The Transcriptome and Its Translation during Recovery from Cell Cycle Arrest in Saccharomyces cerevisiae*

Kyle A. Serikawa‡, Xie Lillian Xu§, Vivian L. MacKay‡, G. Lynn Law‡, Qin Zong¶, Lue Ping Zhao§∥, Roger Bumgarner**, and David R. Morris‡‡

Complete genome sequences together with high throughput technologies have made comprehensive characterizations of gene expression patterns possible. While genome-wide measurement of mRNA levels was one of the first applications of these advances, other important aspects of gene expression are also amenable to a genomic approach, for example, the translation of message into protein. Earlier we reported a high throughput technology approach, for example, the translation of message into protein. Earlier we reported a high throughput technology

The recent sequencing of whole genomes and the development of high throughput methodologies have allowed biological questions to be framed in terms of systems of interacting biological molecules rather than of individual genes. One of the first demonstrations of this powerful new approach was the application of microarrays to analysis of transcription. The path leading from genes to phenotype, however, passes through several other important regulatory steps including translation and protein modification, localization, and degradation. In recent years investigators have uncovered numerous examples of protein levels being regulated via mechanisms operating at the level of translation (1). Expansion of high throughput technologies to encompass other aspects of gene expression such as translation will greatly enhance our ability to describe, in a comprehensive and quantitative way, the flow of information from DNA sequence to phenotype.

Translation state array analysis (TSAA)¹ is a microarray-based technology that allows the genome-wide evaluation of mRNA association with the translational machinery (2–5). In TSAA, intact polyribosomes from cells or tissues are separated according to size by centrifugation through sucrose density gradients. In the most easily implemented and high throughput form of TSAA, the gradients are fractionated and one containing messages with few or no ribosomes (“poorly translated” transcripts) and one containing messages that are better loaded with ribosomes (“well translated” transcripts). RNA isolated from the two pools is labeled with different fluorescent dyes and hybridized to the same array to provide a ratio, the “translation state,” for each gene represented on the array. The translation state ratio gives an estimate of how well a gene is translated. By comparing these ratios between different time points or conditions, TSAA provides a survey of the genome for instances of translational regulation.

In the first application of TSAA (2), filter-based arrays were used to study expression of a limited number of mammalian genes. The current work extended this technology to the whole genome level, using much more sophisticated microarray and analysis technologies, in a study of cell cycle arrest and release in the budding yeast Saccharomyces cerevisiae.

¹ The abbreviations used are: TSAA, translation state array analysis; ORF, open reading frame; uORF, upstream ORF; Tris, Trizma base; NFD, number of false discoveries.
In the process we examined two different but related biological processes: the response of the translational apparatus to extrinsic stimuli (pheromone treatment or temperature shift) and the relationship between transcript level and translation state during an intrinsically regulated process (the cell cycle). Although a rich body of literature already exists for transcriptome composition during cell cycle progression (6, 7), little is understood of the significance of translational control in this process.

EXPERIMENTAL PROCEDURES

Strains and Culture Conditions—All experiments used strain BY1782 (MATa cdc15-2 tyr1 leu2 ura3 his7 gal1) (8) except for the β-galactosidase assays, which were performed in BY2125 (MATa ade2-1 his3-11 leu2-3,112 trp1-1 can1-100 ssd1-d) (9). Both strains were generously provided by Dr. Linda Breeden (Fred Hutchison Cancer Research Center). Cultures were grown at 25 °C in a shaking incubator except during the cdc15 arrest (37 °C) and growth of cells for the β-galactosidase assays (30 °C).

cdc15 Synchronization—Cells were grown in 400 ml of YEPD medium (1% yeast extract, 2% peptone, 2% dextrose) (10) to a density of 5–8 × 10⁶ cells/ml, and then the cultures were shifted to 37 °C for 3 h. To release cells from arrest, flasks were swirled rapidly in an ice-water bath so that the culture temperature reached 25 °C in less than 1 min. Two separate experiments were conducted: one in which samples for probe preparation were taken every 10 min from 0 to 200 min and one in which samples were taken every 20 min from 0 to 60 min. The collection of samples took place over several months in four sets (0–50 min, 60–100 min, 110–150 min, and 160–200 min). At the same time that samples for TSAA were taken, aliquots of cells were also removed for bud counts and total RNA isolation.

α-Factor Synchronization—Cells were grown to a density of 5–8 × 10⁶ cells/ml at which point α-factor was added. The cells were fully arrested after 2 h. To release cells from arrest they were centrifuged for 2 min at 3000 × g and resuspended in 400 ml of fresh medium. Two separate experiments were conducted: one in which samples were taken every 10 min from 0 to 200 min and one in which samples were taken every 20 min from 0 to 60 min. The collection of samples took place over several months in four sets (0–50 min, 60–100 min, 110–150 min, and 160–200 min). At the same time that samples for TSAA were taken, aliquots of cells were also removed for bud counts and total RNA isolation.

