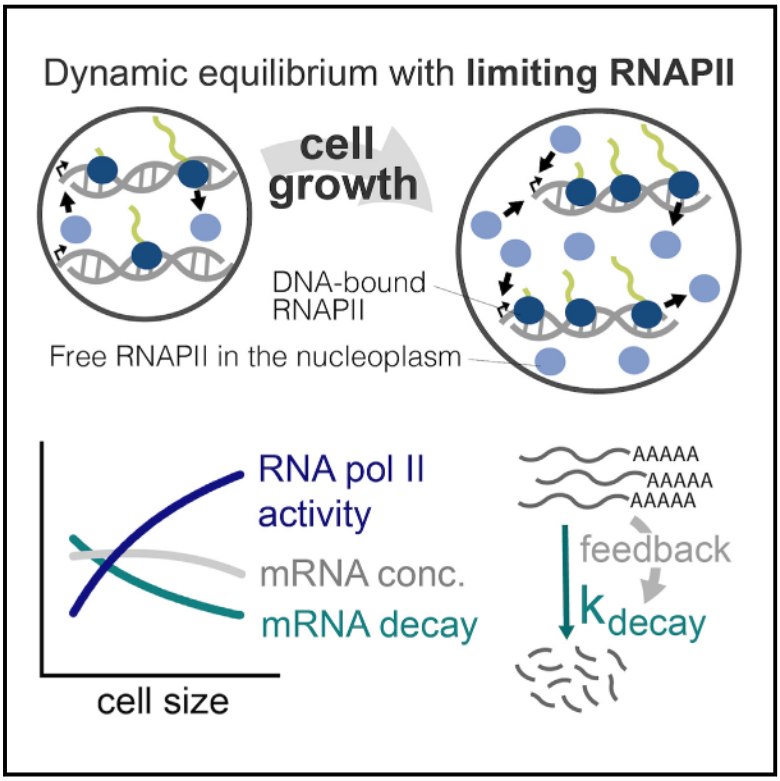
Cell

Article

RNA polymerase II dynamics and mRNA stability feedback scale mRNA amounts with cell size

Graphical abstract



Authors

Matthew P. Swaffer, Georgi K. Marinov, Huan Zheng, ..., William J. Greenleaf, Rodrigo Reyes-Lamothe, Jan M. Skotheim

Correspondence

skotheim@stanford.edu

In brief

mRNA transcription in budding yeast increases with cell size due to the equilibrium kinetics of dosage-limiting RNA polymerase II and the genome. This is then supported by a global increase in mRNA stability in larger cells to better scale mRNA amounts with cell size.

Highlights

Total mRNA must increase as a cell grows to ensure concentration homeostasis

RNAPII is the major dosage-limiting component of the transcriptional machinery in yeast

The size-dependent increase in RNAPII-DNA binding follows dynamic equilibrium kinetics

mRNA turnover decreases in larger cells to support mRNA concentration homeostasis

Swaffer et al., 2023, Cell *86*, 5254–5268
November 22, 2023
https://doi.org/10.1016/j.cell.2023.10.012

Cell Press

Cell Press

Cell

Article

RNA polymerase II dynamics and mRNA stability feedback scale mRNA amounts with cell size

Matthew P. Swaffer,[1,6] Georgi K. Marinov,[2] Huan Zheng,[3] Lucas Fuentes Valenzuela,[1] Crystal Yee Tsui,[1] Andrew W. Jones,[4] Jessica Greenwood,[4] Anshul Kundaje,[2] William J. Greenleaf,[2] Rodrigo Reyes-Lamothe,[3] and Jan M. Skotheim[1,5,7,*]

[1]Department of Biology, Stanford University, Stanford, CA 94305, USA
[2]Department of Genetics, Stanford University, Stanford, CA 94305, USA
[3]Department of Biology, McGill University, Montreal, QC H3G 0B1, Canada
[4]The Francis Crick Institute, London NW1 1AT, UK
[5]Chan Zuckerberg Biohub, San Francisco, CA 94158, USA
[6]Present address: Wellcome Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3BF, UK
[7]Lead contact
*Correspondence: skotheim@stanford.edu
https://doi.org/10.1016/j.cell.2023.10.012

SUMMARY

A fundamental feature of cellular growth is that total protein and RNA amounts increase with cell size to keep concentrations approximately constant. A key component of this is that global transcription rates increase in larger cells. Here, we identify RNA polymerase II RNAPII) as the limiting factor scaling mRNA transcription with cell size in budding yeast, as transcription is highly sensitive to the dosage of RNAPII but not to other components of the transcriptional machinery. Our experiments support a dynamic equilibrium model where global RNAPII transcription at a given size is set by the mass action recruitment kinetics of unengaged nucleoplasmic RNAPII to the genome. However, this only drives a sub-linear increase in transcription with size, which is then partially compensated for by a decrease in mRNA decay rates as cells enlarge. Thus, limiting RNAPII and feedback on mRNA stability work in concert to scale mRNA amounts with cell size.

INTRODUCTION

Cell size has a profound impact on cellular physiology because it sets the scale of subcellular structures, metabolism, surface-to-volume ratios, and most crucially, cellular biosynthesis. It is generally thought that biosynthesis scales with cell size so that total RNA and protein amounts increase with cell size, and their concentrations remain approximately constant over the physiological range of cell sizes (Figure 1A). A key aspect of this biosynthetic scaling is that global transcription rates increase in larger cells[1–10] (Figure 1B). This size-dependent transcriptional scaling is thought to increase global mRNA, rRNA, and tRNA amounts, which drive increased protein synthesis with increased cell size. The importance of this biosynthetic scaling is highlighted by the observation that it only occurs within the physiological range of cell sizes. Beyond this range, both RNA and protein amounts cannot keep pace with the expanding cell volume, their concentrations decline, and the cytoplasm becomes progressively diluted.[6,11] This breakdown in biosynthetic scaling is associated with a collapse in many aspects of cellular physiology, and increasing evidence points to this being a causal driver of cellular aging and senescence.[11–15]

While it is now clear that biosynthetic size scaling is important for cellular functionality, the mechanistic origin of RNA scaling has remained mysterious despite being first reported by foundational radio-labeling experiments in the 1970s.[1–5] This is because transcription increases continuously as a cell grows, even among cells with the same template DNA amount.[6–10] To explain this phenomenon, several groups have proposed that transcription initiation is controlled by a "limiting factor" whose amount increases in proportion to cell size.[6–8,16,17] In one striking illustration of this idea, Padovan-Merhar et al. made a heterokaryon fusion of a large and a small cell and observed that the expression of a GFP reporter transcript from the small cell-derived nucleus reflected the total cell size of the heterokaryon cell.[7] This suggests that the amount of some *trans*-activating limiting factor is responsible for coupling cell size to transcription.

While potential mechanisms through which a limiting factor could drive increased global transcription in larger cells have been explored in theoretical work, they have yet to be tested experimentally.[8,16] In some models, it is the amount of the limiting factor, rather than its concentration, that sets the global transcription rate because it is titrated against the genome. In such models, as a cell grows larger there would be proportionally more of the limiting factor so that it is bound to the genome at higher densities to drive a size-proportional increase in transcription. This increase in transcription does not depend on the DNA concentration or cell ploidy so that similarly sized diploid and

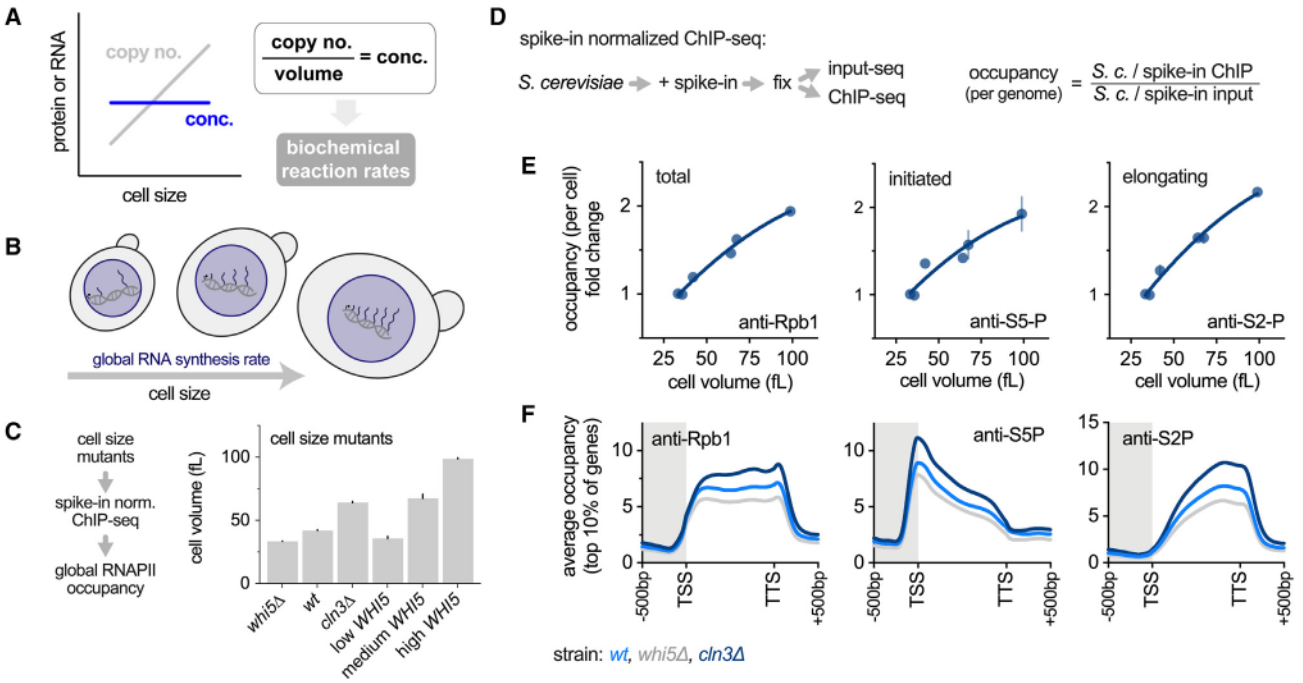


Figure 1. More RNAPII is loaded onto the genome with increasing cell size

(A) Macromolecule amounts need to increase with cell size to keep concentrations constant.

(B) Global RNA synthesis rates increase with cell size.

(C) Cell volume determined by Coulter counter of cell size mutants. Low to high *WHI5* expression is from a dose-responsive beta-estradiol-induced promoter (see also Figure S1F).

(D) Schematic illustrating the spike-in normalized ChIP-seq workflow (see STAR Methods and Figures S1A–S1E for details).

(E) The occupancy per cell of total Rpb1, initiated Rpb1 (anti-S5-P), and elongating Rpb1 (anti-S2-P) in the size mutants shown in (C), plotted as a function of cell size. Each point is the mean (± range) of two biological replicates.

(F) Average Rpb1, Rpb1-S5P, and Rpb1-S2P occupancy across the gene bodies of the top 10% of genes for *WT*, *whi5* , and *cln3* cells (see Figure S1G for similar results in cells expressing low to high *WHI5*). Mean of two biological replicates is shown.

See also Figure S1.

haploid cells are predicted to transcribe at a similar total rate, because the haploid cell would have twice the density of the limiting factor on its genome.[8,16] Several groups have also suggested that some components of the general transcriptional machinery, such as RNA polymerases, could function as a key limiting factor.[6–8,16,17]

Here, we show that more transcriptional machinery is loaded onto the budding yeast genome as cells increase in size. We identify RNA polymerase II (RNAPII) as the major limiting factor for increasing mRNA transcription with cell size. In contrast, other components of the transcriptional machinery are not limiting, and the chromatin environment into which transcriptional machinery is loaded is similar in large and small cells. Importantly, our data are inconsistent with previously proposed titration models and instead support a dynamic equilibrium model that is based on mass action recruitment kinetics of the free nucleoplasmic RNAPII pool to the genome. Our dynamic equilibrium model accurately predicts RNAPII loading at a given size for both haploid and diploid cells. However, we observe that transcription does not increase in direct proportion to cell size, and so the RNAPII dynamic equilibrium alone is insufficient to maintain mRNA concentrations as cells grow. This led us to identify an additional feedback on mRNA turnover, which stabilizes the transcriptome in larger cells to extend the range of sizes over which mRNA concentrations are kept near to constant. Thus, the scaling of mRNA amounts with cell size is driven by both the mass action recruitment kinetics of RNAPII to the genome and feedback regulation on mRNA stability.

RESULTS

RNAPII occupancy on the genome increases with cell size in budding yeast

Previous observations that bulk transcription rates increase with cell size in budding yeast[1] suggest that larger cells should have more RNAPII loaded on their genomes. To test this, we employed a spike-in normalized chromatin immunoprecipitation sequencing (ChIP-seq) methodology to measure RNAPII occupancy (Figures 1C and 1D). This method is quantitative across a wide dynamic range and is robust to variation in the spike-in mixing ratio and cell density (Figures S1A–S1E). While conventional ChIP-seq only resolves differences in the relative binding at different genomic loci, spike-in normalization resolves systematic differences in the global amount of DNA-bound protein (i.e., the relative global occupancy).[18] We isolated the effect of cell size from growth rate by examining yeast cultures of different

average cell sizes but with similar doubling times (Figures 1C and S7B). Spike-in normalized ChIP-seq of these cultures shows that the total amount of RNAPII loaded on the genome increases with cell size (Figure 1E). The initiated and elongating RNAPII populations, identified by S5 and S2 phosphorylation in the C-terminal heptapeptide repeats on Rpb1, increased similarly with cell size (Figure 1E). Moreover, the relative distributions across gene bodies of total, initiated, and elongating RNAPII increase relatively uniformly with cell size (Figure 1F), indicating that initiation is primarily responsible for the size-dependent increase in transcription, as has been suggested in fission yeast.[8]

Cellular and chromatin-bound amounts of transcriptional machinery increase with cell size

The increase in transcription with cell size has previously been attributed to a hypothetical limiting factor.[6–8,16] In such models, the amount of the limiting factor increases in proportion to cell size, resulting in its increased binding to the genome. In principle, this could be any factor essential for transcription. Thus, to identify candidate limiting factors, we quantified the protein amounts of RNAPII and the general transcription factor complexes, which together constitute the RNAPII pre-initiation complex (PIC). The protein amounts of all PIC components that we examined increased in close proportion to cell size (Figure 2A). We next sought to determine how these factors change their association with chromatin in larger cells. To this end, we adapted a chromatin purification technique called ChEP (chromatin enrichment for proteomics)[19] to work in yeast (Figures S2A and S2B) and analyzed protein-chromatin associations in small and large cells by mass spectrometry. The recruitment of all subunits of RNAPII to the genome increased in larger cells, as expected, and so did the initiation and elongation factors that regulate and associate with RNAPII (Figures 2B and S2E). We note that only proteins first validated as being enriched on chromatin were analyzed here to minimize potential confounding background effects (Figures S2C and S2D). Taken together, these experiments show that more transcriptional machinery is expressed and loaded on the genome in larger cells but does not delineate which, if any, component is dosage limiting in a manner that could couple global transcription to cell size.

RNAPII is the major limiting subcomplex of the PIC

To directly test potential limiting factors, we next performed a local perturbation to the nuclear amounts of PIC components, using the anchor-away approach to conditionally deplete half of a targeted factor from the nucleus.[20] To do this, we constructed heterozygous diploids where one of the two alleles of a given PIC subunit is tagged with an FKBP-rapamycin-binding (FRB) domain, allowing us to rapidly and conditionally deplete half of it from the nucleus to a cytoplasmic FKBP12 anchor upon rapamycin treatment (Figure 2C). This revealed that the general transcription factor complexes are not significantly limiting, because a 2-fold reduction in their nuclear amounts results in, at most, a 5%–10% reduction in RNAPII occupancy as measured using spike-in normalized ChIP-seq (Figures 2D and S3B). However, 2-fold depletion of RNAPII from the nucleus led to a much larger reduction of ~40% in occupancy (Figures 2D and S3C–S3H), suggesting that RNAPII is a major limiting factor. Using a FLAG tag to quantify only the non-depleted allele in an *RPB1-FLAG/RPB1-FRB* strain, we confirmed the ~40% value (Figures S3D and S3E). Thus, our conditional depletion analysis shows RNAPII is a major limiting component of the PIC, whereas the other subcomplexes of the PIC are not particularly limiting.

Transient overexpression is sufficient to increase RNAPII loading

If RNAPII is the limiting factor for its own recruitment to the genome, then increasing the amounts of the 12-subunit RNAPII complex should be sufficient to increase polymerase loading. To test this, we first optimized overexpression conditions of individual subunits from an anhydrotetracycline-inducible promoter (*TetPr*). We compared the overexpressed protein amounts (e.g., *TetPr-RPB1-mNeonGreen*) with that of their endogenous counterpart (e.g., *RPB1pr-RPB1-mNeonGreen*) for the 9 RNAPII subunits where C-terminal tagging of the endogenous copy did not cause severe growth defects (Figure S3I). This allowed us to define an induction regime in which all tested subunits are 2- to 3-fold overexpressed 45 min after anhydrotetracycline (atc) addition (Figure 2E). Next, we engineered a single yeast strain (*TetPr-RPB1-12*) in which all 12 subunits are simultaneously and conditionally overexpressed. In this strain, we observed a robust increase in global RNAPII occupancy across the genome after induction, while there was no equivalent change in a control strain expressing mNeonGreen from the *TetPr* integrated at the same loci (Figures 2F and 2G). Importantly, RNAPII overexpression did not alter cell size in these experimental conditions (Figure S3J). Thus, RNAPII is a limiting component of the transcriptional machinery, because RNAPII loading on the genome is sensitive to increases and decreases in its own concentration.

RNAPII occupancy increases sub-linearly with cell size

Having established RNAPII as a major limiting component of the PIC, we sought to examine more quantitatively how this could drive the scaling of transcription with cell size. Prior theoretical work proposed that polymerase could be titrated against the genome.[8,16] In these titration models total protein amounts, rather than cellular concentrations, are critical because the on-rate for binding is so high that nearly all molecules are associated with the genome. This is predicted to result in a directly proportional linear increase in DNA binding with cell size.

To test this genome-titrated hypothesis, we first simplified our experimental setup and measured RNAPII occupancy in cells of different sizes but with the same fixed DNA amount. To do this, we isolated small haploid G1 cells by centrifugal elutriation and arrested them in G1 by turning off expression of the G1 cyclins. By arresting cells for increasing amounts of time, we generated populations of cells with increasing cell size but with the same 1 N DNA content (Figures 3A and S4A–S4C). We also confirmed that the total cellular amounts of RNAPII subunits Rpb1 and Rpb3 increase in proportion to cell size during G1 arrest (Figures 3E and S4M). Spike-in normalized ChIP-seq shows that RNAPII occupancy increases with cell size, but crucially, this increase is not directly linearly proportional to cell size

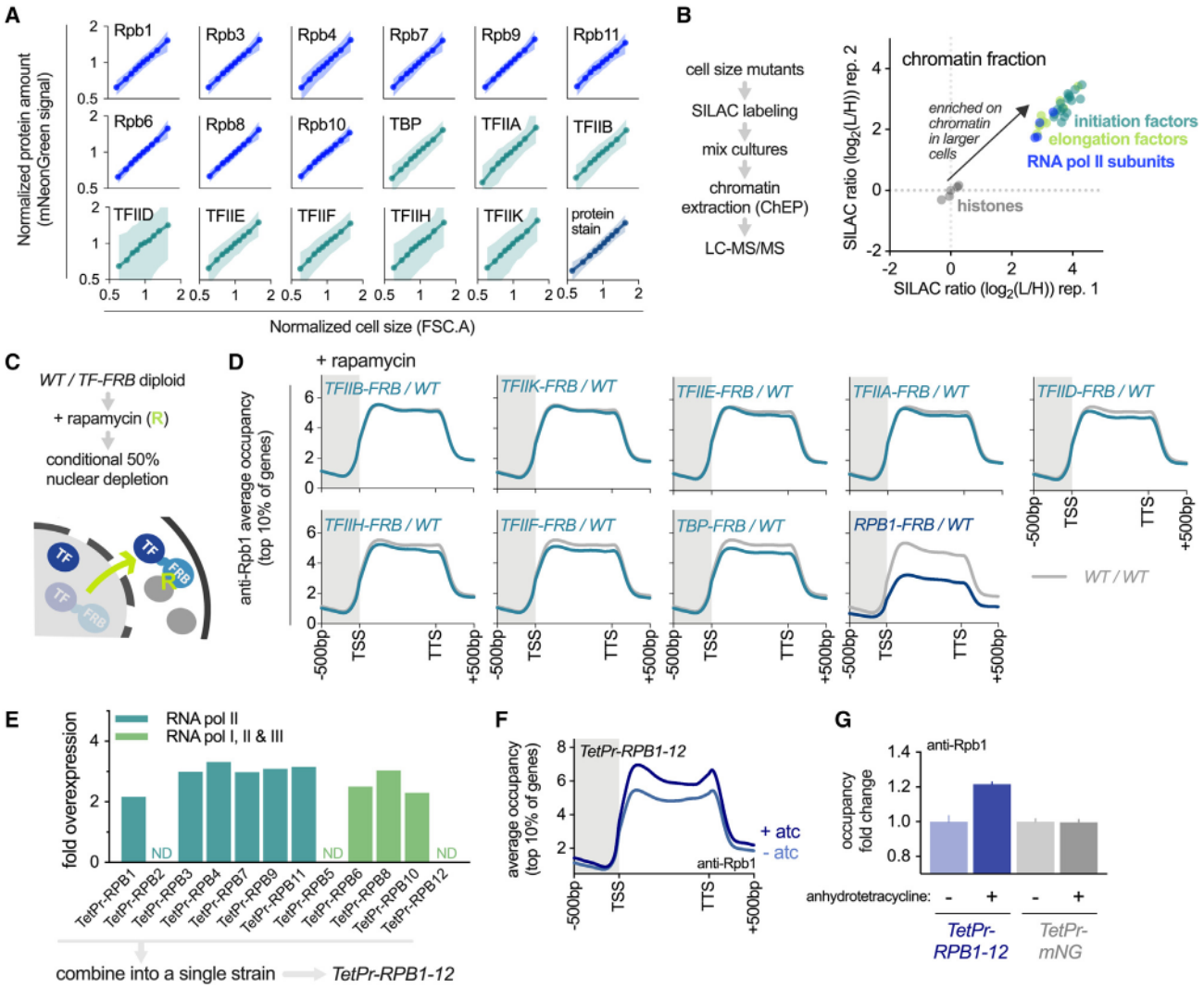


Figure 2. RNAPII is a major limiting component of the transcriptional machinery

(A) Protein amount (mNeonGreen) plotted against cell size (forward scatter) for subunits of the RNAPII pre-initiation complex measured by flow cytometry. The mean (± SD) is shown for each cell size bin. Bottom right panel shows the scaling of total protein with forward scatter.

(B) Chromatin association of transcriptional machinery measured by mass spectrometry. SILAC-labeled cells of different sizes (Figure S2E) were mixed, and chromatin was extracted (ChEP) and analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS) (see Figure S2B for controls). Each axis shows an independent biological replicate for the normalized SILAC ratio of chromatin association between large (L) and small (H) cells. RNAPII subunits, initiation factors, and elongation factors are shown relative to histones. Only chromatin-enriched proteins are shown (see also Figures S2C and S2D).

(C) Schematic illustrating the 50% depletion strategy to test dosage limitation. A pre-initiation complex subunit (TF) is tagged with the FRB domain[20] and crossed with *WT* to make a heterozygous diploid, allowing for conditional 2-fold reduction in nuclear concentrations upon rapamycin treatment.

(D) Average Rpb1 occupancy across gene bodies of the top 10% of genes measured by spike-in normalized ChIP-seq upon 2-fold nuclear depletion of the indicated factor following rapamycin treatment. Rpb1-FRB anchor-away co-depletes Rpb3 (Figure S3A), indicating that the whole RNAPII complex is efficiently co-depleted.

(E) The fold overexpression of individual RNAPII subunits from the TetPr after 45 min of anhydrotetracycline (atc) treatment. Overexpression was determined using C-terminal mNeonGreen-tagged proteins measured by flow cytometry and compared with the endogenously tagged allele of the respective subunit (Figure S3I). *RPB2*, *RPB5*, and *RPB12* were not assessed, as tagging caused growth defects. All 12 subunits were integrated into a single strain to construct the *TetPr-RPB1-12* strain (see STAR Methods for details).

(F) Average Rpb1 occupancy, determined by spike-in normalized ChIP-seq, across the gene bodies of the top 10% of genes for the *TetPr-RPB1-12* strain before (−atc) or 45 min after (+atc) simultaneous overexpression of all 12 RNAPII subunits.

(G) The global Rpb1 occupancy measured by spike-in normalized ChIP-seq after atc induced expression of either all RNAPII subunits (*TetPr-RPB1-12*) or free mNeonGreen (*TetPr-mNG*). Mean (± range) is plotted.

See also Figures S2 and S3.

