
FTR1	TUB2			SUR7
VTC4	PAC2			
SPO77	PRP31			
MNN10	GEA2			
MRH1	SPO77			
SGN1	RLI1			
RPC82	VAS1			
YDR266C	AVO2			
PRR2	ATG26			
NTH1	FMP31			
TEC1	FTR1			
YHR131C	PBS2			
FRS2	PDR17			
ALG6	TEA1			
WSC2	TFG2			
YPL247C	CDC39			
BET4	CBF1			
YML053C	MTH1			
IME2	SEC17			
ATG26				
GEA2				
ARF1				
SET3				
VAS1				
SHE1				
TFG2				
RFA1				
PMT5				
CRT10				
ENT3				
CST6				
TEA1				

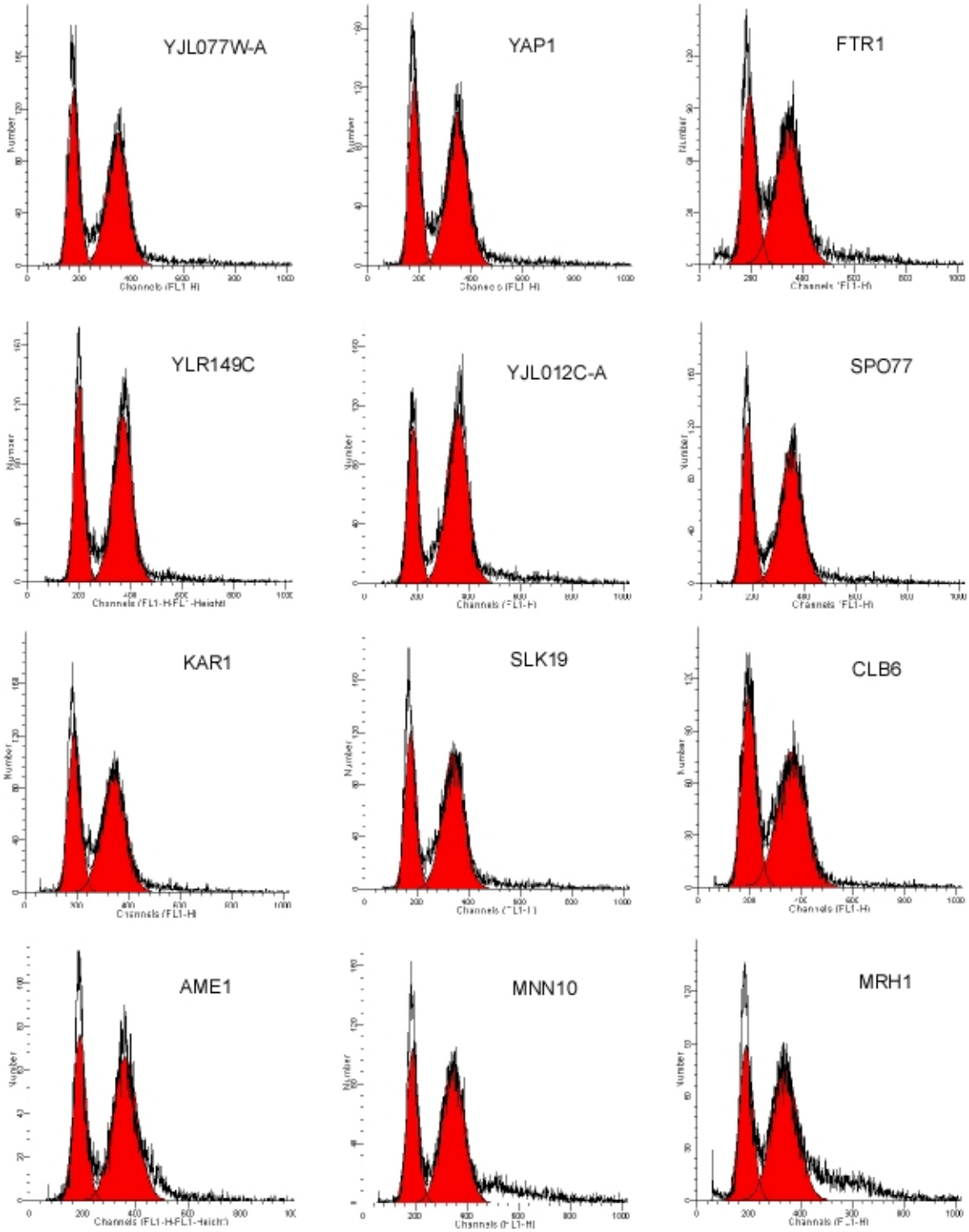
SGT2				
PBS2				
YHR177W				
SGF73				
AVO2				
RPS26B				
YAP1				
RPL14B				
PRR2				