Polysome Preparation and Fractionation—All steps took place in a 4 °C cold room. To prepare polysomes, ~100 ml of culture was poured into a flask containing ~75 ml of crushed frozen YEPD. The flask was swirled rapidly on ice for ~2–3 min and then centrifuged for 2 min at 3000 × g to pellet the cells. The pellet was resuspended in 10 ml of lysis buffer (10 mM Tris-HCl, pH 7.5, 100 mM NaCl, 30 mM MgCl₂, 500 μg/ml heparin, 50 μg/ml cycloheximide) and centrifuged a second time. The pellet was resuspended in 1 ml of lysis buffer and transferred to a 15-ml conical tube containing 2 g of glass beads (~500-μm diameter, Sigma). The tubes were vortexed eight times for 30 s each. Tubes were placed on ice for 30 s between rounds of vortexing. Tubes were centrifuged for 1 min in a clinical centrifuge to pellet the beads, and the supernatant was removed. Lysis buffer (0.5 ml) was added to the beads, the beads were vortexed for 15 s, and the tube was centrifuged again to pellet the beads. The supernatant was removed and added to the first supernatant. The pooled supernatants were centrifuged in a microcentrifuge at top speed (20,000 × g) for 1 min to pellet debris, and the supernatant was transferred to a fresh tube. The A₂₆₀ of the supernatant was measured. For the cdc15 experiments 50–60 absorbance units were loaded per gradient. For the α-factor experiments, 25 absorbance units were loaded per gradient.

The samples were loaded on a 10-ml linear gradient of 7–47% sucrose with a 1-ml 2 m sucrose cushion at the bottom and centrifuged for 2 h in an SW40 Ti swinging bucket rotor at 4 °C and 39,000 rpm. The gradient buffer consisted of 50 mM Tris acetate, pH 7.0, 50 mM NH₄Cl, 12 mM MgCl₂, and 1 mM dithiothreitol. After centrifugation, the gradients were collected into 24 0.5-ml fractions. Fifty microliters of 3 M NaOAc, pH 5.2, and 50 μl of 10% SDS were added to each fraction. During fractionation a running trace of A₂₆₀ was recorded to monitor the polysome profile.

Preparing and Labeling RNA for TSAA—The absorbance trace for each gradient was used to identify the demarcation between fractions containing polysomes with five ribosomes or fewer (the poorly translated pool) and fractions containing polysomes with six or more ribosomes (the well translated pool). These two populations were pooled separately, and the RNA was purified with 2.5 volumes of ethanol and subsequently purified using Qiagen RNeasy midikits following the manufacturer’s protocol. After elution into a volume of 600 μl of DH₂O, the RNA was precipitated by the addition of LiCl to a final concentration of 1 M and resuspended in 100–200 μl of DH₂O. RNA concentration was determined, and the samples were stored at −80 °C.

RNA was reverse transcribed to labeled cDNA with Cy3- and Cy5-labeled dCTP dyes from Amersham Biosciences. Eighty micrograms of RNA was combined with 7 μg of (dT)$_{24}$ with a G/C/A 3’ anchor. The reaction was brought up to a volume of 42 μl with water, heated to 70 °C for 5 min, and placed on ice for 5 min. To the reaction were added 16 μl of 5× reaction buffer (Invitrogen), 8 μl of dithiothreitol (Invitrogen), 4 μl of 20× nucleotides (10 mM dATP, dGTP, and dTTP, 0.5 mM dCTP), 2 μl of RNasin (Promega), and 2 μl of Cy3 or Cy5 dCTP (Amersham Biosciences). The reaction was mixed, centrifuged briefly, and placed at 42 °C for 2–10 min after which 4 μl of SuperScript II reverse transcriptase (Amersham Biosciences) was added.

The samples were incubated at 42 °C for 2 h. Eight microliters of 1 M NaOH, 10 mM EDTA was added to hydrolyze the template, and the reaction was incubated at 42 °C for a further 20 min. The labeled cDNA (target) was purified with a GFX spin column (Amersham Biosciences) and eluted in a total volume of 120 μl (elutions with 50 μl of Tris-EDTA, pH 8.0, 50 μl of DH₂O, and 20 μl of DH₂O successively). The target was centrifuged for 5 min to remove contaminant resin, A₂₆₀ of the supernatant was measured, and either A₅₅₀ (for Cy3) or A₅₅₀ (for Cy5) was used to calculate yield and specific activity.

Slide Hybridization and Quantitation—Custom yeast high density genomic open reading frame (ORF) microarrays were fabricated by the Center for Expression Array Analysis at the University of Washington. Each slide contained two copies of the yeast genome; two slides were hybridized for each time point. For the cdc15 experiment, RNA isolated from the well translated pool was always converted to target with Cy3, and RNA from the poorly translated pool was always converted to target with Cy5. For the α-factor experiment we performed a dye swap experiment in which the cDNA from the well translated pool was labeled with Cy3 for one slide and with Cy5 for the second slide with the corresponding change in labeling of the cDNA from the poorly translated pool.

