# Size-independent mRNA synthesis and chromatin-based partitioning mechanisms generate and maintain constant amounts of protein per cell

Matthew P. Swaffer[1], Devon Chandler-Brown[1], Maurice Langhinrichs[1], Georgi Marinov[2], William Greenleaf[2], Anshul Kundaje[2], Kurt M. Schmoller[1,3], 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] Institute of Functional Epigenetics, Helmholtz Zentrum München, 85764 Neuherberg, Germany

* correspondence: skotheim@stanford.edu

## Summary

Cell size and biosynthesis are inextricably linked. As cells grow, total protein synthesis increases in proportion to cell size so that protein concentrations remain constant. As an exception, the budding yeast cell-cycle inhibitor Whi5 is synthesized in a constant amount per cell cycle, so that it is diluted in large cells to trigger division. Here, we show that this size-independent expression of Whi5 results from size-independent transcription. A screen for similar genes identified histones as the major class of size-independent transcripts during the cell cycle, consistent with histone synthesis being coupled to genome content rather than cell size. However, during asymmetric division size-independent transcription is insufficient for size-independent protein expression and chromatin-binding ensures equal amounts of protein are partitioned to unequally sized cells to maintain size-independent protein amounts. Thus, specific transcriptional and partitioning mechanisms determine size-independent protein expression to control cell size.

A striking feature of cell growth is that total protein and RNA amount per cell increase approximately in proportion to cell volume as a cell increases in size (Fig. 1A) (Fraser and Nurse, 1978, 1979). To achieve this coordinated scaling of macromolecules with cell size, larger cells have proportionally higher transcription and protein synthesis rates (Creanor and Mitchison, 1982; Elliott, 1983; Elliott and McLaughlin, 1979; Elliott et al., 1979; Padovan-Merhar et al., 2015; Sun et al., 2020; Zhurinsky et al., 2010). This size-scaling is generally important because it ensures macromolecule copy number is proportional to cell volume and therefore concentrations are kept constant as a cell grows (Fig. 1B) (Marguerat and Bahler, 2012; Neurohr et al., 2019). Nuclear volume also scales in proportion to cell volume meaning that nuclear concentrations are also expected to be constant (Jorgensen et al., 2007; Neumann and Nurse, 2007).

While it is generally assumed that most individual proteins exhibit this general size-scaling behavior and remain at constant concentration, one striking exception is the cell cycle inhibitor Whi5. Whi5 synthesis is uncoupled from cell size such that a constant amount of Whi5 is made in each cell division cycle independently of cell size, and biosynthetic capacity (Schmoller et al., 2015). Moreover, examination of Whi5 synthesis in a variety of extracellular growth conditions showed that a similar number of molecules were made in all conditions (Qu et al., 2019). This growth and size-independent synthesis pattern for Whi5 ensures its dilution in larger cells to trigger their division and thereby links the basic mechanisms of biosynthesis to cell size control. Thus, cell size determines the pattern of protein biosynthesis, which triggers division in larger cells to, in turn, control cell size. However, the underlying molecular mechanisms determining the relationship between cell size and the expression of individual proteins remain largely unknown.

In principle, any stage of gene expression could be regulated in a manner that results in size-independent protein synthesis (Fig. 1C). Importantly, this is not done using negative feedback because multiple copies of the *WHI5* gene produces a proportional increase in the number of proteins made per cell cycle (Qu et al., 2019; Schmoller et al., 2015). To determine whether Whi5's size-independent behavior originates at the protein or transcript level, we sorted cells of different sizes using a total protein dye as proxy for cell size and performed RNA-seq (Fig. 1D & S1A). Total mRNA content and cell volume were correlated with protein dye intensity confirming it as a good proxy for cell size (Fig. S1B-E). *WHI5* mRNA transcripts per million (TPM) are anticorrelated with cell size, which implies that *WHI5* mRNA concentration is lower in larger cells. In contrast, *MDN1* mRNA TPM, as representative of scaling gene expression, was constant (Fig 1D). When normalized to total mRNA, using an external spike-in, *WHI5* mRNA amount per cell was flat as a function of size (Fig. 1E-F). To corroborate this finding, we performed single-molecule FISH in individual cells while also measuring cell size (Fig 1G). Consistent with the RNA-seq data, *WHI5* mRNA amounts were relatively independent of cell size whereas the *MDN1* counts increased with cell size (Fig. 1H&I, Fig S1A&B).