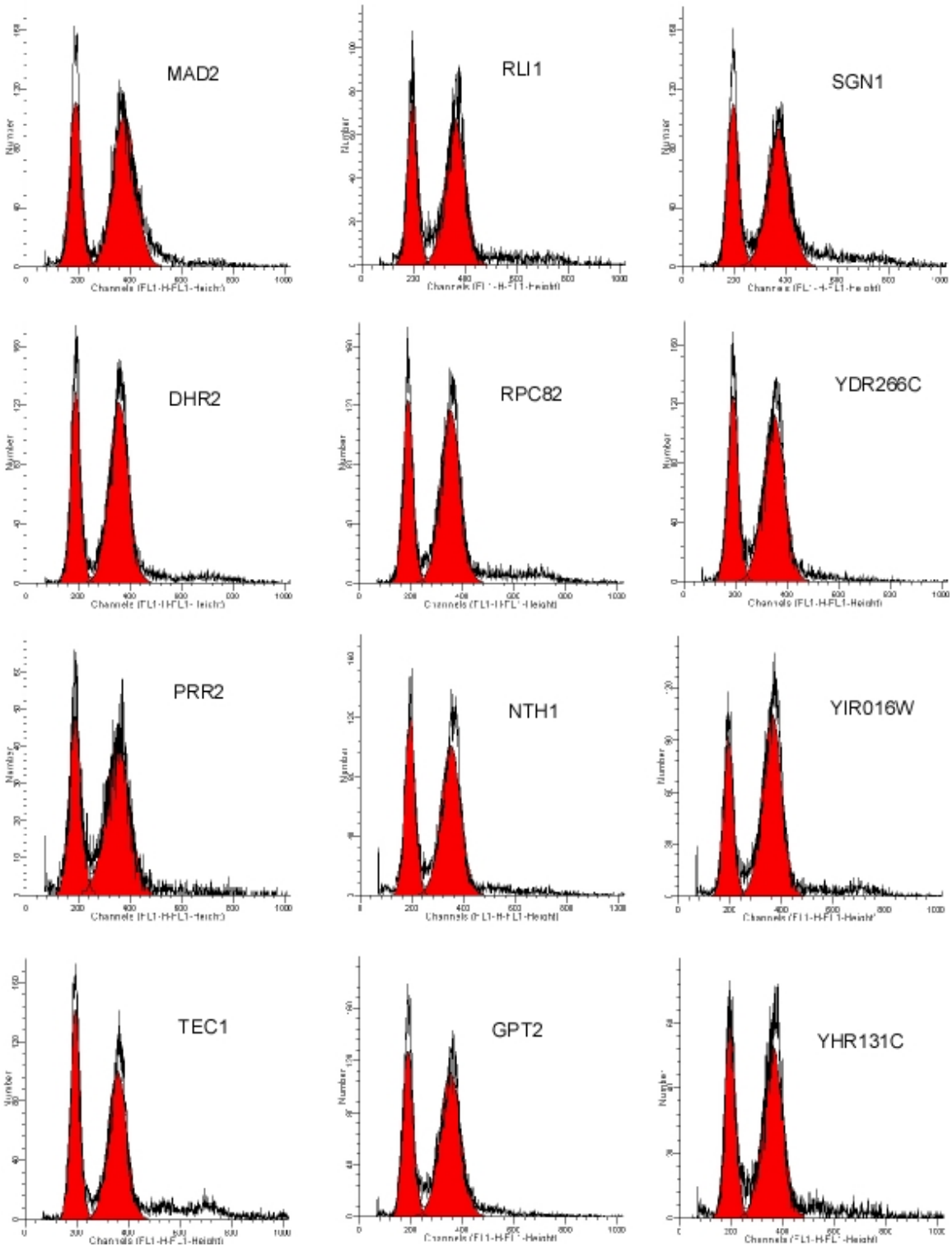

IES3
YOR131C
RPL14B
MIG3
ENO2
TMA64
SHY1
RPA14
IMG1
FMP36
CYT1
YPR152C

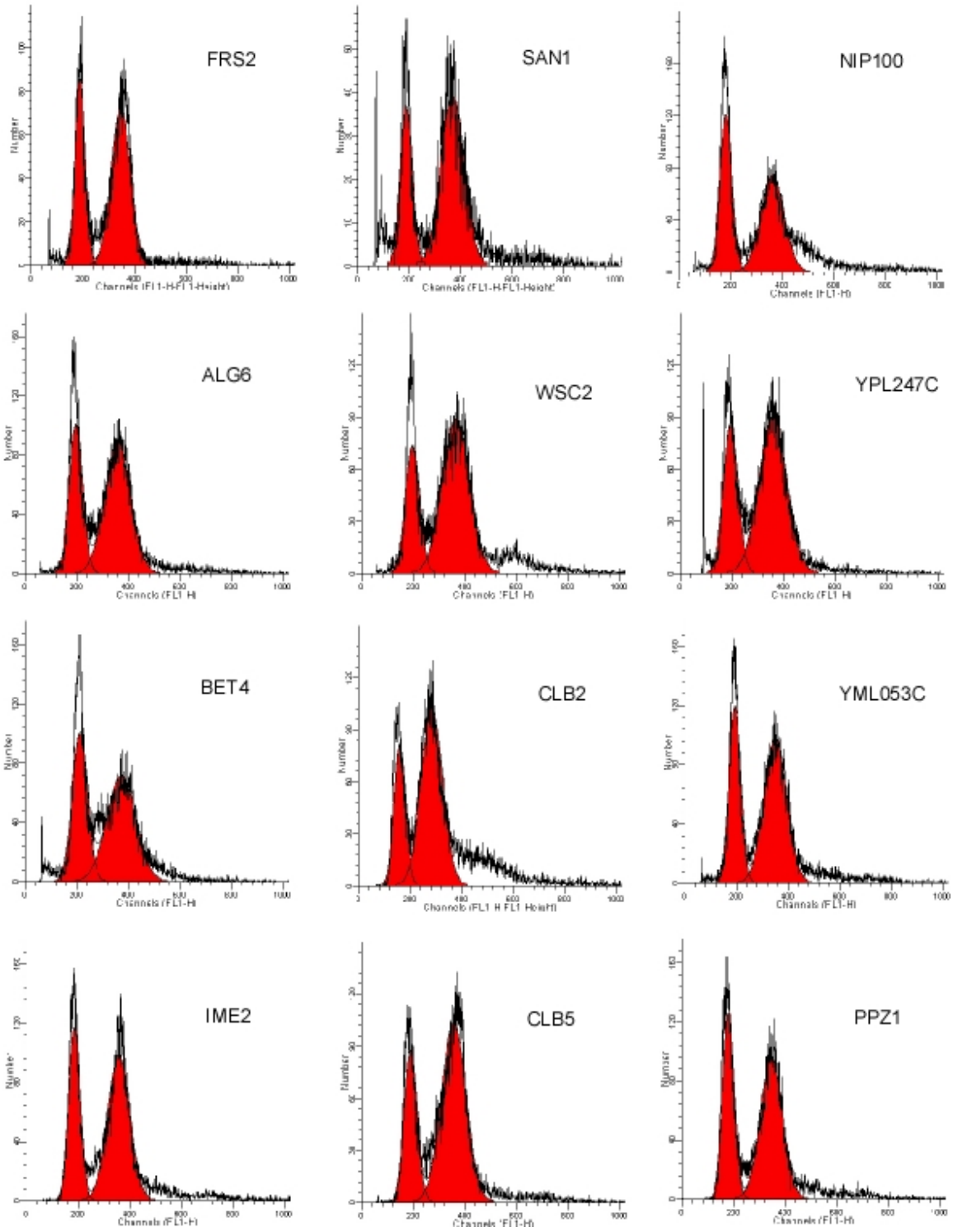

NCB2
GAT4
GOS1
HHO1
TRM5
SKO1
BUL1
ARC1
YLL066W-B

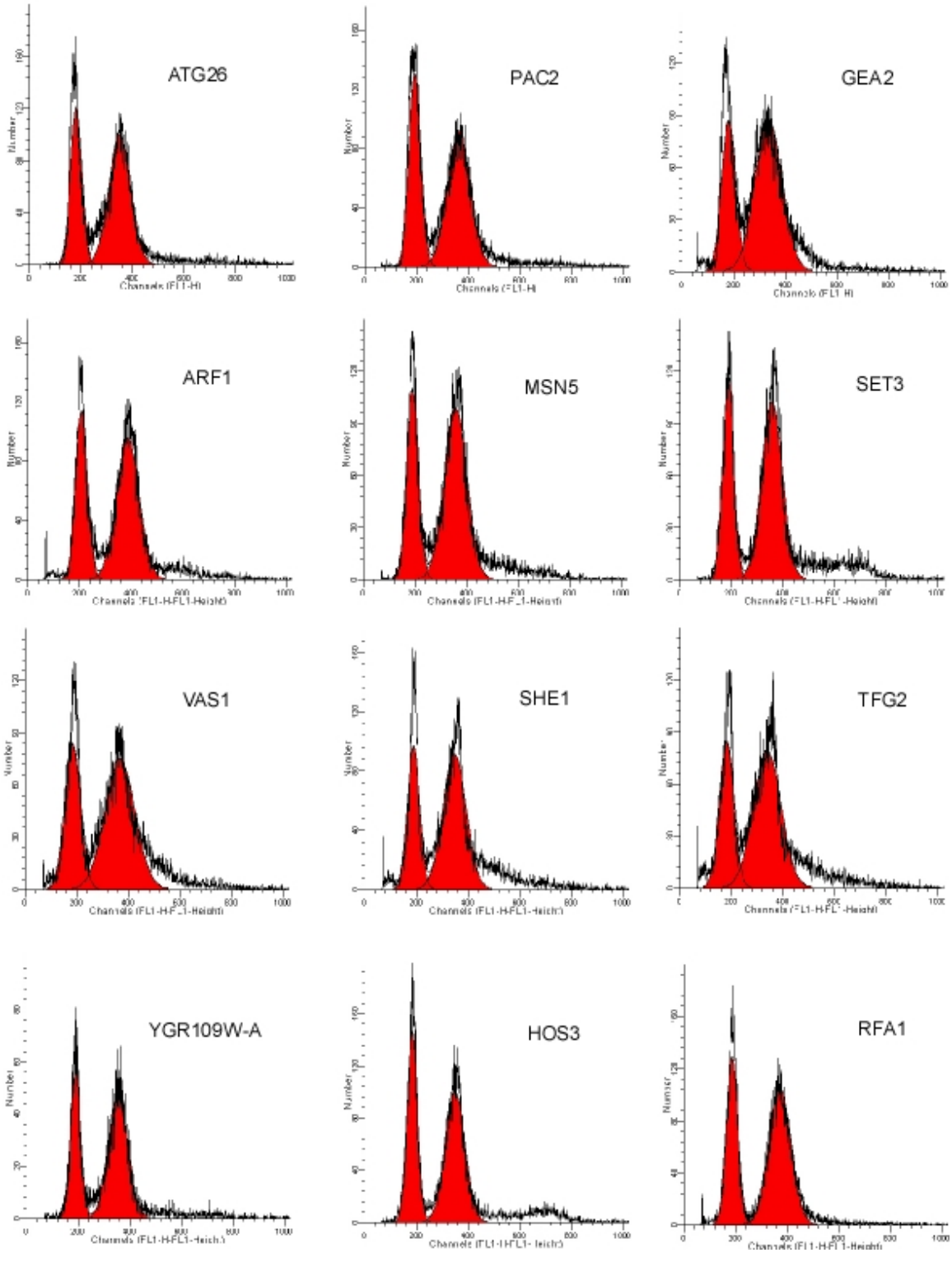

ASK1
YBR131C-A
CDC31
CDC39
YGR206W
SPC2
SPC97
TPM2
SEC17
MYO2
PDR17
MTH1

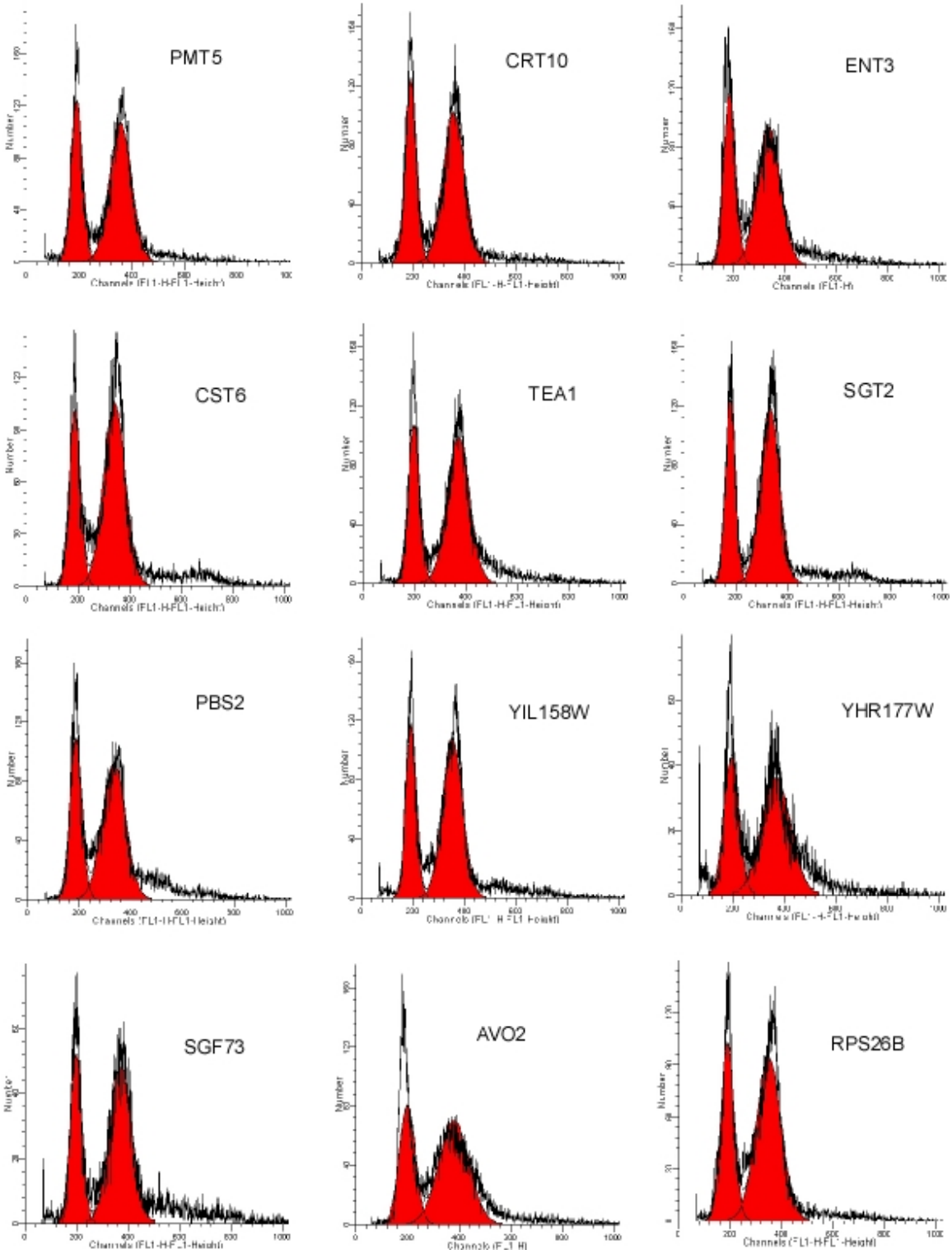

YLR123C
SUR7
SPO13
LEU5
CST9
CLN1
CBF1
FRM2
YPR015C
FMP31
PRP31
NHP10

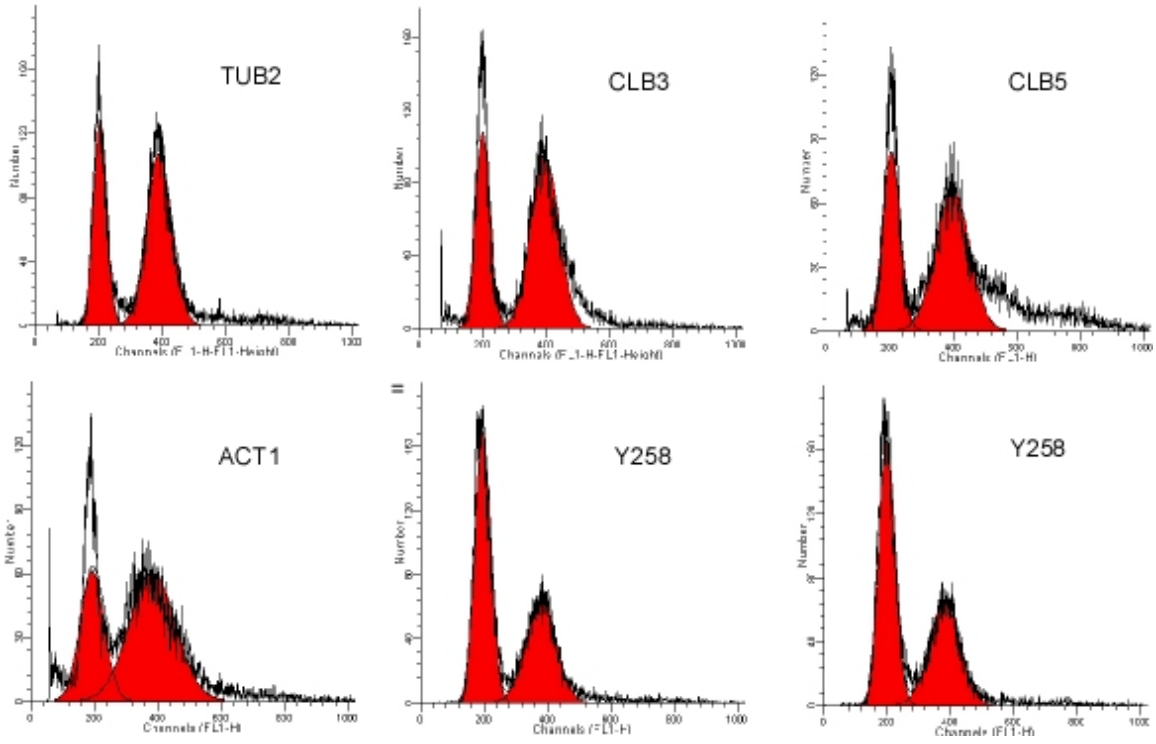

YJL077W-A
YAP1
FTR1
YLR149C
YJL012C-A
SPO77
KAR1
SLK19
CLB6
AME1
MNN10
MRH1


MAD2
RLI1
SGN1
DHR2
RPC82
YDR266C
PRR2
NTH1
YIR016W
TEC1
GPT2
YHR131C

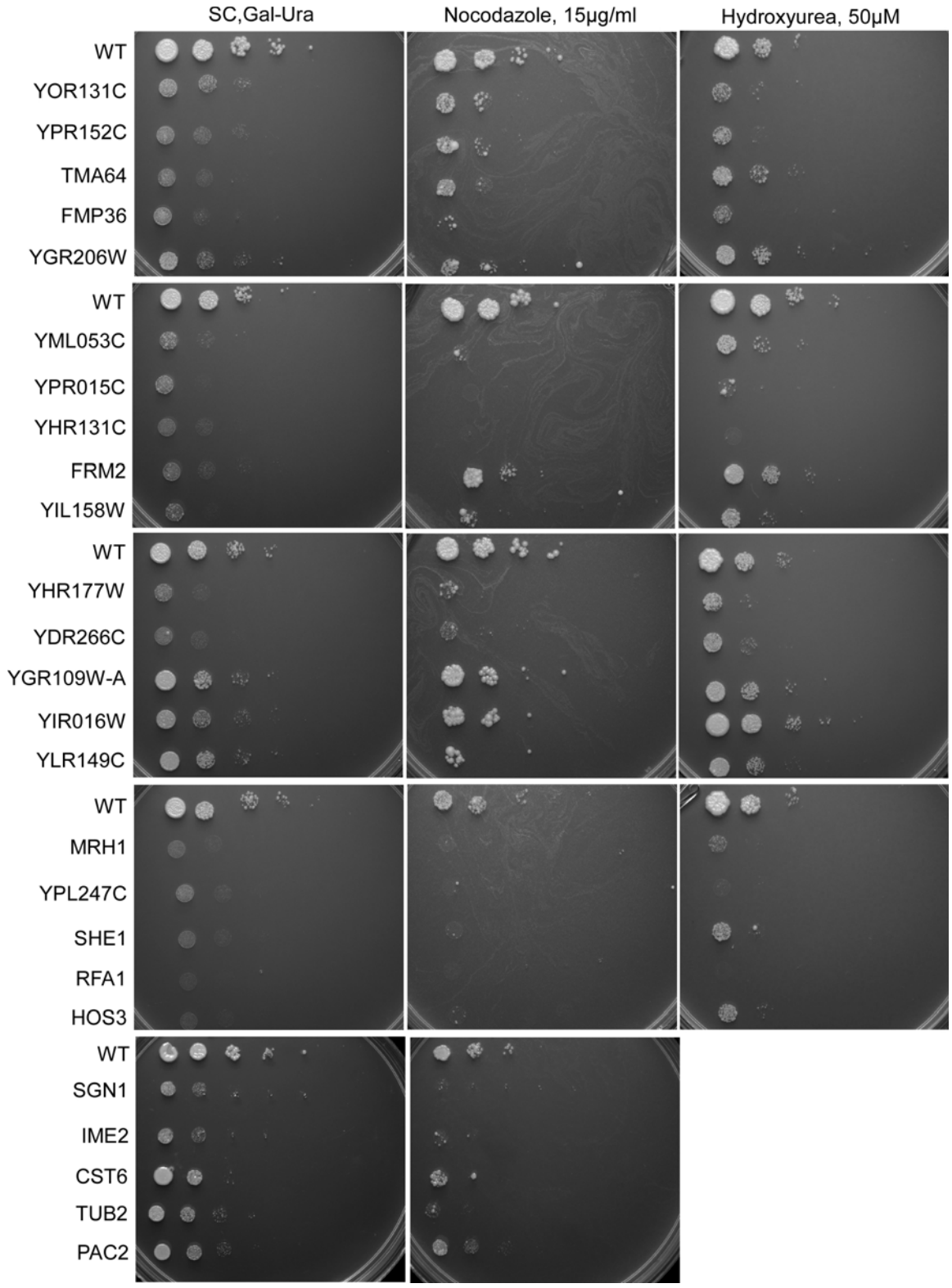

FRS2
SAN1
NIP100
ALG6
WSC2
YPL247C
BET4
CLB2
YML053C
IME2
CLB5
PPZ1

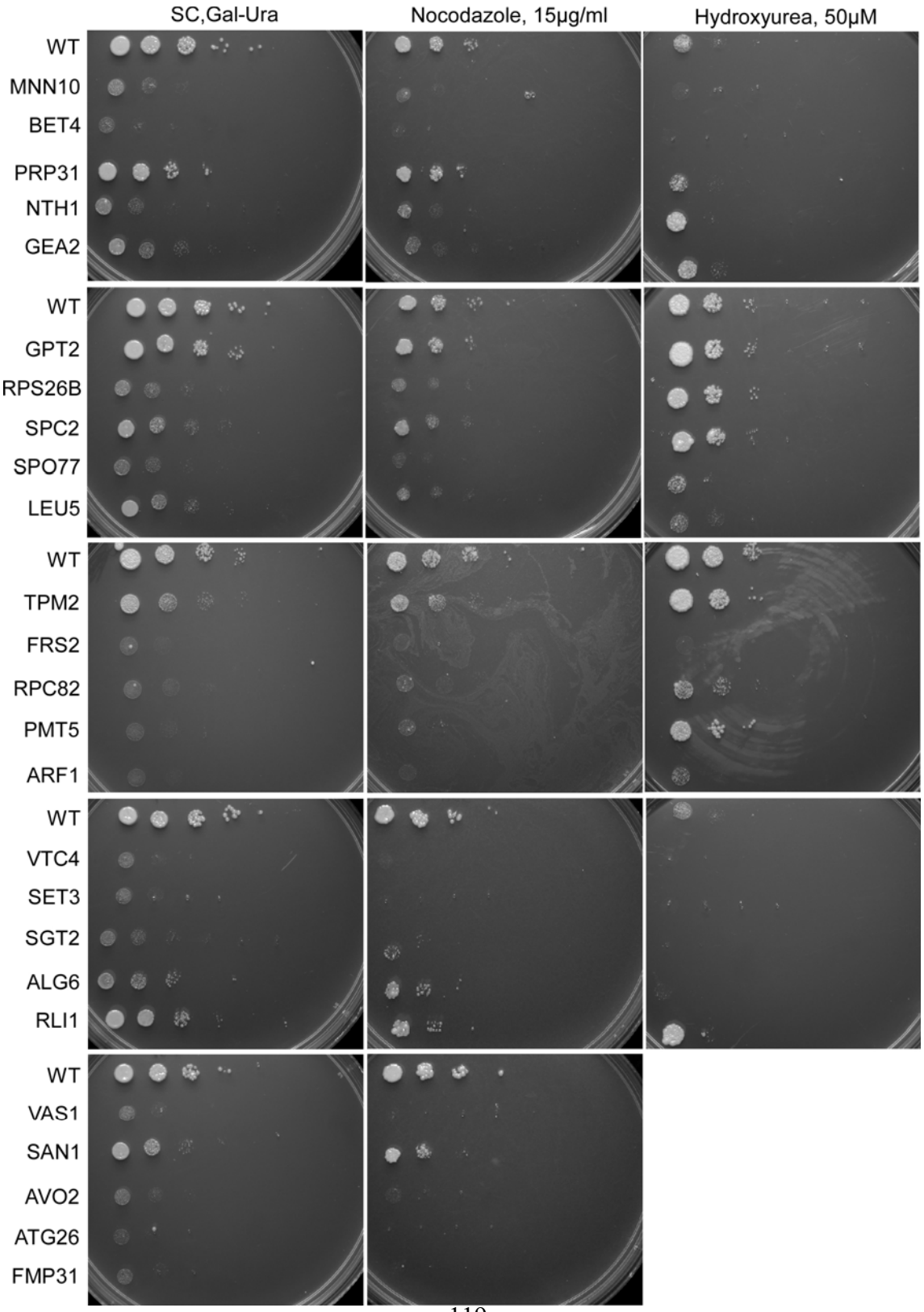

ATG26
PAC2
GEA2
ARF1
MSN5
SET3
VAS1
SHE1
TFG2
YGR109W-A
HOS3
RFA1

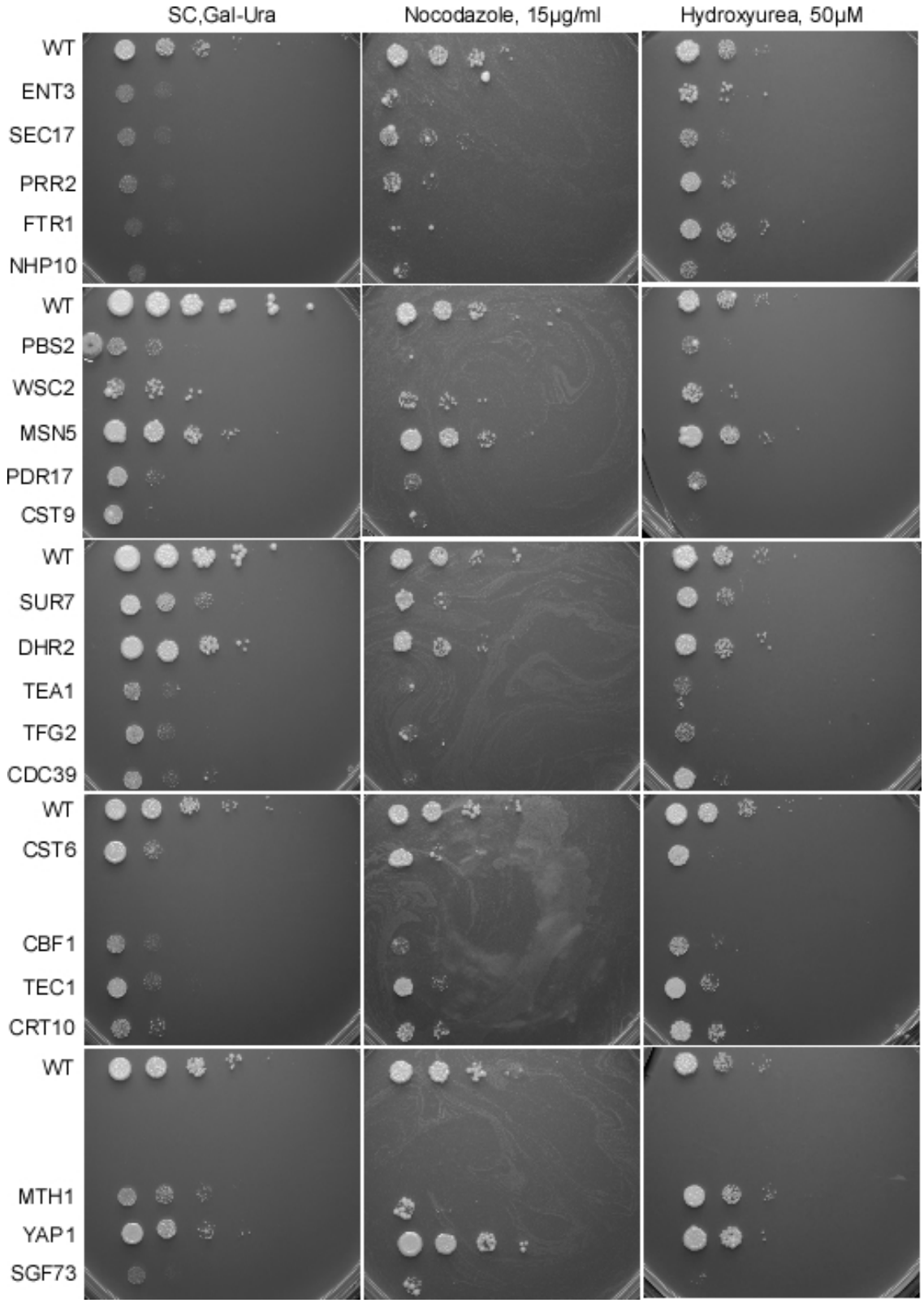

PMT5
CRT10
ENT3
CST6
TEA1
SGT2
PBS2
YIL158W
YHR177W
SGF73
AVO2
RPS26B

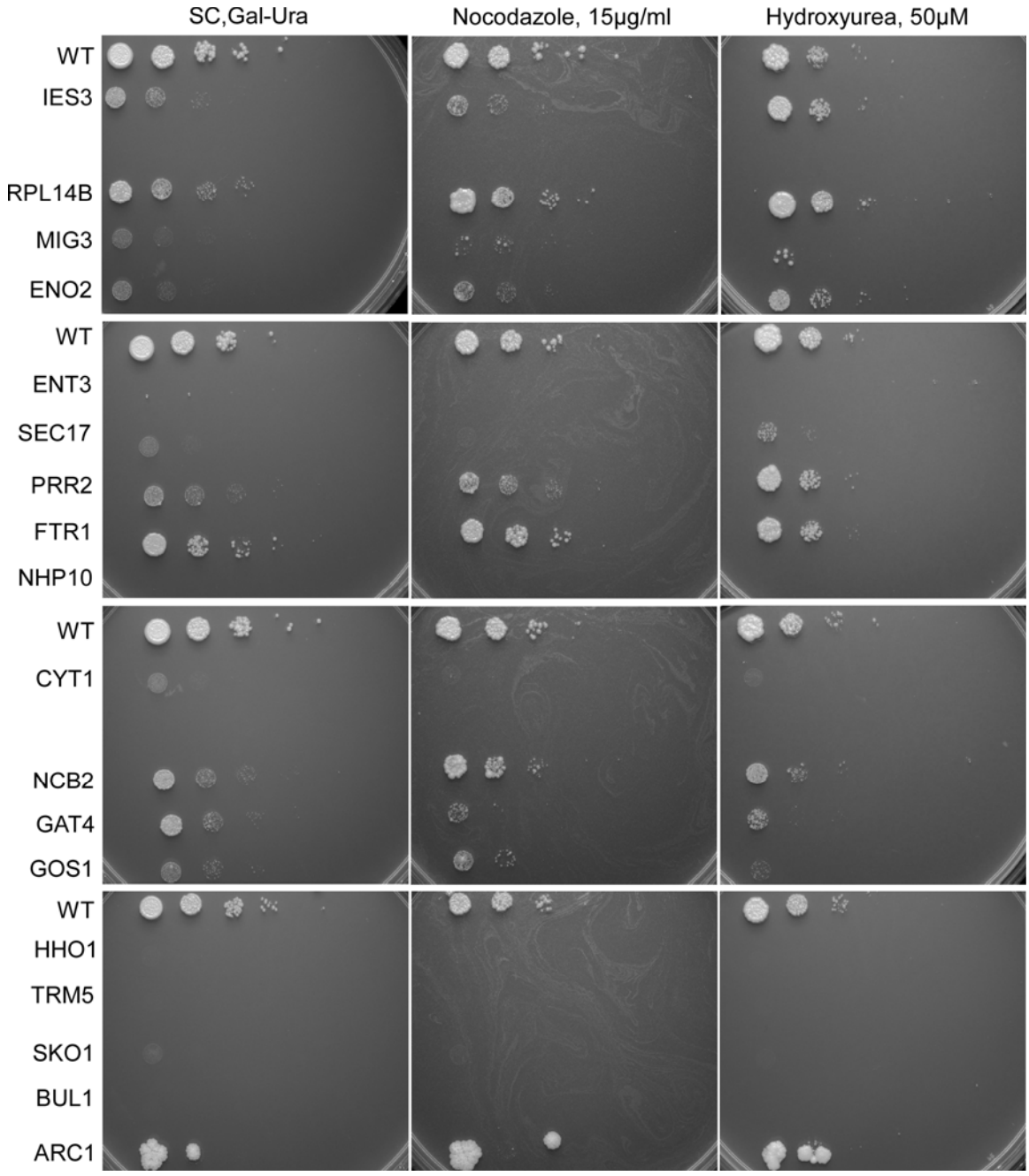


Supplemental Figure 5.1 Flow cytometry histograms of 108 indicated cell cycle genes from the screen. The fluorescence intensity (DNA content) is indicated in x-axis, and the number of cells in a given intensity is plotted in y-axis.


SC,Gal-Ura
Nocodazole, 15µg/ml
Hydroxyurea, 50µM
WT
YOR131C
YPR152C
TMA64
FMP36
YGR206W
WT
YML053C
YPR015C
YHR131C
FRM2
YIL158W
WT
YHR177W
YDR266C
YGR109W-A
YIR016W
YLR149C
WT
MRH1
YPL247C
SHE1
RFA1
HOS3
WT
SGN1
IME2
CST6
TUB2
PAC2


SC,Gal-Ura
Nocodazole, 15µg/ml
Hydroxyurea, 50µM
WT
MNN10
BET4
PRP31
NTH1
GEA2
WT
GPT2
RPS26B
SPC2
SPO77
LEU5
WT
TPM2
FRS2
RPC82
PMT5
ARF1
WT
VTC4
SET3
SGT2
ALG6
RLI1
WT
VAS1
SAN1
AVO2
ATG26
FMP31


SC,Gal-Ura
Nocodazole, 15µg/ml
Hydroxyurea, 50µM
WT
ENT3
SEC17
PRR2
FTR1
NHP10
WT
PBS2
WSC2
MSN5
PDR17