For each hybridization, labeled cDNA containing 20 pmol of Cy3 and 12 pmol of Cy5 were combined, dried down, and resuspended in 35 μl of hybridization buffer (2.5× SSC, 2.5× Denhardt’s reagent, 0.25% SDS, 50% formamide). Slides were dipped rapidly 10 times in nanopure DH₂O and dried with a stream of air. The mixed target was heated to 100 °C for 3 min, quick chilled on ice for 30 s, and then centrifuged briefly. The probe was added to the slide, covered with a
coverslip, placed in a microarray hybridization slide chamber (Corning), and wrapped in foil. The slides were hybridized for 16 h at 42 °C and then washed. Washes consisted of three 5-min washes at 50 °C: 2× SSC, 0.5% SDS; 0.1× SSC, 0.5% SDS; and 0.1× SSC. Slides were then dipped two times in nanopure dH2O and dried with a stream of air.

Slides were scanned at the Center for Expression Array Analysis using an Amersham Biosciences GenII scanner. The resulting images were quantitated using a custom software package available from the Center for Expression Array Analysis.

Data Normalization and Analysis—The TSAA ratio is computed as the ratio of the signal from the well translated pool to the poorly translated pool.

We assumed that the translation state would change for only a small fraction of the genes during release and cell cycle progression and that the overallallocation of mRNA in the well and poorly translated pools, as well as the distribution of the TSAA ratios themselves, would change little during cell cycle progression. Therefore, we first normalized the logarithmically transformed intensity values derived from the well translated and poorly translated pools at each time point using a regression model (11). The sum of the back-transformed normalized ratios or transcript level and computed the corresponding standard errors with estimating equations. The Z scores of the altered level for each gene were calculated by dividing the estimates with their corresponding standard error. The goodness-of-fit of the single-pulse model was measured with R2, which was adjusted by the maximum R2 attained with replicates at each time point. A gene was considered to show a significant cell cycle-dependent pattern, either in TSAA ratio or in transcript level, if the absolute value of the corresponding Z was above 5 and the corresponding R2 was above 0.5. We also combined the α-factor data sets and cdc15 data sets that correspond to the same time point during cell cycle progression and fit the combined data sets with the single-pulse model to identify genes that show similar cell cycle-dependent patterns in both data sets. Data were clustered and visualized using Cluster and Treeview, free software available from rana.lbl.gov/EisenSoftware.htm.

Confirming TSAA Results—Polysomes were prepared from synchronized cell cultures in the same way as described above under “Polysome Preparation and Fractionation.” At the fraction collection step, however, the gradient was fractionated into 12 fractions rather than 24, and the fractions were not pooled. Instead an equal volume (300 μl) of each fraction was precipitated with ethanol. Total RNA was isolated from each aliquot using Tri reagent (Molecular Research Center, Inc.). For Northern analyses, the RNA was electrophoresed, blotted, and probed using the method found at www.fhcrc.org/labs/breeden/Methods/Northern_Blot.html. Probes were synthesized by designing primers to amplify short (200–300-bp) regions of each gene, using PCR on genomic DNA to amplify that fragment, and then gel-purifying the fragment using the Qiaquick gel extraction kit (Qiagen). Probes were labeled with [32P]dCTP using Stratagene’s RmT Random Prime Labeling kit. Intensities of the hybridization signals were quantitated using an Amersham Biosciences Storm PhosphorImager.

For quantitative PCR, RNA samples were isolated from polysome fractions using the RNeasy minikit (Qiagen). An equivalent volume of each sample was reverse-transcribed and examined with quantitative PCR using the Bio-Rad iCycler system according to the manufacturer’s instructions.

Assessing Synchrony—The samples taken for bud counts were immediately adjusted to 0.2% sodium azide. The samples were sonicated to separate cell clumps and then examined under a compound microscope for signs of bud emergence.

RNA was isolated from the samples taken for total RNA using the hot acid-phenol method (14). The RNA was electrophoresed and blotted as described above and probed with radiolabeled CLN2 and ACT1 fragments. The resulting pattern of CLN2 expression was normalized using the ACT1 signal and then compared with the bud counts to confirm that the cultures used for TSAA samples had been synchronized for the cell cycle.

β-Galactosidase Activity Assays—Strain BY2125 was transformed with either the pMHY1 or pMHY3 constructs (15). These constructs contain the GCN4 leader cloned 5’ of the β-galactosidase gene. All endogenous upstream open reading frames (uORFs) were removed from the GCN4 leader, and either the wild type (pMHY1) or an altered (AUG → GUG, pMHY3) S-adenosylmethionine decarboxylase uORF was cloned into the leader (for construction details see Ref. 15). These strains were grown in the absence of uracil to A600 = 1.0 and then harvested for preparation of polysomes and total RNA as described above. The activity of β-galactosidase was measured in the same extracts and normalized to cell mass based on A600 of the culture.

