RNA polymerase II dynamics and mRNA stability feedback determine mRNA scaling with cell size

Matthew P. Swaffer[1], Georgi K. Marinov[2], Huan Zheng[3], Andrew W. Jones[4], Jessica Greenwood[4], Anshul Kundaje[2], Ambrosius P. Snijders[4], William J. Greenleaf[2], Rodrigo Reyes-Lamothe[3], Jan M. Skotheim[1]

[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

Summary

A defining feature of cellular growth is that protein and mRNA amounts scale with cell size so that concentrations remain approximately constant, thereby ensuring similar reaction rates and efficient biosynthesis. A key component of this biosynthetic scaling is the scaling of mRNA amounts with cell size, which occurs even among cells with the same DNA template copy number. Here, we identify RNA polymerase II as a major limiting factor increasing transcription with cell size. Other components of the transcriptional machinery are only minimally limiting and the chromatin environment is largely invariant with size. However, RNA polymerase II activity does not increase in direct proportion to cell size, inconsistent with previously proposed DNA-titration models. Instead, our data support a dynamic equilibrium model where the rate of polymerase loading is proportional to the unengaged nucleoplasmic polymerase concentration. This sublinear transcriptional increase is then balanced by a compensatory increase in mRNA stability as cells get larger. Taken together, our results show how limiting RNA polymerase II and feedback on mRNA stability work in concert to ensure the precise scaling of mRNA amounts across the physiological cell size range.

INTRODUCTION

Cell size has a profound impact on many aspects of cellular physiology because cell size sets the scale of subcellular structures, metabolism, surface-to-volume ratios, and, most crucially, the cellular biosynthetic capacity. Biosynthesis scales with cell size so that total RNA and protein per cell increase in proportion to cell volume such that their concentrations remain approximately constant as cells grow (**Fig. 1A**). Nuclear concentrations are also expected to be constant because the nuclear volume increases in proportion to cell volume (Jorgensen et al., 2007; Neumann and Nurse, 2007). Thus, by scaling biosynthesis to cell size, cells maintain approximately constant concentrations of key enzymes and reactants of the biochemical reactions supporting core cellular processes.

The importance of this biosynthesis scaling with cell size is most clearly illustrated in excessively large cells where it breaks down. Biosynthetic scaling only occurs within a window of physiological cell sizes beyond which both RNA and protein synthesis fail to keep pace with the expanding cell volume so that the cytoplasm becomes increasingly diluted (Neurohr et al., 2019; Zhurinsky et al., 2010). This break down in biosynthetic scaling is associated with a dramatic decline in many aspects of cellular physiology and increasing evidence points to this being a causal driver of cellular ageing and senescence (Cheng et al., 2021; Lanz et al., 2021; Lengefeld et al., 2020; Neurohr et al., 2019)

To support the scaling of macromolecules with cell size, larger cells have higher global transcription and protein synthesis rates (Creanor and Mitchison, 1982; Elliott, 1983; Elliott and McLaughlin, 1979; Elliott et al., 1979; Fraser and Nurse, 1978, 1979; Padovan-Merhar et al., 2015; Sun et al., 2020; Zhurinsky et al., 2010) (**Fig. 1B**). The size-dependent scaling at the transcriptional level is thought to ensure constant concentrations of global mRNA, rRNA and tRNA, which then result in the scaling of protein synthesis with cell size since protein and mRNA turnover rates are thought to be largely independent of cell size (Padovan-Merhar et al., 2015; Zhurinsky et al., 2010). Consistent with the primary importance of transcriptional scaling, the consequences of excessive growth are clearly linked to a DNA-templated process because diploids grow to approximately twice the size of haploids before exhibiting defects associated with excessively large size (Neurohr et al., 2019; Zhurinsky et al., 2010).

While the scaling of transcription with cell size is likely important for maintaining efficient biosynthesis, its mechanistic origin is mysterious because the template DNA content increases abruptly in S phase while transcription increases continuously as a cell grows, even among cells with the same amount of DNA such as cells in the G1 phase of the cell cycle. To explain this phenomenon several researchers have proposed models based on a 'limiting factor' (Lin and Amir, 2018; Marguerat and Bahler, 2012; Padovan-Merhar et al., 2015; Sun et al., 2020; Zhurinsky et al., 2010). In these models, there is a factor which is limiting for transcription and that increases in amount in proportion to cell size so that it's concentration is constant. Crucially, the amount of this limiting factor, rather than its concentration, determines the global transcription rate because it is titrated against the genome so that, for example, similarly sized diploid and haploid cells will transcribe at a similar total rate (Lin and Amir, 2018; Sun et al., 2020). While such a limiting factor model could in principle underlie the size-scaling of transcription, there has been no direct experimental test of this model and no empirical identification of the requisite limiting factor(s).

Here, we identify RNA polymerase II as a major limiting factor for increasing transcription with cell size. In contrast, other components of the transcriptional machinery are only minimally limiting and the chromatin environment into which the transcriptional machinery is loaded is effectively constant regardless of cell size. Importantly, our quantitative ChIP-seq measurements show that RNA polymerase II loading does not scale in direct proportion with cell size, which is inconsistent with previously proposed titration models. Instead, our data support a dynamic equilibrium model that predicts transcription rates as a function of the free RNA polymerase II concentration, cell size, and DNA content. Since transcriptional scaling is not precisely proportional to cell size, achieving a size-independent mRNA concentration requires an additional feedback mechanism on mRNA decay rates, which we then validate experimentally. Thus, the scaling of mRNA amounts with cell size is driven both by RNA polymerase II dynamics and feedback regulation on mRNA stability.

RESULTS

RNA polymerase II occupancy increases with cell size in budding yeast

Previous observations that bulk transcription rates in budding yeast increase with cell size (Elliott and McLaughlin, 1979) suggests that larger cells should have more RNA polymerase II loaded on their genomes than smaller cells. To test this, we employed a spike-in normalized ChIP-seq methodology to measure RNA polymerase II occupancy (**Fig. 1C**). While conventional ChIP-seq can only resolve differences in the relative binding at different genomic loci within a sample, spike-in normalized ChIP-seq also measures systematic differences in occupancy across the genome between samples (Hu et al., 2015). First, we performed controls that confirmed that this approach is quantitative across a wide dynamic range and is robust to variation in the spike-in mixing ratio and cell density (**Fig. S1**; see methods).

After validating our spike-in normalized ChIP-seq methodology, we then measured global RNA polymerase II occupancy in budding yeast cultures of different average cell sizes, but with similar growth rates (**Fig. 1D & S2A**). This demonstrates that the total amount of RNA polymerase II loaded on the genome increases with cell size (**Fig. 1E**). Similarly, the initiated and elongating RNA polymerase II populations, identified by S5 and S2 phosphorylation in the C-terminal heptapeptide repeats on Rpb1, also increased in larger cells (**Fig. 1E**). Importantly the relative distribution across gene bodies of total, initiated (S5-P) and elongating (S2-P) RNA polymerase II also increases uniformly with cell size (**Fig. 1F**), indicating that the initiation rate of RNA polymerase is likely to be primarily responsible (Sun et al., 2020).

Total and chromatin bound fractions of Pre-Initiation Complex (PIC) subunits increase with cell size

The increase in transcription with cell size has previously been attributed to a hypothesised limiting factor model (Lin and Amir, 2018; Padovan-Merhar et al., 2015; Sun et al., 2020; Zhurinsky et al., 2010). In such models, the amount of a factor, which is is limiting for transcription, increases in proportion to the cell size resulting in increased binding to the genome. This drives more polymerase loading and transcriptional initiation in larger cells. Thus, in principle, any factor essential for transcription could serve such a role.

While the idea of a limiting factor has been repeatedly discussed in the literature, this model has not been experimentally tested and crucially there has been no empirical identification of the requisite limiting factor(s). To address this question, we first analysed the protein scaling of multiple subunits of the RNA polymerase II complex, as well as components of each of the general transcription factor complexes, which together constitute the RNA polymerase II pre-initiation complex (PIC). Consistent with the global protein scaling trends the protein amounts of all pre-initiation complex components that we examined increased in close proportion to cell size (**Fig. 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) (Kustatscher et al., 2014) for work in yeast (**Fig. S3A-B**) and analysed protein-chromatin associations in small and large cells by mass spectrometry. This clearly reveals that not only are all subunits of RNA polymerase II recruited to the genome more in larger cells, but so are all the initiation and elongation factors that regulate and associate with RNA polymerase II (**Fig. 2B & S3E**). We note that only proteins first validated as being enriched on chromatin were analysed here to minimise potential confounding background effects (**Fig. S3C-D**). 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.

RNA polymerase II is a major limiting sub-complex of the PIC

To directly test for limiting factors we next performed a local perturbation to the nuclear amounts of each PIC component using the anchor away approach to conditionally deplete a targeted factor from the nucleus (Haruki et al., 2008). We constructed heterozygous diploids where one of the two alleles of a given PIC subunits is FRB tagged, allowing us to rapidly and conditionally deplete ~50% of it from the nucleus upon rapamycin treatment (**Fig. 2C**). To determine how this ~50% depletion affects RNA polymerase II loading, we then performed anti-Rpb1 spike-in normalized ChIP-seq. We observed that each of the general transcription factor complexes are only very partially limiting such that a 50% reduction in their nuclear amounts results in at most a 5-10% reduction in global RNA polymerase II occupancy (**Fig. 3D**). However, when we depleted 50% of the RNA polymerase II complex from the nucleus we observed a far larger ~40% reduction in occupancy (**Fig. 3D**), suggesting that RNA polymerase II is a major limiting factor. Moreover, quantification of only the non-depleted allele in a *RPB1-FLAG/RPB1-FRB* strain confirmed the 40% value and limiting nature of RNA polymerase II complex amounts (**Fig. S4D-E**). Thus, our conditional depletion analysis shows RNA polymerase II is a major limiting sub-complex while PIC subunits are only minimally limiting.

Transient overexpression is sufficient to increase RNA polymerase II loading

Having established that RNA polymerase II is a limiting factor for its own recruitment to the genome, we set about to test if an increase in the amounts of the 12 subunit RNA polymerase II complex would be sufficient to increase polymerase loading. We first optimised the overexpression conditions of individual subunits by comparing the overexpressed protein amounts (*e.g., Tetpr-RPB1-mNeonGreen*) with that of their endogenously expressed counterpart (*e.g., RPB1pr-RPB1-mNeonGreen*) (**Fig. S5**). This allowed us to define

an induction regime in which all tested subunits are 2-3 fold overexpressed from an anhydrotetracycline-inducible promoter (*TetPr*) in a 45 minute window (**Fig. 2E**). Next, we engineered a single yeast strain in which all 12 subunits can be simultaneously and conditionally overexpressed (*TetPr-RPB1-12*). In this strain, we observed a robust increase in global RNA polymerase II occupancy across the genome after anhydrotetracycline treatment while there was no equivalent change in a control strain expressing mNeonGreen from the *TetPr* integrated at the same loci (**Fig. 2F-G**). Thus, RNA polymerase II is a limiting component of the transcriptional machinery and increases or decreases in RNA polymerase II concentrations alone are sufficient to change RNA polymerase II loading on the genome.

RNA polymerase II binding is driven by dynamic equilibrium kinetics

Having established RNA polymerase II as a major limiting component of the PIC we sought to examine more quantitatively how limiting RNA polymerase II can function to scale transcription with cell size. Prior theoretical models had proposed that a limiting factor(s), which we have identified as RNA polymerase II itself (**Fig. 2**), is titrated against the genome (Lin and Amir, 2018; Sun et al., 2020). Crucially, in such models the amount of the limiting factor, rather than its concentration, is critical because nearly all molecules are engaged on the genome.

To test the genome-titrated limiting factor model, we first sought to simplify our experimental set up and measure RNA polymerase II occupancy in cells of different sizes but with a fixed DNA amount. To do this, we isolated small G1 cells by centrifugal elutriation and arrested them in G1 using a temperature sensitive Cdk1 allele (*cdc28-13*). By arresting cells for increasing amounts of time, we generated populations of increasing cell sizes but with the same 1N DNA content (**Fig. 3A-B**). As expected, there is a clear increase in genome-wide RNA polymerase II occupancy with cell size within the physiological cell size range (<100 fL) (**Fig. 3B**). However, the scaling of RNA polymerase II occupancy clearly deviates from a direct linear proportionality, even around 50 fL, in these G1 arrested cells (**Fig. 3B**). Importantly, these trends are not an artefact of the temperature shift used to block the cell cycle (**Fig. S6A-B**), and are also apparent in the RNA polymerase II occupancy data in our series of asynchronous cell size mutants (**Fig. 1**).

The non-linear increase in RNA polymerase II occupancy rules out a simple titration model so we considered two alternative quantitative models to explain how cell size regulates RNA polymerase II occupancy (**Fig. 3C**). The first model is a modification of the titration model in which, at a specific size threshold, the genome becomes saturated and transcription cannot increase further (Lin and Amir, 2018). The second model is a more general mass action equilibrium where the rate at which the free RNA polymerase II, pol_{free}, associates with the transcription initiation sites on the genome, DNA, is determined by mass action kinetics and assumes that the genome does not saturate ($eq.1\&2$).

$$eq.1: \quad \frac{d[pol_{bound}]}{dt} = k_{on} * [DNA] * \left[pol_{free}\right] - k_{off} * \left[pol_{bound}\right]$$

$$eq.2: \quad \frac{dpol_{bound}}{dt} = k_{on} * DNA * \frac{pol_{free}}{V} - k_{off} * pol_{bound}$$

This is assumed to be near equilibrium (*i.e.,* $\frac{dpol_{bound}}{dt} = 0$) as the time scale of transcription is on the order of 1 minute, while cell growth is on the order of 1 hour. Finally, using the fact that the total RNA polymerase II amount is proportional to the cell size ($pol_{bound} + pol_{free} = c * V$), where c is a constant, and that the nuclear volume is approximately 10% of cell volume (Jorgensen et al., 2007; Neumann and Nurse, 2007), we can solve for the amount of DNA bound RNA polymerase II ($eq.3$).

$$eq.3: \quad pol_{bound} = \frac{c\,V}{1 + \frac{k_{off}\,V}{10\,k_{on}\,DNA}}$$

When fit to the data, both the titration-saturation model and the dynamic equilibrium model are consistent with our global RNA polymerase II occupancy measurements (**Fig. 3D**).