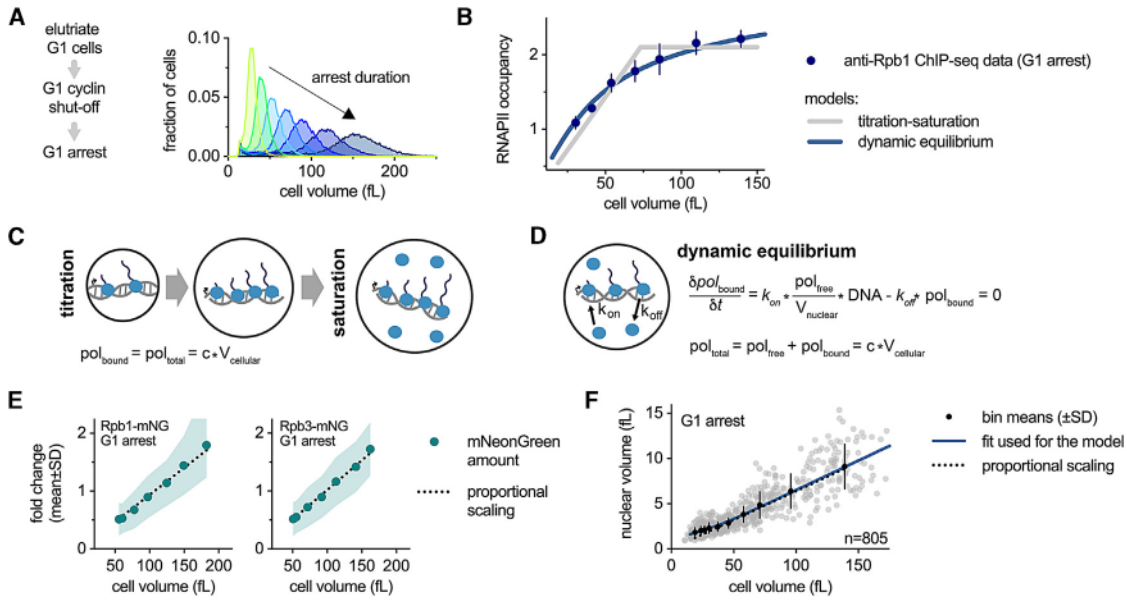


Figure 3. RNAPII recruitment to the genome is not directly proportional to cell size

(A) Cell volume distributions determined by Coulter counter for G1-arrested cells. Small G1 cells were collected by centrifugal elutriation and arrested in G1 by conditional G1 cyclin shut-off to generate populations with increasing cell size (see Figures S4A and S4B for experimental design).

(B) Rpb1 (anti-Rpb1) occupancy per cell, determined by spike-in normalized ChIP-seq, as a function of cell size in the indicated populations of G1-arrested cells shown in (A). Each point shows the mean (± range) of two biological replicates. The fit of two alternative mathematical models are shown.

(C and D) Schematics illustrating two alternative models that can account for the data in (B): the previously proposed titration-saturation model (C) and the dynamic equilibrium model proposed here (D).

(E) Rpb1 and Rpb3 protein amounts as a function of cell size. mNeonGreen amounts were determined by flow cytometry. Cell volume was determined by Coulter counter. Each point corresponds to a different time point after the G1 arrest was initiated (see also Figure S4M).

(F) Nuclear volume as a function of cell volume determined by wide-field fluorescence microscopy (see also Figures S4H–S4J).

See also Figure S4.

(Figure 3B). Importantly, this sub-linear increase is not an artifact of elutriation (Figures S4E and S4F) and is also clearly apparent in our measurements of asynchronously cycling cell size mutants, where a 3-fold increase in cell size results in only a 2-fold increase in RNAPII occupancy (Figures 1E and S4G).

Since the sub-linear increase in RNAPII occupancy rules out a simple titration model, we considered two alternatives. The first model is a modification of the titration model in which the genome becomes saturated at a specific size threshold so that transcription cannot increase further (“titration-saturation”)[16] (Figure 3C). The second model is a “dynamic equilibrium” model (Figure 3D), in which the rate at which the free RNAPII, pol_{free}, associates with the genome, DNA, is determined by mass action kinetics so that

$$\frac{dpol_{bound}}{dt} = k_{on} * DNA * \frac{pol_{free}}{V_{nuclear}} - k_{off} * pol_{bound} = 0,$$

which we have approximated as steady state ($\frac{dpol_{bound}}{dt} = 0$) because the timescale of cell volume changes (hours) is much slower than that of transcription (minutes). To complete the set of equations, we use the observation that the total amount of RNAPII is proportional to cell volume (Figures 2A and 3E) so that

$$pol_{total} = pol_{free} \quad pol_{bound} = c\, V_{cell}$$

These two algebraic equations can then be solved for the amount of DNA-bound RNAPII as a function of cell and nuclear size:

$$pol_{bound} = \frac{c\, V_{cell}}{1 \quad \frac{k_{off}\, V_{nucleus}}{k_{on}\, DNA}}$$

Consistent with prior work, we observed that nuclear volume increases with cell size in G1 but that this increase is initially slower than the cell volume increase[21] (Figures 3F and S4H–S4J). This results in a modestly higher nuclear-to-cell volume ratio in the smallest cells in the range that we are examining (Figure S4I). We used this empirical relationship to substitute for the $V_{nucleus}$ term (see STAR Methods). We then separately fit the dynamic equilibrium model and the titration-saturation model to our RNAPII occupancy measurements, which showed that both models are similarly capable of fitting these data (Figure 3B). We note that the dynamic equilibrium model only has two free parameters that we are fitting. In essence, one parameter is the cellular concentration of RNAP II (c), and the other parameter is the dissociation constant for RNAPII and the genome ($\frac{k_{off}}{k_{on}}$).

RNAPII binding is driven by dynamic equilibrium kinetics

While both models are consistent with the amount of RNAPII bound as a function of cell size, they make very different

predictions about what fraction of RNAPII is associated with the genome. The titration-saturation model predicts that nearly all nuclear RNAPII should be engaged on the genome, resulting in a near-negligible free nucleoplasmic pool. Thus, the genome-bound fraction should be fixed until cells surpass the saturation threshold, at which point the bound fraction should decrease in proportion to the inverse of the cell volume $(1/V)$. In contrast, the fit of the dynamic equilibrium model to our ChIP-seq data predicts that approximately 50% of nuclear RNAPII should be freely diffusing in the nucleoplasm of a 50 fL G1 cell and that the bound fraction should decrease gradually and continuously as cell size increases.

To distinguish between these models, we measured the bound fraction of RNAPII by imaging and tracking single molecules of its subunit Rpb1-HALO. Freely diffusing molecules diffuse rapidly and have a large radius of gyration, whereas chromatin-bound molecules are relatively immobile and have a small radius of gyration. Consistent with expectations, we measured that the majority of H2B histone molecules were tightly bound, and the vast majority of a free nuclear mCitrine-NLS were highly mobile (Figures 4A and 4B). Importantly, only half of the nuclear Rpb1 is freely diffusing, and the other half is chromatin bound (Figures 4A and 4B), consistent with a recent independent single-molecule-based analysis of Rpb1 bound fractions.[22] Moreover, the bound fraction of Rpb1 in individual G1 cells decreases gradually with cell size as predicted by the dynamic equilibrium model (Figure 4C). In contrast, these data show a very poor correspondence to the prediction of the titration-saturation model (Figure 4C). We emphasize that the two free parameters for the dynamic equilibrium model were determined by fitting to the anti-Rpb1 ChIP-seq data (Figure 3B) and are not adjusted in response to the single-molecule imaging data.

In our dynamic equilibrium model, the RNAPII loading rate is determined by the mass action recruitment kinetics of unbound polymerase and its target sites on the genome (Figure 3D). Thus, the amount of RNAPII on the genome should be sensitive to the amount of DNA, as well as cell size. To test this, we determined the effects of increasing the genome copy number by repeating our spike-in normalized Rpb1 ChIP-seq experiment with elutriated G1-arrested diploid cells (Figures 4D and S4A–S4D). These data show that RNAPII occupancy is significantly higher in diploids compared with haploid cells of the same size. Moreover, these data show a very close correspondence to the model prediction for doubling DNA amount (Figures 4E and 4F). Importantly, this prediction is made using no additional free parameters, as we simply double the value of the "*DNA*" term in the model after fitting to the haploid data. Taken together, our quantitative ChIP-seq and single-molecule tracking experiments are inconsistent with any form of titration-based model. Instead, we find that the size-dependent changes in global RNAPII occupancy are best explained via a surprisingly simple dynamic equilibrium model.

Stochastic simulations of RNAPII transcription corroborate the assumptions of the dynamic equilibrium model

The dynamic equilibrium model makes the important simplifying assumptions that RNAPII occupancy can be modeled using simple binding-unbinding kinetics and that promoters and gene bodies do not saturate with polymerase (Figure S5A). While this accurately predicts the global relationship between cell size and RNAPII occupancy, in reality transcription is a highly complex multistep process that occurs on a finite space on the genome. To test the feasibility of our assumptions underpinning the dynamic equilibrium model, we built a more complex granular stochastic stepwise simulation of RNAPII transcription that includes promoter recruitment, initiation, and elongation on a representative gene of finite length (Figures 4G and S5B). We parameterized these simulations using measurements from the budding yeast literature. For example, we used recently reported single-molecule measurements of PIC subunits[22] to define a limited range of feasible parameter combinations for the promoter association and transcriptional initiation rates (see STAR Methods for details) (Figures S5C and S5D). In all cases, our stochastic simulations predicted very similar relative changes in global RNAPII occupancy compared with the dynamic equilibrium model (Figures 4H and S5E). This is because, at the average gene, RNAPII occupancy in the promoter and gene body is in a linear regime with respect to the upstream recruitment kinetics of RNAPII. This makes intuitive sense because the absolute PIC occupancy of an average promoter is estimated to be ~4%, and there are only ~0.4 RNAPII molecules in an average gene body (Figure S5D). Thus, our more granular stochastic simulations revealed that RNAPII molecules are unlikely to interfere with one another on the genome, consistent with the simplifying assumptions underlying the dynamic equilibrium model.

While the dynamic equilibrium model reflects what is likely occurring at most genes, the finite linear space of a gene may impact the most highly expressed genes. To explore this, we modeled genes of different strengths by varying the promoter recruitment rate. Under some, but not all, parameter combinations, this resulted in the top few percent of genes showing a smaller increase in RNAPII occupancy in larger cells, compared with the rest of the genome (Figures 4I, 4J, and S5F). Consistent with this, when we binned genes according to their RNAPII occupancy in our ChIP-seq data, we saw a similar deviation in the most highly expressed genes (Figure 4K). This suggests that the top percentiles of the most highly expressed genes could be starting to approach saturation kinetics. However, it is also possible that this apparent saturation results from a programmed transcriptional response that involves the regulated repression of highly expressed genes. We note that for this analysis we have excluded cell-cycle and stress response genes that might be anticipated to alter their relative expression during a cell-cycle arrest.[11]

Taken together, our data and modeling support the conclusion that RNAPII occupancy is primarily determined via a surprisingly simple mass action dynamic equilibrium. As a cell grows and synthesizes more RNAPII, new RNAPII first enters the free fraction, which increases the free nucleoplasmic concentration. This then drives more polymerase onto the genome until a new equilibrium is established at the increased size. However, this increase in the free concentration is not in direct proportion to cell size, which results in the sub-linear scaling of RNAPII occupancy on the genome that becomes more pronounced as cells grow larger. This sub-linear increase is not due to genome saturation as previously thought. Instead, this is due to a progressively smaller

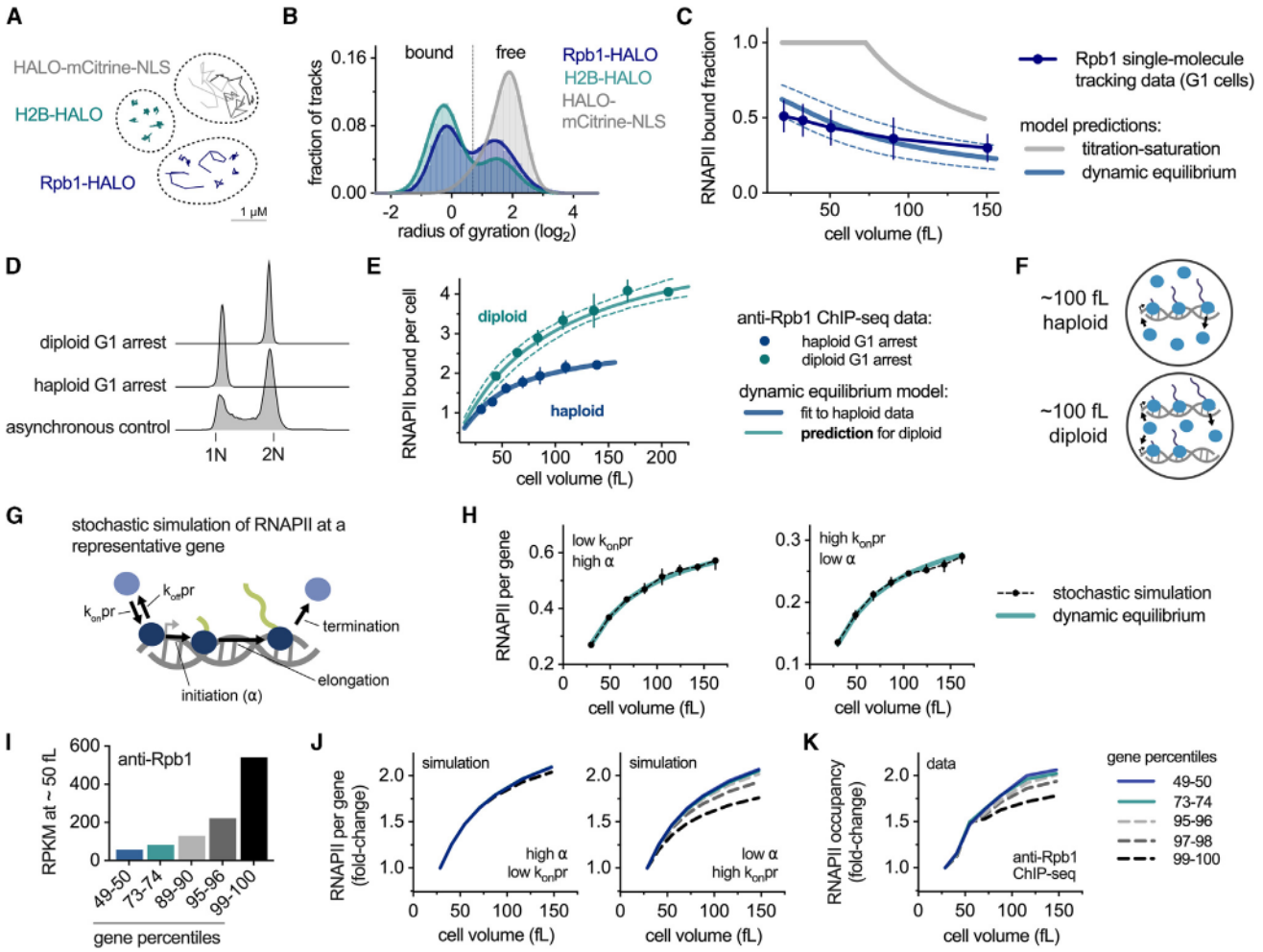


Figure 4. A dynamic equilibrium model explains the size dependence of RNAPII recruitment to the genome

(A–C) Single-molecule imaging was used to determine the DNA-bound fraction of the RNAPII subunit Rpb1. (A) Example single-molecule tracks of the indicated proteins in representative nuclei. (B) Histogram of the radius of gyration, determined from single-molecule imaging, for HALO-mCitrine-NLS fusion proteins (n = 8,057 tracks), H2B-HALO (Htb1; n = 35,978 tracks), and Rpb1-HALO (n = 16,008 tracks). (C) The fraction of DNA-bound Rpb1-HALO molecules, determined from single-molecule imaging, plotted as a function of cell size. Only G1 cells were analyzed. The mean (± SD) for each cell size bin is plotted. A combination of *WT* (n = 182 cells) and *cln3* (n = 87 cells) cells is shown. *cln3* was used to increase the size range. The prediction of the bound fraction from the two alternative models outlined in Figures 3C and 3D is shown. Dashed lines show 90% confidence intervals of the dynamic equilibrium model prediction generated from bootstrapping.

(D–F) Comparison of RNAPII occupancy in haploid and diploid G1-arrested cells. (D) DNA content analysis by flow cytometry of haploid and diploid G1-arrested cells. (E) Rpb1 (anti-Rpb1) occupancy per cell, determined by spike-in normalized ChIP-seq, as a function of cell size in haploid and diploid G1-arrested cells. Each point is the mean (± range) of two biological replicates. The dynamic equilibrium model prediction for diploids is shown and was calculated using the fit to the haploid data. Dashed line shows 90% confidence intervals for the model prediction generated from bootstrapping. Haploid data are also show in Figures 3B and 6A. (F) Summary schematic illustrating the effects of increased ploidy on DNA-bound RNAPII in similarly sized cells.

(G–J) Stochastic simulations of RNAPII on a representative gene. 24 empirically derived feasible parameter sets were simulated (see also Figure S5 and STAR Methods for details). (G) Schematic of the steps modeled in the simulation. (H) Comparison of the simulated RNAPII occupancy at a representative gene with the dynamic equilibrium model. Simulation results from two parameter sets in different regimes of the range of feasible values are shown. (I) The average relative RNAPII occupancy (RPKM) for gene groups binned according to their occupancy in ~50 fL G1-arrested cells. The bin for percentiles 73–74 is closest to the global average of all genes. (J) Fold change in the simulated RNAPII occupancy on a gene representative of each of the gene bins shown in (I). Simulation results from two parameter sets in different regimes of the range of feasible values are shown.

(K) Fold change in the average Rpb1 (anti-Rpb1 ChIP-seq) occupancy with cell size for the different gene groups shown in (I). Data are from the same experiment as the haploid global averages shown in (E).

See also Figures S4 and S5.

fraction of polymerase being depleted from the nucleoplasm onto the genome in larger cells so that there is a smaller and smaller increase in the concentration of free nuclear polymerase that determines loading kinetics as cells enlarge.

The chromatin template is invariant with cell size

A key component of our dynamic equilibrium model is that the on-rate (k_{on}) is independent of cell size so that changes in the loading are only determined by changes in the free

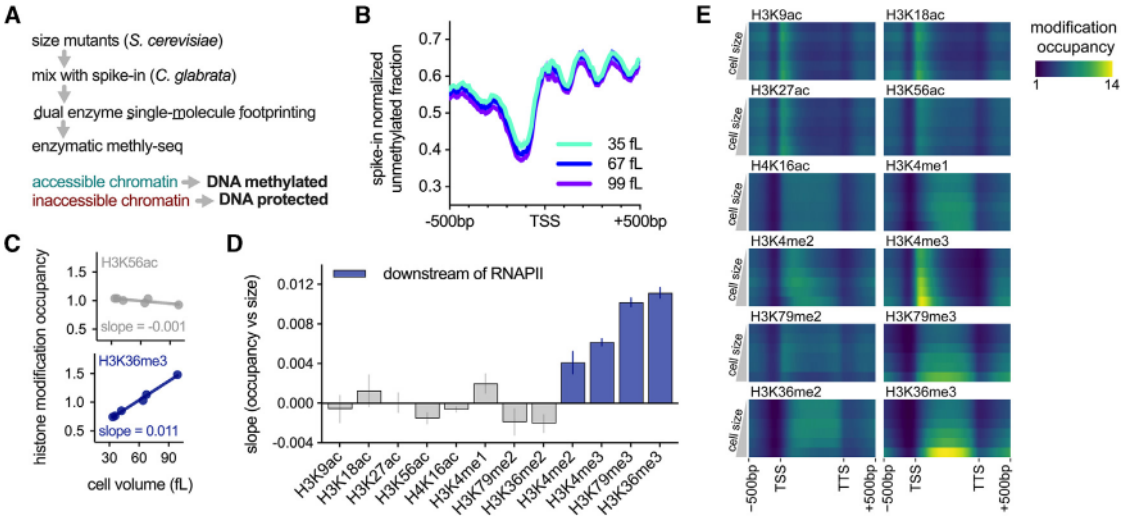


Figure 5. The chromatin environment upstream of transcription is similar in large and small cells

(A) Schematic of the workflow for measuring global chromatin accessibility using spike-in normalized dual-enzyme single-molecule footprinting (dSMF).

(B) Mean unmethylated DNA fraction plotted around transcriptional start sites (TSSs) determined by dSMF in cells of different sizes (see also Figure S6A). Unmethylated DNA corresponds to inaccessible chromatin. Mean (± range) of two technical replicates is shown.

(C–E) Histone modification occupancy measured in the cell size mutants shown in Figure 1C. (C) The occupancy H3K56ac and H3K36me3 plotted against cell size (see also Figure S6B). (D) The slope (± SE) for the linear fit between histone modification occupancy and cell size. Larger slopes correspond to increased modification occupancy in larger cells. Modifications shown in blue are reported to be downstream of RNAPII initiation and/or elongation (see main text). (E) Average occupancy across gene bodies for the indicated modification. Each row corresponds to a different cell size mutant, ordered from small (top) to large (bottom).

See also Figure S6.

nucleoplasmic concentration of RNAPII. One prediction of this is that chromatin accessibility should be similar in large and small cells. To test this, we adapted dual-enzyme single-molecule footprinting (dSMF)[23] for use in yeast. Briefly, this involves treating nuclei with a combination of CpG and GpC methyltransferases so that more accessible DNA is methylated and inaccessible DNA is not. Differences in methylation patterns are then quantified by sequencing and normalized to a spike-in (Figure 5A). When we performed dSMF on cell size mutants spanning an approximately 3-fold range in size, we observed no major differences in chromatin accessibility (Figures 5B and S6A). Consistent with this, we also did not see changes in histone occupancy measured by spike-in normalized ChIP-seq (Figure S4D).

Although the physical accessibility of chromatin is unchanged with size, the chromatin landscape could still be modified by an altered pattern of histone modifications. To test this, we measured a panel of histone modifications implicated in transcription by spike-in normalized ChIP-seq. Most modifications we measured did not change appreciably with cell size. However, four modifications associated with active transcription did increase significantly in larger cells: H3K4me2, H3K4me3, H3K79me3, and H3K36me3 (Figures 5C–5E and S6B). Importantly, these di/tri-methyl marks are known to be deposited downstream of initiation or elongation so that their deposition depends on the recruitment of RNAPII.[24–29] Increased transcriptional frequency is also proposed to broaden the spatial distribution, as well as the absolute intensity, of H3K4me3 as more H3K4me1/2 becomes trimethylated.[29] Indeed, we clearly see this shift in the spatial pattern of H3K4me3 in large cells, further supporting the conclusion that the changes in histone modifications that we observed are a downstream result of increased RNAPII occupancy. We therefore conclude that the chromatin landscape, both in terms of accessibility and histone modifications upstream of RNAPII recruitment, is predominantly invariant with cell size.

mRNA decay feedback partially compensates for sub-linear transcriptional scaling

The equilibrium kinetics of RNAPII produce a sub-linear scaling of RNAPII occupancy that becomes more pronounced as cells get larger. In other words, the increased transcription rate in larger cells is insufficient to increase global mRNA amounts proportionally with cell size. To assess how mRNA amounts are impacted by this, we performed spike-in normalized RNA sequencing (RNA-seq) on a panel of cell size mutants and on cells arrested for increasing amounts of time in G1 after elutriation (Figures 6A and S7A–S7E). As expected, this showed that total mRNA amounts increase monotonically with cell size. However, it also reveals that mRNA amounts increase more than RNAPII occupancy as size increases. Thus, while mRNA concentrations do start to decline in the largest cells, this decline is both smaller and occurs at larger sizes than that predicted from the global transcription rate alone (Figure 6A).

The fact that mRNA amounts increase faster than the transcription rate implies that mRNA turnover rates are reduced in larger cells (Figure 6B). To test this directly, we performed pulse-chase experiments using the uridine analog 5-ethynyl

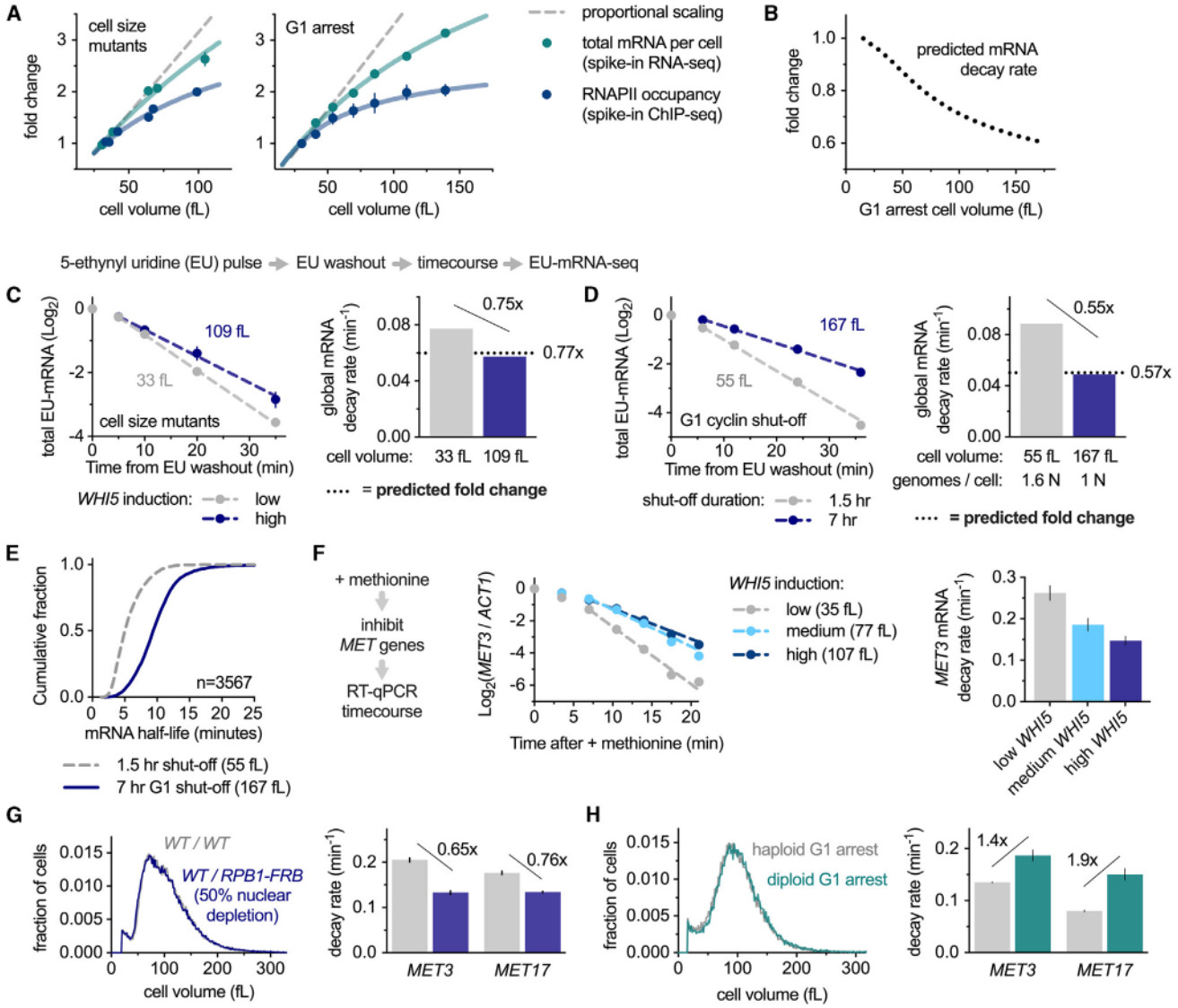


Figure 6. Global mRNA stability increases in larger cells, partially compensating for sub-linear transcriptional scaling

(A) The relative total amount of mRNA per cell, determined by spike-in normalized RNA-seq, in cell size mutants (left, n = 4, mean ± SEM) and during G1 arrest (right, n = 2, mean ± range). Anti-Rpb1 occupancies, determined by spike-in normalized ChIP-seq, from Figures 1E and 3B are shown for comparison. Dashed gray line denotes proportional scaling.