CST9
WT
SUR7
DHR2
TEA1
TFG2
CDC39
WT
CST6
CBF1
TEC1
CRT10
WT
MTH1
YAP1
SGF73



Supplemental Figure 5.2 Growth defects and drug sensitivities caused by the overexpression of indicated cell cycle genes from the screen. Cells containing the control plasmid BG1766 or the indicated cell cycle genes from the screen were spotted on SC-

URA, galactose plates and SC-URA, galactose plates containing hydroxyurea (HU) or nocodazole (NOC) and photographed after growing at 30^{o}C.

Chapter 6: Conclusions and Future Directions

6.1 CONCLUSIONS

The emerging of systems biology greatly broadens the scope of traditional molecular biology studies, accelerates our understanding of genes and proteins function in the existing genomes, and makes it possible to reconstruct the gene interaction network under both the normal and perturbed conditions and to predict and remedy human diseases eventually. Our major goals in studies presented here were focused on two important factors for helping to achieve such goals of systems biology: large-scale genomic study and high-throughput technology. First, we developed new high-throughput technologies for measuring cellular phenotypes on a large scale. Second, we applied newly developed 'cell chip' technologies to systematic morphology and protein localization profiling, revealing new genes and proteins involved in mating pheromone response pathway and evaluating yeast essential genes function in microtubule organization. Third, we used established high-throughput flow cytometry approach to systematically investigate yeast genome for genes regulating cell cycle progression under the overexpression condition, capturing 108 genes (82 are new) involved in cell cycle progression. In conclusion, the inventions provided systems