Total β-galactosidase transcript levels per cell as well as mRNA distribution within the polysome gradient were measured by quantitative PCR with the Bio-Rad iCycler system according to the manu-
facturer’s instructions. Measurements of β-galactosidase transcripts were normalized to levels of the endogenous transcripts from DED1. Measurements of β-galactosidase transcripts within polysome gradient fractions were normalized to a control, synthetic RNA (Utility Control 2, Amersham Biosciences), that was added in equal amounts to each fraction prior to RNA extraction.

RESULTS

Comparison of TSAA Ratio to Translational Efficiency—To validate the TSAA approach, a direct comparison is needed between TSAA ratios and rates of synthesis of known proteins. This comparison cannot be made with steady-state protein levels, however, because stabilities can vary between individual proteins and between different physiological states. We have taken advantage of a well characterized expression system in which the same protein, β-galactosidase, is synthesized from transcripts with different efficiencies of translation (15). Two constructs were examined: pMHY1 in which translation of the β-galactosidase mRNA is repressed by an inhibitory uORF and pMHY3, which has an ~20-fold elevated translational efficiency due to a mutation in the initiator codon of the uORF (15).

Polysome distribution, total mRNA levels, and β-galactosidase activities were measured for yeast lines carrying these two constructs. Cultures of the line carrying pMHY3 showed efficient ribosome loading onto the β-galactosidase transcript because of the absence of a functional uORF (Fig. 1A); fraction 11 in the sucrose gradient corresponds to >10 ribosomes per mRNA based on the absorbance profile (not shown). In contrast, the line carrying pMHY1 showed much more dispersion, and on average lower, ribosome loading than the pMHY3 line. As a control, we analyzed the same polysome gradients for the distribution of another transcript, DED1 (Fig. 1B). In contrast to the β-galactosidase mRNA, DED1 shows no difference between the two lines, consistent with a transcript-specific difference in ribosome loading between pMHY1 and pMHY3.

In TSAA, the polysome distribution is split into two halves: those transcripts with five or fewer ribosomes and those with six or more ribosomes (between fractions 1–7 and 8–12 in this specific experiment). Carrying out this operation with the β-galactosidase experiments and summing the signal from both pools allows us to calculate the TSAA ratio for each construct. The transcripts from pMHY1 and pMHY3 give log₂-transformed TSAA ratios of ~0.49 and 1.20, respectively (Fig. 1C), which correspond to over a 3-fold numerical change in ratio.

For the same cultures, we measured β-galactosidase activities and the levels of the mRNAs. The activity of the same protein was measured in both cultures, protein half-lives do not enter into the calculation, and the steady-state activities should be proportional to the rates of synthesis from the two mRNAs. Dividing the enzyme activities by the relative mRNA levels yields translational efficiencies, which differ by 29-fold between the two constructs (Fig. 1C). Comparing this result with the 3-fold value from TSAA reveals that the TSAA ratio provides a qualitative, but not quantitative, indication of where a message is distributed within the polysome gradient relative to the position of the split (between five and six ribosomes per message). Because of the qualitative nature of the TSAA ratio, its utility lies as a gene-finding tool. The specific relationship between TSAA ratio and the rate of synthesis of the encoded protein depends on the properties, in particular ORF length, of the transcript in question (see “Discussion”).

General Characteristics of TSAA Data—Growing cultures of S. cerevisiae were blocked at different points in the cell cycle either through use of the temperature-sensitive cdc15-2 mutation or by treatment with α-factor. The cdc15-2 allele arrests cells at the restrictive temperature in the telophase of the cell cycle, while α-factor causes an arrest in mid-G_1 phase (16, 17)
After release from the arrest, samples were taken at 10-min intervals for TSAA.

The relationship between TSAA ratio and ORF length for the 1000 most highly expressed genes (based on Cy labeling and hybridization) is displayed in Fig. 2. There is a clear positive correlation between these parameters as would be expected if TSAA ratios accurately reflect the degree of ribosome loading per mRNA molecule. Longer ORFs can accommodate more ribosomes and therefore should have, on average, higher TSAA ratios because more of their mRNAs should be in the well translated pool (six or more ribosomes). A few ORFs show a much higher TSAA value than expected given their size (Fig. 2, upper left quadrant). When these were examined in more detail they were all found to be hypothetical ORFs existing on the opposite strand of longer genes. Given that the microarrays contained both strands for each ORF, it is likely that these high TSAA values are due to expression of the longer rather than the shorter ORF.