(B) The predicted fold change in mRNA turnover rate during a haploid G1 arrest, calculated from the data in (A).

(C) Spike-in normalized EU pulse-chase experiment to determine global mRNA turnover rates in cells with low (33 fL) or high (109 fL) *WHI5* expression. Left: total EU-mRNA after EU washout (chase) following a 1-hour EU pulse. The mean (± range) of two biological replicates is shown. Dashed line shows an exponential fit to the data. Right: global mRNA turnover rate calculated from the fits in the left panel. Dotted black line shows the predicted change based on the data in (A).

(D) As in (C), for cells after G1 cyclin shut-off. After a 1.5-h shut-off, cells are not fully arrested, and so the average number of genomes per cells is 1.6 N. The prediction for the fold-change accounts for both cell size and DNA content per cell. One biological replicate is shown.

(E) Cumulative frequency distributions of individual mRNA half-lives from the experiment shown in (D).

(F) Left: *MET3* mRNA levels relative to *ACT1* after methionine addition in cells with the indicated size. Right: mean (± range) *MET3* mRNA turnover rates (n = 2) (see also Figures S7K–S7M and S8A–S8D).

(G) Comparison between wild-type diploid cells (*WT*/*WT*) and diploid cells in which 50% of the RNAPII has been depleted from the nucleus (*WT*/*RPB1-FRB*) following a 40-min rapamycin treatment. Cell size distributions (left) and *MET3* and *MET17* mRNA decay rates after methionine addition (right) are shown. The mean (± range) of two biological replicates is shown (see also Figure S8E).

(H) As in (G), but a comparison between ~100 fL diploid cells and ~100 fL haploid cells that have been arrested in G1 via G1 cyclin shut-off (see also Figure S8F).

See also Figures S7 and S8.

uridine (EU). We quantified EU-labeled mRNA by spike-in normalized EU-mRNA-seq, following EU washout. This allowed us to directly monitor global mRNA turnover kinetics in unperturbed conditions across cells of different sizes (Figures S7F–S7I). This clearly reveals that in both the cell size mutants and the G1 arrest conditions, EU-labeled mRNA is turned over more slowly in larger cells (Figures 6C-6E). Crucially, when we calculate the fold change in total mRNA decay rate from these data, we find a very close quantitative agreement to a prediction of how the decay rate should change based on our measurements of mRNA amounts per cell and RNAPII occupancy that are shown in Figure 6A.

To validate this global finding that mRNA is more stable in larger cells, using an independent approach, we conditionally repressed the transcription of *MET* genes and then measured *MET3* and *MET17* mRNA by RT-qPCR. This shows a clear decrease in the mRNA decay rate in larger cells (Figures 6F and S8A–S8D). Similar experiments to repress the *GAL* genes *GAL1*, *GAL7*, and *GAL10* also showed the same size-dependent trend (Figures S7K–S7M). Thus, mRNA turnover rates decrease with size to stabilize the transcriptome in larger cells and thereby promote the scaling of mRNA amounts with cell size.

Next, we considered whether the regulation of mRNA turnover was directly coupled to cell size or reflected a form of feedback on mRNA concentration or transcription. To do this, we examined mRNA turnover rates in cells subject to a rapid depletion of half the nuclear RNAPII, which results in a ~40% reduction in RNAPII occupancy on the genome (see Figures 2C and 2D). In these experiments, *MET3* and *MET17* mRNA turnover rates also decreased even though cell size does not change (Figures 6G and S8E). Similarly, we also observed that ~100 fL G1 haploid cells have lower mRNA turnover rates than ~100 fL G1 diploid cells (Figures 6H and S8F), consistent with the lower global transcription rate in haploids compared with diploids of a similar cell size (Figures 4E and 4F). Taken together, these experiments suggest that the compensatory adjustment in mRNA turnover rates in larger sizes is more likely due to feedback on global mRNA concentrations rather than any direct coupling to cell size.

DISCUSSION

How cells scale global mRNA amounts with cell size during growth is a long-standing question in cell biology.[1–8,10] It was previously suggested that the size scaling of mRNA was due to its synthesis rates increasing in direct proportion to cell size. In these models, such size scaling was proposed to be due to an unknown limiting transcription factor being titrated against the genome.[8,16]

Transcriptional scaling occurs via a dynamic equilibrium model for limiting RNAPII

Here, we empirically identified RNAPII as a major limiting factor in budding yeast as polymerase loading onto the genome is sensitive to both increases and decreases in its own concentration. In contrast, a 2-fold reduction in nuclear concentrations of other components of the transcriptional initiation machinery resulted in only minimal reductions in RNAPII loading (Figure 2). While our work shows that RNAPII is limiting for transcription, we also identify important differences from previous versions of the limiting factor model that were based on titration against the genome. We found that RNAPII is not titrated against the genome as about half the RNAPII is freely diffusing in the nucleoplasm, as revealed by single-molecule imaging experiments. We note that a previous study claimed to have shown RNAPII is titrated against the genome in *S. pombe* and is therefore likely to be limiting for transcription. However, this conclusion was based on confocal imaging of the localization of RNAPII subunits, which cannot be used to distinguish the DNA-bound and unbound states.[8]

While our data in budding yeast show that the amount of active RNAPII increases in larger cells, it does not scale in direct proportion with cell size so that larger cells exhibited systematically larger deviations from the amount of RNAPII engaged on the genome predicted by titrated factor models. This observation led us to propose a dynamic equilibrium model in which the loading of RNAPII on the genome is determined by mass action kinetics with a defined on-rate k_{on}, which is independent of cell size (Figures 3 and 4). The size independence of k_{on} is consistent with our measurements showing that there are no major changes in chromatin accessibility as cells grow larger (Figure 5). In the dynamic equilibrium model, newly synthesized RNAPII increases its free nucleoplasmic concentration as a cell grows. This drives more polymerase onto the genome, thereby establishing a new equilibrium between the bound and unbound RNAPII populations at the increased size. The increase in free nucleoplasmic RNAPII concentrations is a combined effect of the underlying equilibrium kinetics and the modest decline in the nuclear-to-cell volume ratio in budding yeast, which occurs as cells increase in size.[21] We note that a similar decline in the nuclear-to-cell volume ratio is observed in other cell types.[30] While our model quantitively predicts the relationships between RNAPII occupancy, cell size, nuclear size, and ploidy (Figures 3 and 4), it is unclear whether RNAPI and RNAPIII activities are governed in a similar way or whether additional regulatory inputs are involved in the scaling of rRNA and tRNA production. Investigating this will be an important next step toward a complete picture of how biosynthesis and cell size are coordinated.

mRNA concentrations are buffered in larger cells by feedback on mRNA turnover

While our dynamic equilibrium model successfully accounts for the size-dependent increase in RNAPII loading that we have quantified, this alone is insufficient to explain the scaling of mRNA amounts. This is because global mRNA concentrations are not appreciably diluted as RNAPII loading first deviates from proportional scaling. This observation led us to identify a reduction in mRNA turnover in larger cells as a secondary process promoting mRNA scaling with size (Figure 6). If mRNA half-lives were independent of cell size, as previously thought,[6–8] mRNA concentrations would decrease faster in larger cells than we observed. Thus, mRNA stability increases in larger cells to partially compensate for the sub-linear size dependence of transcription (Figure 6). Taken together, our data show how limiting RNAPII acts in concert with feedback on mRNA stability to increase mRNA amounts nearly in proportion to cell size so that

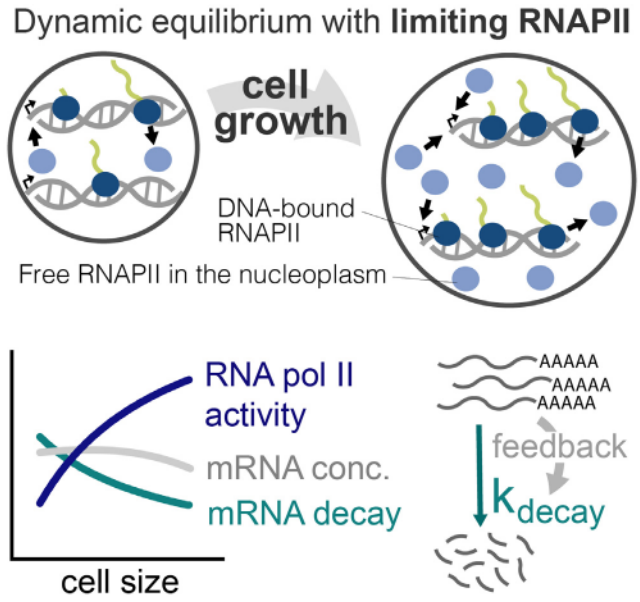


Figure 7. Summary schematic

A dynamic equilibrium drives increase loading of limiting RNAPII onto the genome in larger cells, resulting in a sub-linear increase in transcription with cell size. Feedback on mRNA decay increases mRNA stability in larger cells to partially compensate for the sub-linear transcriptional trend, thereby extending the range of sizes over which mRNA concentrations are close to constant.

concentrations remain closer to constant as cells grow larger (Figure 7).

We propose that the adjustment in mRNA turnover rates in larger cells is driven by a form of feedback. The presence of compensatory modulation of mRNA stability is highly reminiscent of the previous observations, in both yeast and human cells, that pertubations affecting transcription are buffered by a decrease in mRNA turnover.[9,31–37] This is most dramatically seen in a recent study that examined global transcription rates and RNA concentrations in a genome-wide small interfering RNA (siRNA) screen. While perturbations affecting transcription were numerous, these did not result in changes to RNA concentration.[9] Thus, growing larger and perturbation to transcription both elicit a similar feedback effect to buffer mRNA concentrations. Consistent with this, we observed that yeast cells of the same size with quantitative differences in global RNAPII occupancy also show corresponding changes in mRNA stability. Thus, we anticipate that the feedback that maintains mRNA concentrations in growing cells and in response to transcriptional mutations arises from the same underlying molecular mechanism. Indeed, natural variations in cell size may provide the physiological context for which this mRNA stability feedback first evolved.

We note that while previous studies have reported no obvious size-dependent changes in mRNA turnover for a handful of candidate transcripts,[6–8] these studies have relied on transcriptional inhibitors such as thiolutin or actinomycin D. Since global transcriptional inhibition severely reduces mRNA turnover rates, it is not possible to use these data to quantitively evaluate the size-dependent changes in mRNA turnover of unperturbed and actively transcribing cells.

Transcriptional scaling and optimal size ranges for cellular fitness and function

While the mechanisms we identify here work to keep mRNA concentrations close to constant at physiological sizes, the scaling of macromolecule amounts with cell size breaks down when cells grow excessively large. This causes global RNA and protein concentrations to decline, effectively diluting the cytoplasm.[11] Very large cell size is also associated with a decline in many aspects of cellular physiology, including reduced growth rate, attenuated signal transduction, and defective conditional gene expression programs.[11] In addition, the fact that proliferation is attenuated by increased cell size has recently been isolated as a causal contributor to senescence, the tumor suppressive state of irreversible cell-cycle exit characterized by massive cell enlargement.[11,13,14,38,39] Moreover, stem cells are among the tiniest in the body, but they tend to enlarge as an organism ages[40–42] and naturally or artificially enlarged hematopoietic or intestinal stem cells accelerate the aging-associated decline in stem cell function.[40]

The size-dependent patterns of global RNA and protein scaling are likely to be important for determining these optimal and non-optimal size range in different circumstances. However, it has not previously been established why cells cannot scale biosynthesis with cell size above this upper size limit. Our data provide a quantitative explanation for this, as we observed that RNAPII recruitment to the genome does not increase in direct proportion to cell size and eventually reaches a plateau (Figures 1 and 3). The plateau does not arise because of the saturation of the genome with RNAPII, as has been previously suggested,[16] but is instead a natural consequence of the mass action equilibrium kinetics that govern RNAPII loading. As cells grow larger, there is a progressively more pronounced divergence from a proportional relationship between cell size and global transcription. Initially, global mRNA concentrations are not meaningfully affected because the sub-linear transcriptional scaling is buffered by feedback on mRNA decay rates to stabilize the transcriptome in larger cells. However, this balancing act between global transcription and mRNA decay slowly breaks down as the feedback on mRNA stability is no longer able to effectively compensate for the ever-increasing divergence between transcription and cell size. We speculate that the upper cell size limit on cell function corresponds to the point at which increased transcription and RNA turnover regulation can no longer keep RNA concentrations close to constant. Consistent with this, doubling the ploidy increases the size-dependent loading of RNAPII onto the genome (Figures 4D–4F) and increases the upper cell size at which cells can function and grow efficiently.[11] This also provides a simple rationalization as to why polyploidy is commonly coupled to increased cell size in nature; namely, polyploidy is a simple route to promote efficient cellular biosynthesis in bigger cells by scaling transcription with cell size over a wider size range.

Limitations of the study

One unanticipated finding of our work is the role of mRNA degradation alongside transcription for mRNA size scaling. Here, we have focused on the mechanisms responsible for

global transcriptional changes with cell size, and so an important unresolved question is how, mechanistically, is global mRNA degradation modulated to buffer mRNA concentrations? We suggest this is via some form of feedback that might work by sensing cytoplasmic mRNA concentrations. However, we have not tested this directly and are yet to identify which steps of mRNA degradation are affected or which enzymes and regulators are responsible. A second limitation is that our study is focused on the scaling of total mRNA. However, it is also well established that individual transcripts can be uncoupled from these global trends to decrease or increase their concentrations with cell size. These behaviors are known as sub- and super-scaling, respectively.[13,43–45] Our modeling suggests that one feasible mechanism is that the promoters of highly expressed genes could start to saturate in larger cells (Figure 4), which may explain why highly expressed genes such as the histones sub-scale.[43,46] However, this does not satisfactorily explain how low-expressed transcripts such as the cell-cycle inhibitor *WHI5* are sub-scaling[43,47–49] or how super-scaling trends arise.

STAR METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2023.10.012.

ACKNOWLEDGMENTS

We thank Christine Jacobs-Wagner, Jim Ferrell, Kurt Schmoller, Helena Cantwell, Mariam Abdullah, and the Skotheim lab for feedback on the manuscript. We thank Mart Loog for discussion on the equilibrium model and Carl Wu and Vu Nguyen for input on parameterizing the RNAPII simulations. This work was supported by the NSF (MCB 2040908) and the Chan Zuckerberg Biohub. M.P.S. was supported by a Simons Foundation Fellowship of the LSRF and an EMBO Long-Term Postdoctoral Fellowship.

AUTHOR CONTRIBUTIONS

M.P.S. and J.M.S. conceived the study. M.P.S. designed, performed, and analyzed the experiments, except for the single-molecule tracking (H.Z. and R.R.-L.) and the proteomics (J.G. and A.W.J.). G.K.M. processed all sequencing data. L.F.V. wrote the model and simulations. M.P.S. and J.M.S. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 10, 2022
Revised: July 16, 2023
Accepted: October 10, 2023
Published: November 8, 2023

REFERENCES

1. Elliott, S.G., and McLaughlin, C.S. (1979). Regulation of RNA synthesis in yeast. III. Synthesis during the cell cycle. Mol. Gen. Genet. *169*, 237–243. https://doi.org/10.1007/BF00382269.
2. Fraser, R.S., and Nurse, P. (1978). Novel cell cycle control of RNA synthesis in yeast. Nature *271*, 726–730. https://doi.org/10.1038/271726a0.
3. Fraser, R.S., and Nurse, P. (1979). Altered patterns of ribonucleic acid synthesis during the cell cycle: a mechanism compensating for variation in gene concentration. J. Cell Sci. *35*, 25–40.
4. Elliott, S.G. (1983). Coordination of growth with cell division: regulation of synthesis of RNA during the cell cycle of the fission yeast Schizosaccharomyces pombe. Mol. Gen. Genet. *192*, 204–211. https://doi.org/10.1007/BF00327667.
5. Elliott, S.G. (1983). Regulation of the maximal rate of RNA synthesis in the fission yeast Schizosaccharomyces pombe. Mol. Gen. Genet. *192*, 212–217. https://doi.org/10.1007/BF00327668.

6. Zhurinsky, J., Leonhard, K., Watt, S., Marguerat, S., Bähler, J., and Nurse, P. (2010). A coordinated global control over cellular transcription. Curr. Biol. *20*, 2010–2015. https://doi.org/10.1016/j.cub.2010.10.002.

7. Padovan-Merhar, O., Nair, G.P., Biaesch, A.G., Mayer, A., Scarfone, S., Foley, S.W., Wu, A.R., Churchman, L.S., Singh, A., and Raj, A. (2015). Single mammalian cells compensate for differences in cellular volume and DNA copy number through independent global transcriptional mechanisms. Mol. Cell *58*, 339–352. https://doi.org/10.1016/j.molcel.2015.03.005.

8. Sun, X.M., Bowman, A., Priestman, M., Bertaux, F., Martinez-Segura, A., Tang, W., Whilding, C., Dormann, D., Shahrezaei, V., and Marguerat, S. (2020). Size-dependent increase in RNA polymerase II initiation rates mediates gene expression scaling with cell size. Curr. Biol. *30*, 1217–1230.e7. https://doi.org/10.1016/j.cub.2020.01.053.

9. Berry, S., Müller, M., Rai, A., and Pelkmans, L. (2022). Feedback from nuclear RNA on transcription promotes robust RNA concentration homeostasis in human cells. Cell Syst. *13*, 454–470.e15. https://doi.org/10.1016/j.cels.2022.04.005.

10. Basier, C., and Nurse, P. (2023). The cell cycle and cell size influence the rates of global cellular translation and transcription in fission yeast. EMBO J. *42*, e113333. https://doi.org/10.15252/embj.2022113333.

11. Neurohr, G.E., Terry, R.L., Lengefeld, J., Bonney, M., Brittingham, G.P., Moretto, F., Miettinen, T.P., Vaites, L.P., Soares, L.M., Paulo, J.A., et al. (2019). Excessive cell growth causes cytoplasm dilution and contributes to senescence. Cell *176*, 1083–1097.e18. https://doi.org/10.1016/j.cell.2019.01.018.

12. Lengefeld, J., Cheng, C.-W., Maretich, P., Blair, M., Hagen, H., McReynolds, M.R., Sullivan, E., Majors, K., Roberts, C., Kang, J.H., et al. (2020). Cell size is a determinant of stem cell potential during aging. Sci. Adv. *7*, eabk0271. https://doi.org/10.1126/sciadv.abk0271.

13. Lanz, M.C., Zatulovskiy, E., Swaffer, M.P., Zhang, L., Zhang, S., You, D.S., Marinov, G.K., McAlpine, P., Elias, J.E., and Skotheim, J.M. (2021). Increasing cell size remodels the proteome and promotes senescence. Mol. Cell *82*, 3255–3269.e8.

14. Cheng, L., Chen, J., Kong, Y., Tan, C., Kafri, R., and Björklund, M. (2021). Size-scaling promotes senescence-like changes in proteome and organelle content. Preprint at bioRxiv. https://doi.org/10.1101/2021.08.05.455193.

15. Xie, S., Swaffer, M., and Skotheim, J.M. (2022). Eukaryotic cell size control and its relation to biosynthesis and senescence. Annu. Rev. Cell Dev. Biol. *38*, 291–319. https://doi.org/10.1146/annurev-cellbio-120219-040142.

16. Lin, J., and Amir, A. (2018). Homeostasis of protein and mRNA concentrations in growing cells. Nat. Commun. *9*, 4496. https://doi.org/10.1038/s41467-018-06714-z.

17. Marguerat, S., and Bähler, J. (2012). Coordinating genome expression with cell size. Trends Genet. *28*, 560–565. https://doi.org/10.1016/j.tig.2012.07.003.

18. Hu, B., Petela, N., Kurze, A., Chan, K.L., Chapard, C., and Nasmyth, K. (2015). Biological chromodynamics: a general method for measuring protein occupancy across the genome by calibrating ChIP-seq. Nucleic Acids Res. *43*, e132. https://doi.org/10.1093/nar/gkv670.

19. Kustatscher, G., Wills, K.L., Furlan, C., and Rappsilber, J. (2014). Chromatin enrichment for proteomics. Nat. Protoc. *9*, 2090–2099. https://doi.org/10.1038/nprot.2014.142.

20. Haruki, H., Nishikawa, J., and Laemmli, U.K. (2008). The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes. Mol. Cell *31*, 925–932. https://doi.org/10.1016/j.molcel.2008.07.020.

21. Jorgensen, P., Edgington, N.P., Schneider, B.L., Rupes, I., Tyers, M., and Futcher, B. (2007). The size of the nucleus increases as yeast cells grow. Mol. Biol. Cell *18*, 3523–3532. https://doi.org/10.1091/mbc.E06-10-0973.

22. Nguyen, V.Q., Ranjan, A., Liu, S., Tang, X., Ling, Y.H., Wisniewski, J., Mizuguchi, G., Li, K.Y., Jou, V., Zheng, Q., et al. (2021). Spatiotemporal coordination of transcription preinitiation complex assembly in live cells. Mol. Cell *81*, 3560–3575.e6. https://doi.org/10.1016/j.molcel.2021.07.022.

23. Krebs, A.R., Imanci, D., Hoerner, L., Gaidatzis, D., Burger, L., and Schübeler, D. (2017). Genome-wide single-molecule footprinting reveals high RNA polymerase II turnover at paused promoters. Mol. Cell *67*, 411–422.e4. https://doi.org/10.1016/j.molcel.2017.06.027.

24. Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C., Schreiber, S.L., Mellor, J., and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. Nature *419*, 407–411. https://doi.org/10.1038/nature01080.

25. Krogan, N.J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D.P., Beattie, B.K., Emili, A., Boone, C., et al. (2003). Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. Mol. Cell Biol. *23*, 4207–4218. https://doi.org/10.1128/MCB.23.12.4207-4218.2003.

26. Ng, H.H., Robert, F., Young, R.A., and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. Mol. Cell *11*, 709–719. https://doi.org/10.1016/s1097-2765(03)00092-3.

27. Krogan, N.J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M.A., Dean, K., Ryan, O.W., Golshani, A., Johnston, M., et al. (2003). The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. Mol. Cell *11*, 721–729. https://doi.org/10.1016/s1097-2765(03)00091-1.

28. Xiao, T., Hall, H., Kizer, K.O., Shibata, Y., Hall, M.C., Borchers, C.H., and Strahl, B.D. (2003). Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. Genes Dev. *17*, 654–663. https://doi.org/10.1101/gad.1055503.

29. Soares, L.M., He, P.C., Chun, Y., Suh, H., Kim, T., and Buratowski, S. (2017). Determinants of histone H3K4 methylation patterns. Mol. Cell *68*, 773–785.e6. https://doi.org/10.1016/j.molcel.2017.10.013.

30. Malerba, M.E., and Marshall, D.J. (2021). Larger cells have relatively smaller nuclei across the Tree of Life. Evol Lett *5*, 306–314. https://doi.org/10.1002/evl3.243.

31. Sun, M., Schwalb, B., Schulz, D., Pirkl, N., Etzold, S., Larivière, L., Maier, K.C., Seizl, M., Tresch, A., and Cramer, P. (2012). Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. Genome Res. *22*, 1350–1359. https://doi.org/10.1101/gr.130161.111.

32. Schulz, D., Pirkl, N., Lehmann, E., and Cramer, P. (2014). Rpb4 subunit functions mainly in mRNA synthesis by RNA polymerase II. J. Biol. Chem. *289*, 17446–17452. https://doi.org/10.1074/jbc.M114.568014.

33. Warfield, L., Ramachandran, S., Baptista, T., Devys, D., Tora, L., and Hahn, S. (2017). Transcription of nearly all yeast RNA polymerase II-transcribed genes is dependent on transcription factor TFIID. Mol. Cell *68*, 118–129.e5. https://doi.org/10.1016/j.molcel.2017.08.014.

34. Rodríguez-Molina, J.B., Tseng, S.C., Simonett, S.P., Taunton, J., and Ansari, A.Z. (2016). Engineered covalent inactivation of TFIIH-kinase reveals an elongation checkpoint and results in widespread mRNA stabilization. Mol. Cell *63*, 433–444. https://doi.org/10.1016/j.molcel.2016.06.036.

35. Baptista, T., Grünberg, S., Minoungou, N., Koster, M.J.E., Timmers, H.T.M., Hahn, S., Devys, D., and Tora, L. (2017). SAGA is a general cofactor for RNA polymerase II transcription. Mol. Cell *68*, 130–143.e5. https://doi.org/10.1016/j.molcel.2017.08.016.

36. Slobodin, B., Bahat, A., Sehrawat, U., Becker-Herman, S., Zuckerman, B., Weiss, A.N., Han, R., Elkon, R., Agami, R., Ulitsky, I., et al. (2020). Transcription dynamics regulate poly(A) tails and expression of the RNA

degradation machinery to balance mRNA levels. Mol. Cell *78*, 434–444.e5. https://doi.org/10.1016/j.molcel.2020.03.022.