While both models can account for the RNA polymerase II occupancy, they make important additional predictions that distinguish them. First, the titration model predicts that nearly all nuclear RNA polymerase II should be engaged on the genome resulting in a near-negligible free nucleoplasmic pool. As such, 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 ($^1/_V$) (**Fig. 3C**). In constant, the dynamic equilibrium model predicts that a significant fraction of RNA polymerase II should be free at all times. Using the fit of the dynamic equilibrium model to the global trends in our ChIP data (**Fig. 3D**) we estimated that approximately 50% of nuclear RNA polymerase II should be freely diffusing in the nucleoplasm of a 30 fL G1 cell and that the bound fraction should then decrease continuously as size increases.

Thus, to distinguish between these models, we sought to measure the fraction of bound and free RNA polymerase II in single cells using single-molecule imaging. We imaged and tracked single molecules of Rpb1-HALO and estimated their radius of gyration. Freely diffusing molecules diffuse rapidly across the nucleus and have a large radius of gyration, while chromatin-bound molecules are relatively immobile and have a small radius of gyration. Consistent with expectations, the majority of H2B histone molecules are tightly bound and the vast majority of free nuclear mCitrine-NLS fluorescent molecules do not associate with chromatin (**Fig. 3E-F**). Crucially, about half of the nuclear RNA polymerase II subunit Rpb1 is freely diffusing and the other half is bound (**Fig. 3E-F**). Moreover, the bound fraction of Rpb1 in individual G1 cells decreases gradually with cell size in agreement with the prediction of our dynamic equilibrium model (**Fig. 3G**). In contrast, these data show a very poor correspondence to the prediction of the titration-saturation model (**Fig. 3G**). We emphasize that the two free parameters for the equilibrium model were determined by fitting to the anti-Rpb1 ChIP data and not adjusted in response to the single-molecule imaging data.

A second prediction of the titration model that differs from our dynamic equilibrium model is that at larger cell sizes regions of the genome should become saturated with RNA polymerase II. If we assume the saturation occurs because genes get full of polymerase, it should be the genes with the initially higher RNA polymerase II occupancy that get saturated first and should therefore decrease in their relative expression compared to initially lower expressed genes. However, when we examined cells below (~30 fL) and above (~150 fL), the

hypothetical saturation transition, we see that genes behave similarly regardless of their initial RNA polymerase II occupancy (**Fig. 3H-I**).

Taken together, our quantitative ChIP-seq data and single molecule tracking experiments are inconsistent with a simple titration-saturation model. Instead, we propose that the global occupancy of limiting RNA polymerase II is primarily determined via a simple mass action dynamic equilibrium. Because a significant fraction of RNA polymerase II is both bound and unbound, as a cell grows and synthesises more RNA polymerase II this first enters the free fraction increasing the free nucleoplasmic concentration. This 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 resulting in a sub-linear scaling of RNA polymerase II occupancy that becomes more pronounced as cells get larger.

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 rate of loading RNA polymerase II are determined predominantly by changes in the concentration of free nucleoplasmic RNA polymerase II. One direct prediction of this is that the global state of chromatin should not become more or less permissive to polymerase recruitment at different cell sizes. To test this, we adapted the dual-enzyme single-molecule footprinting (dSMF) assay (Krebs et al., 2017) to measure chromatin accessibility in yeast. Briefly this involves treating nuclei with a combination of CpG and GpC methyltransferases so that more accessible DNA is methylated and protected DNA is and not methylated. Differences in methylation patterns are then read out by sequencing and normalized to an external spike-in (**Fig. 4A**). When we performed dSMF on cells ranging from ~30 fL to ~90 fL we observed no major changes in chromatin accessibility between the cells of different sizes (**Fig. 4B & S7A**). Consistent with this, we also do not see changes in histone occupancy measured by spike-in normalized ChIP-seq (**Fig. 3B & S6B**).

Although the physical accessibility of chromatin is not changing with size, the chromatin landscape could still be modified via a changing pattern of histone modifications. To test possibility, we measured a panel of histone marks implicated in transcriptional processes by spike-in normalized ChIP-seq. The majority of modifications we measured were constant as a function of cell size. However, four modifications associated with active transcription did increase significantly in larger cells: H3K4me2, H3K4me3, H3K79me3, H3K36me3 (**Fig. 4C-E**). These di/tri-methyl marks are all modifications known to be deposited downstream of transcriptional initiation or elongation, *i.e.*, their deposition depends on the recruitment of RNA polymerase II (Krogan et al., 2003; Ng et al., 2003; Santos-Rosa et al., 2002; Xiao et al., 2003). This is consistent with our prior observation that more RNA polymerase II is loaded onto the genome in larger cells. We therefore conclude that the chromatin landscape, both in terms of accessibility and histone modifications, is largely invariant with cell size with the exception of modifications that occur as a downstream consequence of increased RNA polymerase II recruitment in larger cells.

mRNA decay feedback compensates for non-linear transcriptional scaling

Next, we considered the implications of our RNA polymerase II measurements for the global scaling of mRNA amounts. Our measurements and model for limiting RNA polymerase II show that this scaling is not proportional with cell size and instead displays a non-linear trend that deviates from the linear proportionality (**Fig. 1 & 3**). This has two possible consequences: either mRNA amounts follow the same trends and mRNA concentrations decline in larger cells or there is a compensatory mechanism that stabilizes mRNA, as size increases, to buffer the non-linear transcriptional scaling and maintain constant mRNA concentrations (**Fig. 5A**).

To test for the presence of size-dependent feedback on mRNA stability, we measured the levels of endogenous transcripts following transcriptional repression. We utilized two distinct treatments – in one we added methionine to the media to shut off the *MET* genes (Rouillon et al., 2000), and in the second we added glucose to shut off *GAL* genes. We then measured mRNA levels of these genes, following shutoff, in the different cell size mutants (**Fig. 5B**) to estimate the size-dependency of mRNA decay. This clearly showed that for all 5 genes mRNA degradation decreases in larger cells (**Fig. 5B & S8B-D**). We then asked how this compared to the decay rate required to maintain a constant mRNA concentration given the size-dependence of the RNA polymerase II activity we previously measured (**Fig. 1**). Strikingly, the relative change in mRNA decay rate for all 5 transcripts we examined aligned well with that predicted by the mRNA decay compensation model (**Fig. 5D**). Thus, mRNA decay rates are precisely adjusted to compensate for the imperfect scaling of RNA polymerase II activity with size to ultimately ensure mRNA concentrations are kept constant as cell size increases.

Although we do not know the precise mechanism for this compensation, we believe that it is due to a feedback on global mRNA concentrations. This is because we observed a similar decrease in mRNA turnover rates in cells subject to an acute RNA polymerase II 50% nuclear depletion (**Fig. 5E-F, S8E & S9**). In this situation cell size is not changing indicating that the compensatory adjustments to mRNA decay rates are not determined directly by cell size but instead as part of a feedback mechanism that operates to buffer global mRNA concentrations.

DISCUSSION

How cells scale global mRNA amounts to remain in proportion to cell size during growth has been a long-standing and fundamental question in cell biology (Elliott, 1983; Elliott and McLaughlin, 1979; Fraser and Nurse, 1978, 1979; Marguerat and Bahler, 2012). It was previously thought that the size-scaling of mRNA was due to mRNA synthesis rates increasing in direct proportion to cell size and was proposed to be due to an unknown limiting transcription factor whose amount increase in proportion to cell size and is titrated against the genome (Marguerat and Bahler, 2012).

Here, we empirically identified RNA polymerase II as a limiting factor by showing that RNA polymerase II loading onto the genome was sensitive to both acute increases and decreases in its concentration, whereas other components of the transcriptional initiation machinery were not significantly limiting (**Fig. 2**). However, our single-molecule imaging revealed that RNA polymerase II is not titrated against the genome as there is a significant pool of RNA polymerase II that is freely diffusing in the nucleoplasm. Moreover, the amount of

active RNA polymerase II did not scale in direct proportion with cell size so that larger cells exhibited systematically larger deviations from the amount of limiting RNA polymerase II engaged on the genome predicted by the titrated factor model. Instead we propose a new dynamic equilibrium model in which the loading of RNA polymerase II on the genome is determined by mass action kinetics with a defined size-independent on-rate (k_{on}) (**Fig. 3**). The size-independence of $k_{on}$is consistent with our measurements showing that there are no size-dependent changes in chromatin accessibility or histone modifications thought to be upstream of transcription (**Fig. 4**). In this model, as a cell grows, newly synthesised RNA polymerase II increases the free nucleoplasmic concentration driving more polymerase onto the genome thereby establishing a new equilibrium at the increased size. However, this increase is not in direct proportion to cell size resulting in a deviation from linear scaling that becomes increasingly pronounced as cells get larger. Our two-parameter dynamic equilibrium model accounts for this precise relationship between the transcription rate and cell size, and predicts the size-dependency of the fraction of RNA polymerase that is bound to the genome and that which is freely diffusing in the nucleus (**Fig. 3**).

While our dynamic equilibrium model can account for size-dependent increases in transcription, it raised the question as to how the global mRNA concentration could be constant when transcription scales sub-linearly with cell size. If mRNA half-lives were independent of cell size as previously thought (Padovan-Merhar et al., 2015; Zhurinsky et al., 2010), this would lead to a size-dependent decrease in mRNA concentration. However, we found that cells avoid this fate via a feedback that regulates mRNA half-lives to compensate for the sublinear size-dependence of transcription (**Fig. 5**). Taken together, our results support a model where scaling mRNA amounts with cell size is due to mass action kinetics of limiting RNA polymerase II to initiate transcription coupled to a feedback mechanism operating on mRNA stability (**Fig. 6**).

The presence of feedback operating on mRNA stability is highly reminiscent of several previous observations in both yeast (Baptista et al., 2017; Plaschka et al., 2015; Rodriguez-Molina et al., 2016; Schulz et al., 2014; Sun et al., 2012; Warfield et al., 2017) and human cells (Berry et al., 2021; Helenius et al., 2011; Slobodin et al., 2020) where mutations affecting transcription rates are buffered by a compensatory decrease in mRNA decay rates. This is most dramatically seen in a recent study that examined transcription rates and global RNA concentrations in a genome-wide siRNA screen. While perturbations affecting transcription were numerous, these did not result in changes to RNA concentration (Berry et al., 2021). Thus, both growing larger and perturbation to transcriptional activity elicit a similar feedback effect to buffer mRNA concentrations.

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 in which this mRNA stability feedback is critical. While it is not known what molecular mechanisms underlying this compensatory feedback, one intriguing possibility is suggested by the fact that the incorporation of mRNA into polysomes increases its stability (Chan et al., 2018; Huch and Nissan, 2014; Schwartz and Parker, 1999). Excess mRNA would then increase the non-ribosome bound fraction to drive increased turnover and reduction of excess mRNA. This suggests that ribosomes *per se*, whose concentration is independent of cell size, could contribute to the molecular basis of the feedback

mechanism ensuring a size-independent, efficient stoichiometry between the two most critical components of protein translation.

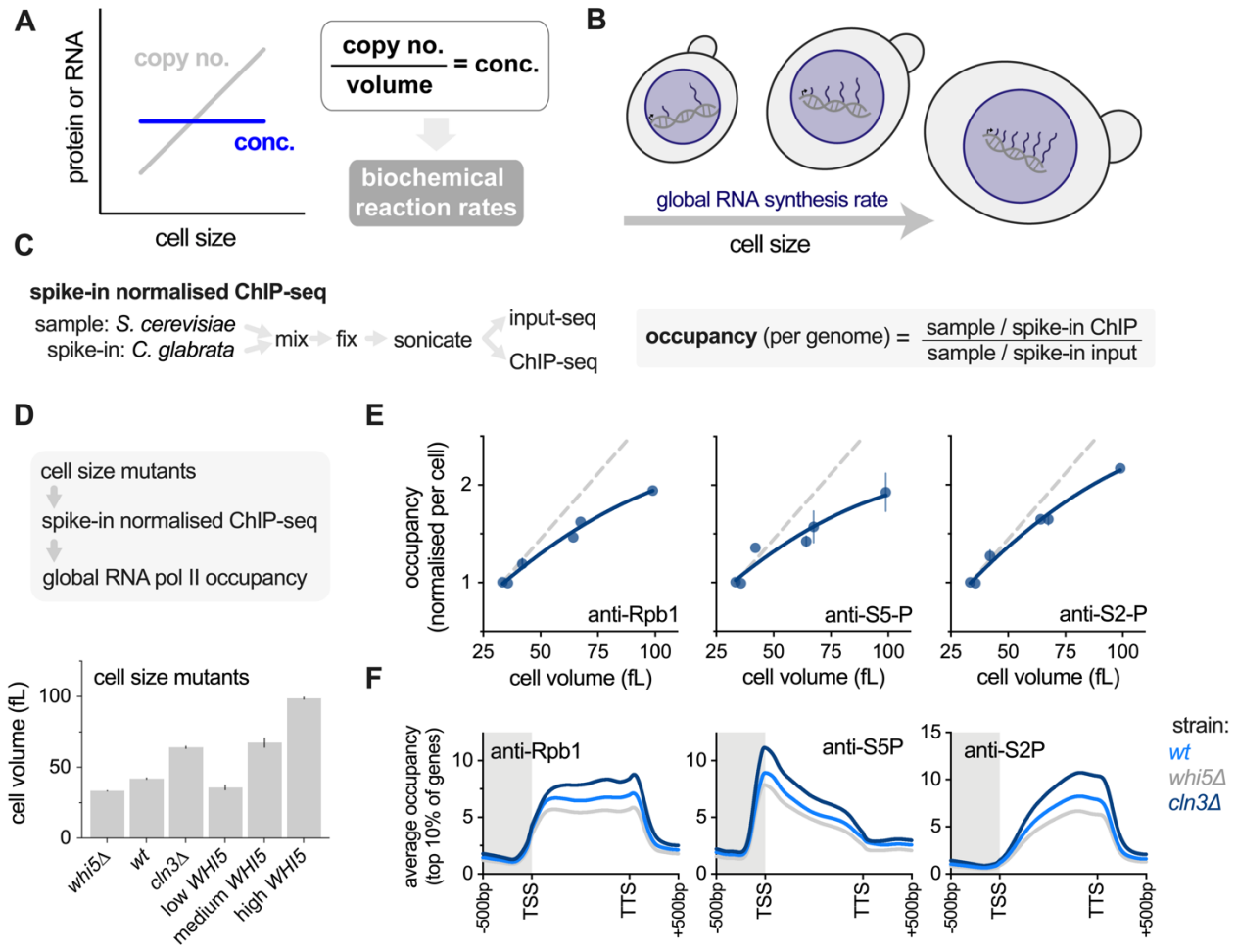

Figure 1 I More RNA polymerase II is loaded onto the genome with increasing cell size

See also Fig. S1-S2

(**A-B**) Schematic showing that (**A**) macromolecule copy number increases in proportion to cell size to keep RNA and protein concentrations constant and that (**B**) global RNA synthesis rates increase with cell size.