The relationship between TSAA ratio and transcript level is explored in Fig. 3A. The yeast genes were ranked by their average expression level and plotted against their TSAA ratios. In general no strong trends were seen in the distribution of TSAA ratios, although a small bias toward higher TSAA ratios was seen for very well expressed genes, perhaps indicating that very highly expressed genes are also optimized for translation. These data suggest that transcript level does not have a profound effect on how efficiently an mRNA is translated. In Fig. 3B, the genes are ranked by expression level and plotted against the coefficient of variation (S.D./mean) of the TSAA ratios. For the majority of genes, coefficients of variation averaged 0.2, suggesting that replicate measurements generally were robust.

By taking a ratio between the well and poorly translated pools, TSAA should detect changes in translation state irrespective of changes in transcript levels. To exemplify this, the log_{10}-transformed intensity values from both pools for two representative genes from the α-factor data set are plotted versus time in Fig. 4. CLN2 is regulated at the transcriptional level during the cell cycle but does not undergo translational regulation, and the intensity values from both pools change in parallel (Fig. 4A). CTS1, in contrast, does undergo translational regulation in the 1st h following release from arrest, and that can be seen from the divergence and convergence of the intensity values over the first 60 min (Fig. 4B). Thereafter the
intensity values for the two pools change in synchrony.

*Estimating Transcript Levels from TSAA*—Transcript level as well as translation state can be obtained from TSAA data sets. Steady-state transcript levels can be extracted from the data by summing the signal from the well and poorly translated pools. To examine how the mRNA levels obtained from TSAA compare with those generated from conventional array analysis, the temporal expression profiles of the 800 cell cycle-regulated genes from a previous analysis (7) were compared with values obtained from the TSAA data sets. A visual representation can be seen in Fig. 5 where the genes have been ordered by timing of expression during the cell cycle. The overall pattern shows good agreement between the data sets. To provide a statistical measure of the agreement between the two methodologies, the correlation coefficient was calculated for the expression values for each gene. At the 80% confidence limit ($p < 0.2$), both arrest methods showed good agreement: for the $\alpha$-factor data set, 58% of genes exceeded the confidence limit; for the cdc15 data set, 42% of genes surpassed the limit. This degree of congruence is of the same magnitude as that observed when different analysis methodologies are applied to the same cell cycle data (13).

To examine the behavior of several well characterized cell cycle genes, the transcript levels for the cyclin genes were extracted from the TSAA data sets and plotted as a function of time (Fig. 6). As expected, periodic increases and decreases in the transcripts of the G$_1$ cyclins (CLN1, CLN2, and CLN3, Fig. 6A), and the G$_2$/M cyclins (CLB5 and CLB6, Fig. 6C) preceded expression of the G$_2$/M cyclins (CLB1 and CLB2, Fig. 6B). These data, taken together, demonstrate that estimates of steady-state mRNAs levels, as well as translation state, can be reliably extracted from TSAA results.

*Responses of mRNA Levels and Translation to Release from Cell Cycle Arrest*—Time points immediately following the release from cell cycle arrest were examined for changes in translation that occurred either after removal of $\alpha$-factor from mating type a cells or after shifting cdc15-2 cultures from 37 °C to 25 °C. Each data set was analyzed using a regression-based modeling methodology designed for time course studies (12). This methodology models changes in microarray data over time as either a linear or a quadratic function, identifying shifts in expression that fit the model. Thus, genes that either increase or decrease, or experience a transient change in translation state, will be identified. To locate as many candidates as possible the analysis was performed over two subsets of data for each synchrony method: the first six time points to identify rapid changes and the first 12 time points to identify more gradual changes in translation state. The stringency threshold for this analysis was selected to give a NFD $< 1$; that is, we would expect less than one false positive among the selected subset of genes. This cut-off likely discards a number of true positives, but for purposes of the current study we have focused on the most likely candidates. The full data set and statistical scores are available at the website tsaa.info/

Initially 85 genes in the $\alpha$-factor data set and 65 genes in the cdc15 data set met the stringency criterion for at least one of the time periods and modeled behaviors. A second, shorter (four- to six-time point) microarray experiment was performed for each type of arrest, and genes that did not behave similarly in both data sets as assessed by visual inspection of the data were discarded from the analysis. Comparing the two data sets for each treatment, 48 and 54 genes in common were identified for the $\alpha$-factor and cdc15 data sets, respectively.

The translation states and transcript levels for these pools were subjected to hierarchical clustering (18) to group together genes that behaved similarly. The patterns of change in ribosome loading and transcript level for these genes are summarized in Figs. 7 and 8. In these clustergrams, the translation states and transcript levels over time for each gene have been translated into color values (yellow for well translated or well expressed relative to the mean and blue for poorly translated or poorly expressed relative to the mean). The TSAA ratio data for the two independent experiments performed for each type of arrest and release are presented side-by-side to allow a visual comparison between the two. There is strong agreement between the two independently
derived data sets for each mode of cell cycle arrest. Comparing the patterns in Figs. 7 and 8, some overall differences in translational response to release from the two types of arrest are evident. After release from α-factor arrest, roughly equal proportions of the mRNA species showed increases and decreases in ribosome loading (Fig. 7). In contrast, after the temperature shift, greater than 80% of the genes decreased in translation after release (Fig. 8). At the present level of resolution (10-min time points), an examination of the behavior of TSAA ratios revealed no great complexities in temporal regulation of translation over time. After release from either mode of arrest the reductions and increases in ribosome loading seemed to behave as single kinetic classes.

