

Supplemental Materials

Molecular Biology of the Cell

Kelliher et al.

Supporting Information Methods to Accompany:

Layers of regulation on cell-cycle gene expression in the budding yeast *Saccharomyces cerevisiae*

Christina M. Kelliher, Matthew W. Foster, Francis C. Motta, Anastasia Deckard, Erik J. Soderblom, M. Arthur Moseley, and Steven B. Haase

RNA-Sequencing data normalization

RNA-Sequencing data for wild-type cells were taken from a previous study [1], and data for *clb1-6* mutant cells were generated in this study. All raw FASTQ files were aligned to the *S. cerevisiae* S288C reference genome (Ensembl build R64-1-1, downloaded in March 2016) using STAR [2]. Reads mapping uniquely to annotated yeast genes were counted using HTSeq-count [3]. RNA-Seq mapping statistics are shown in the table below. In wild-type cells, the average total reads from the time series was ~18.5M. In *clb1-6* cells, the average total reads from the time series was ~23.1M. The average reads mapping uniquely to the yeast genome was ~16M (87%) in wild type and ~20.6M (89%) in *clb1-6* mutants. The average reads mapping uniquely to an annotated gene in the reference genome was ~14.4M (78%) in wild type and ~16M (69%) in *clb1-6*.

<i>S. cerevisiae</i> Experiment	Average Total Reads per TP	Average Unique Reads per TP (STAR)	Average Feature Reads per TP (HTSeq)	Average Percent Mapped (%)	Average Mapped Uniquely (%)	Average Mapped Feature (%)
wild type	$1.85 \times 10^7 \pm 2.51 \times 10^6$	$1.60 \times 10^7 \pm 2.22 \times 10^6$	$1.44 \times 10^7 \pm 2.06 \times 10^6$	97.86 ± 0.31	86.64 ± 1.24	77.67 ± 1.31
<i>clb1-6</i>	$2.31 \times 10^7 \pm 2.17 \times 10^6$	$2.06 \times 10^7 \pm 2.07 \times 10^6$	$1.60 \times 10^7 \pm 1.76 \times 10^6$	96.29 ± 2.15	89.35 ± 2.21	69.42 ± 2.43

Transcript quantification of annotated yeast genes was performed using Cufflinks [4]. Time point samples from the two experiments were normalized together using the CuffNorm algorithm. Normalized fpkm gene expression values (fragments per kilobase

of transcript per million mapped reads, or fpkm units, taken from “genes.fpkm_table”) were used in the analyses presented. To avoid fractional and zero values, 1 was added to every fpkm value in each dataset using the R statistical programming environment [5]. Raw RNA-Sequencing data from wild-type cells are available at the NCBI Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE80474 [1]. Raw RNA-Seq data from *clb1-6* cells from this manuscript are available under accession number GSE104904.

CLOCCS aligning *S. cerevisiae* time series experiments

In order to compare RNA and protein curve shapes, we needed to align time series experiments on a common cell-cycle timeline (Figures 1-5). For all experiments, *S. cerevisiae* cells were synchronized using alpha-factor mating pheromone, and budding index data from wild-type and *clb1-6* were fit using the CLOCCS model [6,7]. At least 200 cells were counted in RNA-Seq experiments (1 replicate per condition) and in mass spectrometry experiments (2 biological replicates per condition). The CLOCCS fit parameters are given in the table below for each experiment, which contains the mean value and 95% confidence interval (in parentheses) for each model parameter. The mean values for cell-cycle period (λ) and recovery time (μ_0) were used to align the wild-type and *clb1-6* time series by converting time points to scaled CLOCCS lifeline points, as described previously [1].