37. Helenius, K., Yang, Y., Tselykh, T.V., Pessa, H.K., Frilander, M.J., and Mäkelä, T.P. (2011). Requirement of TFIIH kinase subunit Mat1 for RNA Pol II C-terminal domain Ser5 phosphorylation, transcription and mRNA turnover. Nucleic Acids Res. *39*, 5025–5035. https://doi.org/10.1093/nar/gkr107.
38. Wilson, G.A., Sava, G., Vuina, K., Huard, C., Meneguello, L., Coulombe-Huntington, J., Bertomeu, T., Maizels, R.J., Lauring, J., Tyers, M., et al. (2021). Active growth signalling promotes cancer cell sensitivity to the CDK7 inhibitor ICEC0942. Preprint at bioRxiv. https://doi.org/10.1101/2021.09.10.459733.
39. Demidenko, Z.N., and Blagosklonny, M.V. (2008). Growth stimulation leads to cellular senescence when the cell cycle is blocked. Cell Cycle *7*, 3355–3361. https://doi.org/10.4161/cc.7.21.6919.
40. Lengefeld, J., Cheng, C.W., Maretich, P., Blair, M., Hagen, H., McReynolds, M.R., Sullivan, E., Majors, K., Roberts, C., Kang, J.H., et al. (2021). Cell size is a determinant of stem cell potential during aging. Sci. Adv. *7*, eabk0271. https://doi.org/10.1126/sciadv.abk0271.
41. Yang, Y.K., Ogando, C.R., Wang See, C., Chang, T.Y., and Barabino, G.A. (2018). Changes in phenotype and differentiation potential of human mesenchymal stem cells aging in vitro. Stem Cell Res. Ther. *9*, 131. https://doi.org/10.1186/s13287-018-0876-3.
42. Li, Q., Rycaj, K., Chen, X., and Tang, D.G. (2015). Cancer stem cells and cell size: a causal link? Semin. Cancer Biol. *35*, 191–199. https://doi.org/10.1016/j.semcancer.2015.07.002.
43. Swaffer, M.P., Kim, J., Chandler-Brown, D., Langhinrichs, M., Marinov, G., Greenleaf, W., Kundaje, A., Schmoller, K.M., and Skotheim, J.M. (2021). Size-independent mRNA synthesis and chromatin-based partitioning mechanisms generate and maintain constant amounts of protein per cell. Mol. Cell *81*, 4861–4875.e7.
44. Chen, Y., Zhao, G., Zahumensky, J., Honey, S., and Futcher, B. (2020). Differential scaling of gene expression with cell size may explain size control in budding yeast. Mol. Cell *78*, 359–370.e6. https://doi.org/10.1016/j.molcel.2020.03.012.
45. Keifenheim, D., Sun, X.M., D'Souza, E., Ohira, M.J., Magner, M., Mayhew, M.B., Marguerat, S., and Rhind, N. (2017). Size-dependent expression of the mitotic activator Cdc25 suggests a mechanism of size control in fission yeast. Curr. Biol. *27*, 1491–1497.e4. https://doi.org/10.1016/j.cub.2017.04.016.
46. Claude, K.L., Bureik, D., Chatzitheodoridou, D., Adarska, P., Singh, A., and Schmoller, K.M. (2021). Transcription coordinates histone amounts and genome content. Nat. Commun. *12*, 4202. https://doi.org/10.1038/s41467-021-24451-8.
47. Schmoller, K.M., Turner, J.J., Kõivomägi, M., and Skotheim, J.M. (2015). Dilution of the cell cycle inhibitor Whi5 controls budding-yeast cell size. Nature *526*, 268–272. https://doi.org/10.1038/nature14908.
48. Schmoller, K.M., Lanz, M.C., Kim, J., Koivomagi, M., Qu, Y., Tang, C., Kukhtevich, I.V., Schneider, R., Rudolf, F., Moreno, D.F., et al. (2022). Whi5 is diluted and protein synthesis does not dramatically increase in pre-Start G1. Mol. Biol. Cell *33*, lt1. https://doi.org/10.1091/mbc.E21-01-0029.
49. Qu, Y., Jiang, J., Liu, X., Wei, P., Yang, X., and Tang, C. (2019). Cell cycle inhibitor Whi5 records environmental information to coordinate growth and division in yeast. Cell Rep. *29*, 987–994.e5. https://doi.org/10.1016/j.celrep.2019.09.030.
50. Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25. https://doi.org/10.1186/gb-2009-10-3-r25.
51. Roberts, A., and Pachter, L. (2013). Streaming fragment assignment for real-time analysis of sequencing experiments. Nat. Methods *10*, 71–73. https://doi.org/10.1038/nmeth.2251.
52. Tinevez, J.Y., Perry, N., Schindelin, J., Hoopes, G.M., Reynolds, G.D., Laplantine, E., Bednarek, S.Y., Shorte, S.L., and Eliceiri, K.W. (2017). TrackMate: an open and extensible platform for single-particle tracking. Methods *115*, 80–90. https://doi.org/10.1016/j.ymeth.2016.09.016.
53. Ewald, J.C., Kuehne, A., Zamboni, N., and Skotheim, J.M. (2016). The yeast cyclin-dependent kinase routes carbon fluxes to fuel cell cycle progression. Mol. Cell *62*, 532–545. https://doi.org/10.1016/j.molcel.2016.02.017.
54. Wosika, V., Durandau, E., Varidel, C., Aymoz, D., Schmitt, M., and Pelet, S. (2016). New families of single integration vectors and gene tagging plasmids for genetic manipulations in budding yeast. Mol. Genet. Genomics *291*, 2231–2240. https://doi.org/10.1007/s00438-016-1249-1.
55. Azizoglu, A., Brent, R., and Rudolf, F. (2021). A precisely adjustable, variation-suppressed eukaryotic transcriptional controller to enable genetic discovery. eLife *10*. https://doi.org/10.7554/eLife.69549.
56. Walters, A.D., Amoateng, K., Wang, R., Chen, J.H., McDermott, G., Larabell, C.A., Gadal, O., and Cohen-Fix, O. (2019). Nuclear envelope expansion in budding yeast is independent of cell growth and does not determine nuclear volume. Mol. Biol. Cell *30*, 131–145. https://doi.org/10.1091/mbc.E18-04-0204.
57. Picelli, S., Björklund, A.K., Reinius, B., Sagasser, S., Winberg, G., and Sandberg, R. (2014). Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. Genome Res. *24*, 2033–2040. https://doi.org/10.1101/gr.177881.114.
58. Corces, M.R., Trevino, A.E., Hamilton, E.G., Greenside, P.G., Sinnott-Armstrong, N.A., Vesuna, S., Satpathy, A.T., Rubin, A.J., Montine, K.S., Wu, B., et al. (2017). An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. Nat. Methods *14*, 959–962. https://doi.org/10.1038/nmeth.4396.
59. Rouillon, A., Barbey, R., Patton, E.E., Tyers, M., and Thomas, D. (2000). Feedback-regulated degradation of the transcriptional activator Met4 is triggered by the SCF(Met30) complex. EMBO J. *19*, 282–294. https://doi.org/10.1093/emboj/19.2.282.
60. Haimovich, G., Medina, D.A., Causse, S.Z., Garber, M., Millán-Zambrano, G., Barkai, O., Chávez, S., Pérez-Ortín, J.E., Darzacq, X., and Choder, M. (2013). Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. Cell *153*, 1000–1011. https://doi.org/10.1016/j.cell.2013.05.012.
61. Hughes, C.S., Moggridge, S., Müller, T., Sorensen, P.H., Morin, G.B., and Krijgsveld, J. (2019). Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. Nat. Protoc. *14*, 68–85. https://doi.org/10.1038/s41596-018-0082-x.
62. Neumann, F.R., and Nurse, P. (2007). Nuclear size control in fission yeast. J. Cell Biol. *179*, 593–600. https://doi.org/10.1083/jcb.200708054.
63. Mason, P.B., and Struhl, K. (2005). Distinction and relationship between elongation rate and processivity of RNA polymerase II in vivo. Mol. Cell *17*, 831–840. https://doi.org/10.1016/j.molcel.2005.02.017.
64. Ho, B., Baryshnikova, A., and Brown, G.W. (2018). Unification of protein abundance datasets yields a quantitative Saccharomyces cerevisiae proteome. Cell Syst. *6*, 192–205.e3. https://doi.org/10.1016/j.cels.2017.12.004.
65. Rhee, H.S., and Pugh, B.F. (2012). Genome-wide structure and organization of eukaryotic pre-initiation complexes. Nature *483*, 295–301. https://doi.org/10.1038/nature10799.
66. Ho, Y.H., Shishkova, E., Hose, J., Coon, J.J., and Gasch, A.P. (2018). Decoupling yeast cell division and stress defense implicates mRNA repression in translational reallocation during stress. Curr. Biol. *28*, 2673–2680.e4. https://doi.org/10.1016/j.cub.2018.06.044.
67. Skrzypek, M.S., Binkley, J., Binkley, G., Miyasato, S.R., Simison, M., and Sherlock, G. (2017). The Candida Genome Database (CGD): incorporation of Assembly 22, systematic identifiers and visualization of high throughput

sequencing data. Nucleic Acids Res. *45*, D592–D596. https://doi.org/10.1093/nar/gkw924.

68. Shipony, Z., Marinov, G.K., Swaffer, M.P., Sinnott-Armstrong, N.A., Skotheim, J.M., Kundaje, A., and Greenleaf, W.J. (2020). Long-range single-molecule mapping of chromatin accessibility in eukaryotes. Nat. Methods *17*, 319–327. https://doi.org/10.1038/s41592-019-0730-2.

69. Lock, A., Rutherford, K., Harris, M.A., Hayles, J., Oliver, S.G., Bähler, J., and Wood, V. (2019). PomBase 2018: user-driven reimplementation of the fission yeast database provides rapid and intuitive access to diverse, interconnected information. Nucleic Acids Res. *47*, D821–D827. https://doi.org/10.1093/nar/gky961.

70. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21. https://doi.org/10.1093/bioinformatics/bts635.

STAR METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-Rpb1 (clone 8wG16)	Millipore	*Cat 05-952; RRID:AB_492629*
anti-Rpb1-S2P (clone 3E10)	Millipore	*Cat 04-1571; RRID:AB_11212363*
anti-Rpb1-S5P (clone 3E8)	Millipore	*Cat 04-1572; RRID:AB_11213421*
anti-FLAG (clone M2)	Sigma-Aldrich	Cat F3165; RRID:AB_259529
anti-H3	Abcam	Cat ab1791; RRID:AB_302613
anti-H4	Abcam	Cat ab10158; RRID:AB_296888
anti-H4K9ac	Abcam	Cat ab4441; RRID:AB_2118292
anti-H3K56ac	Millipore	Cat 07-677; RRID:AB_390167
anti-H3K4me1	Abcam	Cat ab8895; RRID:AB_306847
anti-H3K4me2	Abcam	Cat ab7766; RRID:AB_2560996
anti-H3K4me3	Abcam	Cat ab1012; RRID:AB_442796
anti-H3K79me2	Abcam	Cat ab3594; RRID:AB_303937
anti-H3K79me3	Abcam	Cat ab2621; RRID:AB_303215
anti-H3K36me2	Abcam	Cat ab9049; RRID:AB_1280939
anti-H3K36me3	Abcam	Cat ab9050; RRID:AB_306966
anti-H3K18ac	Abcam	Cat ab1191; RRID:AB_298692
anti-H3K27ac	Millipore	Cat 07-360; RRID:AB_310550
anti-H4K16ac	Millipore	Cat 07-329; RRID:AB_310525
Chemicals, peptides, and recombinant proteins		
Beta-estradiol	MP Biomedicals	Cat 0210165605
Anhydrotetracycline	Cayman Chemical	Cat 10009542
Rapamycin	Cayman Chemical	Cat 13346
Critical commercial assays		
NEBNext Ultra II DNA Library Prep kit	NEB	Cat E7645
NEBNext Enzymatic Methyl-seq Kit	NEB	Cat E7120
NEBNext Poly(A) mRNA Magnetic Isolation Module	NEB	Cat E7490
NEBNext Ultra II Directional RNA Library Prep Kit for Illumina	NEB	Cat E7760
Click-iT Nascent RNA Capture Kit	Invitrogen	Cat C10365
Dynabeads MyOne Streptavidin T1	Invitrogen	Cat 65601
Deposited data		
ChIP-seq data	This study	GEO: GSE242874
RNA-seq data	This study	GEO: GSE242874
Experimental models: Organisms/strains		
S. pombe: PN10597: h- leu1-32::pFS181[adh1 hENT1::leu1+]_ his7336::pJL218[adh1-hsvTK::his7+]	Nurse lab	PN10597
S. pombe: MSsp109: h- Rpb1-3xFlag::KanMX6	This study	MSsp109
C. glabrata: MS99: WT	Doncic lab	MS99
S. Cerevisiae: W303 and derived strains	This study	See Table S1 for strain list
Software and algorithms		
Bowtie (v.1.0.1)	Langmead et al.[50]	N/A
eXpress (version 1.5.1)	Roberts and Pachter [51]	N/A
Trackmate	Tinevez et al.[52]	N/A
Code for the dynamic equilibrium model	This study	https://doi.org/10.5281/zenodo.8327425
Code for the stochastic simulations of RNAPII transcription at a representative gene	This study	https://doi.org/10.5281/zenodo.8327425

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Jan Skotheim (skotheim@stanford.edu).

Materials availability

All plasmids and strains generated in this study are available from the lead contact upon request without restriction.

Data and code availability

All sequencing data associated with this study are available at the GEO repository and assigned accession number GEO: GSE242874. Original code (i.e., the modelling and simulations) has been deposited at Zendo: https://doi.org/10.5281/zenodo.8327425. Any additional information required to reanalyze the data reported in this paper is freely available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Yeast genetics

All *S. cerevisiae* budding strains used in this study are in the *W303* background. Full genotypes of all *S. cerevisiae* strains used in this study are listed in Table S1. Lithium acetate based transformation was used for *S. cerevisiae* and *S. pombe* strain construction.

WHI5 induction strains (MS63, MS73 & MS96) are constructed to express *WHI5* from a beta-estradiol responsive promoter (*LexOPr-WHI5* construct).

Strains used for G1 cyclin shut-off (MS54, MS64, MS65, MS66, MS67 & MS95) are modifications of a previously described G1 cyclin shut-off strain.[53]

Anchor away strains (all strains listed in Table S1 between MS202 and MS299) are derived from previously described anchor away strains.[20] Diploid strains were constructed by crossing the respective parent haploid strains listed in Table S1 and then selecting on SCD -LEU -TRP.

Strains for EU-pulse chase experiments (MS95 & MS96) were constructed by genomic integration of pMS282 into the *URA3* locus. pMS282 was constructed by inserting a *TEF2pr-HSV-TK* and a *TEF1pr-hENT1* sequence into the SIVu backbone.[54] This allows the uptake and incorporation of EU (5-ethynyl uridine) in yeast. In MS96 the URA marker was then removed by Cre-LoxP recombination.

Yeast strains MS627-MS635 were used for the conditional overexpression of individual mNeonGreen tagged RNAPII subunits (Figures 2E and S3I) and were constructed as follows. pMS153, which encodes the WTC846 TET transcription factor system described in Azizoglu et al.,[55] was linearized and integrated into the *HIS3* locus. Each RNA polymerase II subunit was then separately sub-cloned with a C-terminal mNeonGreen tag into the SIVu backbone[54] with an upstream WTC846 *TetPr*[55] and a downstream *ADH1* terminator. Each plasmid (pMS200-pMS208) was then linearized and integrated into the *URA3* locus. 3 out the 12 RNA polymerase II subunits were not tested (*RPB2*, *RPB5* and *RPB12*) because C-terminal tagging of the endogenous gene locus was either inviable or had a significant growth defect.

Yeast strain MS698 was used for the simultaneous conditional overexpression of all RNA polymerase II subunits (Figures 2F, 2G, S3I, and S3J) and was constructed as follows. First a plasmid (pMS152), which encodes the WTC846 TET transcription factor system,[55] was integrated into the *HO* locus to make strain MS46. Four plasmids (pMS226, pMS230, pMS231 and pMS232) were then constructed, each based on a different single integration vector (SIV) backbone[54] with three different RNAPII subunits cloned in tandem into the multi-cloning site with an upstream WTC846 *TetPr*[55] and a downstream *ADH1* terminator. The *RPB3* construct (in pMS226) was also fused via a linker to the mNeonGreen fluorescent protein. Each plasmid was linearized and sequentially integrated into the *URA3*, *HIS3*, *TRP1*, and *LEU2* loci, respectively, of MS46 and then *WHI5* was deleted using a hygromycin resistance cassette. A control strain (MS656) was constructed in the same manner as MS698 but using four integration vectors each encoding mNeonGreen expressed from the WTC846 *TetPr* with an *ADH1* terminator downstream (pMS198, pMS233, pMS234 and pMS235).

Yeast media, culturing conditions and drug treatments

Cells were cultured in synthetic complete media (SC) with either 2% glucose, 2% glycerol + 1% ethanol, or 2% raffinose as a carbon source. Unless stated otherwise, SC + 2% glycerol + 1% ethanol was used for all experiments examining cell size mutants because these conditions increase the range in cell size between mutants. SC + 2% glycerol + 1% ethanol was also used for the elutriation experiments (see STAR Methods section: centrifugal elutriation G1 arrest experiments) and G1 arrest experiments (see STAR Methods section: G1 arrest experiments). For *GAL* shut-off experiments cells were initially cultured in SC + 2% raffinose (see STAR Methods section: RT-qPCR mRNA decay experiments). For EU pulse-chase experiments, cells were initially cultured in SC + 2% glycerol + 1% ethanol but uracil was omitted from the media and 100 nM uridine was added (see STAR Methods section: EU (5-ethynyl uridine) pulse-chase RNA-seq to determine global mRNA turnover rates). All other experiments were performed in SC + 2% glucose. Unless stated otherwise, cells were grown at 30°C, cultured below OD_{600} = 0.45, and collected at an OD_{600} between 0.2 and 0.4.

Low, medium, and high expression of *WHI5* from MS63 or MS96, was induced by growing cells in 0nM, 3.5-10 nM, or 16-30 nM beta-estradiol (MP Biomedicals), respectively, for >36 h at which point cell size is at a steady state. For *TETpr* induction, cells were treated with 50 ng/ml anhydrotetracycline for 45 min. For anchor away experiments, cells were treated with 1 μg/ml rapamycin for either 100 min (all samples in Figure 2D and S3B except *Rpb1-FRB / WT*) or 40 min (*Rpb1-FRB / WT* samples in Figure 2D and all samples in Figures S3C–S3H). The shorter 40-min rapamycin treatment was used to avoid potential compensation in Rpb1 protein amounts following 50% nuclear depletion.

For all SILAC (Stable Isotope Labeling by Amino acids in Cell culture) experiments (Figures 2B and S2C–S2E) SILAC compatible strains (*i.e.*, the *lys1 arg4 CAN1+* background) were grown for more than 10 generations in SILAC media supplemented with light (L-Arginine (unlabeled) & L-Lysine (unlabeled)), medium (L-arginine:HCL [U13C6] and L-lysine:2HCL [4,4,5,5-D4], or heavy (L-arginine:HCL (13C6, 15N4) and L-lysine:2HCL (13C6, 15N2)) amino acid isotopes (Cambridge Isotope Laboratories Inc.).

METHOD DETAILS

Centrifugal elutriation G1 arrest experiments

For G1 arrest elutriation experiments (Figures 3A, 3B, 3F, 4D, 4E, 6A, and S4A–S4J), a G1 cyclin shut-off strain (*cln2 cln3 LexOpr-CLN1*: (strains = MS64, MS67 or MS68) was grown in SC + 2% glycerol + 1% ethanol with 15 nM beta-estradiol. SC + 2% glycerol + 1% ethanol media was used as this allows for the collection of the smallest G1 cells without carbon source removal or cooling cells to 4°C during elutriation. The beta-estradiol is to induce expression of *CLN1* from a synthetic beta-estradiol responsive promoter (*LexOpr-CLN1*). Small G1-arrested cells were then collected as follows. 2 liters of culture were grown at 30°C to OD_{600} ~0.7 and then washed twice on a filter membrane in fresh pre-warmed media lacking beta-estradiol and re-inoculated into 4 liters of media lacking beta-estradiol at 30°C and grown for 2 h. Cells were then collected on a filter membrane and resuspended in 100 ml fresh room temperature media. Cells were then sonicated (3 x 20 seconds, 4 min between sonication cycles) and loaded into a JE 5.0 elutriation rotor fitted for a two-chamber run (Beckman Coulter) in a J6-MI Centrifuge (2.4k rpm, 23°C). The elutriation chambers were pre-equilibrated and run with room temperature SC + 2% glycerol + 1% ethanol. The pump speed was gradually increased until small G1 cells were collected with minimal additional debris. The smallest G1 fractions were then combined and concentrated on a filter membrane, resuspended at 30°C in SC + 2% glycerol + 1% ethanol media, and split into 7 cultures. Each culture was arrested in G1 for different periods of time before fixation and collection for ChIP-seq. At the point of splitting into 7 cultures, the OD_{600} of each culture was adjusted, according to its arrest time, so that each culture was at OD_{600} ~0.3 at the point of fixation and collection. For both the haploid and diploid experiments, two independent biological replicates were performed. A control experiment was also conducted where a *WT* strain (*i.e.*, not the G1 cyclin shut-off strain) was subjected to the same procedure, but after elutriation all cells were recovered from the elutriation rotor resulting in an asynchronous population that does not increase in cell size. Samples were then collected at the same timepoints used for the G1 arrest, this therefore controls for all possible artefactual effects associated with handling of the cultures during elutriation.

G1 arrest experiment

For the G1 arrest experiment without elutriation (Figures 3E, 6D, 6E, 6H, S4M, S7H, S7I, and S8F), a G1 cyclin shut-off strain was grown in SC + 2% glycerol + 1% ethanol with 15 nM beta-estradiol. The beta-estradiol is to induce expression of *CLN1* from a synthetic beta-estradiol responsive promoter (*LexOpr-CLN1*). Cells were grown to OD_{600} ~0.3 and washed 3x on a filter membrane in pre-warmed media lacking beta-estradiol to inhibit G1 cyclin expression and initiate the G1 arrest. Cells were then collected when the mean cell volume reached the indicated size.

Cell size measurements

Cell volume was measured using a Beckman Coulter Z2 counter. Cells were placed on ice, sonicated, and then diluted in 10-20 ml of Isoton II diluent (Beckman Coulter 8546719) before measurement.

DNA content analysis

DNA content was determined by flow cytometry and was used to estimate genome copy number per cell for cell size mutants. 0.4 ml culture was added to 1 ml 100% 4°C ethanol and stored at 4°C. Cells were pelleted (13k rpm, 2 min), washed, and resuspended in 50 mM Sodium Citrate (pH 7.2). Cells were then incubated with 0.2 mg/ml RNAseA (overnight, 37°C) and then treated with 0.4 mg/ml proteinase K (1-2 h, 50°C) before addition of 25 μM Sytox Green (ThermoFisher Scientific). Cells were then sonicated, and DNA-content was analyzed for >10,000 events on either a BD Biosciences FACScan Analyzer or an Attune NxT flow cytometer. Genome copy number per cell was then calculated after gating for single cells (FlowJo) and was used to convert occupancy per genome determined by spike-in normalized ChIP-seq to occupancy per cell (for details see STAR Methods section: spike-in normalized ChIP-seq).

Flow cytometry measurements of protein levels

Protein levels of C-terminally mNeonGreen-tagged proteins (Figures 2A, 2E, 3E, S3I, and S4M) were quantified using an Attune NxT flow cytometer. 1 ml of cells were gently sonicated and placed on ice before acquisition of data from >10,000 cells per sample. Single cells were gated based on FSC and SCC (FlowJo). mNeonGreen intensity was measured in the BL1-A channel and cell

size was approximated using FSC-A. For cycling cells (strains: MS347, MS348, and all strains between MS615 and MS733) the size-dependent autofluorescence background in the BL1-A channel was determined by measuring an untagged background control strain (MS1 or MS47) and fitting a robust linear regression to BL1-A vs. FSC-A. This was then used to subtract the interpolated BL1-A background for each cell based on its FSC-A. For G1 arrested cultures (strains: MS446 & MS447), autofluorescence background was determined from an untagged strain (MS64) also arrested in G1. For total protein measurements, cells were fixed in 70% ethanol (on ice, 30 min), washed in 1x PBS, stained with a total protein dye (Alexa Fluor 488 NHS Ester; 1 μg/ml) for 1 h, washed 3x in 1xPBS, sonicated, and analyzed.

Assessing efficiency of anchor away nuclear depletion using microscopy

Cells expressing mCitrine tagged Rpb1 or Rpb3 (strains: MS222, MS223 & MS226) were imaged on a wide-field epifluorescence Zeiss AXIO Observer.Z1 microscope (63X/1.4NA oil immersion objective and a Colibri LED module). To assess the efficiency of anchor away nuclear depletion, a single *z*-plane was imaged after treatment with rapamycin for 1 generation. mCitrine fluorophores were imaged in the yellow channel (505 nm LED module) and phase contrast images were taken to give cell boundaries. For the images in Figure S3A, the exposure and contrast were manually adjusted in imageJ individually for each image to allow comparison of the relative subcellular localization of the signal and is therefore not appropriate for comparing relative total protein intensities.

Nuclear volume measurements during G1 arrest using microscopy

Cells expressing a *PUS1-mNeonGreen* nuclear marker[56] (strain: MS68) were used to determine the nuclear scaling with cell size in G1 arrest (Figures 3F and S4H–S4J). Small G1 cells were collected and arrested in G1 as described above (see STAR Methods section: centrifugal elutriation G1 arrest experiments). Cells were spun down and imaged live at different times during the arrest corresponding to approximately the same cell sizes as the timepoints for the G1 arrest experiments used to determine RNAPII occupancy presented in Figures 3A and 3B. Cells were imaged on a wide-field epifluorescence Zeiss AXIO Observer.Z1 microscope (63X/1.4NA oil immersion objective and a Colibri LED module) and 11 *z*-slices were taken at intervals of 0.35 μm. mNeonGreen was imaged in the yellow channel (505 nm LED module; 100% intensity; 0.25 second exposure) and phase contrast images were taken to give cell boundaries.