We confirmed the changes seen in TSAA through Northern blot or quantitative PCR analysis of polysome fractions prepared in the same manner as our original probes. Testing of 22 genes, including negative controls, yielded a better than 90% agreement rate with our TSAA data and statistical analysis. Representative distributions across polysome gradients of two genes from each mode of synchronization are shown in Fig. 9. Ribosome loading of two transcripts, SST2 and ARO4, was down-regulated after release from the α-factor and cdc15 blocks, respectively. OLE1 mRNA became better loaded with ribosomes after the temperature shift, while the behavior of CTS1 upon α-factor release was complex. CTS1 first became better loaded at 20 min after release, and then ribosome loading declined by 60 min. This behavior of CTS1 was seen in both sets of α-factor TSAA ratios (Fig. 7).

The genes identified in the two data sets were assigned to functional categories based on the classification in the Munich Information Center for Protein Sequences database (Fig. 10). The likelihood ratio test (19) was used for each category to identify functional classes that differed from expectation based on the frequency of those classes in the yeast genome as a whole. The under- or over-represented categories for the α-factor data set included Protein Fate, Regulation of/Interaction with the Cellular Environment, Transcription, and Unclassified Proteins. For the cdc15 data set, the under- and over-represented categories included Energy, Metabolism, Protein Synthesis, Transcription, and Unclassified Proteins for the temperature shift data set.

The first two categories of α-factor-regulated genes (Protein Fate and Regulation of/Interaction with the Cellular Environment) were over-represented relative to expectations based on the frequency of these genes in the genome as a whole (Fig. 10). This over-representation is in part due to the translational regulation of genes involved in pheromone response, including BAR1, SST2, and FAR1, which fall into one or both of these categories (under the Munich Information Center for Protein Sequences classification scheme, genes

![Fig. 5. Transcript levels from TSAA.](image)
can belong to more than one functional category). The pheromone response genes decrease in translation state after removal of \( \alpha \)-factor, correlating with re-entry into the cell cycle.

For the temperature shift experiment, the categories of Energy, Metabolism, and Protein Synthesis showed significant over-representation relative to expectations. In parallel with overall patterns of translational change after the temperature shift (Fig. 8), the majority of genes in each of these categories underwent reductions in translation state (Fig. 10). Within the categories of both Energy and Metabolism, genes encoding enzymes involved in glycolysis (\( \textit{ENO1}, \textit{ENO2}, \textit{FBA1}, \textit{PGK1}, \textit{TDH1}, \textit{TDH2}, \textit{TDH3}, \text{and} \textit{TPI} \)) predominate. The Protein Synthesis genes included both ribosomal subunits (\( \textit{RPS2}, \textit{RPL4A}, \text{and} \textit{RPL8A} \)) and translation initiation and elongation factors (\( \textit{TIF1}, \textit{TIF2}, \text{and} \textit{TEF2} \)). While not belonging to a significant category, another interesting class of genes whose functions correlated with their change in translation state is the heat shock and stress response genes, for example \( \textit{SSA2}, \textit{HSC82}, \text{and} \textit{HSP82} \), which demonstrated reduced ribosomal loading after the shift to the lower temperature.

For both data sets, the Transcription and Unclassified Pro-
tein categories were under-represented. Given how well studied pheromone response and temperature effects are in yeast, it was perhaps anticipated that fewer Unclassified Proteins than expected were found. Finding Transcription genes to be under-represented was more unexpected. Transcription plays a major role in both the pheromone and temperature responses (20, 21), but the transcription factors that effect the changes in transcription during these responses do not appear to be under translational regulation.

Examining transcript levels and translation states together produced interesting patterns of global regulation. For regulated genes, changes in ribosome loading and transcription often occur in parallel. Following the removal of α-factor, several genes showed parallel changes in ribosome loading and transcript level (Fig. 7, vertical red bars). Following the temperature shift, several genes decreased in ribosome loading and transcription (Fig. 8, vertical red bars). Parallel regulation of genes on both the transcriptional and translational level suggests coordinated control of gene expression. However, this kind of coordinated control is not a general feature of gene regulation since there are also numerous examples of genes for which transcription and translation do not change in parallel (Figs. 7 and 8).