Experiment	μ_0	δ	σ_0	σ_v	λ	β
wild-type, RNA-Seq, replicate 1	31.05 (27.44, 35.86)	11.59 (5.17, 20.56)	13.20 (12.10, 14.28)	0.10 (0.07, 0.12)	68.78 (62.64, 73.09)	0.24 (0.19, 0.28)
wild-type, proteomics, replicate 1	33.14 (30.09, 36.36)	9.75 (4.58, 15.59)	11.97 (10.78, 13.08)	0.08 (0.06, 0.10)	62.92 (59.22, 66.30)	0.24 (0.20, 0.27)
wild-type, proteomics, replicate 2	32.27 (28.78, 36.22)	10.30 (4.68, 17.40)	13.27 (11.83, 14.59)	0.09 (0.07, 0.12)	63.56 (58.68, 67.68)	0.25 (0.20, 0.28)
<i>clb1-6</i> , RNA-Seq, replicate 1**	24.84 (23.08, 26.62)	NA	9.44 (8.15, 11.83)	0.21 (0.19, 0.23)	108.87 (106.83, 111.05)	0.15 (0.15, 0.15)
<i>clb1-6</i> , proteomics, replicate 1	29.24 (27.47, 31.01)	NA	9.52 (8.77, 10.98)	0.22 (0.21, 0.24)	102.36 (100.41, 104.49)	0.15 (0.15, 0.15)
<i>clb1-6</i> , proteomics, replicate 2**	24.84 (23.08, 26.62)	NA	9.44 (8.15, 11.83)	0.21 (0.19, 0.23)	108.87 (106.83, 111.05)	0.15 (0.15, 0.15)

** indicates that yeast samples were derived from the same time series experiment

Parallel Reaction Monitoring (PRM) targeted mass spectrometry method development on asynchronous yeast cells

We initially targeted 48 yeast proteins of interest (including 4 controls for constitutive expression levels: Cic1, Rim11, Taf12, and Vps9). When possible, three different heavy labeled peptides were obtained to quantify expression levels of the target proteins (Supplementary Table 1). PRM method development was applied in triplicate to a sample of total protein from asynchronous wild-type yeast cells (technical replicates, data not shown). For the subsequent time series experiments, we decided to target only 2 peptides for most proteins in order to avoid crowded retention time windows in the PRM method. This finding about retention time “rush hours” was an unexpected obstacle to consider for future experiments when determining the number of targetable peptides per experiment.

We compared the results from asynchronous cells to time series sampling and found that time series analysis improved detection of periodically expressed cell-cycle proteins. Peptide/protein observation counts were made after pruning noisy signals as described (Supplementary Table 1). For time series samples, observations were

counted if the peptide was detected with high confidence in at least one time point sample and shown in the table below.

Experiment	Total Proteins Targeted	Total Proteins Observed	% Detected	Total Peptides Run	Total Peptides Observed (Heavy and Light forms)	% Detected
wild-type, asynchronous, triplicate	48	22	45.83	75	38	50.67
wild-type, replicate 1	48	34	70.83	97	78	80.41
wild-type, replicate 2	48	36	75.00	103	83	80.58
<i>clb1-6</i> , replicate 1	43	31	72.09	82	77	93.90
<i>clb1-6</i> , replicate 2	43	32	74.42	82	76	92.68

This finding demonstrates the utility of PRM/MRM methods to detect low abundance proteins at the times of peak expression from a synchronous population of cells. Our particular interest in transcription factors and other low-level yeast proteins rendered PRM the best suited method for our experiments over other mass spectrometry methods (open platform, phospho-enrichment, etc.).

In this study, we quantified noise in the time series proteomic data (Supplementary Table 2). The noise quantification presented was an average across five quality control samples or across biological replicates. However, when the PRM data were visualized in Skyline [8], we observed that some interfering signals occurred only at individual time points. We hypothesize that this is due to the fact that the composition of the yeast proteome changes over time during the cell cycle. Therefore, the datasets generated by this study could be useful for investigating noise in interfering signals that may change over time.

Yeast strains used in this study

Yeast strains were used for arrest-release time series experiments followed by RNA or protein extraction (for immunoblotting or mass spectrometry). All *S. cerevisiae* strains are derived from BF264-15D and were constructed with standard yeast methods.

Strain ID	Genotype	Source
BF264-15Dau	<i>MATa ade1 his2 leu2-3,112 trp1-1a ura3Δns</i>	[9]
SBY151	<i>MATa bar1</i>	[10]
SBY661	<i>MATa arg4 clb1::URA3 clb2::LEU2 clb3::TRP1 clb4::HIS2 clb5 clb6::ADE1 P_{GAL1}-CLB1::LEU2</i>	[11]
SBY934	<i>MATa bar1::hphMX4 arg4 clb1::URA3 clb2::LEU2 clb3::TRP1 clb4::HIS2 clb5::ARG4 clb6::ADE1 P_{GAL1}-CLB1::LEU2</i>	[12]
SBY1205	<i>MATa bar1 YOX1-13MYC-T_{ADH1}::kanMX6</i>	This study
SBY1258	<i>MATa bar1 CLB2-HA::kanMX6</i>	[13]
SBY1328	<i>MATa bar1 SWI4-13MYC-T_{ADH1}::kanMX6</i>	This study
SBY1340	<i>MATa bar1 YHP1-13MYC-T_{ADH1}::kanMX6</i>	This study
SBY2371	<i>MATa bar1 NRM1-HA3::kanMX6</i>	This study