For each image, the most in focus *z*-plane was determined by manual inspection and used to segment all cells in the image. Automated segmentation of cell boundaries was performed using Cell-ACDC (default settings) before manual correction of segmentation errors in Cell-ACDC. Nuclei were manually segmented from the *PUS1-mNeonGreen* fluorescence using the *napari* package (python). Because the nucleus is not always in the same focal plane as the central plane of the cell, nuclei were individually segmented in the individual *z*-plane in which they were most in focus. Cell volumes were taken directly from the Cell-ACDC output and nuclear volumes were estimated as an ellipsoid where the first two axes are the long and short axes of the nuclear mask, and the third axis is the mean of the two measured axes. For all samples, Coulter counter measurements were collected at the same time as cells were imaged. A linear regression ($y=m^*x$) between the mean cell size of each condition of Coulter counter measurements (x) and microscopy measurements (y) was used to re-scale the microscopy-based volume estimates by a factor m. This ensures the absolute values from the imaging data are directly comparable to the Coulter counter measurements and can therefore be used for fitting the dynamic equilibrium model to the ChIP-seq, for which size information is only from Coulter counter measurements.

Spike-in normalized ChIP-seq

The spike-in normalized ChIP-seq protocol was adapted from Hu et al.[18] *S. cerevisiae* cultures were combined with spike-in cultures (*C. glabrata* or *S. pombe*), mixed, and then immediately fixed (< 5 seconds after mixing) by the addition of formaldehyde to a final concentration of 1% (15 min). Cells were then quenched with 0.125 M glycine (5 min), washed twice in cold PBS, pelleted, snap-frozen, and stored at −80°C.

Sample and spike-in were mixed at a ratio between 1:2 and 1:5 by OD_{600}. For most experiments *C. glabrata* (grown in the same media conditions as *S. cerevisiae*) was used as the spike-in. For experiments where anti-FLAG ChIP was performed, the S. pombe strain MSsp109 (*Rpb1-3xFLAG*) was grown in EMM4S and used as the spike-in. Within each batch of samples, the same spike-in culture was used for all samples and all samples were collected at effectively the same time (a 20-30 second interval between each sample and no more than 8 samples were collected in each batch). The only exception to the above is for the elutriation G1 arrest experiments where different *S. cerevisiae* cultures could not be mixed with the spike-in at the same time (because different samples required different arrest durations). For this experiment, *S. cerevisiae* and *C. glabrata* were therefore separately fixed, quenched, washed, and then mixed in PBS before being pelleted and snap frozen.

For the experiments in Figures 2, 3, and 4, a background sample was included within each batch of experiments to determine the *S. cerevisiae* background-to-spike-in ratio. For anti-Rpb1 ChIP, the background was determined using a sample where TBP was conditionally depleted from the nucleus (MS205 or MS289; *TBP-FRB* treated with rapamycin 60-65 min for SC + 2% glucose or 90 min rapamycin for SC + 2% glycerol + 1% ethanol) to block Rpb1 loading on the genome. For anti-FLAG ChIP, the background was determined using an untagged sample with no FLAG epitope (MS207). Background samples were otherwise collected and processed as other samples. See STAR Methods section: spike-in normalized ChIP-seq analysis for how this is used to calculate the global occupancy value for each ChIP.

Pellets were thawed and lysed in 300 μl FA lysis buffer (50 mM HEPES–KOH (pH 8.0), 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 1 mM PMSF, Roche protease inhibitor) with ~1 ml ceramic beads on a Fastprep-24 (MP Biomedicals).

The entire lysate was then collected and adjusted to 1 ml before sonication with a 1/8" microtip on a Q500 sonicator (Qsonica) for 8-16 min (cycles of 10 seconds on and 20 seconds off). The sample tube was held suspended in a −20°C 80% ethanol bath to prevent sample heating during sonication. Cell debris was then pelleted, and the supernatant was retained for ChIP or input. For each ChIP reaction, 20 μl Protein G Dynabeads (Invitrogen) were blocked (PBS + 0.5% BSA, incubate 40 min at room temperature), pre-bound with 5 μl of antibody in PBS (incubate 40 min at room temperature), and washed 2x with PBS before being incubated with 0.5 ml supernatant (4°C overnight). For RNA polymerase II ChIPs, the supernatant was adjusted so 0.5 ml corresponds to 50-100 ml of cells at OD_{600} ~ 0.35. For histone and histone modification ChIPs, the supernatant was adjusted so 0.5 ml corresponds to 10-25 ml of cells at OD ~0.35. See below for a list of antibodies used. After overnight incubation, Dynabeads were washed (5 min per wash) 2x in FA lysis buffer and 3x in high-salt FA lysis buffer (50 mM Hepes-KOH (pH 8.0), 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 1 mM PMSF). ChIP DNA was then eluted in ChIP elution buffer (50 mM Tris-HCl (pH 7.5), 10 mM EDTA, 1% SDS) at 65°C for 20 min. At the same time, 15 μl of input was mixed directly with 115 μl of ChIP elution buffer. Eluted ChIP DNA or input DNA were then incubated to reverse crosslinks (65°C, 5 h) before treatment with RNAse A (37°C, 1 h) and then Proteinase K (65°C, 2 h). DNA was then purified using the ChIP DNA Clean & Concentrator kit (Zymo Research). Indexed sequencing libraries were generated using the NEBNext Ultra II DNA Library Prep kit (NEB E7645), pooled, and then sequenced by paired-end (2x150bp) Illumina sequencing.

The following antibodies were used for ChIP against the indicated epitopes: anti-Rpb1 clone 8wG16 (mouse, monoclonal, Millipore 05-952), anti-Rpb1-S2-P clone 3E10 (rat, monoclonal, Millipore 04-1571), anti-Rpb1-S5-P clone 3E8 (rat, monoclonal, Millipore 04-1572), anti-FLAG clone M2 (mouse, monoclonal, Sigma-Aldrich F3165), anti-Histone 3 (rabbit, polyclonal, Abcam ab1791), anti-Histone 4 (rabbit, polyclonal, Abcam ab10158), anti-H3K18ac (rabbit, polyclonal, Abcam ab1191), anti-H3K27ac (rabbit, polyclonal, Millipore 07-360), anti-H4K9ac (rabbit, polyclonal, Abcam ab4441), anti-H4K16ac (rabbit, polyclonal, Millipore 07-329), anti-H3K56ac (rabbit, polyclonal, Millipore 07-677), anti-H3K4me1 (rabbit, polyclonal, Abcam ab8895), anti-H3K4me2 (rabbit, polyclonal, Abcam ab7766), anti-H3K4me3 (mouse, monoclonal, Abcam ab1012), anti-H3K79me2 (rabbit, polyclonal, Abcam ab3594), anti-H3K79me3 (rabbit, polyclonal, Abcam ab2621), anti-H3K36me2 (rabbit, polyclonal, Abcam ab9049), anti-H3K36me3 (rabbit, polyclonal, Abcam ab9050).

Dual enzyme single-molecule foot printing dSMF)

The protocol for dSMF to measure chromatin accessibility (Figures 5A and 5B) was based on Krebs et al.[23] and adapted for yeast as follows. ~$1x10^8$ *S. cerevisiae* cells were pelleted and resuspended in 1 ml digestion buffer (1.4 M sorbitol, 40 mM Hepes-KOH (pH 7.5), 0.5 mM $MgCl_2$, 10 mM DTT) and mixed with ~$1x10^7$ *C. glabrata* cells (spike-in) in 100 μl digestion buffer. Cells were pelleted and resuspended in 0.5 ml digestion buffer + 0.5 mg/ml 100T Zymolase (MP biomedicals), and then incubated on a shaker (30°C, 10 min, 500 rpm). Cells were then pelleted (5k rpm, 2 min, 4°C), resuspended in 0.5 ml wash buffer (1.4 M sorbitol, 40 mM Hepes-KOH (pH 7.5), 0.5 mM $MgCl_2$), pelleted again (5k rpm, 2 min, 4°C), resuspended in 0.3 ml ice-cold lysis buffer (10 mM Tris-HCl[pH7.5], 10 mM NaCl, 3 mM $MgCl_2$, 0.1 mM EDTA (pH 7.5), 0.5% NP-40), and incubated on ice (10 min). Nuclei were then pelleted (5k rpm, 4 min, 4°C), resuspended in 0.3 ml nuclei wash buffer (10 mM Tris-HCl (pH 7.5), 10 mM NaCl, 3 mM $MgCl_2$, 0.1 mM EDTA (pH 7.5)), pelleted again, and resuspended in 50 μl 1x reaction buffer (50 mM Tris-HCl (pH 8.5), 50 mM NaCl, 10mM DTT). Nuclei were then treated with M.CviPI (NEB) and M.SssI (NEB) methyltrasferases as follows. 50 μl cells in M.CviPI reaction buffer were mixed with 47 μl M.CviPI reaction mix (50 mM Tris-HCl (pH 8.5), 50 mM NaCl, 10 mM DTT, 0.7 M sucrose, 1.3 mM SAM, 200 U M.CviPI) and incubated (8 min, 30°C) before mixing with 3 μl M.CviPI boost mix (100 U M.CviPI, 42.5 μM SAM), and returning to incubation (7 min, 30°C). 10 μl M.SssI reaction mix (50 mM Tris-HCl (pH 8.5), 50 mM NaCl, 10 mM DTT, 0.11 M $MgCl_2$, 12.8 μM SAM, 60 U M.SssI) was then added to the 100 μl M.CviPI reaction, mixed, and incubated (8 min, 30°C). The reaction was then stopped with the addition of 190 μl lysis buffer (from MasterPure Yeast DNA Purification Kit (Lucigen)), and DNA was extracted with a MasterPure Yeast DNA Purification Kit (Lucigen). Unmethylated cytosines were then deaminated to uracil and indexed libraries were prepared using the NEBNext Enzymatic Methyl-seq Kit (NEB E7120) before being pooled and then sequenced by Illumina paired-end (2x150bp) sequencing.

Spike-in normalized RNA-seq to determine global mRNA amount per cell

To determine global mRNA amounts per cell (Figures 6A and S7E), 1 ml of culture (OD_{600} = 0.2-0.35) was pelleted (30 seconds, 14k rpm), snap-frozen, and stored at −80°C. 4 independent biological replicates were collected for each condition for the cell size mutants, while two independent biological replicates were collected for each condition for the haploid elutriation G1 arrest experiment. RNA samples for the haploid elutriation G1 arrest experiment were taken from the same culture as the ChIP-seq samples (Figures S4A–S4C), while cell size mutants were collected from independent biological experiments (Figures S7A–S7D).

DNA and RNA were separately extracted from each pellet as follows. For each experiment, *C. glabrata* was used as the spike-in. *C. glabrata* was cultured in the same media conditions as *S. cerevisiae*, pelleted, snap-frozen and stored at −80°C. The *C. glabrata* pellet was thawed and resuspended in ice-cold PBS and 50 μl was then added to each *S. cerevisiae* pellet on ice so that the *S. cerevisiae-to-C. glabrata* ratio was approximately 2:1 by OD_{600}. The mixture of *S. cerevisiae* and *C. glabrata* in PBS was then quickly mixed. 20 μl was removed for gDNA extraction (see below) and 300 μl TRI Reagent (Zymo Research) was added immediately to the remaining ~30 μl for RNA extraction.

Cells in TRI Reagent were lysed by bead beating using a Fastprep 24 (4°C, settings: 5.5 m/s, 1 x 35 seconds). Cell debris was pelleted (14k rpm, 1 min) and the supernatant recovered. RNA was then extracted using the direct-zol RNA microprep kit (Zymo

Research). mRNA was enriched using the NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB, E7490) and NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB, E7760) was then used to prepare libraries for paired-end (2x150bp) Illumina sequencing (Novogene). More than 10 million reads were sequenced per sample.

Separately, the 20 μl of cells for gDNA extraction was added to 110 μl YD Digestion Buffer (YeaStar Genomic DNA Kit; Zymo Research) and gDNA was extracted according to Protocol 1 for the YeaStar Genomic DNA Kit (Zymo Research). DNA concentration was then determined by Qubit and gDNA was prepared for sequencing by tagmentation with custom made Tn5.[57] Two tagmentation reactions were performed for each sample, each with a slightly different DNA-to-Tn5 ratio: (A) 4 ng DNA, 4 μl Tn5, 1x TD buffer (50 μl reaction) and (B) 4 ng DNA with 2 μl Tn5, 1x TD buffer (25 μl reaction). Tagmentation reactions were incubated (30 min at 37°C), and then cleaned up (MinElute PCR Purification Kit, Quiagen). Tagmented DNA was then amplified (7 cycles) with NEBNext High-Fidelity 2X PCR Master Mix (M0541) and indexed as previously reported.[58] Libraries were then cleaned up using AMpure XP beads (Beckman Coulter), pooled, and sequenced by paired-end (2x150bp) Illumina sequencing (Novogene) at >1 million reads per sample.

EU 5-ethynyl uridine) pulse-chase RNA-seq to determine global mRNA turnover rates

Strains expressing *TEF2pr-HSV-TK* and *TEF1pr-hENT1* (MS95 & MS96) were used for pulse-chase experiments to allow rapid and efficient uptake of the uridine analogue EU (5-ethynyl uridine). MS95 was used to compare global mRNA turnover between low and high *WHI5* induction conditions (Figures 6C and S7G). MS96 was used to compare global mRNA turnover between 1.5 hr and 7 hr G1 cyclin shut-off (Figures 6D, 6E, S7H, and S7I). Cells were grown in SC + 2% glycerol + 1% ethanol - uracil + 100 μM uridine. 5 μM EU (MS95) or 10 μM EU (MS96) was added to cells at OD_{600} ~0.2 for 1 h (*i.e.*, the pulse). Cells were then washed 3x in pre-warmed media lacking EU on a filter membrane (total wash duration = 3-4 min) and re-inoculated into pre-warmed media lacking EU at OD_{600} ~0.3 in a shaking water bath (*i.e.*, the chase). Samples were then collected at timepoints up to 36 min after the wash was completed: at the respective timepoint, 1 ml cells were pelleted (30 seconds, 14k rpm), snap-frozen, and stored at −80°C. For each culture, a no-EU background sample was also collected before the EU pulse. Two independent biological replicates were performed for low and high *WHI5* induction experiments, and 1 biological replicate was performed for the 1.5 h and 7 h G1 cyclin shut-off experiments. As a spike-in, the *S. pombe* strain PN10597 (h- *adh1pr-HSV-TK adh1pr-hENT1*) was cultured in EMM media to OD_{600} ~0.2, 200 μM EU was then added for 2 h before 12ml cells were pelleted, snap-frozen, and stored at −80°C.

RNA was then extracted and processed as follows. See Figure S7F for workflow schematic. The *S. pombe* pellet was resuspended in TRI Reagent (Zymo Research) and was added to each *S. cerevisiae* pellet so that the *S. cerevisiae-to-S. pombe* ratio was approximately 3:2 by OD_{600}. Cells in TRI Reagent were lysed by bead beating using a Fastprep 24 (4°C, settings: 5.5 m/s, 1 x 35 seconds). Cell debris was pelleted (14k rpm, 1 min) and the supernatant recovered. RNA was then extracted using the direct-zol RNA microprep kit (Zymo Research). 5 μg RNA was then biotinylated by click reaction using the Click-iT Nascent RNA Capture Kit (Invitrogen). Biotinylated RNA was then precipitated, resuspended in 50 μl nuclease free water, and used for mRNA enrichment with the NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB, E7490). mRNA was eluted from the Poly(A) beads in 14 μl Tris buffer. 12.3 μl mRNA was used to enrich EU-mRNA and 1 μl was used directly for RNA-seq as the input sample.

EU-mRNA was enriched using the Click-iT Nascent RNA Capture Kit (Invitrogen C10365) with the following modifications to the kit protocol. 5 μl Dynabeads MyOne Streptavidin T1 (Invitrogen 65601) per sample were blocked and washed as follows. Beads were washed 1x 50 μl wash buffer 2 (kit component J) and then resuspended in the 50 μl wash buffer 2 + 5% (v/v) Denhardt's reagent and incubated at room temperature (10 min). Blocked beads were then washed 3x in 50 μl wash buffer 2 and resuspended in 5 μl wash buffer 2. 12.5 μl Click-iT RNA binding buffer (kit component G) and 0.2 μl RNaseOUT Recombinant Ribonuclease Inhibitor (Invitrogen 10777019) was then added to 12.3 μl mRNA, mixed, and incubated at 70°C (5 min). 5 μl of blocked and washed beads were then added to each 25 μl mRNA reaction, mixed, and incubated on a rolling mixer at room temperature (30 min) to bind biotinylated mRNA to the Streptavidin beads. Beads were then washed 4x with 100 μl wash buffer 1 (kit component I), 2x 100 μl SDS wash buffer (1% SDS, 5 mM TRIS-HCL, 1 mM EDTA), 4x with 100 μl wash buffer 2 and resuspended in 5 μl wash buffer 2. The 5 μl of beads was immediately used as the input for cDNA synthesis on the beads using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB, E7760). cDNA synthesis reactions were mixed by pipetting every 10 min on the thermocycler to keep Streptavidin beads in suspension. Beads were then removed prior to the cDNA clean-up step. For the input mRNA samples, 1 μl of the biotinylated mRNA (*i.e.*, no EU enrichment) was used directly for cDNA synthesis and library prep. ~10 million reads (EU-enrichment) or ~3 million reads (input) were then sequenced by paired-end (2x150bp) Illumina sequencing (Novogene).

RT-qPCR mRNA decay experiments

To quantify *MET3* and *MET17* decay rates (Figures 6F and S8), cells were grown in SC + 2% glycerol + 1% ethanol media lacking methionine. 1 mM methionine was then added to repress MET genes.[59] To quantify *GAL1*, *GAL7* and *GAL10* decay rates (Figures S7K–S7M), cells were grown in SC media + 2% raffinose before addition of 2% galactose for 75 min. Cells were then washed (1x) and resuspended in SC + 4% glucose.[60] 1 ml of culture was pelleted (13k rpm, 30 seconds, 4°C) at the indicated time points after either methionine addition (for *MET3* and *MET17*) or after the start of the wash in SC + 4% glucose (for *GAL1*, *GAL7* and *GAL10*). The pellet was immediately snap frozen in liquid nitrogen at t + 1 min and stored at -80°C. Cells were subsequently thawed in 300 μl TRI Reagent (Zymo Research) and lysed by bead-beating using a Fastprep 24 (4°C, settings: 5.5 m/s, 1 x 30 seconds). Cell debris was pelleted (13k rpm, 2 min) and the supernatant recovered. RNA was then purified using the direct-zol RNA microprep kit (Zymo Research R2061) and cDNA was then synthesized using 800 ng of RNA with iScript Reverse Transcription Supermix for

RT-qPCR (BioRad 1708841). cDNA was used for qPCR with iTaq Universal SYBR Green Supermix (BioRad 1725121). Two biological replicates were performed per experiment and 2-4 technical replicates were performed per biological replicate. All values were normalized to *ACT1* and then rescaled so that the first time point is set = 1. Decay rates were calculated from a one-phase exponential decay fit to the data (Prism 6; weighting = 1/Y, plateau = 0). For the purposes of fitting, only timepoints between 6 min and 21 min were used because 6 min is the earliest timepoint at which transcription was determined to be fully inactivated at *MET3* and *MET17* by anti-Rpb1 ChIP-qPCR (Figure S8D).

Single-molecule imaging microscopy

For the experiments shown in Figures 4A–4C, 1 ml culture (SC + 2 % glucose) was pelleted (4k rpm, 1 min) and resuspended in 0.5 ml fresh media. JF-PA549 (Janelia Farms photoactivatable 549) dye was added at a final concentration of 75 nM, except for MS162 (HTB1-HALO), where a concentration of 10 nM was used to compensate for the higher protein copy number. Cultures were incubated with dye (30°C, mixing at 550 rpm) for 40 min. Cells were then washed (3x) in fresh media to remove unbound dye and resuspended in 20 μl media, 4 μl of which was placed on an agarose pad. The agarose pad was made by mixing 0.5 ml 2% agarose Optiprep mixture (20 mg agarose in 1 ml Optiprep (Sigma), heated to 90°C) with 0.5 ml 2x media. Approximately 110 μl of this mixture was placed within a Gene Frame (Thermo Scientific), with excess being removed with a KimWipe. Prior to imaging, we waited ∼15 min to let any remaining unbound dye be released from the cells. Coverslips were cleaned using 2% VersaClean detergent solution overnight. The coverslips were then washed with MilliQ water 3 times, sonicated in acetone for 30 min, washed with MilliQ water 3 times, washed in methanol (flame excess from coverslips), and then placed in a Plasma Etch plasma oven for 10 min.

Imaging was done at 23°C using a Leica DMi8 inverted microscope with a Roper Scientific iLasV2 (capable of ring total internal reflection fluorescence (TIRF)) and an Andor iXon Ultra 897 EMCCD camera. An Andor ILE combiner was used and the maximum power from the optical fiber was 100 mW for the 405 nm wavelength, and 150 mW for the 488 nm and 561 nm wavelengths. The iLasV2 was configured for HILO (ring highly inclined and laminated optical sheet), selective illumination and single-molecule sensitivity. Metamorph software was used to control acquisition. A Leica HCX PL APO 100x/1.47 oil immersion objective was used with 100 nm pixel size. Z-stacks were determined using a PIano piezo Z controller. Single-particle photoactivated localization microscopy (sptPALM) experiments were performed by using continuous activation of molecules with low power (0.1% – 10% in Metamorph software) 405 nm light to photo activate ∼1 molecule/cell, with simultaneous fast-exposure (10 ms) illumination with 561 nm light (70% in Metamorph software) to image molecules. A bright field and a 561 nm z-stack of 10 μm (0.5 μm step size) were taken and used to identify the unbudded G1 phase cells and to quantify nuclear area using the Pus1-GFP nuclear marker.

Chromatin enrichment for Proteomics ChEP)

For SILAC (Stable Isotope Labeling by Amino acids in Cell culture) experiments (Figures 2B and S2C–S2E), SILAC compatible strains were labeled with amino acid isotopes (for details see STAR Methods section: Yeast media, culturing conditions and drug treatments) before collection. The ChEP chromatin enrichment protocol was based on Kustatscher et al.,[19] and adapted for yeast as follows. Cells were fixed by addition of formaldehyde to a final concentration of 1%, shaken for 15 min, quenched with 0.125 M glycine (5 min), washed twice in cold PBS, pelleted, snap-frozen, and then stored at −80°C. Each pellet contained cells from 500 ml culture at OD_{600} ∼ 0.4. Pellets were thawed and lysed in 300 μl FA lysis buffer (50 mM HEPES–KOH (pH 8.0), 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, 1 mM PMSF, 1x Roche protease inhibitor) with ∼1 ml ceramic beads on a Fastprep-24 (MP Biomedicals). The entire lysate was then collected and adjusted to 1.5 ml.

For the cytoplasmic fraction (CYTO), 50 μl lysate was adjusted to 0.5 ml, pelleted (4°C, 14k rpm, 10 min), and the supernatant was then taken as the cytoplasmic fraction. At this stage, protein concentration was determined by Bradford before Laemmli sample buffer was added to a 1x final concentration and the sample was then boiled to reverse crosslinks (99°C, 30 min). For whole cell extract (WCE), 50 μl lysate was adjusted to 0.5 ml and then sonicated with a 1/8 microtip on a Q500 sonicator (Qsonica) for 5 min (Amp = 25%, 10 seconds on, 20 seconds off) to solubilize chromatin. Laemmli sample buffer was added to a 1x final concentration, and the sample was then boiled to reverse crosslinks (99°C, 30 min) before pelleting (4°C, 14k rpm, 10 min). The supernatant was then taken as the WCE. WCE protein concentration was assumed to be the same as was determined for the cytoplasmic fraction for the same sample.

For chromatin fractions, the lysate was pelleted (4°C, 14k rpm, 30 min) and supernatant was removed and discarded. The pellet was retained and either processed with extraction method A (high purity, low yield) or extraction method B (high purity, low yield). For extraction method A, the fuzzy translucent top layer (high purity chromatin) of 3 pellets was resuspended in 300 μL FA lysis buffer without disturbing the lower opaque portion of the pellets and combined to a single fresh tube that was adjusted to 1.45 ml. For extraction method B, 1 entire pellet (top and bottom layer) was resuspended in FA lysis buffer to a final volume of 1.45 ml. Extractions A and B were then processed identically for all subsequent steps. Chromatin was first treated with 100 μg/ml RNaseA (37°C, 15 min) and then pelleted (4°C, 14k rpm, 30 min). The supernatant was then discarded before the pellet was resuspended in 300 μL SDS buffer (4 % SDS, 10 mM EDTA, 25 mM Tris-HCl (pH 7.5), 1 mM PMSF, 1x Roche protease inhibitor) and incubated at room temp (10 min). 1 ml urea buffer (8 M urea, 1 mM EDTA, 10 mM Tris-HCl (pH 7.5), 1 mM PMSF, 1x Roche protease inhibitor) was then added and the sample was gently mixed. The sample was then pelleted (18°C, 14k rpm, 30 min) and the supernatant was discarded. The pellet was resuspended in 300 μl SDS buffer mixed with 1 ml urea buffer before being again pelleted (18°C, 14k rpm, 30 min). The supernatant was discarded, and the pellet was resuspended in 300 μl SDS buffer and then adjusted to 1.45 ml with SDS

buffer before being again pelleted (18°C, 14k rpm, 30 min). The supernatant was again discarded and 0.6 ml Storage buffer (10% glycerol, 1 mM EDTA, 10 mM Tris-HCl (pH 7.5), 25 mM NaCl, 1 mM PMSF, 1x Roche protease inhibitor) was added on top of the pellet. DNA was then resuspended and sheared by sonication (1/8" microtip on a Q500 sonicator (Qsonica) for 10 min (Amp = 25%, 10 seconds on, 20 seconds off)). The sonicated sample was then pelleted (4°C, 14k rpm, 10 min) and the supernatant was retained as the chromatin fraction. Protein concentration was determined by Bradford before Laemmli sample buffer was added to a 1x final concentration and the sample was then boiled to reverse crosslinks (99°C, 30 min).