(**C**) Schematic illustrating the workflow for spike-in normalized ChIP-seq and how the global relative occupancy is calculated (see Methods and Fig. S1 for details).

(**D**) Mean cell volume determined by Coulter counter of cell size mutants used to analyse the RNA polymerase II subunit Rpb1 occupancy in (E-F). Low, medium, and high *WHI5* expression is from a beta-estradiol responsive promoter induced with 0mM, 10mM or 30mM beta-estradiol. Cultures are grown to steady state before collection and beta-estradiol concentration has no major impact on cell growth (Fig. S2A-B).

(**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 (D) plotted as a function of cell size. Each point shows the mean (±range) of two biological replicates.

(**F**) 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 *WT, whi5Δ, and cln3Δ* cells. See Fig. S2C for low, medium and high *WHI5* conditions. Mean of two biological replicates is shown.

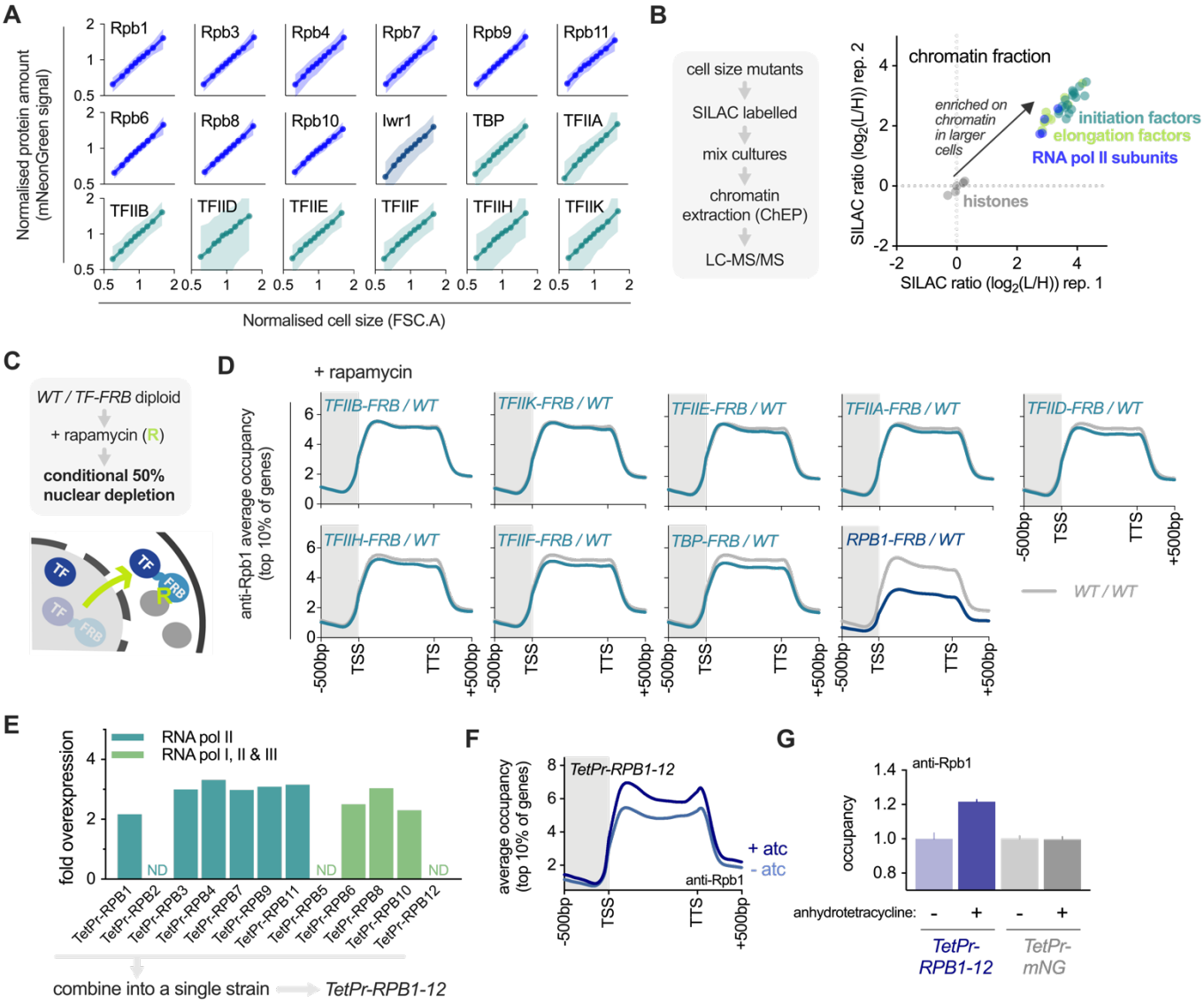

Figure 2 | RNA polymerase II is a major limiting component of the transcriptional machinery

See also Fig. S3-S5

(**A**) Protein amount (mNeonGreen signal) plotted against cell size (forward scatter) for subunits of the RNA polymerase II transcriptional machinery determined by flow cytometry. The mean (±SD) is shown for each cell size bin.

(**B**) Chromatin association of transcriptional machinery measured by mass spectrometry of chromatin fractions. SILAC-labelled cells of different sizes (Fig. S3E) were mixed and then chromatin was extracted by ChEP and analysed by LC-MS/MS (see methods for details and see Fig. S3B for ChEP controls). Each axis shows an independent biological replicate for the normalized SILAC ratio of chromatin association between large (L) and small (H) cells. RNA polymerase II subunits, initiation factors, and elongation factors are shown in comparison to histones. Only chromatin-enriched proteins are shown (Fig. S3C-D).

(**C**) Schematic illustrating the 50% conditional depletion of RNA polymerase II Pre-Initiation Complex (PIC) subunits to test for dosage-limitation. A PIC subunit (TF) is tagged with the FRB domain (Haruki et al., 2008) then crossed with a *WT* haploid strain to form a heterozygous diploid where one of the two alleles is FRB tagged, allowing for conditional depletion of 50% of that subunit from the nucleus.

(**D**) Average Rpb1 occupancy across gene bodies of the top 10% of genes measured by spike-in normalized ChIP-seq in wild-type diploid cells (*WT/WT*) or diploids where 50% of the indicated PIC sub-complex was depleted from the nucleus upon

rapamycin treatment. Rpb1-FRB anchor away results in efficient co-depletion of Rpb3 (Fig. S4A) indicating the whole RNA polymerase II complex is efficiently co-depleted. Nuclear depletion is efficient and near-complete (i.*e.*, ~50%) because ChIP against total Rpb1 and ChIP against a FLAG tagged non-depleted allele give a similar result (Fig. S4C-E).

(**E**) The fold overexpression of individual RNA polymerase II subunits from the TetPr after 45 minutes of anhydrotetracycline (atc) treatment. Expression levels from the TetPr was quantified using C-terminal mNeonGreen tagged proteins measured by flow-cytometry, which was then compared to the endogenously tagged allele of the respective subunit to calculate the fold overexpression (Fig. S5A). This was not determined (N.D.) for *RPB2*, *RPB5* and *RPB12* because tagging resulted in inviable strains. All 12 subunits were then integrated into a single strain (see Methods for details) to construct the *TetPr-RPB1-12* strain used in (F-G). 45 minute atc treatment results in approximately 2-3x overexpression of each subunit.

(**F**) Average Rpb1 occupancy across the gene bodies of the top 10% of genes for the *TetPr-RPB1-12* strain before (-atc) or 45 minutes after (+atc) simultaneous overexpression of all RNA polymerase II subunits.

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

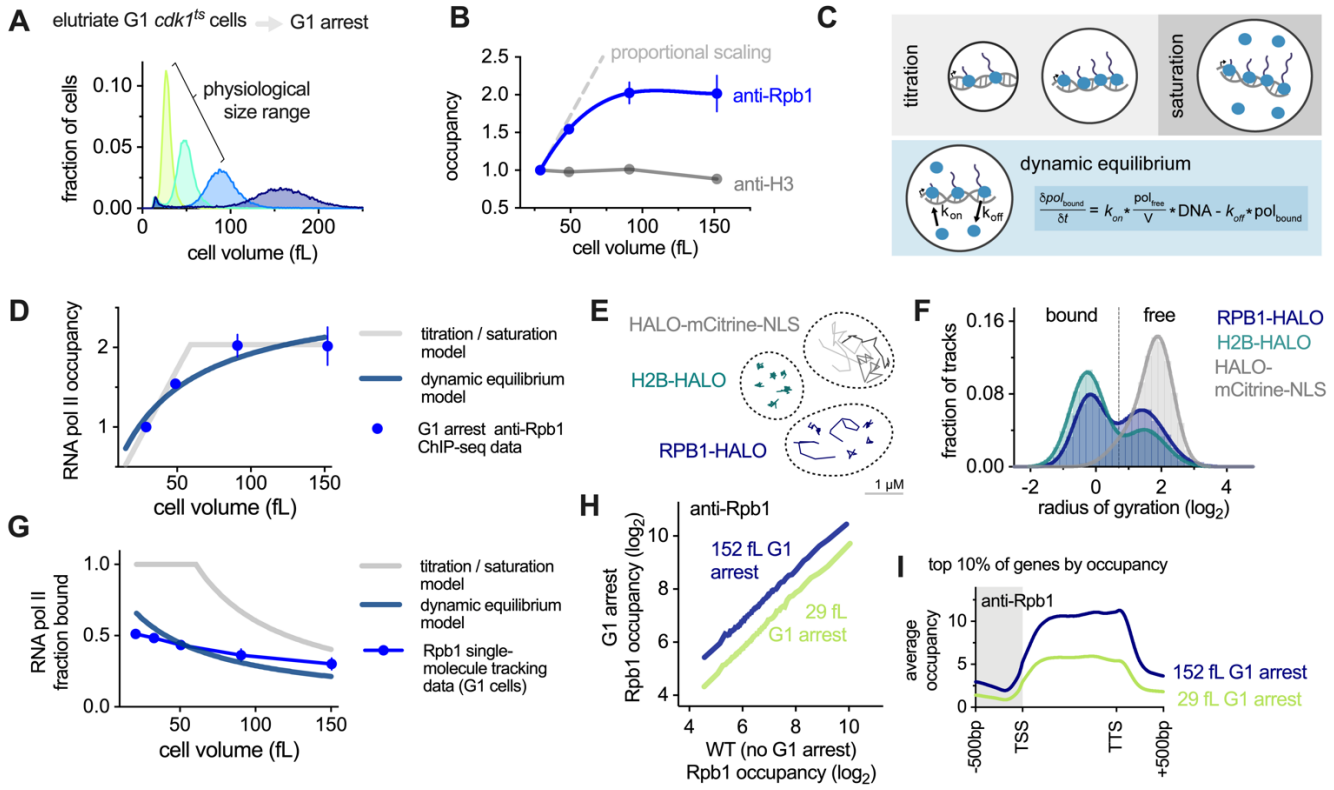

Figure 3 I A dynamic equilibrium model predicts how cell size determines the recruitment of RNA polymerase II to the genome

See also Fig. S6

(**A**) Cell volume distributions determined by Coulter counter for G1 cells of different sizes that were used to analyse Rpb1 occupancy in (B, H, and I). Small G1 cells harbouring a temperature sensitive *CDK1* allele (*cdc28-13*) were collected by centrifugal elutriation and then arrested in G1 at the restrictive temperature for increasing time to generate populations of G1 cells of increasing size, which were then collected for spike-in normalized ChIP-seq.

(**B**) Rpb1 (anti-Rpb1) and Histone 3 (anti-H3) occupancy per genome plotted as a function of cell size in the indicated populations of G1 cells shown in (A). Each point shows the mean of two biological replicates (±range). Dashed grey line shows proportional scaling with cell size – *i.e.*, twofold increase in cell size = twofold increase in occupancy. A control for the effects of the temperature shift is plotted in Fig. S6B.

(**C**) Schematics illustrating the previously proposed titration/saturation model, and the dynamic equilibrium model proposed here.

(**D**) The global RNA polymerase II occupancy in (B) compared with the fit of two different mathematical models. Both models are in close agreement with the average global RNA polymerase II data.

(**E**) Example single molecule tracks in representative nuclei of the indicated genotype.

(**F**) Histogram of the radius of gyration determined by single molecule imaging for HALO-mCitrine-NLS fusion proteins (n = 8057 tracks), Htb1 (n = 35978 tracks), and Rpb1 (n = 16008 tracks). The bimodal histogram of Rpb1 radii of gyration indicate chromatin-bound and freely diffusing fractions.

(**G**) The fraction of bound Rpb1 molecules plotted as a function of cell size. Only G1 cells were analyzed. The mean for each cell size bin is plotted. Data shown are a combination of *WT* (n=182 cells) and *cln3Δ* (n=87 cells) cells. *cln3Δ* were used to increase the range of G1 cell sizes.

(**H**) Rpb1 occupancy in G1 arrested cells of different sizes (y-axis) plotted against asynchronous *WT* cells subjected to the same temperature shift (x-axis). Genes were binned (n=100 bins) according to Rpb1 occupancy in *WT* and the bin means are shown.

(**I**) Average Rpb1 occupancy across the gene bodies for the top 10% of genes, based on RNA polymerase II occupancy in *WT*, in the indicated for populations of G1 cells (29 fL or 152 fL) shown in (A). Mean of two biological replicates is shown.