References

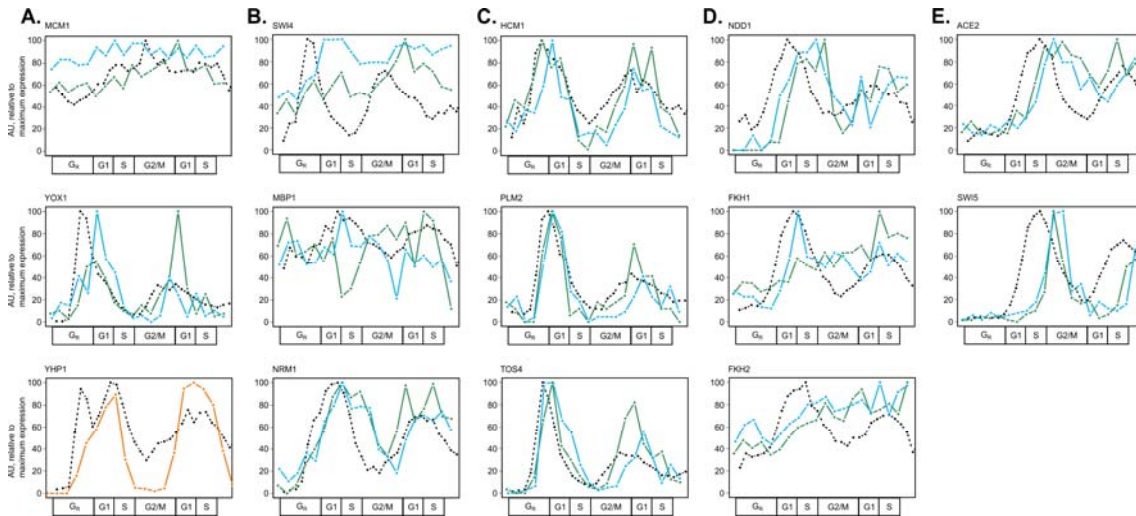
1. Kelliher CM, Leman AR, Sierra CS, Haase SB. Investigating Conservation of the Cell-Cycle-Regulated Transcriptional Program in the Fungal Pathogen, *Cryptococcus neoformans*. *PLOS Genet.* 2016;12(12):e1006453.
2. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29(1):15–21.
3. Anders S, Pyl PT, Huber W. HTSeq-A Python framework to work with high-throughput sequencing data. *Bioinformatics.* 2015;31(2):166–9.
4. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn JL, Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat*

- Biotechnol. 2013;31(1):46–53.
5. R Core Development Team. R: a language and environment for statistical computing, 3.3.2 ed. [Internet]. R Foundation for Statistical Computing. 2017. Available from: <https://www.r-project.org>
 6. Orlando DA, Lin CY, Bernard A, Iversen ES, Hartemink AJ, Haase SB. A probabilistic model for cell cycle distributions in synchrony experiments. *Cell Cycle*. 2007;6(4):478–88.
 7. Orlando DA, Iversen ES, Hartemink AJ, Haase SB. A branching process model for flow cytometry and budding index measurements in cell synchrony experiments. *Ann Appl Stat*. 2009;3(4):1521–41.
 8. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, Kern R, Tabb DL, Liebler DC, MacCoss MJ. Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics*. 2010;26(7):966–8.
 9. Reed SI, Hadwiger JA, Lőrincz AT. Protein kinase activity associated with the product of the yeast cell division cycle gene CDC28. *Proc Natl Acad Sci*. 1985;82(12):4055–9.
 10. Cole GM, Reed SI. Pheromone-induced phosphorylation of a G protein beta subunit in *S. cerevisiae* is associated with an adaptive response to mating pheromone. *Cell*. 1991;64(4):703–16.
 11. Haase SB, Reed SI. Evidence that a free running oscillator drives G1 events in the budding yeast cell cycle. *Nature*. 1999;401(6751):394–7.
 12. Chee MK, Haase SB. B-Cyclin/CDKs regulate mitotic spindle assembly by phosphorylating kinesins-5 in budding yeast. *PLOS Genet*. 2010;6(5):e1000935.
 13. Bristow SL, Leman AR, Simmons Kovacs LA, Deckard A, Harer J, Haase SB. Checkpoints couple transcription network oscillator dynamics to cell-cycle

progression. *Genome Biol.* 2014;15(9):446.