biologists additional useful tools for generating new large-scale data focusing on cellular features at a single cell level, and the new scientific discoveries contributed to system biology society with valuable information that is critical in deciphering the mechanisms of essential pathways such as cell polarity and cell cycle. Given the fact that budding yeast has high homology to higher eukaryotic organisms, it is likely that present studies in yeast would provide insight into the understanding of the corresponding pathways in higher eukaryotes such as human.

6.2 FUTURE DIRECTIONS

Systems biology, although it is not a hypothesis-driven study, gives rise to many interesting hypotheses that need to be validated in traditional ways in molecular biology. We are intrigued by the interesting phenomenon that surfaced from the cell cycle screen and try to investigate some of them in further details. The future plans can be done as the following manners. First, both ChIP-chip and mRNA profiling experiments could be carried out under over-expression condition for 17 transcription factor whose overproduction affected cell cycle progression. It will give insight into what are the regulatory roles of transcription factors under overexpression condition. The mRNA profiling experiment of *SKO1* that is currently underway will help understanding the roles of *SKO1* in regulating cell cycle progression and the mechanism underlying shmoo formation when overproduced. Second, chromosome-loss color sectoring assays will provide direct evidence for the roles of *YPR015C* and *SHE1* in chromosome segregation. In addition to the scientific questions described above, we anticipate applying our 'cell chip' technologies to broader biological questions and higher eukaryotic systems.