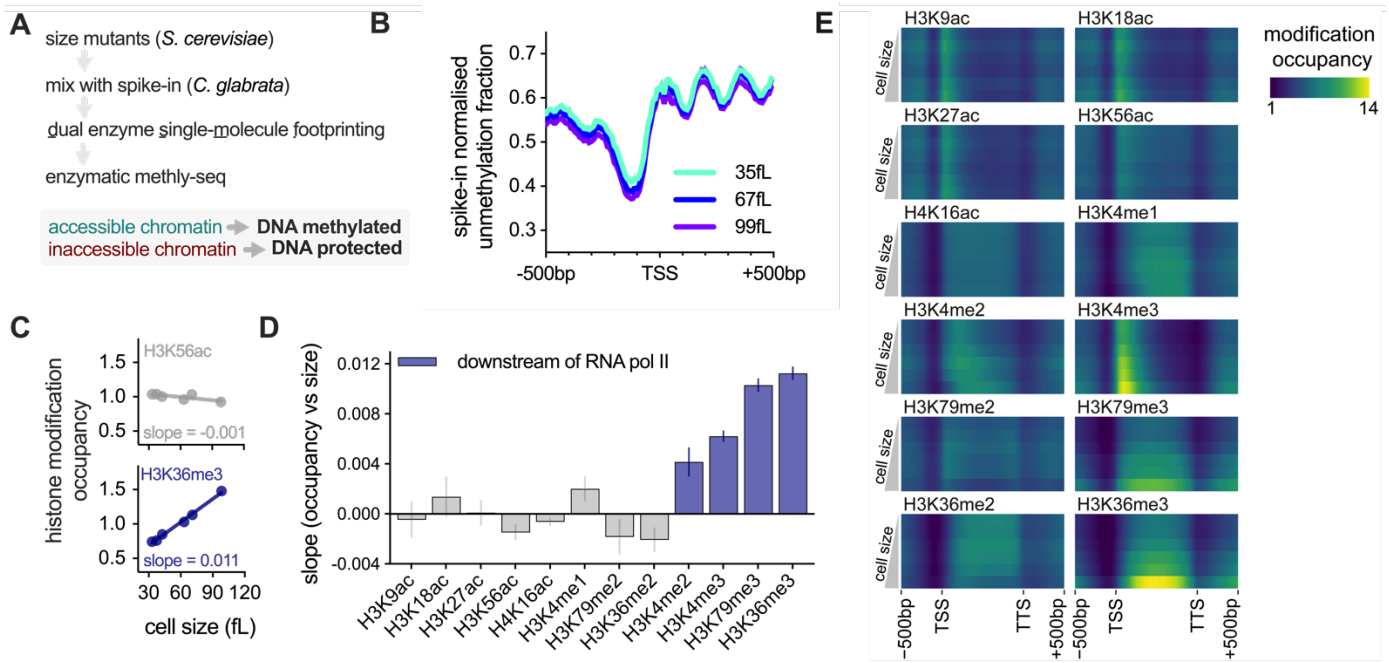

Figure 4 | The upstream chromatin environment does not show major changes with cell size

See also Fig. S7

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

(**B**) Mean unmethylated DNA fraction plotted around transcriptional start sites (TSS) following dSMF in cell size mutants (low *WHI5* = 35 fL, medium *WHI5* = 67 fL, high *WHI5* = 99 fL; Fig. S7A). Unmethylated DNA corresponds to inaccessible chromatin and methylated DNA corresponds to accessible chromatin. Mean (±range) is plotted.

(**C-E**) Histone modification occupancy measured in the different cell size mutants shown in Fig. 1D.

(**C**) The occupancy for two example modifications (top: H3K56ac; bottom: H3K36me3) plotted against cell size. See also Fig. S7B.

(**D**) The slope (±SE) for the linear fit between histone modification occupancy and cell size. See (C) for example fits for H3K56ac and H3K36me3. Larger slopes correspond to increased modification occupancy with cell size. Modifications shown in blue are reported to be deposited downstream RNA polymerase 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) cell size.

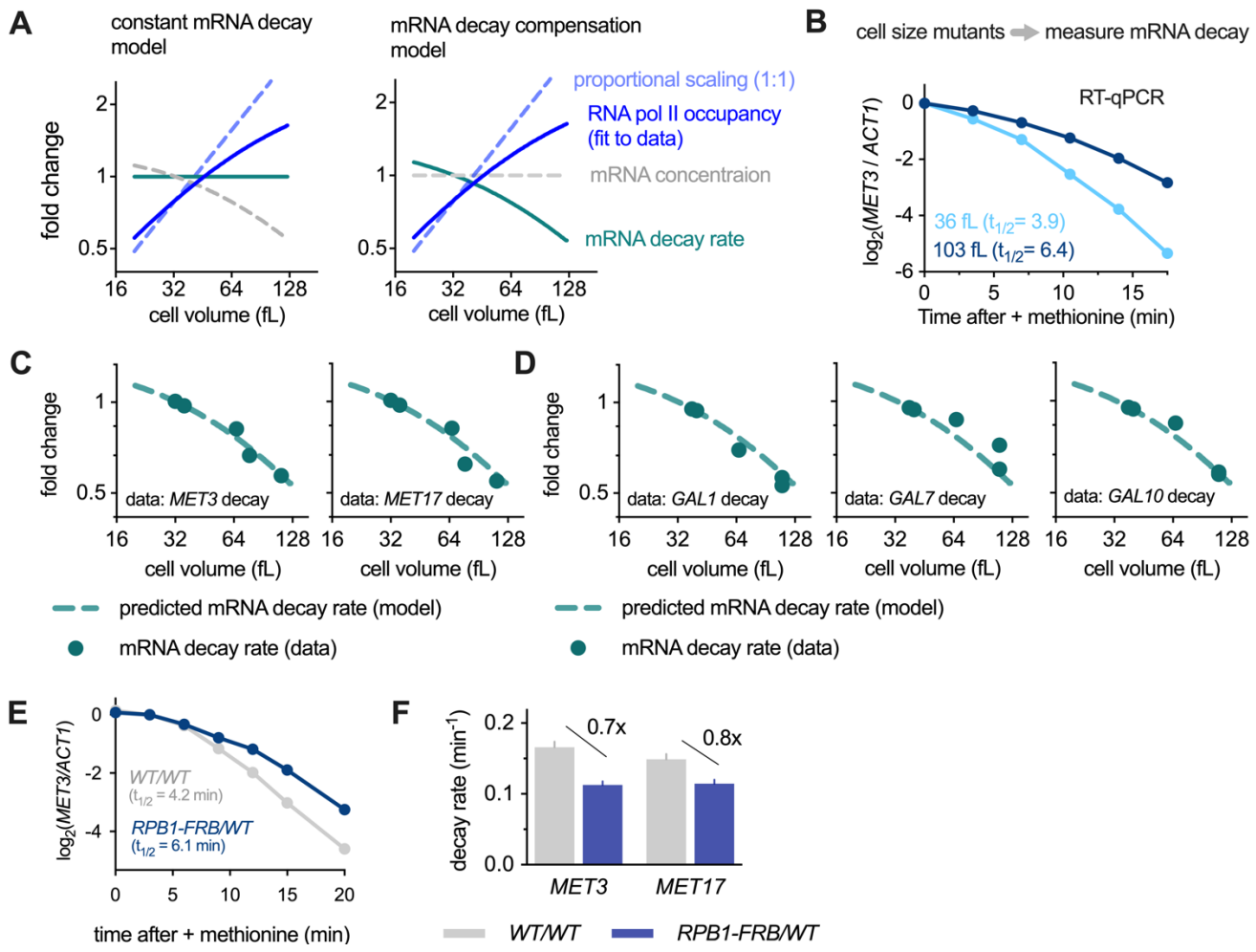

Figure 5 I mRNA decay feedback compensates for the imperfect size-scaling of RNA polymerase II activity

See also Fig. S8&S9

(**A**) Schematics illustrating the predicted trends in global mRNA concentration and mRNA decay rate as a function of cell size that would result from the empirically observed transcription rate for either (left) no compensation or (right) a feedback on mRNA decay rates that buffers mRNA concentrations as global RNA polymerase II activity deviates from the proportional scaling trend. The empirical RNA polymerase II occupancy trend is a fit of the dynamic equilibrium model (Fig. 3C) to the anti-Rpb1 ChIP-seq data of the cell size mutants presented in Fig. 1C.

(**B**) *MET3* mRNA levels relative to *ACT1* mRNA after methionine addition in small (low *WHI5*) and large (high *WHI5*) cells. See also Fig.S8A-C.

(**C**) Fold change in the *MET3* and *MET17* mRNA decay rates following transcriptional shutoff after methionine addition between cells of different sizes (see methods for details). Dashed line indicates the predicted fold-change in the mRNA decay rate from the mRNA decay compensation model shown in (A).

(**D**) Fold change in *GAL1, GAL7* and *GAL10* mRNA decay rate following their transcriptional shutoff after glucose addition between cells of different sizes (see methods for details). Dashed line indicates the predicted fold-change in the mRNA decay rate from the mRNA decay compensation model shown in (A).

(**E**) Wild-type diploid cells (*WT/WT*) or diploid cells (*RPB1-FRB/WT*) where 50% of the RNA polymerase II is depleted from the nucleus upon rapamycin treatment are analysed. *MET3* mRNA levels relative to *ACT1* mRNA were measured following methionine addition 40 minutes after rapamycin treatment. See also Fig S8D.

(**F**) *MET3* and *MET17* decay rate in *WT/WT* or *RPB1-FRB/WT* cells treated as described in (E). Mean (±SEM) of three biological replicates is shown.

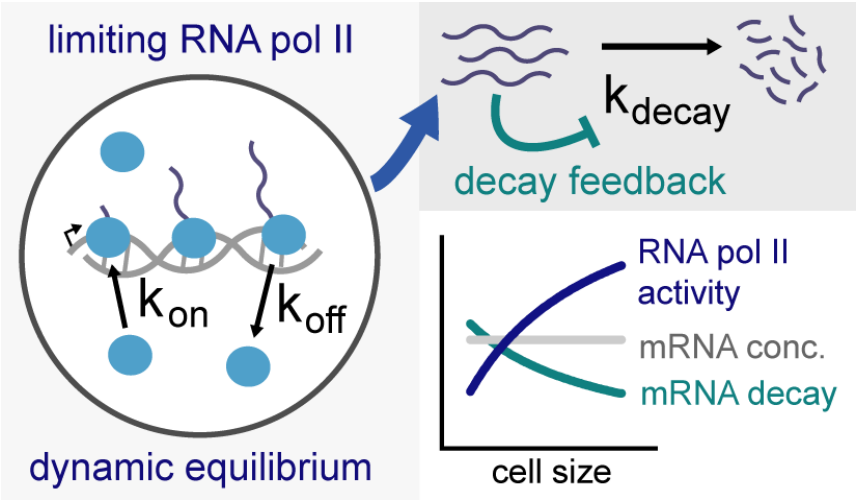

Figure 6 | Summary model
RNA polymerase II as a major limiting factor for coupling global transcription with cell size, whilst other components of the transcriptional machinery are only minimally limiting and the chromatin environment is largely invariant as a function of cell size. A dynamic equilibrium drives more limiting RNA polymerase II onto the genome in larger cells, resulting in a sub-linear increase in transcription rate with cell size. As a consequence of this imperfect sub-linear scaling, an additional feedback mechanism on mRNA decay rates operates to keep mRNA concentrations approximately constant as cell size increases.

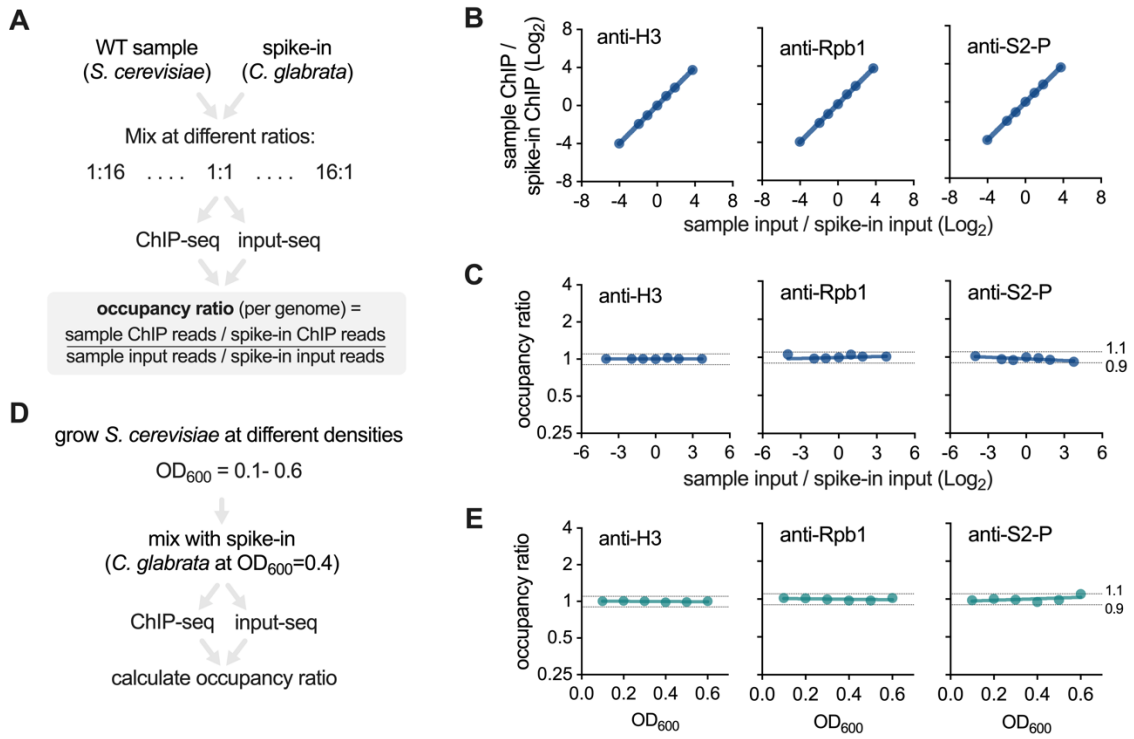

Figure S1 I Controls for spike-in normalized ChIP-seq

Related to Fig 1C

(**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 reflected in proportional changes in the *S. cerevisiae* to *C. glabrata* ChIP ratio (**B**) and in no changes to the input normalized occupancy ratio (**C**). ChIP and occupancy ratios are expressed the fold-change relative to the middle sample. This control is conceptually the same as to the reported for Scc1 ChIP by Hu et al. 2015 (Hu et al., 2015). This control also 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-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 where grown to different cell densities (OD_{600}) mixed with approximately the same cell number of *C. glabrata*. The sample was then fixed and processed for spike-in normalized ChIP-seq. The occupancy ratio for anti-Rpb1, anti-S2P, or anti-H3 is constant as a function of cell densities (**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.