For the immunoblotting analysis presented in Figure S2B, ~ 5 μg of the chromatin fraction B was loaded per lane alongside ~25 μg CYTO and WCE. For the proteomics analysis in Figures S2C and S2D, chromatin and WCE with opposite SILAC labels were mixed in an approximately 1:2 protein amount ratio before processing ~ 100 μg for LC-MS/MS (for details see STAR Methods section: LC-MS/MS sample preparation and data acquisition). For the proteomics analysis in Figure 2B, heavy, medium and light labeled cultures were mixed prior to fixation in an approximately 1:1:1 cell number ratio. Two biological replicates were performed for chromatin extraction B. ~50 μg of each replicate was then analyzed by LC-MS/MS (for details see STAR Methods section: LC-MS/MS sample preparation *and data acquisition*).

Immunoblotting

Protein samples were resolved on a Bolt 4-12% Bis-Tris protein gel (Invitrogen) and transferred to a nitrocellulose membrane with the iBlot 2 dry blotting system (Invitrogen). The following primary antibodies were used for western-blotting at 1/1,000 dilution: anti-V5 clone SV5-Pk1 (mouse, monoclonal, BioRad, MCA1360), anti-Rpb3 clone 1Y26 (mouse, monoclonal, BioLegend 665004), anti-FRB (rabbit, polyclonal, Enzo ALX-215-065-1), anti-Beta tubulin (rabbit, polyclonal, Abcam ab15568), anti-Histone 3 (rabbit, polyclonal, Abcam ab1791) and anti-Histone 4 (rabbit, polyclonal, Abcam ab10158). Anti-GAPDH clone GAR1 (mouse, monoclonal, ThermoFisher MA5-15738) was used for western-blotting at 1/2,500 dilution. Primary antibodies were detected using the following fluorescently labeled secondary antibodies at 1/10,000 dilution: IRDye 800CW goat anti-Mouse (Licor), Alexa Fluor 680 Donkey anti-Mouse (Invitrogen), Alexa Fluor 680 donkey anti-Rabbit (Invitrogen) and Alexa Fluor 790 Goat anti-Rabbit (Invitrogen). Membranes were then imaged using a LI-COR Odyssey CLx.

LC-MS/MS sample preparation and data acquisition

Each protein sample was reduced with 5 mM dithiothreitol for 25 min at 56°C, alkylated with 10 mM iodoacetamide (30 min, room temperature, dark), and then quenched with 7.5 mM DTT. Samples were digested and cleaned using SP3 on-bead methodology[61] with the modification that 50 mM HEPES (pH 8.5) was used in place of ammonium bicarbonate. Briefly, proteins were bound to the SP3 beads (10:1 beads:protein (w/w)) in 50% ethanol (v/v) and then washed three times in 80% ethanol prior to resuspension in 50 mM HEPES (pH 8.5) with 1:40 (trypsin:protein (w/w)) overnight at 37°C. The peptides were then fractionated using the High pH Reversed-Phase Peptide Fractionation Kit (Pierce) and dried under vacuum centrifugation. Peptides were subsequently resuspended in 0.1% trifluoroacetic acid and analyzed on an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher) coupled to an UltiMate 3000 HPLC system for online liquid chromatographic separation. Each run consisted of a 160 min gradient elution from a 75 μm x 50 cm C18 column.

Modelling – Dynamic equilibrium model

The dynamic equilibrium model assumes that the genome does not saturate and that the rate at which the free RNAPII, pol_{free}, associates with the genome, *DNA*, is determined by mass action kinetics so that:

$$\frac{d[pol_{bound}]}{dt} = k_{on} * [DNA] * [pol_{free}] - k_{off} * [pol_{bound}],$$

which can be rewritten in terms of amounts:

$$\frac{dpol_{bound}}{dt} = k_{on} * DNA * \frac{pol_{free}}{V_{nucleus}} - k_{off} * pol_{bound}$$

Because the time scale of transcription is more than an order of magnitude faster than that of appreciable cell growth, we can set this at steady state so that:

$$0 = k_{on} * DNA * \frac{pol_{free}}{V_{nucleus}} - k_{off} * pol_{bound}$$

As the total amount of RNAPII is proportional to cell size (Figures 3E and S4M) we can write:

$$pol_{total} = pol_{bound} \quad pol_{free} = c * V_{cell}$$

$$pol_{free} = c * V_{cell} - pol_{bound}$$

Substituting this into the above equation gives:

$$k_{on} * DNA * \frac{c * V_{cell} - pol_{bound}}{V_{nucleus}} - k_{off} * pol_{bound} = 0$$

This can be rearranged to give the amount of DNA bound RNAPII as a function of cell and nuclear size:

$$pol_{bound} = \frac{c\ V_{cell}}{1 \quad \frac{k_{off}\ V_{nucleus}}{k_{on}\ DNA}}$$

Measuring the relationship between cell volume and nuclear volume during G1 arrest (Figures 3F and S4H–S4J) shows that nuclear volume increases continuously as cell size increases. However, we found this increase deviates from a directly proportional linear scaling at smaller sizes, consistent with prior measurements reported by Jorgensen et al..[21] To account for the larger nuclear fraction at small cell sizes, we used a monotonically decreasing function for the nuclear to cell volume ratio, $f_n(V_{cell})$, whose parameters were fit to our data:

$$f_n(V_{cell}) = A \exp(-\gamma V_{cell}) \quad B$$

Using this empirically determined function, we can then express the amount of DNA bound RNAPII as a function of cell size:

$$pol_{bound} = \frac{c\ V_{cell}}{1 \quad \frac{k_{off}\ f_n(V_{cell})V_{cell}}{k_{on}\ DNA}}$$

Given that DNA amount is constant during G1 arrest, and that cell and nuclear volume are empirically determined, the dynamic equilibrium model only has two free parameters: (i) the cellular concentration of RNAP II (c) and (ii) the dissociation constant ($k_d = \frac{k_{off}}{k_{on}}$) for the interaction of RNAPII with DNA. To derive values for c and k_d we fit the above equation to our haploid RNAPII ChIP-seq data after substituting the fitted nuclear fraction relationship and setting the *DNA* term to 1 (Figure 3B). Fitting was done using *curve_fit()* from *scipy* (Python). When predicting the RNAPII occupancy in cycling cells or in G1-arrested diploids the values for c and k_d from the haploid fit were used, and *DNA* was set to 1.46 (cycling cells; Figure S4G) or 2 (diploid; Figure 4E). We computed confidence intervals by bootstrapping on 100 subsamples of 7 points. We can then predict the bound fraction of RNAPII as a function of cell size so that:

$$f_{bound} = \frac{pol_{bound}}{pol_{total}} = \frac{1}{1 \quad \frac{k_{off}\ f_n(V_{cell})V_{cell}}{k_{on}\ DNA}}$$

We note that if instead we assume a directly proportional relationship between cell size and nuclear volume ($V_{nucleus} = 0\ 07 * V_{cell}$), as is the case in organisms such as fission yeast,[62] the solution for DNA bound RNAPII is:

$$pol_{bound} = \frac{c\ V_{cell}}{1 \quad \frac{k_{off}\ 0\ 07\ V_{cell}}{k_{on}\ DNA}}$$

This constant nuclear fraction model also fits our ChIP-seq data very closely (Figure S4K) and gives only marginally less precise predictions of the bound fraction of RNAPII (Figure S4L).

Modelling - Stochastic simulations of RNAPII transcription

The dynamic equilibrium model considers RNAPII occupancy as the product of simple binding and unbinding events and assumes the genome size is so large that the amount of bound RNAPII never meaningfully saturates the genome (Figure S5A). To explore the validity of these simplifying assumptions, we wrote a more detailed stochastic model that captures the salient feature of the transcriptional cycle on a representative gene of a finite length (Figure S5B) as detailed below.

In this more complex model, a promoter is bound with RNAPII at the rate $k_{on}\ pr * pol_{free}$ and can then either dissociate with a rate $k_{off}\ pr$ or initiate transcription with a rate . Once initiated, RNAPII elongation is modelled as a series of 35 bp steps along a 1 kb gene-body were each RNAPII has a 35 bp footprint. RNAPII then dissociates from the gene in the final termination step. We draw from standard implementation and results of Totally Asymmetric Simple Exclusion Processes (TASEPs) to simulate this system. We note that our version of this model is inspired by and conceptually related to that presented in Sun et al.,[8] but here we have modelled promoter-association as a reversible step and parameterized our model using empirically determined values from the budding yeast literature as outlined below. This allowed us to be more specific about the parameter regime we limit our model to and thus to be more confident in the biological relevance of its predictions.

Elongation rate

We used an elongation rate of 2 kb/min from Mason and Struhl,[63] which was calculated from high temporal resolution Rpb1 ChIP time course data following inhibition of a *GAL1pr-YLR454* gene construct.

Termination rate

Nguyen et al.[22] reported an average RNAPII dwell time of 21-26 seconds based on single molecule tracking (21-23 seconds) and FRAP (26 seconds) experiments. We therefore assumed that termination time in yeast for the average transcriptional event is negligible so that RNAPII dissociates once it reaches the end of the gene-body.

Promoter on-rate (k_{on} pr), promoter off-rate (k_{off} pr), and initiation rate)

To explore the feasible range of k_{on} pr, k_{off} pr, and values, we simulated the steady-state gene body occupancy and the promoter occupancy of RNAPII at different combinations of k_{on} pr and values. For each parameter combination, k_{off} pr was then derived by constraining the average promoter residence time to 1.7 seconds based on the dwell time of the shortest-lived pre-initiation complex (PIC) subunit (TFIIF) from single-molecule tracking experiments reported by Nguyen et al.[22]

We selected 24 empirically feasible combinations of k_{on} pr, k_{off} pr, and (Figures S5C and S5D) on which subsequent analysis was performed. These parameter sets were identified by simulating the steady-state number of RNAPII molecules per promoter ($pol_{bound(promoter)}$) and RNAPII molecules per gene body ($pol_{bound(gene\ body)}$) across a range of parameter sets. We then compared this to empirical estimates of $pol_{bound(promoter)}$ and $pol_{bound(gene\ body)}$, which we calculated using data from Nguyen et al.[22] and Ho et al.[64] as follows:

$$fraction\ stably\ bound\ to\ DNA = (bound\ fraction) * (stable\ fraction\ of\ the\ bound\ molecules)$$

$$pol_{bound(promoters)} = \begin{pmatrix} PIC\ subunit \\ molecules\ per\ cell \end{pmatrix} * \begin{pmatrix} PIC\ subunit\ fraction \\ stably\ bound\ to\ DNA \end{pmatrix}$$

$$pol_{bound(promoters\ \ gene\ bodies)} = \begin{pmatrix} RNAPII\ subunit \\ molecules\ per\ cell \end{pmatrix} * \begin{pmatrix} Rpb1\ fraction \\ stably\ bound\ to\ DNA \end{pmatrix}$$

	bound fraction Nguyen et al.[22]	*stable fraction of the bound molecules* Nguyen et al.[22]	*molecules per cell* Ho et al.[64]
TFIIB (Sua7)	0.21	0.15	3264
TFIIE (TFA1)	0.33	0.26	5351
Rpb1	0.48	0.4	20891
Rpb9	N.D.	N.D.	5537

Assuming 6,000 PIC binding sites per genome,[65] we can then convert from per cell to per gene to determine the gene-body bound population as

$$pol_{bound(gene\ body)} = pol_{bound(promoter\ \ gene\ body)} - pol_{bound(promoter)}$$

Because estimates of molecules per cell vary between PIC and RNAPII subunits, we used the above values to define a feasible upper and lower limit as follows:

$$pol_{bound(promoter)} lower\ limit = 0\ 017 \text{ per pomoter } \text{ using TFIIB occupancy and copy number}$$

$$pol_{bound(promoter)} upper\ limit = 0\ 055 \text{ per pomoter } \text{ using TFIIE occupancy and copy number}$$

$$pol_{bound(promoter\ \ gene\ body)}\ upper\ limit = 0\ 67 \text{ RNAPII per gene } \text{ using Rpb1 copy number}$$

$$pol_{bound(promoter\ \ gene\ body)} lower\ limit = 0\ 177 \text{ RNAPII per gene } \text{ using Rpb9 copy number Rpb9 has the lowest estimated copy number among RNAPII subunits and was therefore used to define the lower limit}$$

Simulation parameters

Unless stated otherwise, simulations were performed with the following parameters. A 1 kb gene was represented by a 29-site strand, each site representing the 35 bp footprint of an RNAPII molecule. Similarly, the elongation rate was taken to be 0.94 steps/second,

corresponding to the rate at which an RNAPII molecule moves 35 bp. For each parameter set, the time step was set adaptively to accurately model the stochastic process and the number of iterations corresponded to 1000 binding events.

Comparison between simulations of RNAPII transcription on a representative gene and the dynamic equilibrium model

We then set out to determine if the size-dependent kinetics in global RNAPII recruitment described by the dynamic equilibrium model can be recapitulated by our stochastic simulations of an average transcriptional event. We modeled changes in cell size by varying the nucleoplasmic RNAPII concentration (pol_{free}) as predicted for each size by the dynamic equilibrium model. We then asked how the occupancy of RNAPII in the gene body responds. We repeated this across the 24 parameter combinations for $k_{on}\ pr$, $k_{off}\ pr$, and . For all the parameter sets we analyzed, the simulations give the same relative change in gene body occupancy as the dynamic equilibrium model (Figure S5E). This is because in all cases $k_{on}\ pr$ and values are well below the point that promoters or gene bodies start to saturate, and so transcription rates are in a linear regime with respect to polymerase recruitment rates, which are set by $k_{on}\ pr * [pol_{free}]$ Intuitively this makes sense because the absolute occupancy of the promoter is in the range of ~2-6% and there are only on average ~0.2-0.7 RNAPII molecules in the entire gene body. Thus, our granular stochastic simulation – that models the stepwise processes of promoter recruitment, transcriptional initiation, and elongation on a gene of finite length – supports the simplifying assumptions of the dynamic equilibrium for describing global RNAPII binding kinetics in a typical gene.

Analysis of simulation predictions for highly expressed genes

While the above results demonstrate the global occupancy of RNAPII can be precisely modelled according to kinetics of the dynamic equilibrium model, it only represents the average transcriptional unit in the genome. More highly expressed genes can have significantly higher PIC and RNAPII occupancies than considered in the average scenario, which could lead to saturation at large cell sizes. To assess this, we binned genes from our ChIP-seq data according to their RNAPII occupancy at ~50 fL in the elutriation G1 arrest experiment. To reduce the possibility that we are assessing the effects of some gene-class specific transcriptional programs, we excluded two groups of genes that are expected to change during G1 arrest in a programmed manner: cell cycle genes defined in Swaffer et al.[43] and the environmental stress response genes defined in Ho et al.[66] For each of the initial parameter sets, we interpolated the representative $k_{on}\ pr$ for each gene bin using their RNAPII occupancy at ~50 fL and then varied the free RNAPII concentration, as above for the average gene. This revealed different behaviors between different parameter sets for the most highly expressed genes (Figures 4J and S5F). In some parameter sets, there is close to no difference between genes of different occupancies as cell size increases. However, in others there is a modest saturation effect in the highest few percentiles of genes. This saturation effect is always principally due a high $k_{on}\ pr$ that causes promoter saturation before gene body saturation. We observe a similar trend in our data where the top few percentiles of genes do not increase their occupancy in larger cells as much as the global average. This may indicate these top few percent of genes, while not fully saturated, are subject to a non-linear saturating effect. However, it is also possible such changes emerge from a programmed transcriptional response instead of saturation, which would be consistent with several other feasible parameter sets we defined in which $k_{on}\ pr$ is low enough that highly expressed genes are still within the linear range and are, therefore, not subject appreciable saturation.

QUANTIFICATION AND STATISTICAL ANALYSIS

Spike-in normalized ChIP-seq analysis

Unless otherwise specified, all analyses described were carried out using custom-written Python scripts (https://github.com/georgimarinov/GeorgiScripts). Combined *S. cerevisiae* and spike-in (*C. glabrata* or *S. pombe*) genome FASTA files were created using the sacCer3 genome assembly for *S. cerevisiae* and the assemblies for *C. glabrata* (obtained from the Candida Genome Database (CGD)[67]) or *S. pombe* (ASM294v2). Combined gene annotations were created using *S. cerevisiae* gene models updated using transcript-end mapping data as previously described[68] and either the *C. glabrata* gene models available from CGD or the *S. pombe* gene models available from PomBase.[69]

Demultipexed FASTQ files were then mapped to the relevant combined genome indexes as 2 36mers using Bowtie (v.1.0.1)[50] with the following settings: -v 2 -k 2 -m 1 –best –strata. For subsequent analyses, reads that map to the *S. cerevisiae* or spike-in mitochondrial genomes were ignored and reads that map to both the *S. cerevisiae* and spike-in genomes were ignored.

To determine the global occupancy ratio, the sample-to-spike-in ChIP ratio of reads was calculated and then divided by the sample-to-spike-in input ratio for the same sample. For example, when using *C. glabrata* as the spike in:

$$\text{Occupancy per genome} = \left(\frac{\text{ChIP reads}_{S.\ cerevisiae}}{\text{ChIP reads}_{C.\ glabrata}}\right) \Big/ \left(\frac{\text{Input reads}_{S.\ cerevisiae}}{\text{Input reads}_{C.\ glabrata}}\right)$$

The global occupancy ratio is therefore expressed per genome. For the data in Figures 1E, 4E, 6A, and S4D, the occupancy per genome was converted to occupancy per cell by multiplying by the average genomes per cell determined by flow cytometry (see DNA content analysis section above). For histone modification occupancies (Figures 5C–5E and S6B), the occupancy per genome was calculated as above and then normalized to the anti-H3 or anti-H4 occupancy ratio for the same sample. For the anti-Rpb1 and

anti-FLAG ChIP experiments in Figures 2D, 2F, 2G, 3B, 4E, S3B–S3H, S4D, and S4F a control strain was used to determine the background signal (see STAR Methods section: spike-in normalized ChIP-seq), which was subtracted from the global occupancy ratio.

RPM (Reads Per Million) normalized read coverage genome browser tracks showing the 50 bp mid-point of each mapped fragment (*i.e.*, midpoint ± 25 bp) were generated for the *S. cerevisiae* genome using custom-written Python scripts. Subsequent analysis uses these 50 bp mid-point read coverage tracks to calculate per gene occupancies and metagene profiles. Per gene occupancy was calculated as follows. RPKM (Reads Per Kb per Million reads) values were calculated for the gene body (TSS-to-TTS) for all verified ORFs. ChIP RPKM values were then multiplied by the global ChIP occupancy factor described above to give the per gene occupancy value. Metagene profiles showing the average occupancy across gene bodies were generated as follows. Average RPKM values were calculated for the 500 bp upstream of the TSS, 500 bp downstream of the TTS, and gene bodies (TSS-to-TTS; rescaled to be 1 kb). RPKM values were then multiplied by the global occupancy factor described above and smoothed using a 50bp averaging window. The top 10% of genes (verified ORFs only) was determined based on anti-Rpb1 RPKM ChIP and was used where indicated. For the histone modification metagene heatmaps (Figure 5E), values were re-scaled to the median value of all samples displayed on a given heatmap to aid comparison between modifications. For the G1 arrest ChIP data in Figures 3, 4, and S4 the global occupancy ratios were also normalized to occupancy ratios of asynchronous control samples that were collected at the same timepoints but where not arrested in G1 to control for non-specific effects of elutriation procedure.

Dual enzyme single-molecule foot printing dSMF) analysis

Combined *S. cerevisiae* and *C. glabrata* genome and transcriptome files were created as described above (see STAR Methods section: spike-in normalized ChIP-seq analysis). FASTQ files were trimmed of adapter using cutadapt (version 0.16) and Trim Galore (version 0.4.4) with the following settings: –clip_R1 9 –clip_R2 9 –three_prime_clip_r1 6 –three_prime_clip_r2 6 –paired. Trimmed reads were then mapped to the combined genome indexes using the bwameth package (https://github.com/brentp/bwa-meth). Duplicate reads were removed using the MarkDuplicates programs (picard-tools-1.99). Methylation calls were extracted using MethylDackel (https://github.com/dpryan79/MethylDackel) with the following settings: –CHG –CHH. The MethylDackel output was used to create genome browser tracks showing methylation levels and to make metagene plots which were then normalized to the global methylation fraction determined for the spike-in *C. glabrata* genome.

Spike-in normalized RNA-seq analysis

A combined *S. cerevisiae* and *C. glabrata* genome file and gene annotations were generated as described above (see STAR Methods section: spike-in normalized ChIP-seq analysis). For the purposes of RNA-seq data quality evaluation and genome browser track generation, reads were aligned against the combined genome and annotated set of splice junctions using the STAR aligner (version 2.5.3a; settings: –limitSjdbInsertNsj 10000000 –outFilterMultimapNmax 50 –outFilterMismatchNmax 999 –outFilterMismatchNoverReadLmax 0.04 –alignIntronMin 10 –alignIntronMax 1000000 –alignMatesGapMax 1000000 –alignSJoverhangMin 8 –alignSJDBoverhangMin 1 –sjdbScore 1 –twopassMode Basic –twopass1readsN -1).[70] Read mapping statistics and genome browser tracks were generated using custom Python scripts. For quantification purposes, reads were aligned as 2x50mers in transcriptome space against an index generated from the combined gene annotation model using Bowtie (Langmead et al.[50]; version 1.0.1; settings: -e 200 -a -X 1000). Alignments were then quantified using eXpress (version 1.5.1)[51] as TPM (Transcripts Per Million transcripts). The sum of the TPM for all genes for a given species (RNAseq TPM) was calculated after filtering genes for verified ORFs (www.yeastgenome.org) and removing genes with TPM < 0.1 in any sample. The relative amount of mRNA per cell for a given culture is then calculated as:

$$\text{mRNA amount per genome} = \left(\frac{\text{RNAseq TPM}_{S.\ cerevisiae}}{\text{RNAseq TPM}_{C.\ glabrata}}\right) \Big/ \left(\frac{\text{gDNAseq reads}_{S.\ cerevisiae}}{\text{gDNAseq reads}_{C.\ glabrata}}\right)$$

$$\text{mRNA amount per cell} = (\text{mRNA amount per genome}) * (\text{genomes per cell})$$

See also Figure S7E for the experimental workflow. gDNAseq reads are calculated from gDNA-seq data, which was processed in the same way as described above for ChIP-seq data (see STAR Methods section: spike-in normalized ChIP-seq analysis), and the average ratio from two technical replicates was used. For cell size mutants, genomes per cell is determined by flow cytometry based DNA content analysis (Figures S7C and S7D). For the G1 arrest cultures, genomes per cell is always constant. We note this is a relative measurement that allows comparisons between samples processed in the sample batch with the same spike-in, but it does not provide absolute mRNA amounts per cell. For the data presented in Figure 6A all values are rescaled so the smallest samples = 1.

EU 5-ethynyl uridine) pulse-chase RNA-seq analysis

A combined *S. cerevisiae* and *S. pombe* genome file and gene annotations was used to align reads and calculate TPMs for EU-enriched and input samples as described above (see STAR Methods section: spike-in normalized RNA-seq analysis). The sum of the

TPM for all genes for a given species (mRNAseq TPM) was calculated after filtering genes for verified ORFs (www.yeastgenome.org) and removing genes with TPM < 0.1 in any sample. The relative total amount of EU labeled mRNA in each sample is then calculated as follows:

$$\text{total EU mRNA} = \left(\frac{\text{EU mRNAseq TPM}_{S\ cerevisiae}}{\text{EU mRNAseq TPM}_{S\ pombe}}\right) \Big/ \left(\frac{\text{input mRNAseq TPM}_{S\ cerevisiae}}{\text{input mRNAseq TPM}_{S\ pombe}}\right)$$

The total background signal, determined using the no-EU sample, was then subtracted and all values were then normalized so that the first timepoint is set = 1. Global mRNA turnover rates (Figures 6C and 6D) were calculated from a one-phase exponential decay fit to the total EU mRNA data (Prism 6; weighting = 1/Y, plateau = 0), excluding the first time point to permit a short lag after the chase. Turnover rates for individual transcripts were calculated by first renormalizing *S. cerevisiae* TPM values to the spike-in reads and then fitting an exponential decay fitting an exponential decay in R using the *nls_multstart()* function to fit the data to the *SSasymp()* model with the asymptote constrained to be > 0.

Single-molecule imaging analysis

Aside from tracking, all analysis was performed using custom Matlab code. Tracking was performed using Trackmate.[52] First, molecules were localized in each frame using a Laplacian of Gaussian (LoG) method with an estimated diameter of 5 pixels. An intensity threshold was chosen that was low enough to still detect molecules that were moving out of the focal plane and were diffusing quickly. After localization, tracks were formed using the Linear Assignment Problem algorithm by linking molecules in consecutive frames. The linking distance was set to 5 pixels. A gap frame of 3 was used to allow for missed localization. The gap-linking distance was set to 5 pixels more than the linking distance. Linking also had a cost of 0.3 for the ''Quality'' parameter to ensure that correct molecules were linked. Tracks with fewer than five localizations were discarded.

Only unbudded G1 cells, determined from brightfield images, were retained for subsequent analysis. The cell outlines were segmented by a custom-made Matlab code and then manually curated using the ''Freehand'' function in Matlab. Cell volumes were estimated from the cell masks assuming ellipsoid geometry. The volumes of cells were estimated by adding up the cross-section volume at each orthogonal pixel layer of the major axis of cells. Maximum intensity projections of z-stacks in 488nm were then used to segment the nuclear regions using the Pus1-GFP nuclear marker and tracks outside the nucleus were discarded. The radius of gyrations (RoG) was then calculated for each track using the following equation:

$$RoG = \sqrt{\frac{1}{n}\sum_{i}^{n}\left((x_i - \overline{X})^2 \quad (y_i - \overline{Y})^2\right)}$$

where n is the number of localizations in each track, x_i and y_i are the coordinates of the track for each step i, and $\overline{X}$ and $\overline{Y}$ are the mean of the x, y coordinates of the track, respectively. To classify molecules as bound or unbound, a one-dimensional Gaussian mixture model (GMM) was fitted to the $\log_2$-transformed RoG distribution. The initial values for fitting were inferred from the log-transformed RoG distribution of H2B1-HALO and mCitrine-HALO-NLS, which have the majority of molecules bound and unbound, respectively. To determine how many Gaussian groups exist in the RoG distribution, we fitted the distribution with different GMMs, where total gaussian groups varied from 1 to 4. The Bayesian information criterion (BIC) was calculated for each GMM fitting. The model with 2 gaussian groups was selected on the basis that it had the lowest BIC. The Gaussian group with the lowest average RoG value was classified as the group of bound molecules. The tracks inside the chromatin bound fraction in each nucleus was then determined by dividing the number of bound tracks by the number of total tracks. Nuclei with < 30 tracks were discarded from this analysis.