References

Aderem, A. (2005). Systems biology: its practice and challenges. Cell *121*, 511-513.

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. Nature *422*, 198-207.

Aguilar-Uscanga, B., and Francois, J. M. (2003). A study of the yeast cell wall composition and structure in response to growth conditions and mode of cultivation. Lett Appl Microbiol *37*, 268-274.

Ayscough, K. R., and Drubin, D. G. (1998). A role for the yeast actin cytoskeleton in pheromone receptor clustering and signalling. Curr Biol *8*, 927-930.

Baba, M., Baba, N., Ohsumi, Y., Kanaya, K., and Osumi, M. (1989). Three-dimensional analysis of morphogenesis induced by mating pheromone alpha factor in Saccharomyces cerevisiae. J Cell Sci *94 (Pt 2)*, 207-216.

Bagnat, M., and Simons, K. (2002). Cell surface polarization during yeast mating. Proc Natl Acad Sci U S A *99*, 14183-14188.

Ballensiefen, W., and Schmitt, H. D. (1997). Periplasmic Bar1 protease of Saccharomyces cerevisiae is active before reaching its extracellular destination. Eur J Biochem *247*, 142-147.

Benachour, A., Sipos, G., Flury, I., Reggiori, F., Canivenc-Gansel, E., Vionnet, C., Conzelmann, A., and Benghezal, M. (1999). Deletion of GPI7, a yeast gene required for addition of a side chain to the glycosylphosphatidylinositol (GPI) core structure, affects GPI protein transport, remodeling, and cell wall integrity. J Biol Chem *274*, 15251-15261.

Bjorklund, M., Taipale, M., Varjosalo, M., Saharinen, J., Lahdenpera, J., and Taipale, J. (2006). Identification of pathways regulating cell size and cell-cycle progression by RNAi. Nature *439*, 1009-1013.

Blagoev, B., and Pandey, A. (2001). Microarrays go live - new prospects for proteomics. Trends Biochem Sci *26*, 639-641.