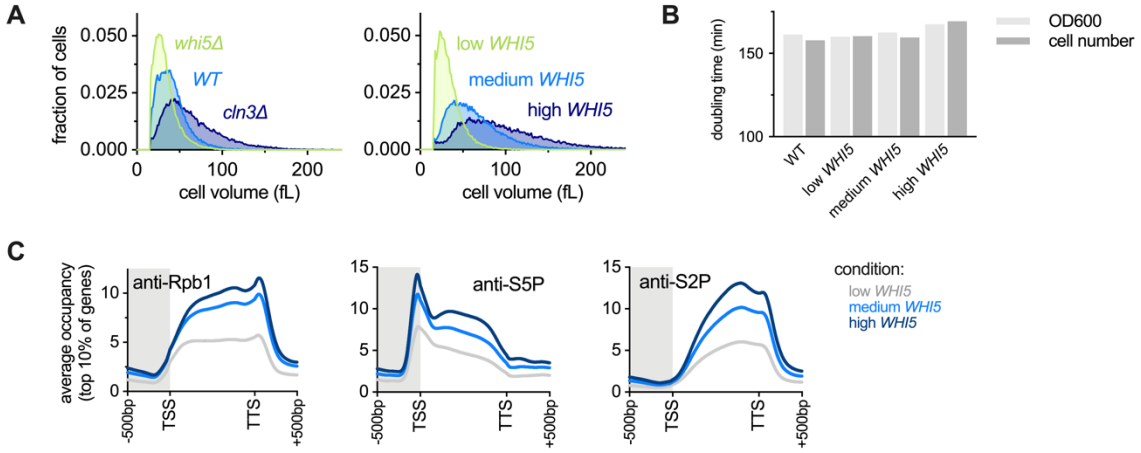

Figure S2 I Cell size mutant and G1 arrest experiments for Rpb1 spike-in normalized ChIP-seq

Related to Fig. 1D-F

(**A**) Cell size distributions determined by Coulter counter for the cell size mutants used for the experiment in Fig. 1D-F. Expression of low, medium, and high *WHI5* expression was induced with 0mM, 10mM, and 30mM beta-estradiol, respectively.

(**B**) Doubling time calculated from either OD_{600} or cell number measurements for cells expressing low, medium, and high levels of *WHI5* induced from a beta-estradiol responsive promoter with 0mM, 10mM, and 30mM beta-estradiol, respectively. Cultures were grown to steady-state before the growth curve measurements were started.

(**C**) 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 Fig.1B.

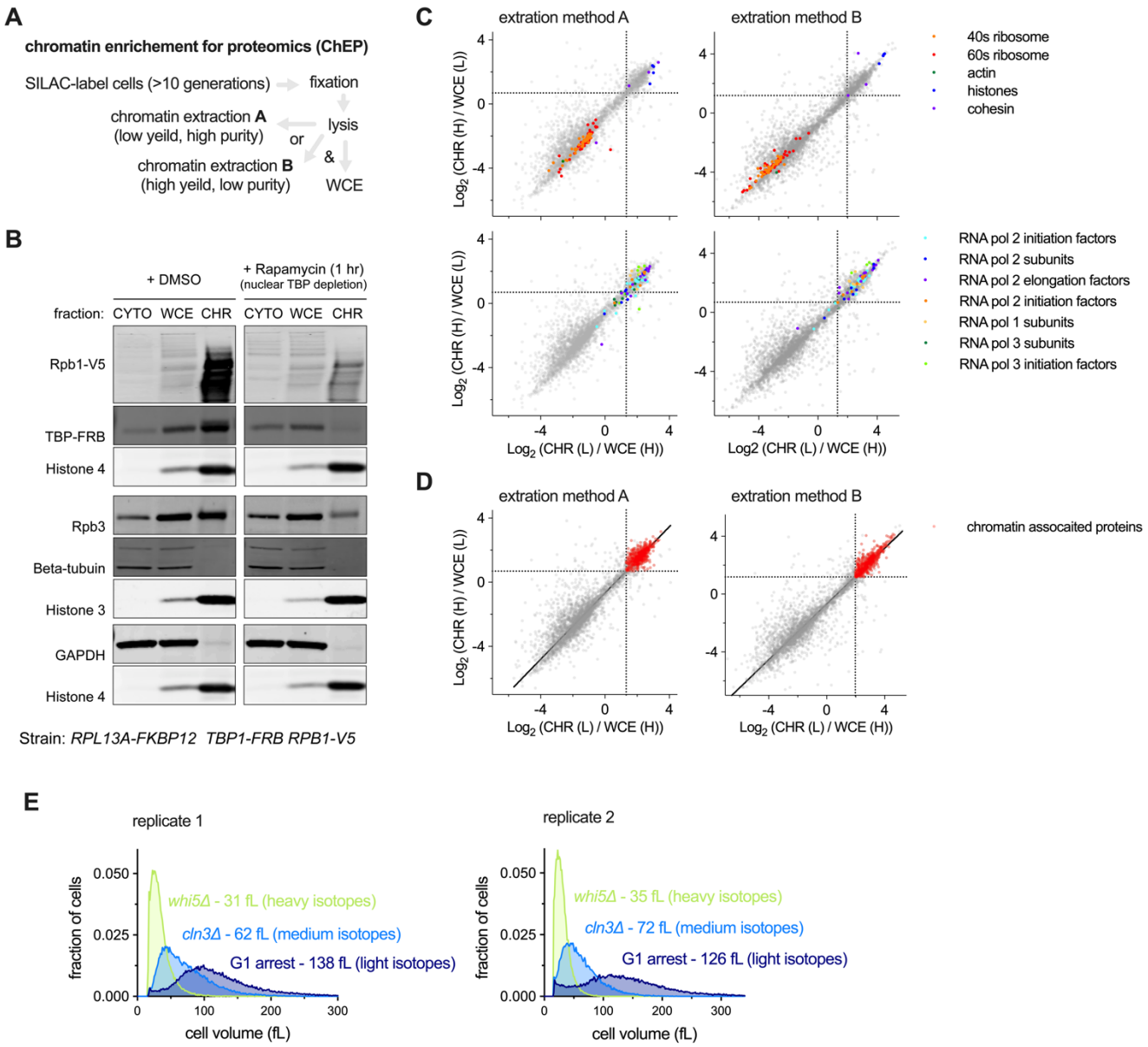

Figure S3 I Chromatin enrichment for proteomics (ChEP) workflow and controls

Related to Fig. 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) are used. See Methods for details.

(**B**) Control experiment showing that ChEP can quantify changes in protein chromatin association. To test this, TPB and RNA polymerase II chromatin association was conditionally prevented using a TBP-FRB strain where TBP is conditionally depleted from the nucleus upon rapamycin treatment thus preventing RNA polymerase II recruitment to the genome. Upon rapamycin treatment for 1 hour, TBP and RNA polymerase II (Rpb1 and Rpb3) have reduced chromatin association measured by ChEP compared to the DMSO control treatment. Thus, ChEP can be used to quantify changes in chromatin association. We note that, under these conditions, RNA polymerase II is still present in the nucleus but is not associated with chromatin (See Fig. S4A), indicating that ChEP is unlikely to be subject to major background due to non-specific chromatin interactions.

(**C-D**) Experiment to identify proteins enriched by ChEP. Cells were SILAC label 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 where mixed and analysed 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). Expected cytoplasmic factors (ribosomes and actin), and expected chromatin-associated proteins (RNA polymerase subunits and co-factors, histones and cohesion) are shown in (**C**) which demonstrates these chromatin extractions enrich for expected chromatin-associated factors. Proteins defined as chromatin associated are shown in red in (**D**). A robust regression line is 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 methods for details).

(**E**) Cell size distributions determined by Coulter counter for the SILAC labelled cultures used in the experiment in Fig. 2B. All strains have a SILAC-compatible genetic background (see methods for details).

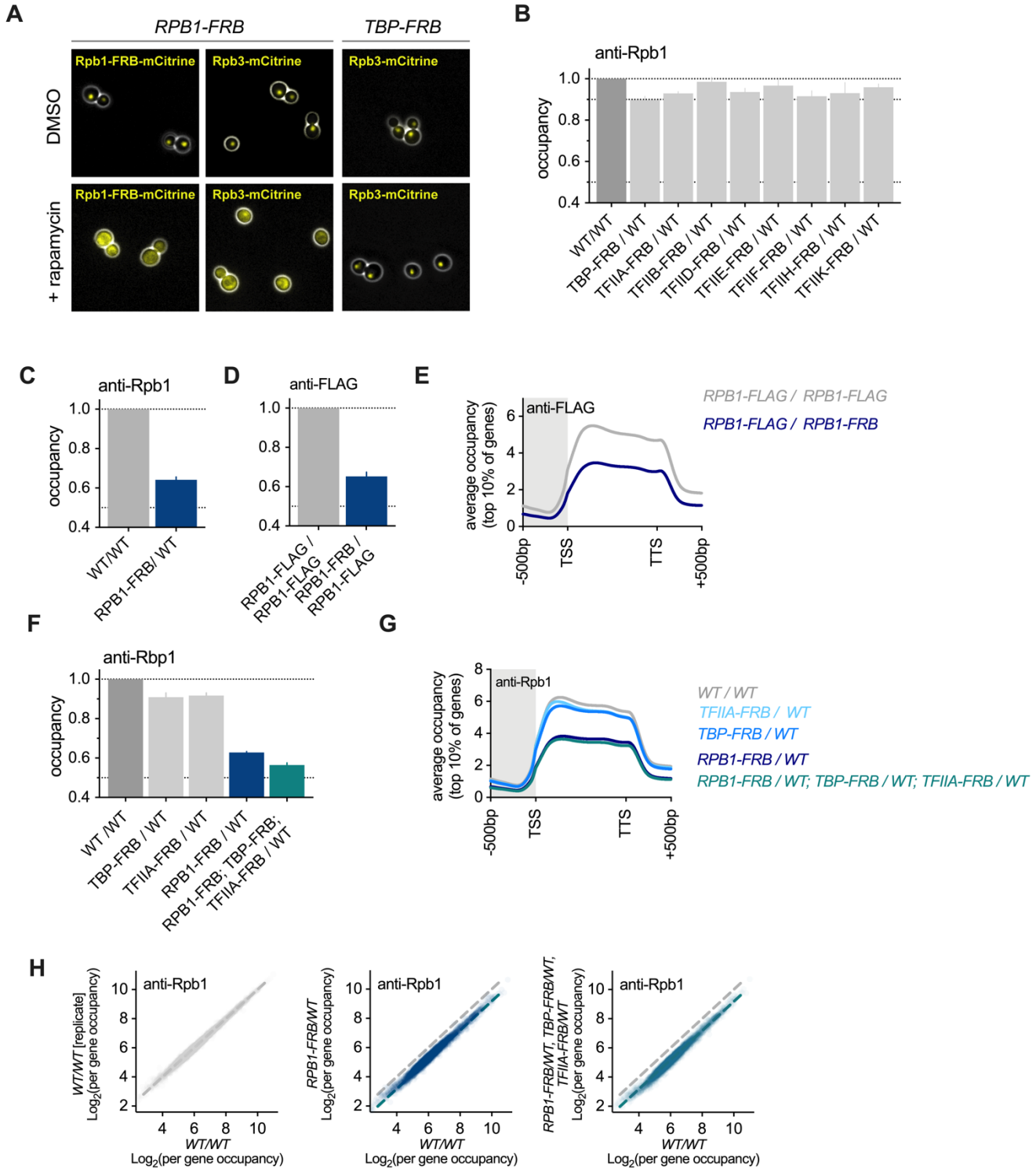

Figure S4 I RNA polymerase II 50% depletion and over-expression experiments supporting data

Related to Fig. 2C-D

(**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

RNA polymerase II complex is likely to be efficiently co-depleted. Treatment of TBP-FRB cells does not alter Rpb3 localization, indicating that different sub-complexes of the Pre-Initiation-Complex do not appear to co-deplete one another.

(**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 RNA polymerase II Pre-Initiation Complex subunit was depleted from the nucleus following a 100 minute 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 Fig. 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 RNA polymerase II was depleted from the nucleus by a 40 minute 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 Fig. 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 RNA polymerase II Pre-Initiation Complex subunits were depleted from the nucleus by a 40 minute rapamycin treatment. Data are also show in an experiment with the simultaneous 50% nuclear depletion TBP, TFIIA, and RNA polymerase II. 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 RNA polymerase II, or (right) in cells after the simultaneous 50% nuclear depletion TBP, TFIIA, and RNA polymerase II. 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).

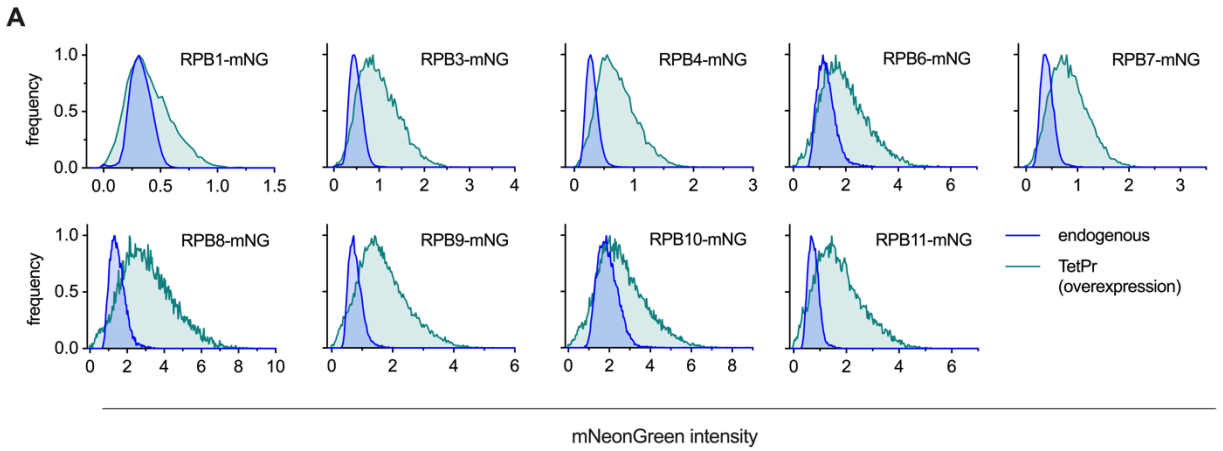

Figure S5 I RNA polymerase II over-expression experiments supporting data

Related to Fig. 2J-M

(**A**) Histogram (mode normalized) of the mNeonGreen signal measured by flow-cytometry for the indicated RNA polymerase II 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 treatment (45 minutes). The fold overexpression that is induced upon *Tet* promoter expression was calculated by comparing the median intensity of the overexpressed construct to the endogenous allele and is shown in Fig. 2E.