Analysis of ChEP LC-MS/MS data

MaxQuant (version 1.5.0.13) was used for all LC-MS/MS data processing. The data were searched against a UniProt extracted *S. cerevisiae* proteome FASTA file. SILAC comparison between WCE and chromatin fractions (see Figures S2C and S2D and STAR Methods section chromatin enrichment for proteomics (ChEP)) was used to determine a high confidence list of chromatin-enriched proteins (n=564). Chromatin-enriched proteins were defined as proteins enriched by chromatin extraction methods A and B compared to WCE in both biological replicates. These 564 proteins and the threshold used to define enrichment are shown in Figure S2D. Normalization of SILAC ratios was not applied because there is no *a priori* assumption that the median of the distribution corresponds to no change. Instead, raw SILAC ratios are plotted and the thresholds for each replicate were defined so that thresholds intersect on the linear regression line between the two replicates (Figure S2D) at a point that separates known chromatin associated factors from abundant cytoplasmic factors (Figure S2C). The distribution of SILAC ratios for the experiment plotted in Figure 2B was normalized so the average histone protein corresponds to a SILAC ratio = 1 and as such these data can be interpreted as protein bound per genome, given histone occupancy per genome is approximately constant as a function of cell size, as determined by histone Chip-Seq (Figure S4D).

Supplemental figures

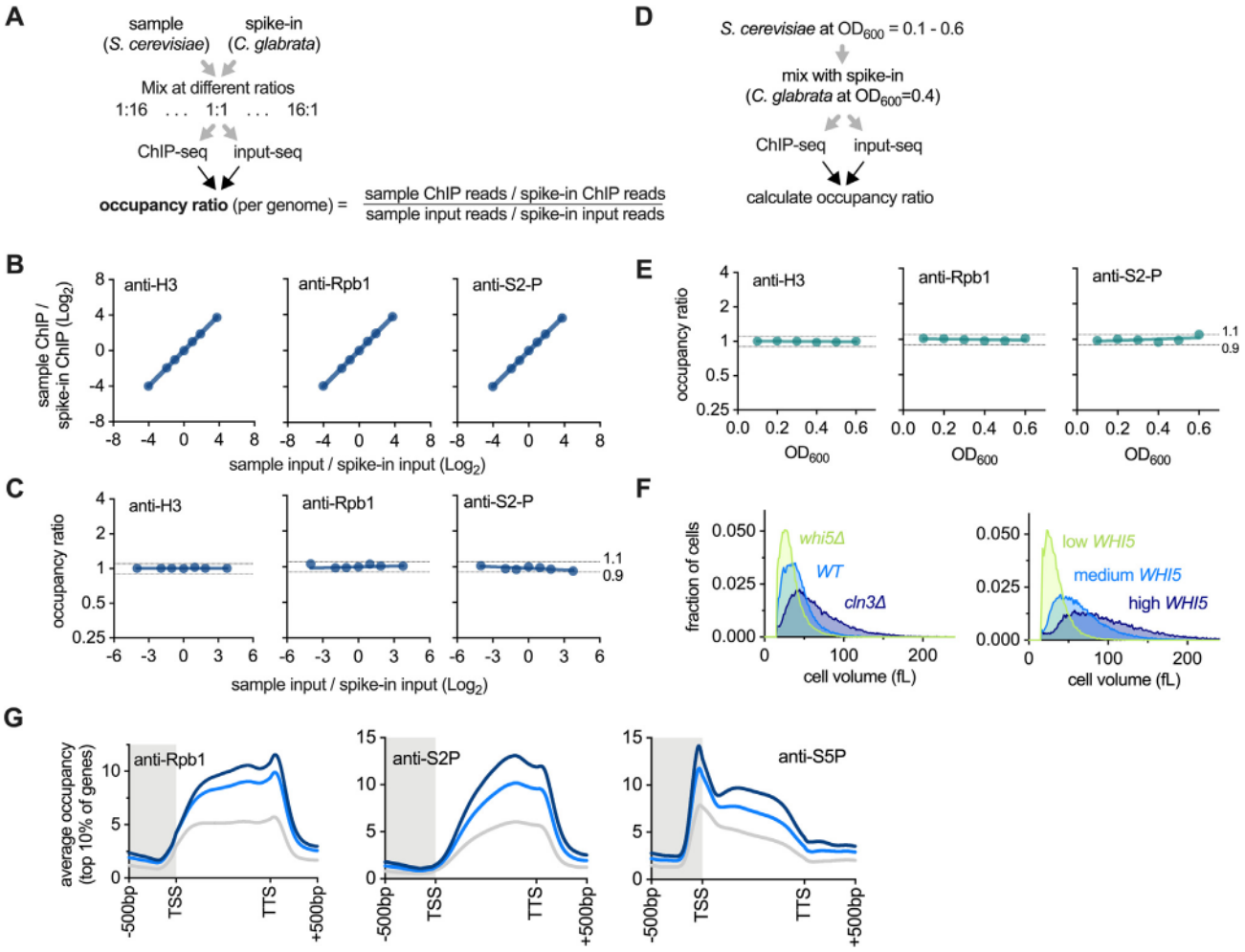


Figure S1. Controls and supporting data for the spike-in normalized ChIP-seq of Rpb1 in cell size mutants, related to Figure 1

(A–C) Mixing-ratio controls for spike-in normalized ChIP-seq. The design of the mixing-ratio control experiment is shown in (A): *S. cerevisiae* and *C. glabrata* samples were mixed in 7 different ratios before anti-Rpb1, anti-S2P, or anti-H3 ChIP-seq. Input samples were also sequenced to determine the mixing ratio. This experiment demonstrates the good dynamic range and linearity of spike-in normalized ChIP-seq, because differences in the input ratio of *S. cerevisiae* to *C. glabrata* are proportional to the changes in the *S. cerevisiae* to *C. glabrata* ChIP ratio (B) so that the input normalized occupancy ratio was constant (C). ChIP and occupancy ratios are expressed as the fold-change relative to the middle sample. This control is effectively the same as that reported for Scc1 ChIP-seq by Hu et al.[18] and demonstrates that different *S. cerevisiae* samples do not need to be mixed at identical ratios with spike-in cells, because variations in mixing ratio are linearly accounted for in the input normalized occupancy ratio.

(D and E) Cell density control for spike-in normalized ChIP-seq. The design of the cell density control experiment is shown in (D): *S. cerevisiae* cells were grown to different cell densities (OD_{600}) and mixed with approximately the same number of *C. glabrata* cells. The sample was then fixed and processed for spike-in normalized ChIP-seq. The occupancy ratio for anti-Rpb1, anti-S2P, and anti-H3 is constant as a function of cell density (E), showing that *S. cerevisiae* samples do not need to be collected at identical cell densities to compare samples. All samples in the rest of the study were collected between OD_{600} = 0.2 and OD_{600} = 0.4, unless stated otherwise.

(F) Cell size distributions determined by Coulter counter for the cell size mutants used in the experiment presented in Figures 1C–1F. Low, medium, and high levels of *WHI5* expression were induced with increasing concentrations of beta-estradiol, respectively.

(G) Average occupancy across the gene bodies of the top 10% of genes for total Rpb1, initiated Rpb1 (anti-S5-P), and elongating Rpb1 (anti-S2-P) in cells expressing low, medium, and high levels of *WHI5*. Global occupancy measurements for the same data are shown in Figure 1E.

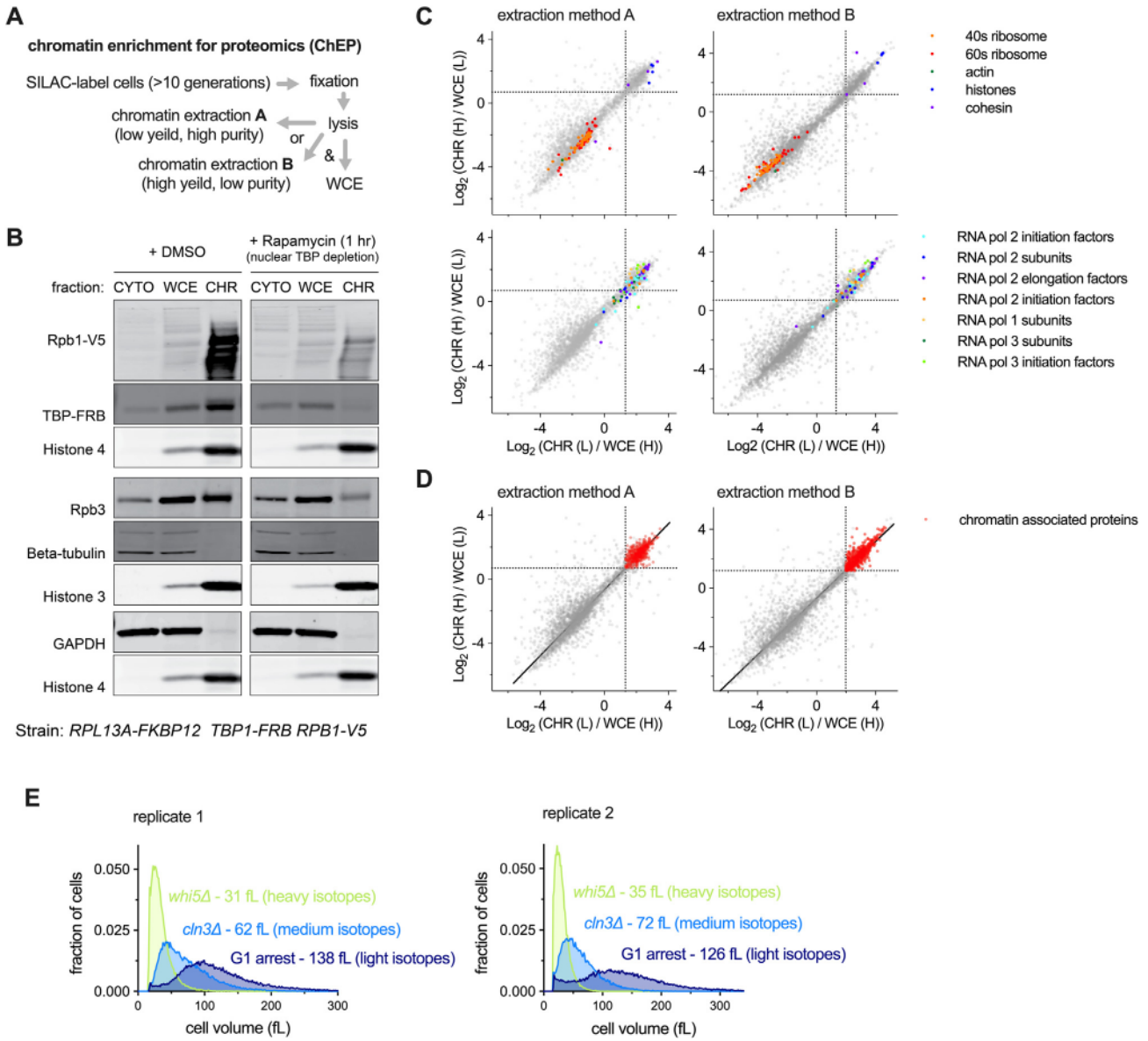


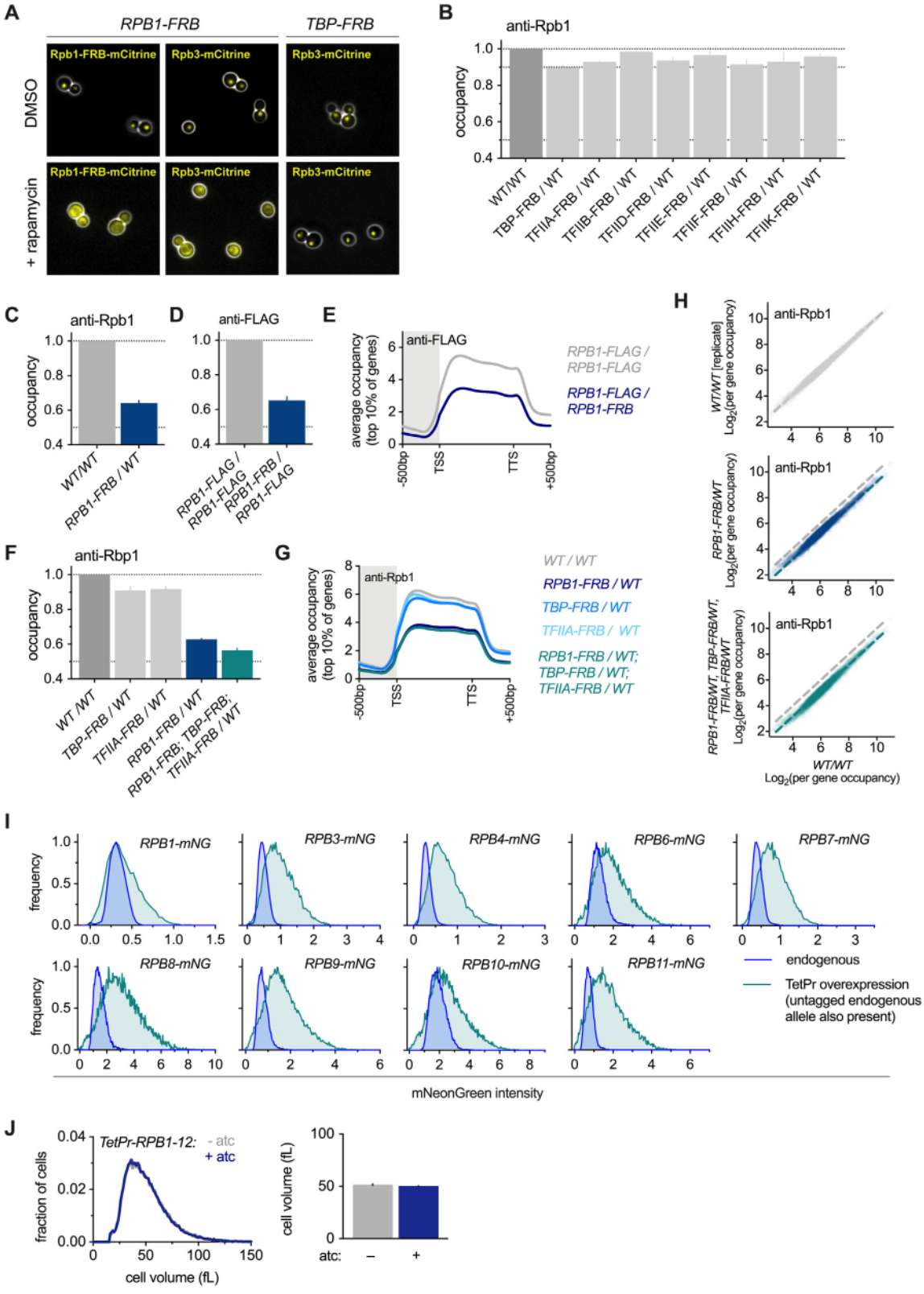
Figure S2. Supporting data and controls for chromatin enrichment for proteomics ChEP), related to Figure 2B

(A) Schematic of the workflow for chromatin extraction by ChEP. Both the high-purity/low-yield extraction (method A) and the low-purity/high-yield extraction (method B) were used (see STAR Methods for details).

(B) Control experiment showing that ChEP can quantify changes in protein chromatin association. To test this, TPB and RNAPII chromatin association was conditionally prevented using a TBP-FRB strain where TBP is conditionally depleted from the nucleus upon rapamycin treatment, thus preventing RNAPII recruitment to the genome. Upon rapamycin treatment for 1 h, TBP and RNAPII (Rpb1 and Rpb3) have reduced chromatin association measured by ChEP, compared with the DMSO control treatment. Thus, ChEP can be used to quantify changes in chromatin association. We note that under these conditions RNAPII is still present in the nucleus but is not associated with chromatin (see Figure S3A), indicating that ChEP signal is not dominated by large amounts of non-specific chromatin interactions, at least in the case of RNAPII.

(C and D) Experiment to identify proteins enriched by ChEP. Cells were SILAC labeled with heavy (H) or light (L) isotopes, and three extractions were recovered from each: whole-cell extract (WCE), chromatin extraction A, and chromatin extraction B. WCE and chromatin with opposite SILAC labels were mixed and analyzed by LC-MS/MS. Each axis shows an independent biological replicate. Left-hand panels show chromatin extraction A (high purity, low yield). Right-hand panels show chromatin extraction B (low purity, high yield). Known cytoplasmic factors (ribosomes and actin) and chromatin-associated proteins (RNA polymerase subunits, transcription factors, histones, and cohesion) are shown in (C), validating that ChEP enriches chromatin-associated factors. Proteins defined as chromatin associated are shown in red in (D). A robust linear regression is shown in black. Chromatin-associated proteins were defined as proteins enriched in the chromatin fraction of both replicates for both extraction method A and extraction method B. The thresholds used to define chromatin-associated proteins are shown as dashed black lines (see STAR Methods for details).

(E) Cell size distributions determined by Coulter counter for the SILAC-labeled cultures used in the experiment in Figure 2B. All strains have a SILAC-compatible genetic background (see STAR Methods for details).



(legend on next page)

Figure S3. Supporting data for PIC subunits depletion and RNAPII overexpression experiments, related to Figures 2C–2G

(A) Microscopy images (phase contrast in white, mCitrine signal in yellow) of the indicated genotypes treated with DMSO or rapamycin for 1 generation before imaging. Treatment of Rpb1-FRB cells leads to nuclear depletion of Rpb3, indicating the whole RNAPII complex is likely to be efficiently co-depleted. Treatment of TBP-FRB cells does not alter Rpb3 localization, indicating that depletion of one subcomplex of the pre-initiation complex do not result in the co-depletion of other subcomplexes.

(B) The global Rpb1 occupancy measured by spike-in normalized ChIP-seq in wild-type diploid cells (*WT/WT*) or diploids where 50% of the indicated RNAPII pre-initiation complex subunit was depleted from the nucleus following a 100-min rapamycin treatment. Mean (± SEM) is plotted. The corresponding average Rpb1 occupancy across the gene bodies of the top 10% of genes in these samples is shown in Figure 2D.

(C) The global Rpb1 occupancy measured by spike-in normalized ChIP-seq in wild-type diploid cells (*WT/WT*) or diploids where 50% of the RNAPII was depleted from the nucleus by a 40-min rapamycin treatment. Mean (± SEM) is plotted. The corresponding average Rpb1 occupancy across the gene bodies of the top 10% of genes in these samples is shown in Figure 2D.

(D) The global Rpb1-FLAG occupancy measured by spike-in normalized anti-FLAG ChIP-seq in wild-type diploid cells (*RPB1-FLAG/RPB1-FLAG*) or diploids where one allele is FLAG tagged and the other is depleted from the nucleus upon rapamycin treatment (*RPB1-FRB/RPB1-FLAG*). Mean (± SEM) is plotted. These data indicate that the nuclear depletion of Rpb1 is efficient and near complete (i.e., ~50%) because ChIP against total Rpb1 in (C) and ChIP against the non-depleted allele in (D) give similar values.

(E) The average Rpb1 occupancy across the gene bodies of the top 10% of genes for the samples shown in (D).

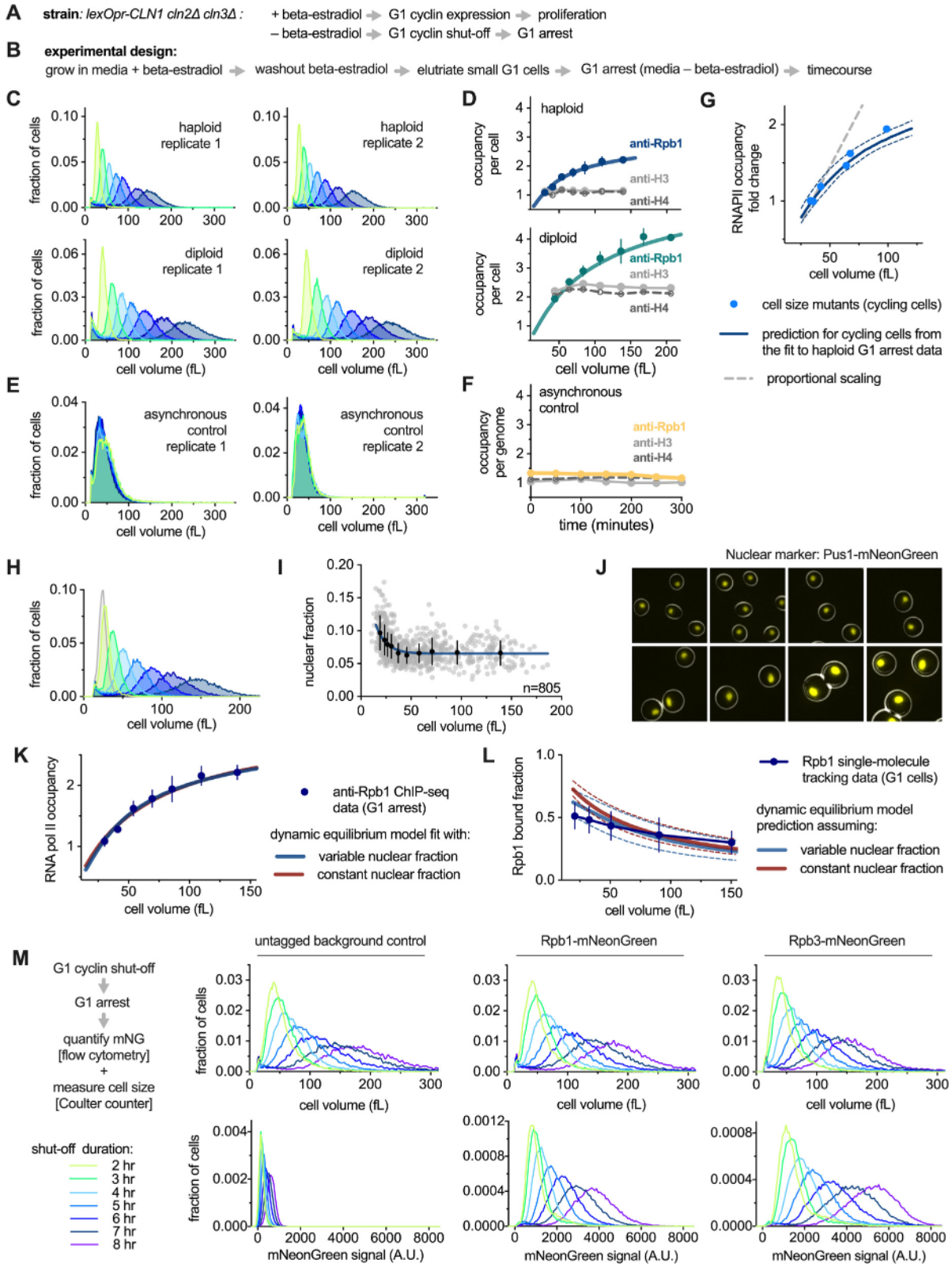
(F) The global Rpb1 occupancy measured by spike-in normalized ChIP-seq in wild-type diploid cells (*WT/WT*) or diploids where 50% of the indicated RNAPII pre-initiation complex subunits were depleted from the nucleus by a 40-min rapamycin treatment. Data are also show in an experiment with the simultaneous 50% nuclear depletion TBP, TFIIA, and RNAPII. Mean (± SEM) is plotted.

(G) The average Rpb1 occupancy across the gene bodies of the top 10% of genes for the samples shown in (F).

(H) The mean per gene Rpb1 occupancy measured (left) in wild-type cells, (center) in cells after 50% nuclear depletion of RNAPII, or (right) in cells after the simultaneous 50% nuclear depletion of TBP, TFIIA, and RNAPII. Each point corresponds to a single gene and shows the Rpb1 occupancy between TSS and TTS. Dashed lines show the average trends for the respective samples. Global occupancy values for the same samples are shown in (F).

(I) Histogram (mode normalized) of the mNeonGreen signal measured by flow cytometry for the indicated RNAPII subunit C-terminal fusion with mNeonGreen (mNG). Fusion proteins were either expressed from their endogenous promoter (blue; endogenous) or overexpressed from a *Tet* promoter alongside the untagged endogenous copy of the same gene (green; TetPr). Expression from the *TetPr* was induced by anhydrotetracycline (atc) treatment for 45 min. The fold overexpression that is induced upon atc addition was calculated by comparing the median intensity of the overexpressed construct to the endogenous allele and is shown in Figure 2E.

(J) Cell size distribution (left) and mean cell size (right) of the *TetPr-RPB1-12* strain. Cell size was measured before (−atc) and 45 min after (+atc) the simultaneous induction of all 12 RNAPII subunits.



(legend on next page)

Figure S4. Supporting data and additional controls for the G1 arrest experiments, related to Figures 3 and 4A–4F

(A and B) Schematic of the experimental design to generate populations of G1-arrested cells of increasing cell size. (A) A G1-cyclin shut-off strain was used that allows the conditional inactivation of G1 cyclin expression upon removal of beta-estradiol from the media. (B) Small G1 cells were then collected by centrifugal elutriation, and populations of increasing cell size were generated as G1 arrest duration increases (see STAR Methods for details).

(C) Cell volume distributions determined by Coulter counter for G1-arrested haploid (top) and diploid (bottom) cultures of increasing cell size used to analyze Rpb1, histone 3, and histone 4 occupancy in (D) and Figures 3 and 4. Cell populations with larger sizes correspond to cells arrested for longer times.

(D) Rpb1 (anti-Rpb1), histone 3 (anti-H3), and histone 4 (anti-H4) occupancy per cell, plotted as a function of cell size in G1-arrested haploids (top) and diploids (bottom). The mean (± range) of two biological replicates is shown for anti-Rpb1, whereas only a single replicate of anti-H3 and anti-H4 is shown. The anti-Rpb1 data are also plotted in Figures 3B, 4E, and 6A.