The TSAA data can also be examined through a functional rather than statistical approach. Translation states for two functional groups of genes, the heat shock genes and the pheromone response genes, are organized through hierarchical clustering in Fig. 11. One benefit of examining data in this way is to allow identification of potential instances of translational regulation that may not have passed the initial rigorous statistical standards but might be biologically relevant. For the heat shock genes, seven new genes in addition to the three already identified appeared to be decreasing in ribosome loading after the shift to the lower temperature (HSP78, HSP104, SSA4, HSP60, HSP26, HSP12, and SSE2). Of the pheromone response genes, four previously unidentified genes appeared to decrease in ribosome loading after the removal of α-factor (FUS3, GPA1, STE2, and AFR1).

Responses of mRNA Translation to Cell Cycle Progression—Samples for TSAA were gathered out to 190–200 min after release from arrest to test whether any genes fluctuated in ribosome loading as a function of the cell cycle. To identify genes that were translationally regulated with progression through the cell cycle, we used the single-pulse model regression modeling methodology (13). As with our analysis of arrest-specific genes, this technique models the expected behavior of a gene and examines how well that gene fits the model. This analysis revealed only 10 genes that appeared to cycle in translation state during the cell cycle.
Of the six genes we examined in greater detail, only two (MNN1 and GAS3) could then be confirmed by Northern analysis. These data suggest that translational regulation does not play a major role during cell cycle progression.

**DISCUSSION**

**Studying Translation with Microarrays**—One of the goals in developing TSAA was to create a robust, high throughput tool for the discovery of genes undergoing regulation at the translational level. The current study achieved this goal in that over 100 genes were identified as potentially undergoing translational regulation under the two physiological transitions studied. These measurements have proven to be quite robust, showing good reproducibility for most genes, especially those at a median level of expression and higher. Even in experiments performed on different days by different researchers the translation state ratio generally proved to be very similar for each gene.

In the formulation of TSAA used in this study, polysome gradients were divided into two pools from which RNA was extracted and labeled for microarray hybridization. After normalization of the intensities, the ratio of the well translated to the poorly translated signal was calculated for each gene. The behavior of this translation state ratio in genome-wide data sets was as expected based on our current understanding of translation mechanisms. Shorter messages, for example, tended to have ratios less than 1 due to physical limitations on numbers of ribosomes loaded on the ORF. Messages with longer ORFs, in contrast, were more disperse in their translation state ratios, suggesting diversity of ribosome spacing on different mRNA species with similar ORF lengths. In addition, significant changes in TSAA ratio over time, or in response to environmental changes, generally corresponded to a visible shift in distribution of that mRNA in the polysome gradient as judged by independent analysis.

Although the current formulation of TSAA provides a powerful gene discovery tool, utilizing the derived translation state ratios as quantitative estimates of gene expression into protein is more problematic. Given two mRNA species of the same ORF length, a large difference in TSAA ratio would generally suggest that one mRNA is loading ribosomes, and thus being translated, at a higher rate than the other (22). Although this qualitative assessment was shown to be valid in this study, a quantitative interpretation of differences in this ratio is not easily made. This is because of the loss of information that results from dividing the polysome gradient into only two pools. For example, in the current set of experiments, all messages with five ribosomes or fewer were placed
in the poorly translated pool, which fails to take into account that smaller messages (500–600 bases or less) could actually be loaded to the physical maximum with just five ribosomes. These considerations also limit the efficacy of TSAA as a gene-finding tool since shifts in ribosome loading that result in an mRNA staying within a single pool would be undetected. There are modifications of the current methodology that could provide a more accurate assessment of ribosome loading and thus more accurate comparisons between genes. Partitioning the gradient into more pools and interrogating arrays with each pool would provide a much more detailed picture of message distribution and movement within the polysome gradient. The greater the number of pools, the more accurate the representation of precise polysome distributions for every mRNA in the cell will be. In the extreme, such a “high resolution” TSAA approach would require considerable computational and experimental resources but would provide a rich description of mRNA translation under a given cellular state. However, as a gene-finding tool, expanding the current approach to three or four pools across the sucrose gradient would considerably decrease the number of false negatives.

**Changes in Ribosome Loading after Release from Cell Cycle Arrest**—When cells were released from arrest either by removal of α-factor or through temperature shift, a significant number of transcripts responded with a change in translation state. Most of the genes regulated in this way seemed to be in response of the cells to the exogenous stimuli rather than to re-entry of the cells into the cell cycle. Regulation at the translational level is clearly suited to controlling rapid fluctuations in levels of key proteins in response to environmental changes.

The identity of many of the genes that changed in translation state after release from arrest followed directly from the particular method of arrest used (α-factor or temperature shift). In the α-factor experiment, a number of genes involved in pheromone response and mating underwent translational down-regulation after removal of α-factor, including AGA1, BAR1, SST2, FAR1, and PRM1. AGA1 functions in agglutination of mating cells (23), and BAR1 and SST2 are involved in...
modulating the α-factor response (24, 25). FAR1 inhibits the CDC28 kinase in a-type cells exposed to α-factor (26). PRM1 is a pheromone-regulated gene that codes for a membrane protein that facilitates plasma membrane fusion (27). For all of these genes, protein levels would be expected to decrease following removal of α-factor, and translation seems to be a major site of regulation. The genes increasing in translation state after release included a few involved in cell cycle processes, including RSC2, part of the chromatin-remodeling complex (28).