Blattner, F. R., Plunkett, G., 3rd, Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F.*, et al.* (1997). The complete genome sequence of Escherichia coli K-12. Science *277*, 1453-1474.

Borneman, A. R., Chambers, P. J., and Pretorius, I. S. (2007). Yeast systems biology: modelling the winemaker's art. Trends Biotechnol *25*, 349-355.

Burja, A. M., Dhamwichukorn, S., and Wright, P. C. (2003). Cyanobacterial postgenomic research and systems biology. Trends Biotechnol *21*, 504-511.

Burke, D., Dawson, D., and Stearns, T. (2000). Methods in Yeast Genetics: A Cold Spring Harbor Laboratory Course Manual (Cold Spring Harbor, NY: CSHL Press).

Butcher, E. C., Berg, E. L., and Kunkel, E. J. (2004). Systems biology in drug discovery. Nat Biotechnol *22*, 1253-1259.

Butty, A. C., Pryciak, P. M., Huang, L. S., Herskowitz, I., and Peter, M. (1998). The role of Far1p in linking the heterotrimeric G protein to polarity establishment proteins during yeast mating. Science *282*, 1511-1516.

Chant, J. (1999). Cell polarity in yeast. Annu Rev Cell Dev Biol *15*, 365-391.

Chenevert, J., Corrado, K., Bender, A., Pringle, J., and Herskowitz, I. (1992). A yeast gene (BEM1) necessary for cell polarization whose product contains two SH3 domains. Nature *356*, 77-79.

Collins, F. S., Green, E. D., Guttmacher, A. E., and Guyer, M. S. (2003). A vision for the future of genomics research. Nature *422*, 835-847.

Conrad, C., Erfle, H., Warnat, P., Daigle, N., Lorch, T., Ellenberg, J., Pepperkok, R., and Eils, R. (2004). Automatic identification of subcellular phenotypes on human cell arrays. Genome Res *14*, 1130-1136.

Coros, C. J., Landthaler, M., Piazza, C. L., Beauregard, A., Esposito, D., Perutka, J., Lambowitz, A. M., and Belfort, M. (2005). Retrotransposition strategies of the Lactococcus lactis Ll.LtrB group II intron are dictated by host identity and cellular environment. Mol Microbiol *56*, 509-524.

Culotti, J., and Hartwell, L. H. (1971). Genetic control of the cell division cycle in yeast. 3. Seven genes controlling nuclear division. Exp Cell Res *67*, 389-401.

Dahan, O., and Kupiec, M. (2002). Mutations in genes of Saccharomyces cerevisiae encoding pre-mRNA splicing factors cause cell cycle arrest through activation of the spindle checkpoint. Nucleic Acids Res *30*, 4361-4370.

Dohlman, H. G., and Thorner, J. W. (2001). Regulation of G protein-initiated signal transduction in yeast: paradigms and principles. Annu Rev Biochem *70*, 703-754.

Dulic, V., and Riezman, H. (1990). Saccharomyces cerevisiae mutants lacking a functional vacuole are defective for aspects of the pheromone response. J Cell Sci *97 (Pt 3)*, 517-525.

Elion, E. A. (2000). Pheromone response, mating and cell biology. Curr Opin Microbiol *3*, 573-581.

Espinet, C., de la Torre, M. A., Aldea, M., and Herrero, E. (1995). An efficient method to isolate yeast genes causing overexpression-mediated growth arrest. Yeast *11*, 25-32.

Euskirchen, G. M. (2002). Nnf1p, Dsn1p, Mtw1p, and Nsl1p: a new group of proteins important for chromosome segregation in Saccharomyces cerevisiae. Eukaryot Cell *1*, 229-240.

Fields, S. (1990). Pheromone response in yeast. Trends Biochem Sci *15*, 270-273.

Finger, F. P., Hughes, T. E., and Novick, P. (1998). Sec3p is a spatial landmark for polarized secretion in budding yeast. Cell *92*, 559-571.

Fitch, I., Dahmann, C., Surana, U., Amon, A., Nasmyth, K., Goetsch, L., Byers, B., and Futcher, B. (1992). Characterization of four B-type cyclin genes of the budding yeast Saccharomyces cerevisiae. Mol Biol Cell *3*, 805-818.

Fleischer, T. C., Weaver, C. M., McAfee, K. J., Jennings, J. L., and Link, A. J. (2006). Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes. Genes Dev *20*, 1294-1307.

Fodor, S. P., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T., and Solas, D. (1991). Light-directed, spatially addressable parallel chemical synthesis. Science *251*, 767-773.

Gehrung, S., and Snyder, M. (1990). The SPA2 gene of Saccharomyces cerevisiae is important for pheromone-induced morphogenesis and efficient mating. J Cell Biol *111*, 1451-1464.

Gelperin, D. M., White, M. A., Wilkinson, M. L., Kon, Y., Kung, L. A., Wise, K. J., Lopez-Hoyo, N., Jiang, L., Piccirillo, S., Yu, H.*, et al.* (2005). Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. Genes Dev *19*, 2816-2826.

Ghaemmaghami, S., Huh, W. K., Bower, K., Howson, R. W., Belle, A., Dephoure, N., O'Shea, E. K., and Weissman, J. S. (2003). Global analysis of protein expression in yeast. Nature *425*, 737-741.

Giaever, G., Chu, A. M., Ni, L., Connelly, C., Riles, L., Veronneau, S., Dow, S., Lucau-Danila, A., Anderson, K., Andre, B.*, et al.* (2002). Functional profiling of the Saccharomyces cerevisiae genome. Nature *418*, 387-391.

Griffin, T. J., Gygi, S. P., Ideker, T., Rist, B., Eng, J., Hood, L., and Aebersold, R. (2002). Complementary profiling of gene expression at the transcriptome and proteome levels in Saccharomyces cerevisiae. Mol Cell Proteomics *1*, 323-333.

Guo, Z., Guilfoyle, R. A., Thiel, A. J., Wang, R., and Smith, L. M. (1994). Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports. Nucleic Acids Res *22*, 5456-5465.

Haase, S. B., and Reed, S. I. (2002). Improved flow cytometric analysis of the budding yeast cell cycle. Cell Cycle *1*, 132-136.

Hall, J. A., and Maloney, P. C. (2001). Transmembrane segment 11 of UhpT, the sugar phosphate carrier of Escherichia coli, is an alpha-helix that carries determinants of substrate selectivity. J Biol Chem *276*, 25107-25113.

Hanna, D. E., Rethinaswamy, A., and Glover, C. V. (1995). Casein kinase II is required for cell cycle progression during G1 and G2/M in Saccharomyces cerevisiae. J Biol Chem *270*, 25905-25914.

Hanna, J. S., Kroll, E. S., Lundblad, V., and Spencer, F. A. (2001). Saccharomyces cerevisiae CTF18 and CTF4 are required for sister chromatid cohesion. Mol Cell Biol *21*, 3144-3158.

Hartwell, L. H. (1971a). Genetic control of the cell division cycle in yeast. II. Genes controlling DNA replication and its initiation. J Mol Biol *59*, 183-194.

Hartwell, L. H. (1971b). Genetic control of the cell division cycle in yeast. IV. Genes controlling bud emergence and cytokinesis. Exp Cell Res *69*, 265-276.

Hartwell, L. H. (1973). Three additional genes required for deoxyribonucleic acid synthesis in Saccharomyces cerevisiae. J Bacteriol *115*, 966-974.

Hartwell, L. H. (1974). Saccharomyces cerevisiae cell cycle. Bacteriol Rev *38*, 164-198.

Hartwell, L. H., Culotti, J., and Reid, B. (1970). Genetic control of the cell-division cycle in yeast. I. Detection of mutants. Proc Natl Acad Sci U S A *66*, 352-359.

Hoyt, M. A., Macke, J. P., Roberts, B. T., and Geiser, J. R. (1997). Saccharomyces cerevisiae PAC2 functions with CIN1, 2 and 4 in a pathway leading to normal microtubule stability. Genetics *146*, 849-857.

Hu, Z., Killion, P. J., and Iyer, V. R. (2007). Genetic reconstruction of a functional transcriptional regulatory network. Nat Genet *39*, 683-687.

Huffaker, T. C., Thomas, J. H., and Botstein, D. (1988). Diverse effects of beta-tubulin mutations on microtubule formation and function. J Cell Biol *106*, 1997-2010.

Hughes, T. R., Mao, M., Jones, A. R., Burchard, J., Marton, M. J., Shannon, K. W., Lefkowitz, S. M., Ziman, M., Schelter, J. M., Meyer, M. R.*, et al.* (2001). Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. Nat Biotechnol *19*, 342-347.

Hughes, T. R., Robinson, M. D., Mitsakakis, N., and Johnston, M. (2004). The promise of functional genomics: completing the encyclopedia of a cell. Curr Opin Microbiol *7*, 546-554.

Huh, W. K., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S., and O'Shea, E. K. (2003). Global analysis of protein localization in budding yeast. Nature *425*, 686-691.

Ideker, T., Galitski, T., and Hood, L. (2001). A new approach to decoding life: systems biology. Annu Rev Genomics Hum Genet *2*, 343-372.