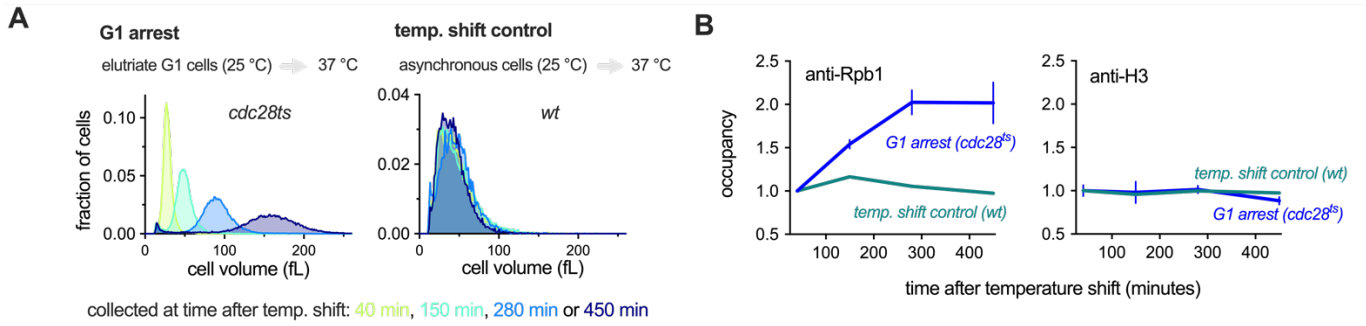

Figure S6 I G1 arrest experiment temperature shift controls

Related to Fig. 3A-C & G-H

(**A**) Cell size distributions determined by Coulter counter for cells used to analyse Rpb1 and Histone 3 occupancy in (B). The left hand panel shows G1 arrested cells of different sizes, and the right hand panel shows an asynchronous control that was subject to the same temperature shift as the G1 arrested cells. Small G1 cells harboring a temperature sensitive *CDK1* allele (*cdc28-13*) were collected by centrifugal elutriation and then arrested in G1 at the restrictive temperature for increasing amounts of time to generate populations of G1 cells of increasing cell size. The data in the left hand panel is also plotted in Fig. 3A.

(**B**) Rpb1 (anti-Rpb1; left panel) and Histone 3 (anti-H3; right panel) occupancy per genome plotted as a function of time after temperature shift for the cells shown in (A). The anti-Rpb1 data (left panel) are also plotted in Fig. 3B against cell size. The temperature shift control shown here demonstrates that the increase in Rpb1 occupancy during the G1 arrest is not due to the temperature shift.

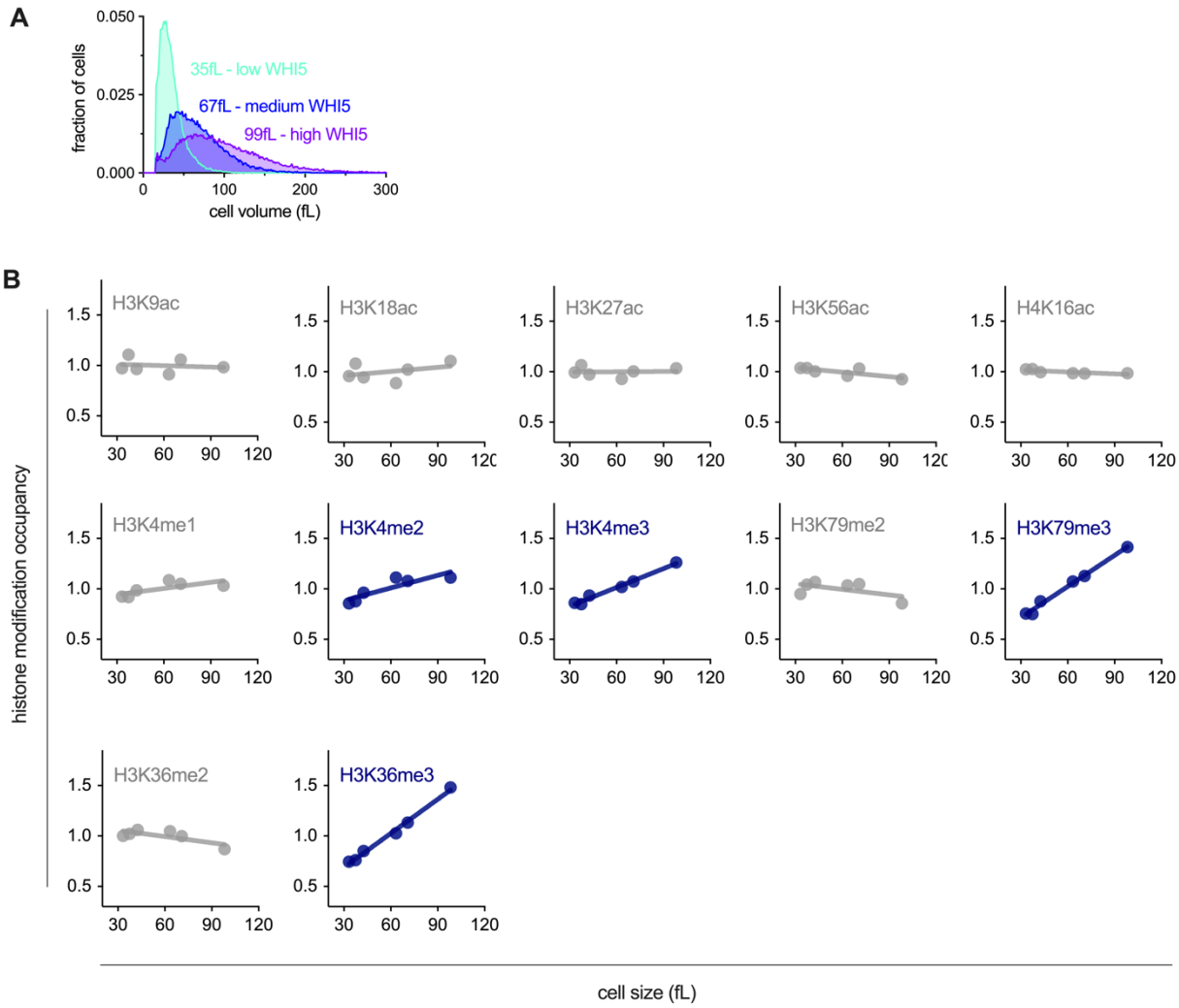

Figure S7 I Analysis of chromatin as a function of cell size supporting data

Related to Fig. 4

(**A**) Cell size distributions determined by Coulter counter for the cultures used in the dSMF experiment in Fig. 4A-B to measure chromatin accessibility in cells of different sizes.

(**B**) Histone modification occupancy for the indicated modification measured in cell size mutants shown in Fig. 1D & S1A. H3K56ac and H3k36me3 data are also shown in Fig. 4C. Line shows the linear fit used to calculate the slope values shown in Fig. 4D.

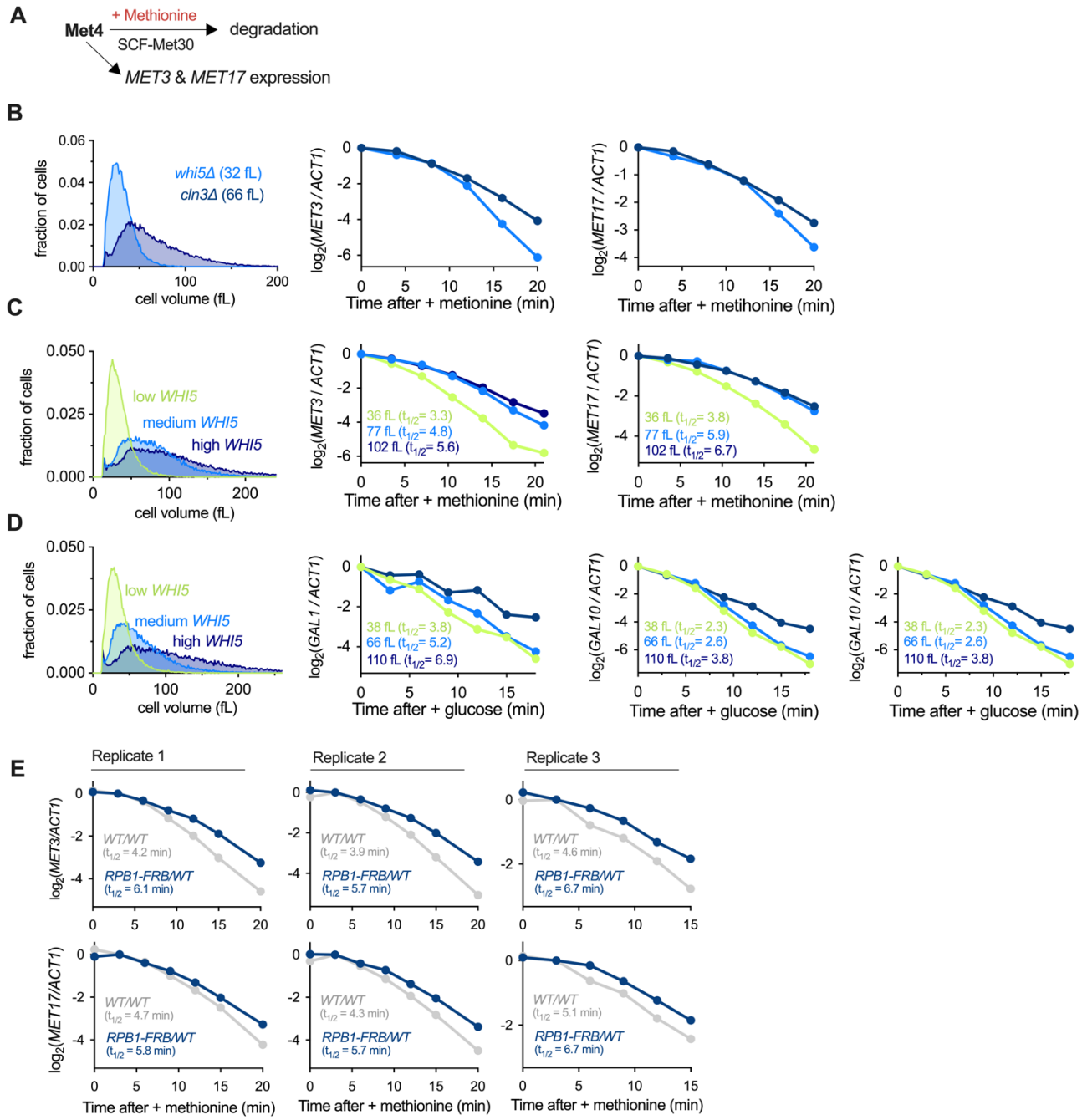

Figure S8 I mRNA decay compensation supporting data

Related to Fig. 5

(**A**) Schematic showing methionine responsive repression of *MET* genes including *MET3* and *MET17* (Rouillon et al., 2000).

(**B-C**) Example cell size distributions (left panel), *MET3* and *MET17* mRNA levels (centre and right panels) after transcriptional inactivation by methionine addition in cultures of different sizes. mRNA levels are determined by rt-qPCR normalised to *ACT1* transcript levels. Decay rate estimates from the one phase exponential decay fits to these data are shown in Fig.5C.

(**D**) Example cell size distributions (left panel), *GAL1, GAL7* and *GAL10* mRNA levels (centre and right panels) after transcriptional inactivation by galactose wash and glucose addition in cultures of different sizes. mRNA levels are determined

by rt-qPCR normalised to *ACT1* transcript levels. Decay rate estimates from the one phase exponential decay fits to these data are shown in Fig.5D.

(**E**) *MET3* (top) and *MET17* (bottom) mRNA levels relative to *ACT1* mRNA after transcriptional inactivation by methionine addition in *WT/WT* diploids or *RPB1-FRB/WT* heterozygous diploids. Methionine was added 40 minutes after rapamycin treatment. Rapamycin treatment is to deplete 50% RNA polymerase II from the nucleus. The bottom middle panel is also shown in Fig. 5E.