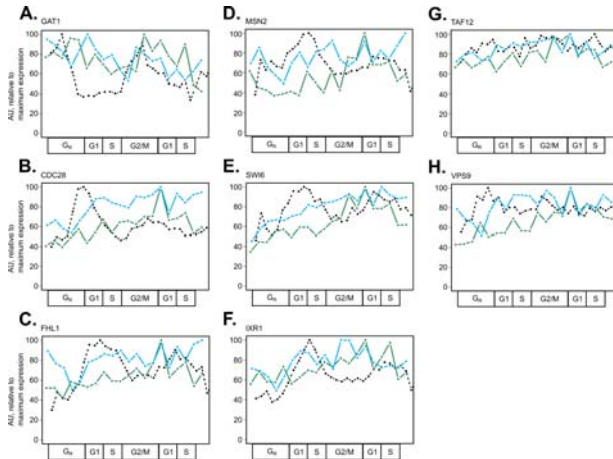
Supplementary Information

Supplementary File 1. Supporting Information Methods. This file contains further details on RNA-Sequencing mapping percentages, CLOCCS alignment of time series experiments, PRM targeted mass spectrometry method development, and yeast strains used in this study.

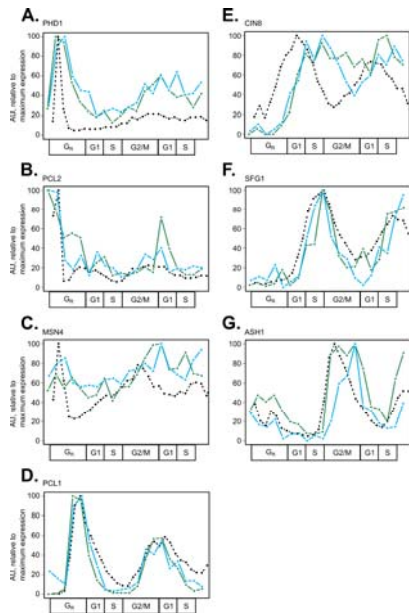


Supplementary Figure 1. Many core cell-cycle TFs cycle at the protein level, but a subset did not match periodic mRNA dynamics and displayed more stable protein expression (from Figure 1 and Supplementary Table 3). Core cell-cycle TFs are organized by phase of expression and peak activity in regulating their target genes. RNA expression and peptide light/heavy ratios were scaled to the maximum value for each gene or protein ([0, 100] linear scale). Line plots for mRNA expression (black, dashed) and biological replicates of wild-type peptide expression (replicate 1 in green, replicate 2 in blue) are shown for: M/G1 phase TFs MCM1_1, YOX1_1, and Yhp1-13MYC (from Figure 2) (**A**), G1/S phase TFs SWI4_1, MBP1_1, and NRM1_1 (**B**), S phase TFs

HCM1_1, PLM2_2, and TOS4_2 (C), S/G2/M phase TFs NDD1_1, FKH1_1, and FKH2_1 (D), and M phase TFs ACE2_1 and SWI5_3 (E).