Iyer, V. R., Eisen, M. B., Ross, D. T., Schuler, G., Moore, T., Lee, J. C., Trent, J. M., Staudt, L. M., Hudson, J., Jr., Boguski, M. S.*, et al.* (1999). The transcriptional program in the response of human fibroblasts to serum. Science *283*, 83-87.

Jacobs, C. W., Adams, A. E., Szaniszlo, P. J., and Pringle, J. R. (1988). Functions of microtubules in the Saccharomyces cerevisiae cell cycle. J Cell Biol *107*, 1409-1426.

Katzmann, D. J., Odorizzi, G., and Emr, S. D. (2002). Receptor downregulation and multivesicular-body sorting. Nat Rev Mol Cell Biol *3*, 893-905.

Keasling, J. D., Bertsch, L., and Kornberg, A. (1993). Guanosine pentaphosphate phosphohydrolase of Escherichia coli is a long-chain exopolyphosphatase. Proc Natl Acad Sci U S A *90*, 7029-7033.

Kilmartin, J. V., and Adams, A. E. (1984). Structural rearrangements of tubulin and actin during the cell cycle of the yeast Saccharomyces. J Cell Biol *98*, 922-933.

Kirschner, M. W. (2005). The meaning of systems biology. Cell *121*, 503-504.

Knop, M., Pereira, G., Geissler, S., Grein, K., and Schiebel, E. (1997). The spindle pole body component Spc97p interacts with the gamma-tubulin of Saccharomyces cerevisiae and functions in microtubule organization and spindle pole body duplication. Embo J *16*, 1550-1564.

Knop, M., Pereira, G., and Schiebel, E. (1999). Microtubule organization by the budding yeast spindle pole body. Biol Cell *91*, 291-304.

Krogan, N. J., Peng, W. T., Cagney, G., Robinson, M. D., Haw, R., Zhong, G., Guo, X., Zhang, X., Canadien, V., Richards, D. P.*, et al.* (2004). High-definition macromolecular composition of yeast RNA-processing complexes. Mol Cell *13*, 225-239.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Lashkari, D. A., DeRisi, J. L., McCusker, J. H., Namath, A. F., Gentile, C., Hwang, S. Y., Brown, P. O., and Davis, R. W. (1997). Yeast microarrays for genome wide parallel genetic and gene expression analysis. Proc Natl Acad Sci U S A *94*, 13057-13062.

Lechner, T., Carrozza, M. J., Yu, Y., Grant, P. A., Eberharter, A., Vannier, D., Brosch, G., Stillman, D. J., Shore, D., and Workman, J. L. (2000). Sds3 (suppressor of defective silencing 3) is an integral component of the yeast Sin3[middle dot]Rpd3 histone deacetylase complex and is required for histone deacetylase activity. J Biol Chem *275*, 40961-40966.

Lee, I., Date, S. V., Adai, A. T., and Marcotte, E. M. (2004). A probabilistic functional network of yeast genes. Science *306*, 1555-1558.

Lee, I., Li, Z., and Marcotte, E. M. (2007). An Improved, Bias-Reduced Probabilistic Functional Gene Network of Baker's Yeast, Saccharomyces cerevisiae. PLoS ONE *2*, e988.

Lee, T. I., Rinaldi, N. J., Robert, F., Odom, D. T., Bar-Joseph, Z., Gerber, G. K., Hannett, N. M., Harbison, C. T., Thompson, C. M., Simon, I.*, et al.* (2002). Transcriptional regulatory networks in Saccharomyces cerevisiae. Science *298*, 799-804.

Leeuw, T., Wu, C., Schrag, J. D., Whiteway, M., Thomas, D. Y., and Leberer, E. (1998). Interaction of a G-protein beta-subunit with a conserved sequence in Ste20/PAK family protein kinases. Nature *391*, 191-195.

Levsky, J. M., Shenoy, S. M., Pezo, R. C., and Singer, R. H. (2002). Single-cell gene expression profiling. Science *297*, 836-840.

Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., Garvik, B. M., and Yates, J. R., 3rd (1999). Direct analysis of protein complexes using mass spectrometry. Nat Biotechnol *17*, 676-682.

Liu, X. S. (2007). Getting started in tiling microarray analysis. PLoS Comput Biol *3*, 1842-1844.

Lockhart, D. J., and Winzeler, E. A. (2000). Genomics, gene expression and DNA arrays. Nature *405*, 827-836.

Long, R. M., Elliott, D. J., Stutz, F., Rosbash, M., and Singer, R. H. (1995). Spatial consequences of defective processing of specific yeast mRNAs revealed by fluorescent in situ hybridization. Rna *1*, 1071-1078.

Mendenhall, M. D., and Hodge, A. E. (1998). Regulation of Cdc28 cyclin-dependent protein kinase activity during the cell cycle of the yeast Saccharomyces cerevisiae. Microbiol Mol Biol Rev *62*, 1191-1243.

Miles, J., and Formosa, T. (1992). Evidence that POB1, a Saccharomyces cerevisiae protein that binds to DNA polymerase alpha, acts in DNA metabolism in vivo. Mol Cell Biol *12*, 5724-5735.

Miller, R. K., Cheng, S. C., and Rose, M. D. (2000). Bim1p/Yeb1p mediates the Kar9p-dependent cortical attachment of cytoplasmic microtubules. Mol Biol Cell *11*, 2949-2959.

Mnaimneh, S., Davierwala, A. P., Haynes, J., Moffat, J., Peng, W. T., Zhang, W., Yang, X., Pootoolal, J., Chua, G., Lopez, A.*, et al.* (2004). Exploration of essential gene functions via titratable promoter alleles. Cell *118*, 31-44.

Moir, D., Stewart, S. E., Osmond, B. C., and Botstein, D. (1982). Cold-sensitive cell-division-cycle mutants of yeast: isolation, properties, and pseudoreversion studies. Genetics *100*, 547-563.

Narayanaswamy, R., Niu, W., Scouras, A. D., Hart, G. T., Davies, J., Ellington, A. D., Iyer, V. R., and Marcotte, E. M. (2006). Systematic profiling of cellular phenotypes with spotted cell microarrays reveals mating-pheromone response genes. Genome Biol *7*, R6.

Nehlin, J. O., Carlberg, M., and Ronne, H. (1992). Yeast SKO1 gene encodes a bZIP protein that binds to the CRE motif and acts as a repressor of transcription. Nucleic Acids Res *20*, 5271-5278.

Nelson, W. J. (2003). Adaptation of core mechanisms to generate cell polarity. Nature *422*, 766-774.

Nern, A., and Arkowitz, R. A. (1998). A GTP-exchange factor required for cell orientation. Nature *391*, 195-198.

Newberry, K. J., and Brennan, R. G. (2004). The structural mechanism for transcription activation by MerR family member multidrug transporter activation, N terminus. J Biol Chem *279*, 20356-20362.

Nilsen, T., Yan, A. W., Gale, G., and Goldberg, M. B. (2005). Presence of multiple sites containing polar material in spherical Escherichia coli cells that lack MreB. J Bacteriol *187*, 6187-6196.

Olson, K. A., Nelson, C., Tai, G., Hung, W., Yong, C., Astell, C., and Sadowski, I. (2000). Two regulators of Ste12p inhibit pheromone-responsive transcription by separate mechanisms. Mol Cell Biol *20*, 4199-4209.

Ouspenski, II, Elledge, S. J., and Brinkley, B. R. (1999). New yeast genes important for chromosome integrity and segregation identified by dosage effects on genome stability. Nucleic Acids Res *27*, 3001-3008.

Palkova, Z., Vachova, L., Valer, M., and Preckel, T. (2004). Single-cell analysis of yeast, mammalian cells, and fungal spores with a microfluidic pressure-driven chip-based system. Cytometry A *59*, 246-253.

Pandey, A., and Mann, M. (2000). Proteomics to study genes and genomes. Nature *405*, 837-846.

Park, H. O., Sanson, A., and Herskowitz, I. (1999). Localization of bud2p, a GTPase-activating protein necessary for programming cell polarity in yeast to the presumptive bud site. Genes Dev *13*, 1912-1917.

Pascual-Ahuir, A., Posas, F., Serrano, R., and Proft, M. (2001a). Multiple levels of control regulate the yeast cAMP-response element-binding protein repressor Sko1p in response to stress. J Biol Chem *276*, 37373-37378.

Pascual-Ahuir, A., Serrano, R., and Proft, M. (2001b). The Sko1p repressor and Gcn4p activator antagonistically modulate stress-regulated transcription in Saccharomyces cerevisiae. Mol Cell Biol *21*, 16-25.

Perlman, Z. E., Slack, M. D., Feng, Y., Mitchison, T. J., Wu, L. F., and Altschuler, S. J. (2004). Multidimensional drug profiling by automated microscopy. Science *306*, 1194-1198.

Perutka, J., Wang, W., Goerlitz, D., and Lambowitz, A. M. (2004). Use of computer-designed group II introns to disrupt Escherichia coli DExH/D-box protein and DNA helicase genes. J Mol Biol *336*, 421-439.