(E) Cell size distributions determined by Coulter counter for asynchronous control cultures used to analyze Rpb1, histone 3, and histone 4 occupancy in (F).

(F) Rpb1 (anti-Rpb1), histone 3 (anti-H3), and histone 4 (anti-H4) occupancy per genome, plotted as a function of cell size in asynchronous control cultures. The mean (± range) of two biological replicates is shown for anti-Rpb1, whereas only a single replicate of anti-H3 and anti-H4 is shown.

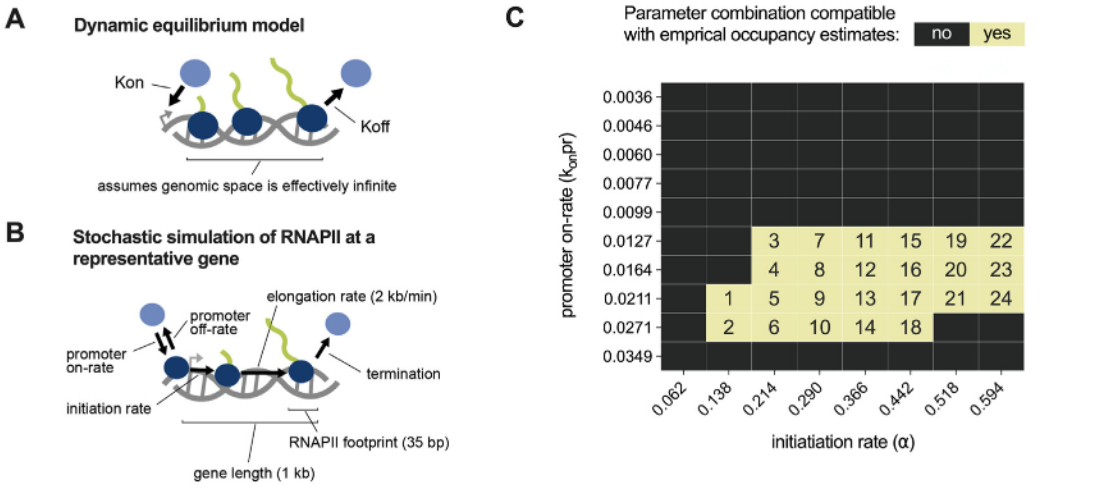
(G) Fold change in RNAPII occupancy in cell size mutants, compared with the prediction from the dynamic equilibrium model that was fit to data from G1-arrested cells. Error bars show 90% confidence intervals for model prediction determined by bootstrapping. Dashed gray line indicates proportional scaling.

(H–J) Cells expressing a nuclear marker (*PUS1-mNeonGreen*) were imaged at different sizes after elutriation and G1 arrest to determine how nuclear volume scales with cell volume. The experimental setup is the same as that outlined in (B). (H) Cell volume distributions determined by Coulter counter for G1-arrested *PUS1-mNeonGreen* cells of increasing cell size. (I) The nuclear fraction of the cell volume as a function of cell volume determined by wide-field fluorescence microscopy. Data from all 8 time points shown in (A) are pooled for this analysis. Bin mean (± SD) are shown in black. The blue line is the fit used to substitute nuclear volume for cell volume in the dynamic equilibrium model (see STAR Methods for details). Figure 3F shows the same data plotted as nuclear volume. (J) Representative images of Pus1-mNeonGreen signal and cell outlines.

(K–M) Comparison of dynamic equilibrium fit and prediction with either a variable nuclear fraction (blue) or a fixed nuclear fraction (red). (K) Rpb1 (anti-Rpb1) occupancy per cell as a function of cell size during G1 arrest. The fit of the dynamic equilibrium model assuming variable nuclear fraction (blue) or a fixed nuclear fraction (red) is shown. Mean (± range) of two biological replicates is shown. Data and variable nuclear fraction fit are also shown in Figures 3B, 4E, and 6A.

(L) The fraction of bound Rpb1 molecules plotted as a function of cell size in G1. Mean (± SD) of each bin is shown. The predictions of the dynamic equilibrium model from the fits in (E) are shown for both a variable nuclear fraction (blue) or a fixed nuclear fraction (red). Error bars for the fits show the 90% confidence interval generated by bootstrapping. Data and variable nuclear fraction fit are also shown in Figure 4C.

(M) The scaling of RNAPII subunit protein amounts with cell size was determined in G1-arrested cells following G1 cyclin shut-off. Cell volume distributions determined by Coulter counter (top) and total mNeonGreen signal distributions determined by flow cytometry (bottom). Three strains were analyzed: untagged cells (left), *RPB1-mNeonGreen* (center), and *RPB3-mNeonGreen* (right) (see Figure 3E for the quantitative comparison of protein amount with cell size).

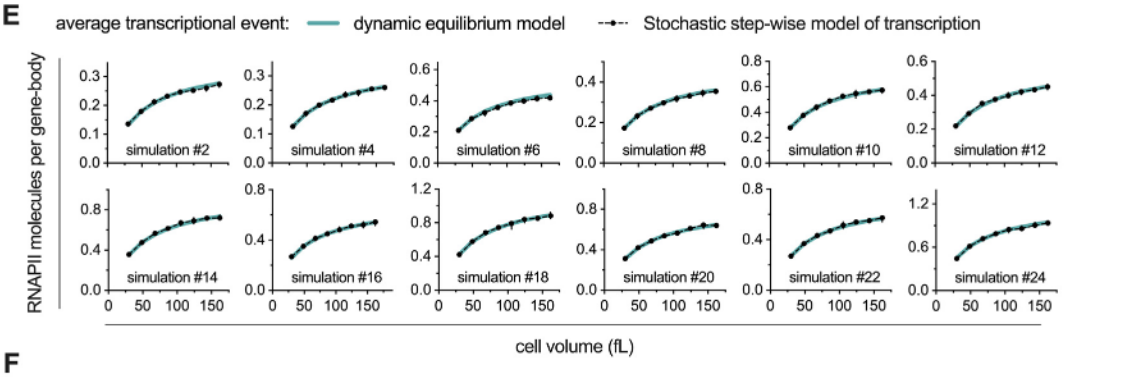


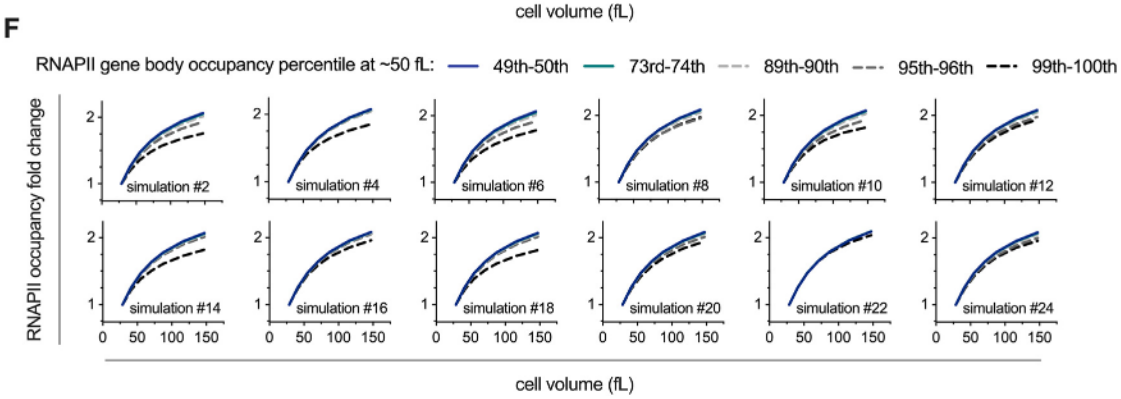
D

feasible parameter sets

	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12
promoter on-rate ($k_{on}pr$)	0.021	0.027	0.013	0.016	0.021	0.027	0.013	0.016	0.021	0.027	0.013	0.016
promoter off-rate ($k_{off}pr$)	0.450	0.450	0.374	0.374	0.374	0.374	0.298	0.298	0.298	0.298	0.222	0.222
initiation rate (α)	0.138	0.138	0.214	0.214	0.214	0.214	0.290	0.290	0.290	0.290	0.366	0.366
promoter occupancy	0.036	0.046	0.021	0.027	0.035	0.046	0.022	0.027	0.035	0.045	0.021	0.027
gene body occupancy	0.144	0.185	0.131	0.171	0.221	0.281	0.184	0.231	0.302	0.387	0.233	0.298

	#13	#14	#15	#16	#17	#18	#19	#20	#21	#22	#23	#24
promoter on-rate ($k_{on}pr$)	0.021	0.027	0.013	0.016	0.021	0.027	0.013	0.016	0.021	0.013	0.016	0.021
promoter off-rate ($k_{off}pr$)	0.222	0.222	0.146	0.146	0.146	0.146	0.070	0.070	0.070	0.000	0.000	0.000
initiation rate (α)	0.366	0.366	0.442	0.442	0.442	0.442	0.518	0.518	0.518	0.594	0.594	0.594
promoter occupancy	0.035	0.046	0.021	0.028	0.036	0.043	0.021	0.027	0.035	0.021	0.028	0.035
gene body occupancy	0.386	0.498	0.278	0.354	0.472	0.581	0.322	0.429	0.543	0.380	0.486	0.612





(legend on next page)

Figure S5. Supporting data for stochastic simulation of RNAPII transcription, related to Figures 4G–4K

(A) Schematic of the steps modeled in the dynamic equilibrium model. RNAPII association with the genome is simplified to a binding (k_{on}) and an unbinding step (k_{off}) and assumes that the genome does not saturate with RNAPII.

(B) Schematic of the steps included in the stochastic simulation of RNAPII at a representative gene. RNAPII association with the genome is modeled as a multistep process on a gene of finite length on which a single RNAPII molecule has a 35-bp footprint. RNAPII binds a single site in the promoter with rate k_{on}pr and then either initiates with rate or dissociates from the promoter with rate k_{off}pr. Once initiated, RNAPII moves through the gene body (1 kb) at a fixed elongation rate before terminating and dissociating from the gene.

(C and D) The model shown in (B) is parameterized using only values from the budding yeast literature. In the case of k_{on}pr, k_{off}pr, and , 24 feasible parameter sets were selected. Parameter sets were deemed feasible only if the absolute number of RNAPII molecules per gene (gene body occupancy) and per promoter (promoter occupancy) falls within the empirically estimated range of these values. See STAR Methods for details on how the estimated ranges were calculated.

(C) Feasible matrix showing the parameter sets that are compatible with the empirical estimates of gene body and promoter occupancy. Numbers 1–24 are assigned to each feasible parameter set and match the annotations in (D)–(F).

(D) Table of k_{on}pr, k_{off}pr, and values for feasible parameter sets, as well as the absolute gene body occupancy and promoter occupancy per gene for each parameter set.

(E) RNAPII occupancy per gene at different sizes predicted by the stochastic simulation of RNAPII at a representative gene and the fit of the dynamic equilibrium model. 12 of the 24 feasible parameter combinations are shown (even numbered parameter sets). Because the dynamic equilibrium model is derived from fitting to data that are expressed as relative changes, the values for dynamic equilibrium model fit were re-scaled so that at ~50 fL it matches the absolute RNAPII molecules per gene body of each simulation.

(F) Fold change in the simulated RNAPII occupancy on model genes representative of the genes in the indicated empirical RNAPII occupancy percentiles (higher percentiles correspond to more highly occupied genes). 12 of the 24 feasible parameter combination are shown as in (E).

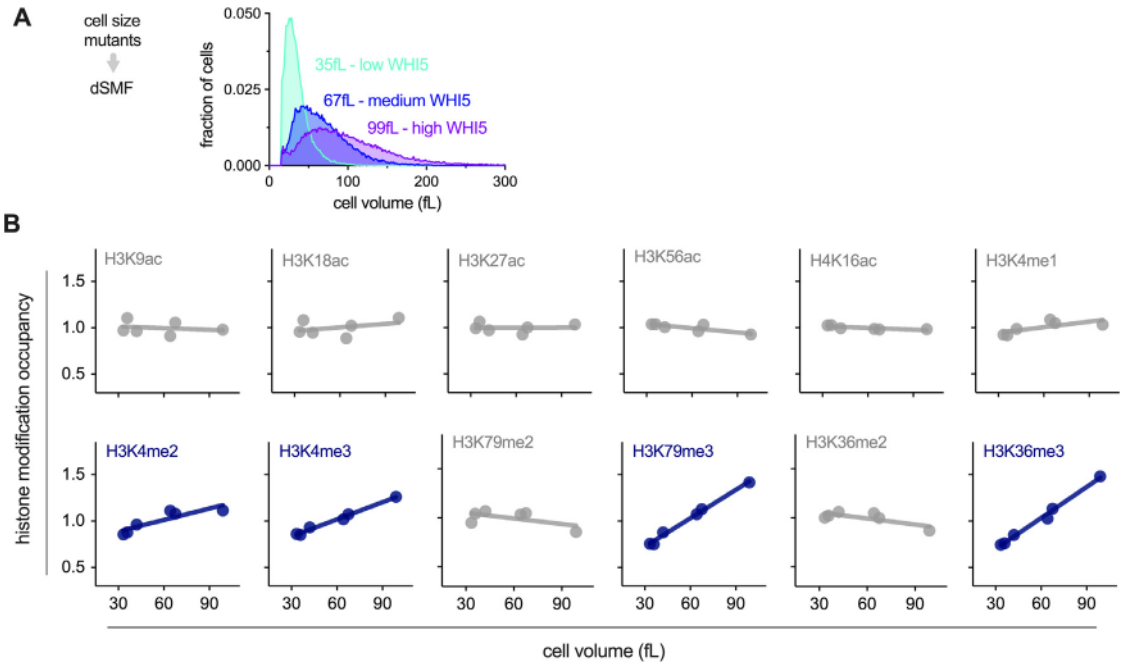
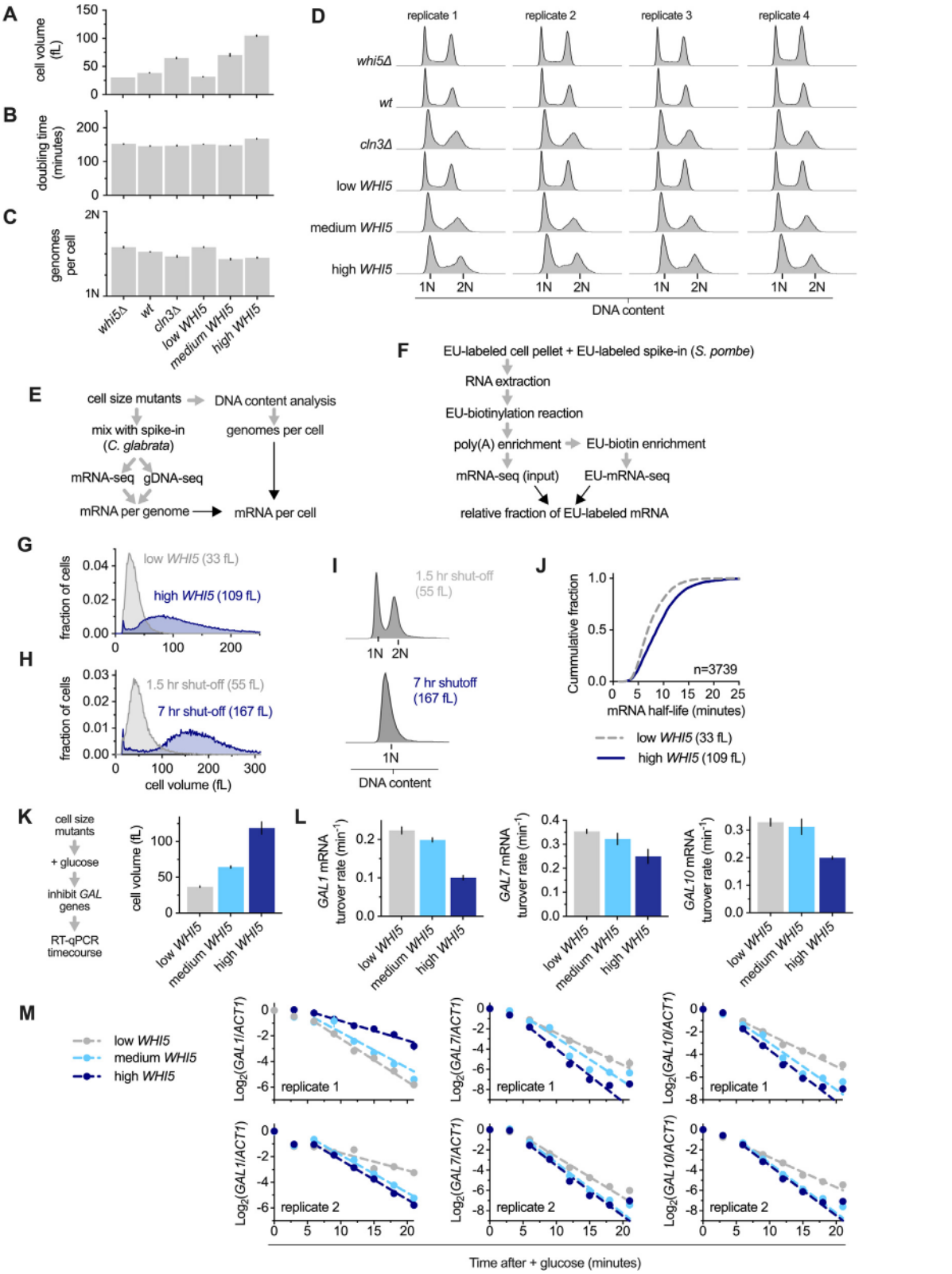


Figure S6. Supporting data for the analysis of chromatin as a function of cell size, related to Figure 5

(A) Cell volume distributions determined by Coulter counter for the cultures used in the dSMF experiment in Figures 5A and 5B to measure chromatin accessibility in cells of different sizes.

(B) Global histone modification occupancy of the data shown in Figures 5C-5E. H3K56ac and H3k36me3 data are also shown in Figure 5C. Line shows the linear fit used to calculate the slope values shown in Figure 5D.



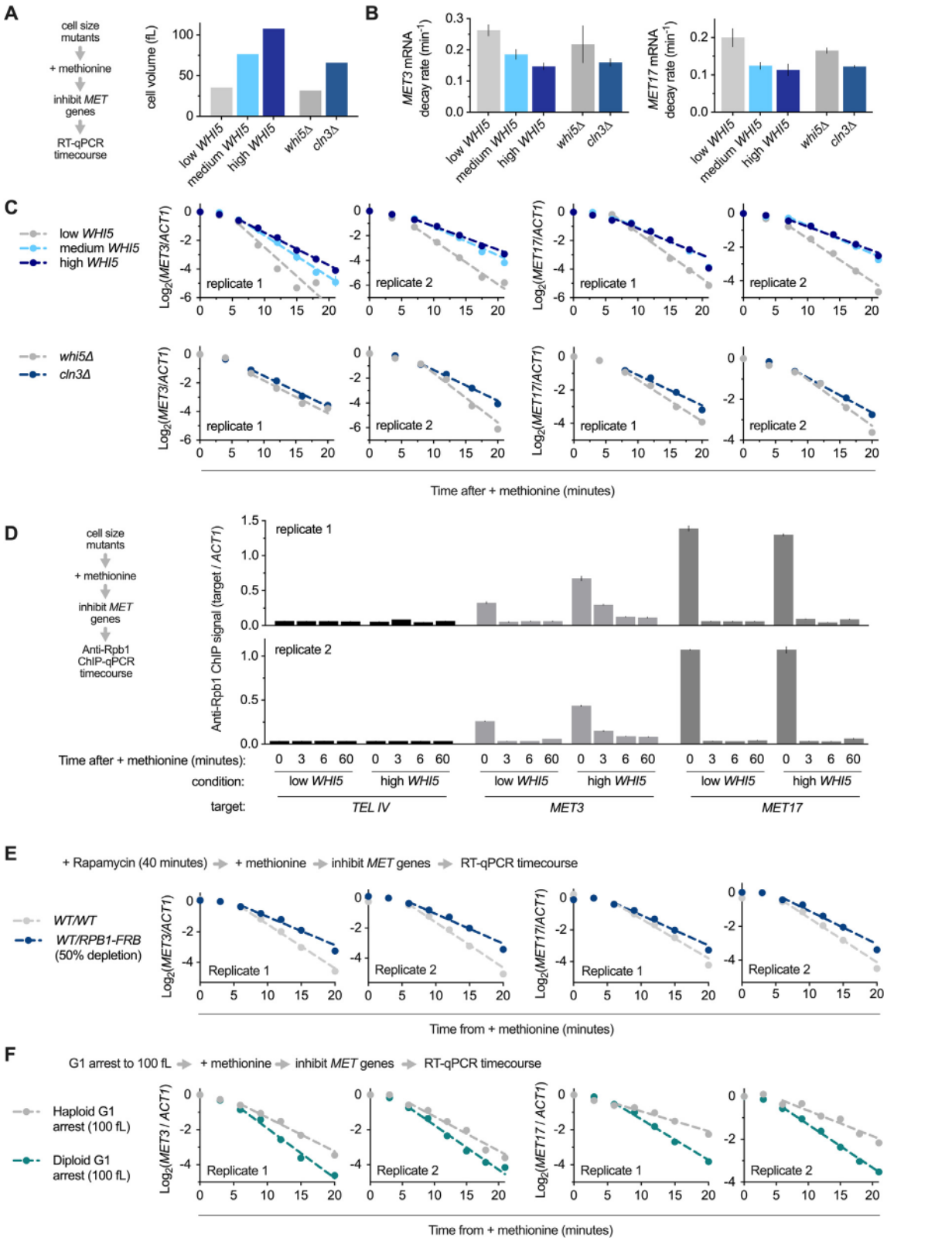
(legend on next page)

Figure S7. Supporting data for measurements of size-dependent mRNA amount and mRNA decay rates, related to Figure 6

(A–E) Cell size mutants were collected for spike-in normalized RNA-seq and gDNA-seq to determine the relative global mRNA amount per cell. Four biological replicates were performed. (A) Mean cell volume determined by Coulter counter for the indicated cell size mutants (n = 4). (B) Doubling time, determined from OD_{600} growth curves, of the indicated cell size mutants (n = 3). (C) Average genomes per cell, calculated from DNA content measurements shown in (D), of the indicated cell size mutants (n = 4). (D) DNA content analysis, determined by flow cytometry, of the indicated cell size mutants. (E) Experimental workflow for performing spike-in normalized RNA-seq to determine the relative global mRNA amount per cell (see STAR Methods for details).

(F–I) Cell size mutants and G1 cyclin shut-off cells were used to measure global mRNA turnover rates by EU pulse-chase, followed by EU-mRNA-seq. (F) Experimental workflow for processing EU-labeled cell pellets to determine the EU-labeled fraction of mRNA (see STAR Methods for details). (G) Cell volume distribution, determined by Coulter counter, of the indicated cell size mutants. Average of two biological replicates is shown. (H) Cell volume distribution, determined by Coulter counter, of cells 1.5 or 7 h after removal of beta-estradiol to induce G1 cyclin shut-off. One biological replicate is shown. (I) DNA content analysis, determined by flow cytometry, of the samples in (G). (J) Cumulative frequency distributions of individual mRNA half-lives from the EU pulse-chase experiment shown in (G) and Figure 6C.

(K–M) *GAL1*, *GAL7*, and *GAL10 mRNA* decay rates were determined in small (low *WHI5*), medium (medium *WHI5*), and large (high *WHI5*) cells by RT-qPCR, following transcriptional inhibition of these genes by the rapid replacement of galactose from the media with glucose (see STAR Methods for details). Two independent biological replicates were performed. (K) Mean cell volume in the indicated cell size mutant. (L) Mean (± range) *GAL1*, *GAL7*, and *GAL10* mRNA decay rates calculated from the decay curves in (M). (M) *GAL1* (left), *GAL7* (center), and *GAL10* (right) mRNA levels relative to *ACT1* mRNA, determined by RT-qPCR, after transcriptional inactivation by galactose replacement with glucose. Dashed line shows an exponential fit to the data.



(legend on next page)

Figure S8. Supporting data for mRNA decay rate measurements of *MET3* and *MET 7*, related to Figures 6F–6H

(A–C) *MET3* and *MET17* turnover rates were determined by RT-qPCR following methionine addition (which inactivates *MET3* and *MET17* transcription) in the indicated cell size mutants. Two independent biological replicates were performed. Part of these data are also shown in Figure 6F. (A) Mean cell volume, determined by Coulter counter, for the indicated cell size mutant. (B) Mean (± range) *MET3* and *MET17* mRNA decay rates calculated from the decay curves in (C). (C) *MET3* (left) and *MET17* (right) mRNA levels relative to *ACT1* mRNA determined by RT-qPCR after methionine addition. Dashed line shows an exponential fit to the data.

(D) *MET3* and *MET17* transcriptional inhibition kinetics following methionine addition were determined in small (low *WHI5*) and large (high *WHI5*) cells by anti-Rpb1 ChIP-qPCR. Both *MET3* and *MET17* reach steady-state repression within 6 min in both conditions. For this reason, all fitting for mRNA decay curves excludes time points before 6 min.

(E) *MET3* (left) and *MET17* (right) mRNA levels relative to *ACT1* mRNA were determined by RT-qPCR after methionine addition in *WT/WT* diploids or *RPB1-FRB/WT* heterozygous diploids. Methionine was added 40 min after rapamycin treatment. Rapamycin treatment depletes ~50% of the RNAPII from the nucleus (see Figures 2C and 2D). Two independent biological replicates were performed. Dashed line shows an exponential fit to the data. Cell volume distributions and the decay rates, calculated from these fits, are shown in Figure 6G.

(F) *MET3* (left) and *MET17* (right) mRNA levels relative to *ACT1* mRNA were determined, by RT-qPCR after methionine addition, in haploids and diploids arrested in G1 until they reached ~100 fL. Dashed line shows an exponential fit to the data. Two independent biological replicates were performed. Cell volume distribution and the decay rates, calculated from these fits, are shown in Figure 6H.