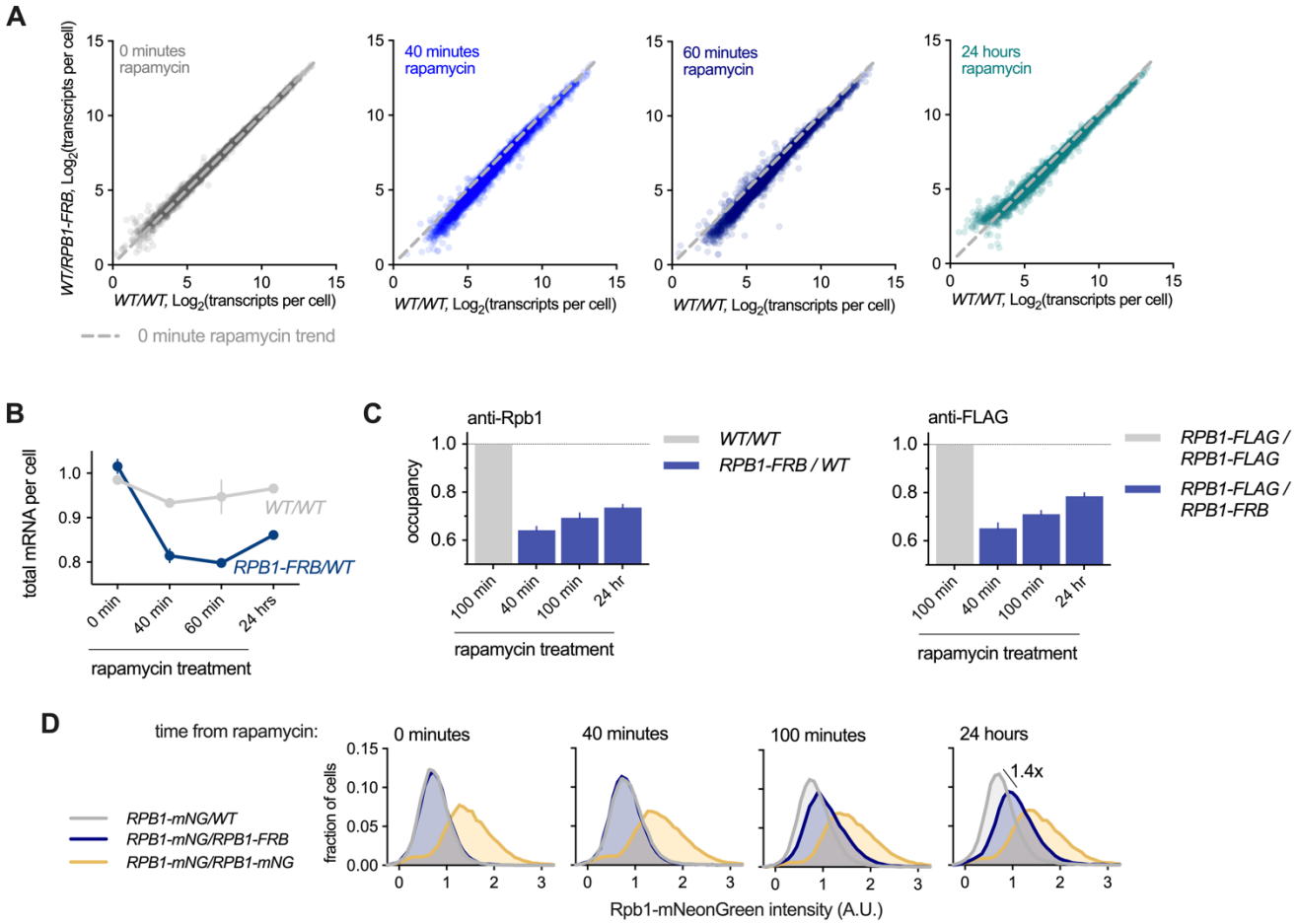

Figure S9 I Additional feedback is engaged following conditional 50% RNA polymerase II depletion

(**A**) Per gene mRNA levels in *WT/WT* (x-axis) or *RPB1-FRB/WT* (y-axis) at the indicated time after rapamycin treatment. Mean of two biological replicates is plotted.

(**B**) Total mRNA per cell determined by spike-in normalized RNA-seq in *WT/WT* or *RPB1-FRB/WT* cells at the indicated time after rapamycin treatment. Mean (±range) of two biological replicates is plotted.

(**C**) The global Rpb1 occupancy at the indicated time after rapamycin treatment measured by (left) anti-Rpb1 spike-in normalized ChIP-seq for *WT/WT* or *RPB1-FRB/WT* cells or (right) anti-FLAG spike-in normalized ChIP-seq in diploids where both alleles are FLAG tagged (*RPB1-FLAG/RPB1-FLAG*) or one allele is FLAG tagged (*RPB1-FLAG/RPB1-FRB*) and the other is FRB tagged for conditional nuclear depletion. 40-minute time point data is also shown in Fig. S4C-D. Mean (±SEM) is shown.

(**D**) Rpb1-mNeonGreen intensity determined by flow cytometry after 50% nuclear depletion of RNA polymerase II (*RPB1-mNG/RPB1-FRB*) at the indicated time after rapamycin treatment. Wild-type diploid cells (*i.e.,* no nuclear depletion of Rpb1) with one (*RPB1-mNG/WT*) or both (*RPB1-mNG/RPB1-mNG*) *RPB1* alleles fused to mNeonGreen are shown for comparison.

Author contributions. MPS designed, performed and analysed all experiments except for single molecule tracking experiments performed and analysed by HZ and proteomics data acquisition performed by AJ. Sequencing data analysis was performed by both MPS and GM. JMS wrote the dynamic equilibrium model. JMS, RRL, AK, AS and WG supervised the work. MPS and JMS wrote the manuscript.

Acknowledgments. We would like to thank members of the Kurt Schmoller, Mart Loog and members of the Skotheim lab for useful discussion and feedback. This work was supported by the NIH (GM092925 and GM115479), the HHMI-Simons (JMS, Faculty Scholars Program). MPS was supported by a Simons Foundation Fellowship of the Life Sciences Research Foundation and an EMBO Long-Term Postdoctoral Fellowship.

Declaration of interests. The authors declare no conflicts of interest.

References

Baptista, T., Grunberg, 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 e135.

Berry, S., Müller, M., and Pelkmans, L. (2021). Nuclear RNA concentration coordinates RNA production with cell size in human cells. bioRxiv.

Chan, L.Y., Mugler, C.F., Heinrich, S., Vallotton, P., and Weis, K. (2018). Non-invasive measurement of mRNA decay reveals translation initiation as the major determinant of mRNA stability. Elife *7*.

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. bioRxiv.

Creanor, J., and Mitchison, J.M. (1982). Patterns of protein synthesis during the cell cycle of the fission yeast Schizosaccharomyces pombe. J Cell Sci *58*, 263-285.

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.

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.

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.

Elliott, S.G., Warner, J.R., and McLaughlin, C.S. (1979). Synthesis of ribosomal proteins during the cell cycle of the yeast Saccharomyces cerevisiae. J Bacteriol *137*, 1048-1050.

Fraser, R.S., and Nurse, P. (1978). Novel cell cycle control of RNA synthesis in yeast. Nature *271*, 726-730.

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.

Haimovich, G., Medina, D.A., Causse, S.Z., Garber, M., Millan-Zambrano, G., Barkai, O., Chavez, S., Perez-Ortin, J.E., Darzacq, X., and Choder, M. (2013). Gene expression is circular: factors for mRNA degradation also foster mRNA synthesis. Cell *153*, 1000-1011.

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.

Helenius, K., Yang, Y., Tselykh, T.V., Pessa, H.K., Frilander, M.J., and Makela, 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.

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.

Huch, S., and Nissan, T. (2014). Interrelations between translation and general mRNA degradation in yeast. Wiley Interdiscip Rev RNA *5*, 747-763.

Hughes, C.S., Moggridge, S., Muller, 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.

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.

Krebs, A.R., Imanci, D., Hoerner, L., Gaidatzis, D., Burger, L., and Schubeler, D. (2017). Genome-wide Single-Molecule Footprinting Reveals High RNA Polymerase II Turnover at Paused Promoters. Mol Cell *67*, 411-422 e414.

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.

Kustatscher, G., Wills, K.L., Furlan, C., and Rappsilber, J. (2014). Chromatin enrichment for proteomics. Nat Protoc *9*, 2090-2099.

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.

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. bioRxiv.

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. bioRxiv.

Lin, J., and Amir, A. (2018). Homeostasis of protein and mRNA concentrations in growing cells. Nat Commun *9*, 4496.

Lock, A., Rutherford, K., Harris, M.A., Hayles, J., Oliver, S.G., Bahler, 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.

Marguerat, S., and Bahler, J. (2012). Coordinating genome expression with cell size. Trends Genet *28*, 560-565.

Neumann, F.R., and Nurse, P. (2007). Nuclear size control in fission yeast. J Cell Biol *179*, 593-600.

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 e1018.

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.

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.

Picelli, S., Bjorklund, 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.

Plaschka, C., Lariviere, L., Wenzeck, L., Seizl, M., Hemann, M., Tegunov, D., Petrotchenko, E.V., Borchers, C.H., Baumeister, W., Herzog, F.*, et al.* (2015). Architecture of the RNA polymerase II-Mediator core initiation complex. Nature *518*, 376-380.

Roberts, A., and Pachter, L. (2013). Streaming fragment assignment for real-time analysis of sequencing experiments. Nat Methods *10*, 71-73.

Rodriguez-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.

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.

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.

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.

Schwartz, D.C., and Parker, R. (1999). Mutations in translation initiation factors lead to increased rates of deadenylation and decapping of mRNAs in Saccharomyces cerevisiae. Mol Cell Biol *19*, 5247-5256.

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.

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.

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 e435.

Sun, M., Schwalb, B., Schulz, D., Pirkl, N., Etzold, S., Lariviere, 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.

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 e1217.

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.

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 e115.

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.

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.

Zhurinsky, J., Leonhard, K., Watt, S., Marguerat, S., Bahler, J., and Nurse, P. (2010). A coordinated global control over cellular transcription. Curr Biol *20*, 2010-2015.

Materials and methods

Yeast genetics

All *S. cerevisiae* budding strains used in this study are in the *W303* background. Standard procedures were used for *S. cerevisiae* and *S. pombe* strain construction. Full genotypes of all strains used in this study are listed in **Table S1** and all plasmids used in this study are listed in **Table S2**. Diploid strains were constructed by crossing the respective parent haploid strains listed in **Table S1** and then streaking out on SCD -LEU -TRP selection plates.

Yeast strains MS627-635, used for the conditional overexpression of individual mNeonGreen tagged RNA polymerase II subunits (**Fig. 2E & S5**) were constructed as follows. pMS153, which encodes the WTC846 TET transcription factor system described in Azizoğlu et al., 2020 was linearized and integrated into the *HIS3* locus. Then each RNA polymerase II subunit was separately sub-cloned with a C-terminal mNeonGreen tag into the *URA3* SIV backbone (Wosika et al., 2016) downstream of the WTC846 *TETpr* (Azizoğlu et al., 2020) and upstream of the *AHD1* terminator. Each plasmid (pMS200-pMS208) was then linearized and integrated into the *URA3* locus. 3 out the 12 RNA polymerase II subunits where 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 MS699, used for the simultaneous conditional overexpression of all RNA polymerase II subunits (**Fig. 2F-G**), was constructed as follows. First, pMS152, which encodes the WTC846 TET transcription factor system described in Azizoğlu et al., 2020, was integrated into the HO locus. Four plasmids (pMS226, pMS230, pMS231 and pMS232) were then constructed that were each based on a different single integration vector (SIV) backbone (Wosika et al., 2016). For each of the plasmids pMS226, pMS230, pMS231, and pMS232, three different RNA polymerase II subunits where cloned in tandem into the into the SIV multi-cloning site with each subunit expressed from the WTC846 *TETpr* (Azizoğlu et al., 2020) with *ADH1* terminator downstream. The *RPB3* construct (in pMS226) was also fused via a linker to the mNeonGreen fluorescent protein. pMS226, pMS230, pMS231, and pMS232 were linearized and integrated into the *URA3*, *HIS3*, *TRP1* and *LEU2* loci, respectively, of MS46 and then *WHI5* was deleted. A control strain (MS656) was constructed in the same manner as MS669, but using four integration vectors each encoding mNeonGreen expressed from the WTC846 *TETpr* with an *ADH1* terminator downstream (pMS198, pMS233, pMS234 and pMS235).

Yeast culturing conditions

Cells were grown in synthetic complete media (SC) with either 2% glucose, 2% glycerol + 1% ethanol or 2% raffinose as a carbon source. SC + 2% glycerol + 1% ethanol was used for all experiments examining cell size mutants (**Fig. 1C-D, 2B, 2H-I, 4, 5B-C S2, S3, S7 & S8B-C**), as this increases the dynamic range in size between mutants. SC + 2% glycerol + 1% ethanol was also used for the elutriation experiment (**Fig. 3A-B & S6,** See *Centrifugal elutriation* section below). SC + 2% raffinose was used for *GAL* induction and shutoff experiments (**Fig. 5E & S8D;** see *qPCR mRNA decay experiments* section below). All other experiments were

performed in SC + 2% glucose. Unless stated otherwise, cells were grown at 30°C, kept 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 (**Fig. 1D**), which carries an integrated copy of a *LexOPr-WHI5* construct, was induced by growing cells for >36 hours in 0nM, 10nM, or 30nM beta-estradiol, respectively, at which point cell size is at a steady state. Rapamycin was used at a final concentration of 1 µg/ml. Cells were collected 100 minutes (all samples in **Fig. S4B** and all samples in **Fig. 2D** except *Rpb1-FRB / WT*) or 40 minutes (*Rpb1-FRB / WT* samples in **Fig. 2D** and all samples in **Fig. S4C-H**) after rapamycin treatment. The shorter 40 minute rapamycin treatment due to longer term compensation in Rpb1 protein amounts following 50% nuclear depletion (**Fig. S9D**). Anhydrotetracycline was used at a final concentration of 50 ng/ml. Drug treatment duration is as specified in associated figure legends.

For the SILAC G1 arrest sample in **Fig. 2B & S3E,** MS54 was grown in the presence of 20nM beta-estradiol, which induces expression of *CLN1* and is essential for proliferation in this strain background lacking the other G1 cyclins (*cln2Δ cln3Δ*). To initiate the G1 arrest, cells were washed on a filter membrane (3x) in pre-warmed media lacking beta-estradiol and then returned to media lacking beta-estradiol.

For all SILAC experiments (**Fig. 2B & S3C-E**) SILAC compatible strains (*i.e.*, the *lys1Δ arg4Δ CAN1+* background) were grown for more than 10 generations in SILAC media supplemented with light (L-Arginine (unlabelled) & L-Lysine (unlabelled)), 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.).

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 (13 krpm, 2 minutes), washed, and resuspended in 50 mM Sodium Citrate (pH = 7.2). They were then incubated with 0.2 mg/ml RNAse A (overnight, 37 °C) and then treated with 0.4 mg/ml proteinase K (1-2 hour, 50°C) before addition of 25 µM Sytox Green (ThermoFisher Scientific). Cells were then sonicated and DNA-content was analysed for >10000 events on a FACScan Analyzer (BD Biosciences). 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 (see spike-in normalized ChIP-seq methods section for details).

Live cell flow cytometry measurements

Protein levels of C-terminally mNeonGreen-tagged proteins (**Fig. 2A&E, S5 & S9D**) were quantified using an Attune NxT flow cytometer. Strains 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. The size-dependent background of the BL1-A signal was determined by measuring an untagged background control strain 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.

Centrifugal elutriation

For the G1 arrest elutriation experiment (**Fig. 3A-B & S6**), *cdk1^{ts}* (*cdc28-13*) cells where grown in SC + 2% glycerol + 1% ethanol because these conditions allow for the collection of the smallest G1 cells and does not require media carbon source removal or cooling to 4 °C during elutriation. G1 cells were collected and then arrested in G1 as follows. 4 litres of culture were grown at 25°C (*cdk1^{ts}* permissive temperature) to OD_{600} ~0.65. Cells were then collected on a filter membrane and resuspended in 50ml fresh media. Cells were then sonicated (3 x 20 seconds, 4 minutes 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 (2.4 krpm, 23 °C). The elutriation chambers were pre-equilibrated and run with room temperature SC +2% glycerol +1% ethanol media. The pump speed was gradually increased until small G1 cells with minimal debris were collected. The smallest G1 fractions were then combined and concentrated on a filter membrane, resuspended at 37°C (*cdk1^{ts}* restrictive temperature) in SC + 2% glycerol +1% ethanol media, and split into 4 cultures placed in a 37 °C shaking water-bath. Each culture was arrested in G1 at 37°C for different periods of time before fixation and collection for ChIP-seq. At the point of splitting into 4 cultures, the OD_{600} of each culture was adjusted, according to its arrest time, so that each culture would be at a similar OD_{600} at the point of collection.