Pruyne, D., and Bretscher, A. (2000). Polarization of cell growth in yeast. I. Establishment and maintenance of polarity states. J Cell Sci *113 (Pt 3)*, 365-375.

Read, E. B., Okamura, H. H., and Drubin, D. G. (1992). Actin- and tubulin-dependent functions during Saccharomyces cerevisiae mating projection formation. Mol Biol Cell *3*, 429-444.

Reinders, J., Zahedi, R. P., Pfanner, N., Meisinger, C., and Sickmann, A. (2006). Toward the complete yeast mitochondrial proteome: multidimensional separation techniques for mitochondrial proteomics. J Proteome Res *5*, 1543-1554.

Roberts, C. J., Nelson, B., Marton, M. J., Stoughton, R., Meyer, M. R., Bennett, H. A., He, Y. D., Dai, H., Walker, W. L., Hughes, T. R.*, et al.* (2000). Signaling and circuitry of multiple MAPK pathways revealed by a matrix of global gene expression profiles. Science *287*, 873-880.

Robinson, M. D., Grigull, J., Mohammad, N., and Hughes, T. R. (2002). FunSpec: a web-based cluster interpreter for yeast. BMC Bioinformatics *3*, 35.

Rochfort, S. (2005). Metabolomics reviewed: a new "omics" platform technology for systems biology and implications for natural products research. J Nat Prod *68*, 1813-1820.

Saito, T. L., Ohtani, M., Sawai, H., Sano, F., Saka, A., Watanabe, D., Yukawa, M., Ohya, Y., and Morishita, S. (2004). SCMD: Saccharomyces cerevisiae Morphological Database. Nucleic Acids Res *32 Database issue*, D319-322.

San-Segundo, P. A., and Roeder, G. S. (2000). Role for the silencing protein Dot1 in meiotic checkpoint control. Mol Biol Cell *11*, 3601-3615.

Schafer, K. A. (1998). The cell cycle: a review. Vet Pathol *35*, 461-478.

Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science *270*, 467-470.

Schwartz, K., Richards, K., and Botstein, D. (1997). BIM1 encodes a microtubule-binding protein in yeast. Mol Biol Cell *8*, 2677-2691.

Segall, J. E. (1993). Polarization of yeast cells in spatial gradients of alpha mating factor. Proc Natl Acad Sci U S A *90*, 8332-8336.

Shaw, S. L., Maddox, P., Skibbens, R. V., Yeh, E., Salmon, E. D., and Bloom, K. (1998). Nuclear and spindle dynamics in budding yeast. Mol Biol Cell *9*, 1627-1631.

Shaw, S. L., Yeh, E., Maddox, P., Salmon, E. D., and Bloom, K. (1997). Astral microtubule dynamics in yeast: a microtubule-based searching mechanism for spindle orientation and nuclear migration into the bud. J Cell Biol *139*, 985-994.

Silva, J. M., Mizuno, H., Brady, A., Lucito, R., and Hannon, G. J. (2004). RNA interference microarrays: high-throughput loss-of-function genetics in mammalian cells. Proc Natl Acad Sci U S A *101*, 6548-6552.

Simon, V. R., Swayne, T. C., and Pon, L. A. (1995). Actin-dependent mitochondrial motility in mitotic yeast and cell-free systems: identification of a motor activity on the mitochondrial surface. J Cell Biol *130*, 345-354.

Sobel, S. G., and Snyder, M. (1995). A highly divergent gamma-tubulin gene is essential for cell growth and proper microtubule organization in Saccharomyces cerevisiae. J Cell Biol *131*, 1775-1788.

Sopko, R., Huang, D., Preston, N., Chua, G., Papp, B., Kafadar, K., Snyder, M., Oliver, S. G., Cyert, M., Hughes, T. R.*, et al.* (2006). Mapping pathways and phenotypes by systematic gene overexpression. Mol Cell *21*, 319-330.

Spencer, F., Gerring, S. L., Connelly, C., and Hieter, P. (1990). Mitotic chromosome transmission fidelity mutants in Saccharomyces cerevisiae. Genetics *124*, 237-249.

Steen, H., and Pandey, A. (2002). Proteomics goes quantitative: measuring protein abundance. Trends Biotechnol *20*, 361-364.

Stevenson, G., Andrianopoulos, K., Hobbs, M., and Reeves, P. R. (1996). Organization of the Escherichia coli K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid. J Bacteriol *178*, 4885-4893.

Stevenson, L. F., Kennedy, B. K., and Harlow, E. (2001). A large-scale overexpression screen in Saccharomyces cerevisiae identifies previously uncharacterized cell cycle genes. Proc Natl Acad Sci U S A *98*, 3946-3951.

Takahashi, K., Hattori, A., Suzuki, I., Ichiki, T., and Yasuda, K. (2004). Non-destructive on-chip cell sorting system with real-time microscopic image processing. J Nanobiotechnology *2*, 5.

Tong, A. H., Lesage, G., Bader, G. D., Ding, H., Xu, H., Xin, X., Young, J., Berriz, G. F., Brost, R. L., Chang, M.*, et al.* (2004). Global mapping of the yeast genetic interaction network. Science *303*, 808-813.

Trueheart, J., Boeke, J. D., and Fink, G. R. (1987). Two genes required for cell fusion during yeast conjugation: evidence for a pheromone-induced surface protein. Mol Cell Biol *7*, 2316-2328.

Ussery, D. W., and Jensen, L. J. (2005). Systems biology: in the broadest sense of the word. Environ Microbiol *7*, 482-483.

Valtz, N., and Herskowitz, I. (1996). Pea2 protein of yeast is localized to sites of polarized growth and is required for efficient mating and bipolar budding. J Cell Biol *135*, 725-739.

Valtz, N., Peter, M., and Herskowitz, I. (1995). FAR1 is required for oriented polarization of yeast cells in response to mating pheromones. J Cell Biol *131*, 863-873.

van den Hazel, H. B., Pichler, H., do Valle Matta, M. A., Leitner, E., Goffeau, A., and Daum, G. (1999). PDR16 and PDR17, two homologous genes of Saccharomyces cerevisiae, affect lipid biosynthesis and resistance to multiple drugs. J Biol Chem *274*, 1934-1941.

Weckwerth, W. (2003). Metabolomics in systems biology. Annu Rev Plant Biol *54*, 669-689.

Westerhoff, H. V., and Palsson, B. O. (2004). The evolution of molecular biology into systems biology. Nat Biotechnol *22*, 1249-1252.

Wigge, P. A., and Kilmartin, J. V. (2001). The Ndc80p complex from Saccharomyces cerevisiae contains conserved centromere components and has a function in chromosome segregation. J Cell Biol *152*, 349-360.

Winey, M., Mamay, C. L., O'Toole, E. T., Mastronarde, D. N., Giddings, T. H., Jr., McDonald, K. L., and McIntosh, J. R. (1995). Three-dimensional ultrastructural analysis of the Saccharomyces cerevisiae mitotic spindle. J Cell Biol *129*, 1601-1615.

Winsor, B., and Schiebel, E. (1997). Review: an overview of the Saccharomyces cerevisiae microtubule and microfilament cytoskeleton. Yeast *13*, 399-434.

Winzeler, E. A., Shoemaker, D. D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J. D., Bussey, H.*, et al.* (1999). Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science *285*, 901-906.

Yu, L., Castillo, L. P., Mnaimneh, S., Hughes, T. R., and Brown, G. W. (2006). A survey of essential gene function in the yeast cell division cycle. Mol Biol Cell *17*, 4736-4747.

Zhao, J., and Lambowitz, A. M. (2005). Inaugural Article: A bacterial group II intron-encoded reverse transcriptase localizes to cellular poles. Proc Natl Acad Sci U S A *102*, 16133-16140.

Zheng, L., Chen, Y., and Lee, W. H. (1999). Hec1p, an evolutionarily conserved coiled-coil protein, modulates chromosome segregation through interaction with SMC proteins. Mol Cell Biol *19*, 5417-5428.

Zheng, Y., Bender, A., and Cerione, R. A. (1995). Interactions among proteins involved in bud-site selection and bud-site assembly in Saccharomyces cerevisiae. J Biol Chem *270*, 626-630.

Ziauddin, J., and Sabatini, D. M. (2001). Microarrays of cells expressing defined cDNAs. Nature *411*, 107-110.

Vita

Wei Niu was born in Changle, China, the daughter of Qinglin Niu and Huilan Zhao. After graduating from Changle high school, Changle, China, in 1994, she entered Shandong Normal University, Jinan, China, in September 1994. She received a degree of Bachelor of Science in Biology in July 1998. In September 1998, she entered the Institute of Botany of Chinese Academy of Sciences, Beijing, China. After three-year graduate study, she received a degree of Master of Science in Botany in July 2001. In August 2002, she was enrolled in the Cell and Molecular Biology graduate program at the University of Texas at Austin. In May of 2003, she joined the research group of Dr. Edward Marcotte to study systems biology in the budding yeast *Saccharomyces cerevisiae*. She received a Master of Arts in August of 2005.

Permanent address: 227 Xinchang Street APT 1-201, Changle, Shandong 262400
P. R. China

This dissertation was typed by the author.