Within the temperature shift data set, several heat shock genes underwent rapid translational down-regulation after the shift to 25 °C. Heat shock genes have been found to be translationally regulated in other systems including Drosophila melanogaster, Leishmania, and humans (29–31). Finding control of the heat shock genes at the translational level in yeast as well as in these other widely diverse organisms underscores the importance for cells of tight control of the stress response.

Another functional group that displayed decreased ribosome loading following the temperature shift contained genes involved in glycolysis and gluconeogenesis (ENO1, ENO2, FBA1, PGK1, TDH1, TDH2, TDH3, and TPI1). Like the heat shock genes, transcript levels of these genes also decreased following a shift from a higher to a lower temperature in parallel with genes from several other classes (oxidative stress, starvation, etc.) (20). These genes have been suggested to be part of an environmental stress response (20). Only a subset of the 283 environmental stress response genes show translational changes in parallel with the transcriptional changes (data not shown), a subset that includes the heat shock and glycolytic genes. The remainder of the environmental stress response genes shows no change in translation state.

**Correlation between Transcript Level and Translation State**—From these data, it appears that there is often, but not always, a correlation between changes in ribosome loading and changes in transcript levels. Combining these two levels of regulation could provide large, coordinated changes in protein levels for some genes in response to certain environmental changes. Alternatively, in the case of some genes, temporal differences in the two modes of regulation could provide either fine tuning of expression or differently timed waves of protein expression. For example, in mammalian cells transcription and translation of the immediate response gene nur77 seem to have different temporal patterns in fibroblasts responding to growth factor (2).

One possibility for a mechanism linking transcript level and translation efficiency could relate to the average age of the population of particular transcripts. As mRNAs age, their poly(A) tails, which are required for optimal translation, become shorter (32, 33). When transcription of a gene increases, the average age of the resulting mRNA population becomes younger, and therefore the mRNAs are possibly better translated. The converse would be true when transcription of a gene is down-regulated. This scenario cannot uniformly hold, however, since not all transcripts that change in level show concomitant changes in translation state. The lack of cell
cycle-regulated changes in translation is one particularly strong example. However, in addition to the putative influence of transcript age, there are many other factors that could mediate coordinated changes of translation state of classes of genes. For example, there are several classes of cis-acting elements that regulate translation, including uORFs, internal ribosome entry sites, secondary structure, and binding sites for regulatory proteins (34–37). It may be that classes of genes that are translationally co-regulated contain conserved motifs of these types that contribute to their control.

Translational Control in the Cell Cycle—Surprisingly few genes appear to change in translation state during the yeast cell cycle. Only two genes were confirmed as showing clear cyclic changes in ribosome loading in both the cdc15 and α-factor synchronized cells. Genes by gene studies in other organisms have documented several examples of cell cycle-regulated changes in translation, including the discovery of cell cycle-dependent internal ribosome entry sites (38, 39) and the cell cycle-dependent accumulation of p27Kip1 in mammalian tissue culture cells (40). One possible explanation for the lack of translational regulation in S. cerevisiae may lie in differences between the cell cycle in budding yeast versus other systems. Many of the reported instances of translational control occur in the G2 phase of the cell cycle in mammalian systems and Schizosaccharomyces pombe. In those systems the G2 phase is longer and has a greater role in cell cycle regulation than G2 in S. cerevisiae, which may account for the general lack of translational control during the cell cycle for S. cerevisiae.

Modeling the Cell—The development of high throughput genomic and proteomic tools has had a profound effect on how we approach and analyze biological problems and systems. These tools provide a quantitative leap in data gathering, generating huge, sometimes daunting, collections of information about the behavior of various molecular components of the cell. The questions surrounding these new technologies concern how best to utilize and synthesize the copious amounts of data to provide a qualitative change in our understanding of the workings of the cell.

In the field of gene expression, it seems that such a synthesis is within reach. Technologies now exist that can monitor several different aspects of how a gene is expressed besides transcript level and translation state, including splicing (41), mRNA decay (42), mRNA subcellular localization and sequestration (43, 44), and steady-state protein levels (45, 46). It seems likely that, as analysis methods evolve and further experiments are performed, it will be possible to develop an integrated model that predicts and describes the behavior of every gene product from the time an RNA is first synthesized to the point at which it and its respective proteins are broken down and relegated back to the building blocks from which they came.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

†† To whom correspondence should be addressed: Dept. of Biochemistry, Box 357350, University of Washington, Seattle, WA 98195. Tel.: 206-543-1694; Fax: 206-543-4822; E-mail: dmorris@ u.washington.edu.

REFERENCES
20. Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B.,
mRNA Translation during Recovery from Arrest