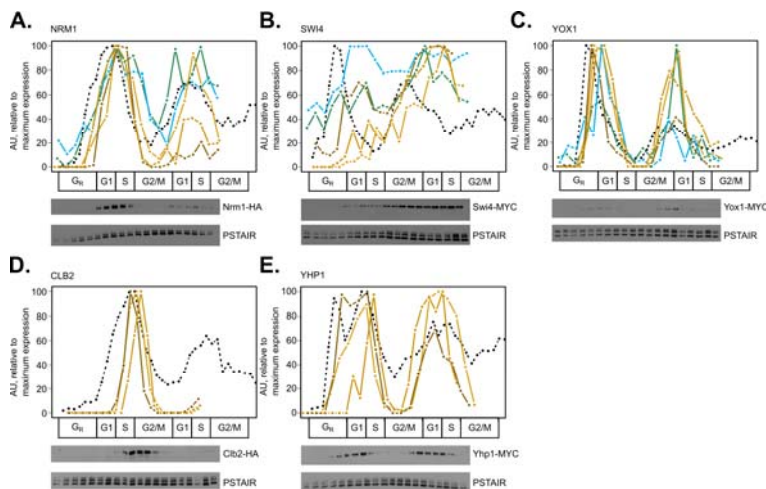


Supplementary Figure 2. Additional cell-cycle regulators have relatively stable protein expression despite cycling mRNA transcripts (from Figure 1 and Supplementary Table 3). RNA expression and peptide light/heavy ratios were scaled to the maximum value for each gene or protein ([0, 100] linear scale). Line plots for mRNA expression (black, dashed) and biological replicates of wild-type peptide expression (replicate 1 in green, replicate 2 in blue) are shown in the same order as the Figure 1 heatmap for: GAT1_1 (A), CDC28_2 (B), FHL1_1 (C), MSN2_1 (D), SWI6_2 (E), and IXR1_2 (F). Two controls for constitutive expression are also shown for TAF12_1 (G, localized to the nucleus) and VPS9_2 (H, localized to the cytoplasm).

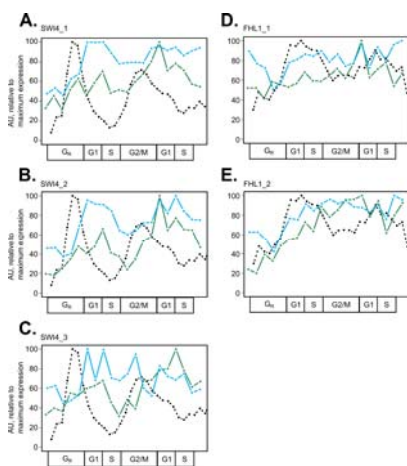


Supplementary Figure 3. Additional cell-cycle regulators did cycle along with their periodic mRNA dynamics (from Figure 1 and Supplementary Table 3). RNA

expression and peptide light/heavy ratios were scaled to the maximum value for each gene or protein ([0, 100] linear scale). Line plots for mRNA expression (black, dashed) and biological replicates of wild-type peptide expression (replicate 1 in green, replicate 2 in blue) are shown in the same order as the Figure 1 heatmap for: PHD1_1 (A), PCL2_2 (B), MSN4_2 (C), PCL1_1 (D), CIN8_1 (E), SFG1_2 (F), ASH1_2 (G).

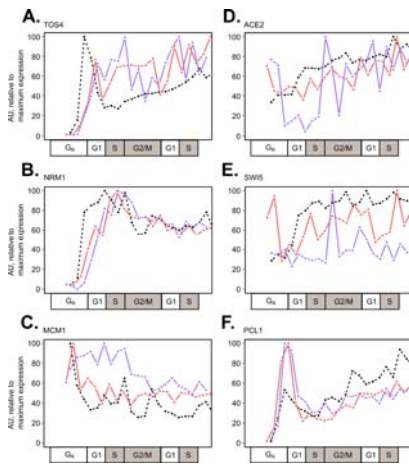


Supplementary Figure 4. Method validation and supplementation of targeted mass spectrometry by comparing to time series immunoblots (from Figure 2). Line plots for mRNA expression (black lines, dashed), three replicates of Western Blot protein expression (gold lines, solid), and two replicates of peptide expression (replicate 1 in green solid, replicate 2 in blue solid) were aligned on a common cell-cycle timeline using CLOCCS. Wild-type cells expressing Nrm1-HA3 (**A**), Swi4-13MYC (**B**), Yox1-13MYC (**C**), Clb2-HA (**D**), or Yhp1-13MYC (**E**) were grown in 2% YEPD media, synchronized by alpha-factor mating pheromone, released into YEPD, and monitored over about 2 cell cycles. Samples were collected every 7 minutes for total protein extraction. Protein immunoblots were normalized to Cdc28/Pho85 (PSTAIR; constitutive levels over the cell cycle) with ImageJ. To assess reproducibility between PRM and immunoblotting, three replicates of Western Blot data were compared to targeted mass spectrometry peptide data for NRM1_1 (**A**), SWI4_1 (**B**), and YOX1_1 (**C**). Transcript expression, peptide light/heavy ratios, and Western Blot data were scaled to maximum expression for each gene or protein ([0, 100] linear scale).