Microscopy

Cells were imaged on a wide-field epifluorescence Zeiss Observer Z1 microscope (63X/1.4NA oil immersion objective and a Colibri LED module). A single *z*-stack was taken and cell boundaries were identified using phase contrast images. mCitrine fluorophores were imaged in the yellow channel (505 nm LED module). For the images in **Fig. S4A,** the exposure and contrast were manually adjusted in imageJ individually for each image to allow comparison of the relative subcellular localisation of the signal and is therefore not appropriate for comparing relative protein intensities.

Spike-in normalized ChIP-seq

The spike-in normalized ChIP-seq protocol was adapted from Hu et al. 2015 (Hu et al., 2015). *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 minutes). Cells were then quenched with 0.125 M glycine (5 minutes), washed twice in cold PBS, pelleted, snap-frozen, and stored at –80 °C.

Sample and spike-in were mixed in 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, S. *pombe* with a Rpb1-3xFLAG (grown in EMM4S) was 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 a given 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 Figure 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 (MS289; *TBP-FRB* treated with rapamycin for 60-65 minutes) to prevent 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 *Spike-in normalized ChIP-seq analysis* section 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 16 minutes (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 retained for ChIP or input. For each ChIP reaction, 20 μl Protein G Dynabeads (Invitrogen) were blocked (PBS + 0.5% BSA, incubate 40 minutes at room temperature), pre-bound with 5 μl of antibody in PBS (incubate 40 minutes 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 minutes per wash) 2x in FA lysis buffer and 3x in high-salt FA lysis buffer (50 mM HepesKOH 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 TrisHCl pH 7.5, 10 mM EDTA, 1% SDS) at 65°C for 20 minutes. At the same time, 15 μl of input was mixed directly with 115 μl of ChIP elution buffer. Eluted ChIP DNA and input DNA were then incubated to reverse crosslinks (65°C, 5hr), before treatment with RNAse A (37 °C, 1 hour) and then Proteinase K (65°C, 2 hours). 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 Cat # E7645), pooled, and then sequenced by Illumina paired-end (2x150bp) sequencing.

The following antibodies were used for ChIP against the indicated epitopes: anti-Rpb1 clone 8wG16 (mouse, monoclonal, Millipore # 05-952-I), anti-Rpb1-S2-P clone 3E10 (rat, monoclonal, Millipore, #04-1571-1), anti-Rpb1-S5-P clone 3E8 (rat, monoclonal, Millipore #04-1572-I), 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).

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) (Skrzypek et al., 2017)) or *S. pombe (*ASM294v2). Combined gene annotations were created using *S. cerevisiae* gene models updated using transcript-end mapping data as previously described (Shipony et al., 2020) and either the *C. glabrata* gene models available from CGD or the *S. pombe* gene models available from PomBase (Lock et al., 2019).

Demultipexed FASTQ files were then mapped to the relevant combined genome indexes as 2×36mers using Bowtie (v.1.0.1) (Langmead et al., 2009) 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. The global occupancy ratio is therefore expressed per genome. For the data in **Fig. 1E** 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). The Histone modification occupancy (**Fig. 4C-E & S7B**) was calculated as above and then normalized to the anti-H3 or anti-H4 global occupancy ratio for the same sample. For the experiments in **Fig. 2, Fig. S4** and **Fig. S9C** the *S. cerevisiae* background-to-spike-in ratio was determined (see *Spike-in normalized ChIP-seq* section above for details on what samples were used to determine the background ratio). For these experiments, the background ratio was then subtracted from the global occupancy ratio. The global occupancy ratios were then re-scaled to be expressed as a fold change with respect to a reference control sample.

RPM (Reads Per Million) normalized read coverage genome browser tracks showing the 50bp mid-point of each mapped fragment (*i.e.*, midpoint ± 25bp) were generated for the *S. cerevisiae* genome using custom-written python scripts. Subsequent analysis uses these 50bp mid-point read coverage tracks to calculate per gene occupancies and metagene profiles. Per gene occupancy was calculated as follows. RPKM 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 500bp upstream of the TSS, 500bp downstream of the TTS and gene bodies (TSS-to-TTS; rescaled to be 1kb) and then multiplied by the global occupancy factor described above and were 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 (**Fig. 4E**), values were re-scaled to the median value of all samples displayed on a given heatmap to aid comparison between modifications.

Dual enzyme single-molecule foot printing (dSMF)

The protocol for dSMF to measure chromatin accessibility (**Fig. 4A-B**) was based on Krebs et al., (2017) (Krebs et al., 2017) 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[pH7.5], 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 minutes, 500 rpm). Cells where then pelleted (5k rpm, 2 minutes, 4°C), resuspended in 0.5 ml wash buffer (1.4 M sorbitol, 40 mM Hepes-KOH[pH7.5], 0.5 mM $MgCl_2$), pelleted again (5k rpm, 2 minutes, 4°C), resuspended in 0.3 ml ice-cold lysis buffer (10 mM Tris-HCl[pH7.5], 10 mM NaCl, 3mM $MgCl_2$, 0.1mM EDTA[pH7.5], 0.5% NP-40), and incubated on ice (10 minutes). Nuclei were then pelleted (5k rpm, 4 minutes, 4 °C), resuspended in 0.3 ml nuclei wash buffer (10 mM Tris-HCl[pH7.5], 10 mM NaCl, 3mM $MgCl_2$, 0.1mM EDTA[pH7.5]), pelleted again, and resuspended in 50 μl 1x reaction buffer (50 mM Tris-HCl[pH8.5], 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[pH8.5], 50 mM NaCl, 10 mM DTT, 0.7 M sucrose, 1.3 mM SAM, 200 U M.CviPI) and incubated (8 minutes, 30°C) before mixing with 3 μl M.CviPI boost mix (100 U M.CviPI, 42.5 μM SAM), and returning to incubation (7 minutes, 30 °C). 10 μl M.SssI reaction mix (50 mM Tris-HCl[pH8.5], 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 minutes, 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.

Dual enzyme single-molecule foot printing (dSMF) data analysis

Combined *S. cerevisiae* and *C. glabrata* genome and transcriptome files were created as described above (see *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 the 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.

qPCR mRNA decay experiments

To quantify *MET3* and *MET17* decay rates, cells were grown in media lacking methionine and 1 mM methionine was then added to repress MET genes (Rouillon et al., 2000). To quantify *GAL1*, *GAL7* and *GAL10* decay rates, cells were grown in SC media + 2% raffinose before addition of 2% galactose for 75 minute and then washed (1x) and resuspended in SC + 4% glucose (Haimovich et al., 2013). 1 ml of culture was pelleted (13 krpm, 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 minute 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: 5.0 m/s, 1 x 30 seconds). Cell debris was pelleted (13 krpm, 2 minutes) and the supernatant recovered. RNA was then purified using the direct-zol RNA microprep kit (Zymo Research #R2061). cDNA was then synthesized using 800ng of RNA with iScript Reverse Transcription Supermix for RT-qPCR (BioRad #1708841) and used for qPCR for *MET3*, *MET17* and *ACT1* using iTaq Universal SYBR Green Supermix (BioRad #1725121). 2-3 biological replicates were performed per experiment and 3-4 technical replicates were performed per biological replicate. Decay rates (**Fig. 5C-D & F**) were calculated from a one-phase exponential decay fit to the data (Prism 6; weighting = 1/Y). 0-minute timepoints (collected before methionine or glucose addition) were excluded from the decay rate calculation assuming a time-delay before inhibition starts, which also resulted in an improved fit.

Spike-in normalized RNA-seq

For RNA-seq, 1ml of culture was pelleted (13 krpm, 30 seconds, 4°C), snap frozen in liquid nitrogen, and stored at -80°C. Spike-in *C. glabrata* cells were collected similarly. Subsequently, the spike-in was thawed in ice-cold PBS and 50 µl (~1/3 of the *S. cerevisiae* cell number per sample) was taken and used to resuspend the *S. cerevisiae* cell pellet (on ice).

15 µl of the 50 µl *S. cerevisiae* and *C. glabrata* mix was mixed directly with 110 µl digestion buffer from the YeaStar Genomic DNA Kit (Zymo Research #2002) and processed for genomic DNA extraction using the YeaStar Genomic DNA Kit. Genomic DNA libraries were then constructed as follows: 4-5ng of gDNA was incubated with 4µl home-made Tn5 transposase (30 min, 37°C) (Picelli et al., 2014), cleaned up using a MinElute PCR purification Kit (QIAGEN), and amplified with NEBNext High-Fidelity 2X PCR Master Mix (NEB #M0541L) to insert indexes before clean up with AMPure XP beads (Beckman Coulter #A63881).

The remaining 35 µl of the *S. cerevisiae* and *C. glabrata* mix was processed for RNA extraction as follows. Cells were mixed directly with 300 µl TRI Reagent (Zymo Research) and lysed by bead beating using a Fastprep 24 (4°C, settings: 5.0 m/s, 1 x 30 seconds). Cell debris was then pelleted (13 krpm, 2 minutes) and the supernatant recovered. RNA was then purified using the direct-zol RNA microprep kit (Zymo Research).

800ng of RNA was then used for rRNA depletion and subsequent library construction using the Zymo-Seq RiboFree Total RNA Library Kit (Zymo Research #R3000). Pooled libraries were sequenced by paired-end (2x150 bp) Illumina sequencing with > 20 million reads sequenced per sample.

RNA-seq data processing

Combined *S. cerevisiae* and *C. glabrata* genome and transcriptome files were created as described above (see *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) (Dobin et al., 2013). 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 annotation described above using Bowtie (version 1.0.1; settings: -e 200 -a -X 1000). Alignments were then quantified using eXpress (version 1.5.1) (Roberts and Pachter, 2013) for effective read count values and TPM (Transcripts Per Million transcripts).

To determine relative mRNA amounts per cell, the TPM for all *S. cerevisiae* genes was normalized to the total TPM of *C. glabrata* reads and then normalized to the *S. cerevisiae*-to-*C. glabrata* gDNA ratio of reads (*i.e.*, the mixing ratio). Total relative mRNA amounts were then calculated by summing the normalized values for all *S. cerevisiae* genes.

Single molecule imaging: microscopy

1 ml culture (SC + 2 % glucose) was pelleted (4k rpm, 1 minute) 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 minutes. 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 minutes to let any remaining unbound dye be released from 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 minutes, washed with MilliQ water 3 times, washed in methanol (flame excess from coverslips), and then placed in a Plasma Etch plasma oven for 10 minutes.

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 Plano 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) 405nm 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, quantify nuclear area using the Pus1-GFP nuclear marker. Nuclear area can then be used to estimate cell volume.

Single molecule imaging: analysis

Aside from tracking, all analysis was performed using custom matlab code. Tracking was performed using Trackmate (Tinevez et al., 2017). 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 for segmenting the nuclear regions using the Pus1-GFP nuclear marker and tracks outside the nucleus were discarded. The radius of gyrations (RoG) were then calculated for each track using the following equation:

$$RoG = \sqrt{\frac{1}{n}\sum_{i}^{n}((x_i - \bar{X})^2 + (y_i - \bar{Y})^2)}$$

where n is the number of localisations in each track, and (x_i, y_i) are the coordinates of the track for each step i, and ($\bar{X}$,$\bar{Y}$) 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 log2-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.

Chromatin enrichment for Proteomics (ChEP)

For SILAC experiments, SILAC compatible strains were labelled with amino acid isotopes (see *Yeast culturing conditions* for details) before collection. The ChEP chromatin enrichment protocol was based on Kustatscher et al., (2014) (Kustatscher et al., 2014) and adapted for yeast as follows. Cells were fixed by addition of formaldehyde to a final concentration of 1%, shaken for 15 minutes, quenched with 0.125 M glycine (5 minutes), 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 minutes), 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 minutes (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 minutes). 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 minutes) 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 purify 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 and 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 minutes) and then pelleted (4 °C, 14k rpm, 30 minutes). The supernatant was then discarded and the pellet was resuspended in 300 μL SDS buffer (4 % SDS, 10 mM EDTA, 25 mM TrisHCl pH 7.5, 1 mM PMSF, 1x Roche protease inhibitor) and incubated at room temp (10 minutes). 1 ml urea buffer (8M urea, 1mM EDTA, 10mM TrisHCl 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 minutes) 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 minutes). The supernatant was discarded and the pellet was resuspended in 300 μl SDS buffer and then adjusted to 1.45 ml with SDS buffer and again pelleted (18°C, 14k rpm, 30 minutes). The supernatant was again discarded and 0.6 ml Storage buffer (10% glycerol, 1 mM EDTA, 10 mM TrisHCl 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 minutes (Amp = 25%, 10 seconds on, 20 seconds off)). The sonicated sample was then pelleted (4 °C, 14k rpm, 10 minutes) 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 **Fig. S3B**, ~ 5 μg of the chromatin fraction B was loaded per lane alongside ~25 μg CYTO and WCE. For the proteomics analysis in **Fig. S3C-D**, 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 *LC-MS/MS sample preparation and data acquisition* below). For the proteomics analysis in **Fig. 2B**, heavy, medium and light labelled 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 analysed by LC-MS/MS (for details see *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 labelled 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 minutes at 56°C, alkylated with 10 mM iodoacetamide (30 minutes, room temperature, dark) and then quenched with 7.5 mM DTT. Samples were digested and cleaned using SP3 on-bead methodology (Hughes et al., 2019) with the variation 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 analysed 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 minute gradient elution from a 75 μm x 50 cm C18 column.

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 **Fig. S3C-D** and 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 **Fig. S3D**. 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 (**Fig.S3D**) at a point that separates known chromatin associated factors from abundant cytoplasmic factors (**Fig.S3C**). The distribution of SILAC ratios for the experiment plotted in **Fig. 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 (**Fig. 3B**).