Supplementary Figure 5. Two proteins, Swi4 and Fhl1, are represented by multiple peptides with variable RNA-protein similarity scores (from Supplementary Table 3). RNA expression and peptide light/heavy ratios were scaled to the maximum value for

each gene or protein ([0, 100] linear scale). Line plots for mRNA expression (black, dashed) and biological replicates of wild-type peptide expression (replicate 1 in green, replicate 2 in blue) are shown for: SWI4_1 (A), SWI4_2 (B), SWI4_3 (C), FHL1_1 (D), FHL1_2 (E). FHL1_1 replicate 1 (D, green) and SWI4_3 replicate 2 (C, blue) had TAKT similarity scores > 0.01 for the RNA-peptide pair.



Supplementary Figure 6. A subset of cell-cycle proteins is correlated with mRNA dynamics in *clb1-6* mutant cells (from Supplementary Table 3). RNA expression and peptide light/heavy ratios were scaled to the maximum value for each gene or protein ([0, 100] linear scale). Line plots for mRNA expression (black, dashed) and biological replicates of *clb1-6* mutant peptide expression (replicate 1 in red, replicate 2 in purple) are shown for: TOS4_2 (A), NRM1_1 (B), MCM1_1 (C), ACE2_2 (D), SWI5_1 (E), and PCL1_1 (F). In the cell-cycle timeline for *clb1-6* cells, S and G_{2/M} phases are shown as gray boxes to indicate that B-cyclin mutant cells are physically arrested at the G₁/S border.

Supplementary Table 1. Summary of 49 *S. cerevisiae* proteins and 149 SIL peptides ordered for this study. Targeted mass spectrometry was used to measure protein expression levels for cell-cycle regulators of interest (columns 1-2 list protein

names and column 3 denotes cell-cycle function). For each protein, the amino acid length (column 4), estimated protein concentrations from previous studies (column 5 from [36]; column 6 from [38]; column 7 from [39]; column 8 from [41]), and estimated half-lives (column 9 from [85]; column 10 from [86]) are listed. Wild-type protein expression dynamics have been investigated previously by immunoblotting for a subset of our proteins of interest, and literature references are shown for these individual experiments (column 11). When possible, we utilized multiple heavy labeled peptides (JPT SpikeTides TQL) for identification and quantitation (column 12). The location of each peptide relative to the N-terminal methionine amino acid is given (column 13). Matched non-interfering transitions for each light/heavy pair were manually selected in Skyline [80] based on matching of relative fragment intensities to a spectral library as well as the coefficient of variation of replication injections of a QC pool (columns 16-19).

Supplementary Table 2. Quantification of noise from targeted mass spectrometry datasets using both quality control samples and biological replicate time series experiments.

For each time series experiment, quality control (QC) samples were composed of an equal mixture of each time point sample, and five QC runs were evenly distributed throughout the mass spectrometry assays (Materials and Methods). The light/heavy ratio measurements for the five QC samples were obtained from Skyline for each dataset. Averages and standard deviations were calculated across the five QC samples, and the coefficient of variation (CoV; standard deviation / mean; low values are the least noisy) is shown (column 3). Protein expression values at each time point (light/heavy ratios) were also compared between the biological replicate experiments using the interpolated time series data (see Supplementary Table 3 legend). For each interpolated lifeline point, the CoV was calculated for each corresponding pair of expression values in replicates 1 and 2, and the average CoV is shown across the time

series (standard deviation / mean; low values are the least noisy). A Signal to Noise Ratio was also computed across the time series between the two biological replicates (high values are the least noisy). The top 25% of peptides with the least noisy values are highlighted in green (QC CoV: top 40 peptides; time series CoV and Signal to Noise Ratio: top 20 peptides highlighted). Wild-type peptides are shown in the first Excel tab, *clb1-6* peptides are shown in the second Excel tab.

Supplementary Table 3. TAKT similarity measures for the 46 high confidence *S. cerevisiae* protein measurements compared to their respective mRNA transcripts (from Figure 1). To place transcriptome (33 genes) and proteome (44 peptides & 2 immunoblotting) data on an identical time scale, experiments were aligned to a cell-cycle timeline with the CLOCCS algorithm. Data points were then interpolated to 30 samples along the cell-cycle timeline interval from 50-250 lifeline points (i.e. all time courses were sampled *in silico* approximately every 6.9 points). Interpolation calculations were done in the R Statistical Programming Environment using the function `approxfun` (arguments: `method="linear", rule=2`) [87]. Any negative values resulting from the interpolation were set to zero. Interpolated mRNA and peptide curves were then compared using the TAKT algorithm. Peptide-RNA pairs with empirical p-values less than 0.05 are shown in blue. Scores from 2/3 biological replicates are shown in the table for triplicate Western blot experiments (Supplementary Figure 4) for Clb2 (p = 0.0002, 0.0004, and 0.0006) and Yhp1 (p = 0, 0.0001, 0.2773).

Supplementary Table 4. Previously annotated and predicted protein degradation mechanisms for cell-cycle proteins of interest. Degradation pathways targeting specific proteins of interest are shown with literature references. The amino acid sequences of all proteins were also tested for various protein instability metrics: N-end

rule [88], PEST sequence predictor [89], F-box predictor from the SMART database [90], and APC/C sequence predictor from the GPS-ARM tool [91]. The four proteins used as controls for constitutive expression during the cell cycle (Rim11, Taf12, Vps9, and Cic1) are not shown, as they are not dynamically expressed. Proteins annotated as “unknown” did not match any bioinformatics sequence predictors and did not have experimental evidence for targeted protein destruction.

Supplementary Table 5. Estimated time delays between mRNA and protein curves during the wild-type and *clb1-6* mutant cell cycle.

Interpolated data for RNA and protein expression (described in the Supplementary Table 3 legend) were truncated to the first cell cycle on the CLOCCS timeline ([50, 200] lifeline points). Time delays were then calculated for the highest confidence RNA-peptide pairs (Figure 1) using the TAKT algorithm. Proteins in this table were highlighted as follows: SCF targets in green (Ash1, Cln2, Hcm1, Sfg1, Sic1, Swi5, and Yox1), APC/C targets in purple (Cdc20, Cin8, Clb2 Western Blot data, Clb5, Fkh1, Nrm1, and Pds1), targeted by both in blue (Ndd1, Tos4, and Yhp1 Western Blot data), and unknown degradation targeting with no highlight (Ace2, Cdc28, Fhl1, Fkh2, Gat1, Ixr1, Mbp1, Mcm1, Msn2, Msn4, Pcl1, Pcl2, Phd1, Plm2, Swi4, and Swi6) (from Supplementary Table 4). The average mRNA-to-peptide delay time from the high confidence list was 13.7 ± 12.8 minutes (assuming a 70-minute cell cycle, 19.5 ± 18.3 cell-cycle timeline points).

Supplementary Table 6. Wild-type normalized datasets used in this study.

Light/heavy ratio measurements for the time series samples were obtained from Skyline for wild-type replicates 1 and 2 (Excel Tab 1). Peptides are labeled according to Supplementary Table 1 along with a replicate 1 or 2 dataset label. Light/heavy ratio

measurements were also scaled to maximum expression for each peptide ([0, 100] linear scale) in replicates 1 and 2 (Excel Tab 2). Light/heavy ratios were interpolated to 30 samples along the cell-cycle timeline interval from 50-250 points using the R function `approxfun` [87] (see Supplementary Table 3 legend). Triplicate Western Blot experiments for Clb2 and Yhp1 are also shown in the intercalated data (Supplementary Figure 4). Any negative values resulting from the interpolation were set to zero, and interpolated data are shown (Excel Tab 3).

Supplementary Table 7. *clb1-6* normalized datasets used in this study. Light/heavy ratio measurements for the time series samples were obtained from Skyline for *clb1-6* replicates 1 and 2 (Excel Tab 1). Peptides are labeled according to Supplementary Table 1 along with a label for replicate 1 or 2. Light/heavy ratio measurements were scaled to maximum expression for each peptide ([0, 100] linear scale) in replicates 1 and 2 (Excel Tab 2). Light/heavy ratios were interpolated to 28 samples along the cell-cycle timeline interval from 50-250 points using the R function `approxfun` [87] (see Supplementary Table 3 legend). Any negative values resulting from the interpolation were set to zero, and interpolated data are shown (Excel Tab 3).