Because *WHI5* mRNA is a cell cycle regulated transcript that peaks in S phase (Fig. S3C, (Pramila et al., 2006)) and cell size will partially correlate with cell cycle stage, it is possible the apparent size-independent behavior of *WHI5* mRNA (Fig. 1E-H) is simply a consequence of the cell cycle rather than cell size *per se*. To control for this possibility, we isolated cells in early G1 by centrifugal elutriation and arrested them in G1 for increasing amounts of time to generate populations of cells of increasing sizes. Cells were then released from the G1 arrest resulting in cultures of cells synchronously traversing the entire cell cycle but at different sizes (Fig. 2A&B & Fig. S2). RNA-seq across the time-course allowed us to compare gene expression in cells of different sizes at precisely the same cell cycle stage (Fig. 2C). *WHI5* TPM are indeed lower in larger cells as *WHI5* expression peaks in S phase. The total amount of *WHI5* expression across the entire cell cycle can then be estimated as the area under the curve, which again shows an anticorrelation between *WHI5* mRNA concentration and cell size (Fig. 2D). Consistent with this, if we restrict our smFISH analysis to cells in early S/G2/M, when *WHI5* expression peaks, the number of *WHI5* transcripts are still uncorrelated with cell size (Fig. 2E & S3E&F). Taken together, this group of experiments suggests that *WHI5* transcription or mRNA degradation is responsible for the size-independent expression of Whi5 protein.

Next, we sought to test if Whi5 size-independent expression originates at the level of transcription. If this were the case, then the *WHI5* promoter should be both necessary and sufficient for size-independent protein expression. To test if the *WHI5* promoter is sufficient, we compared the size-dependency of Whi5-mCitrine protein synthesis expressed from its endogenous promoter with that of a reporter mCitrine also expressed from the *WHI5* promoter (Fig. 2F&G). Both Whi5-mCitrine and the mCitrine reporter are synthesized in a size-independent manner. By contrast, when we expressed Whi5-mCitrine from a scaling promoter (*ACT1pr*) its synthesis rate now does increase with cell size (Fig. 2H). Together, these experiments demonstrate that *WHI5* is transcribed in a size-independent manner and that the *WHI5* promoter is both necessary and sufficient for the size-independent synthesis of Whi5.

Having shown that size-independent expression of *WHI5* is due to a transcriptional mechanism, we sought to determine which other cellular processes are similarly uncoupled from cell size. To do this we sought to determine which other sets of genes are similarly regulated by analyzing our RNA-seq experiments comparing cells of different sizes. We found 16 transcripts that behaved similarly to *WHI5* in both experiments (Figure S4A). These genes are significantly enriched for GO terms related to chromatin and revealed histones as a major class of size-independent genes as 9 of the 16 identified genes encode histones (Fig. 3A). As for *WHI5*, histone mRNAs decreased in larger cells (Fig. 3B&C and Fig. S2B). Again, this is not simply a consequence of their cell cycle regulated transcription because the same trend was observed in cells of different sizes synchronously progressing through the entire cell cycle (Fig. 3D&E and Fig. S4C). We further confirmed the size-independent expression of histone transcripts by comparing the mRNA levels in 1,484 strains

containing a single gene deletion (Kemmeren et al., 2014; O'Duibhir et al., 2014) with the cell size of the respective gene deletion strain (Hoose et al., 2012; Jorgensen et al., 2002; Ohya et al., 2005; Soifer and Barkai, 2014). This revealed a clear anti-correlation between cell size and histone mRNA levels (Fig. 3F and S4D&E).

The size-independent expression of histone mRNAs suggests that histone protein expression is coordinated with genome content rather than cell size and predicts that histone protein synthesis should also be size-independent. To examine this, we first analyzed two published datasets of flow cytometry measurements across the collection of strains in which each individual open reading frame was fused to GFP. We compared the relationship between GFP fluorescence (protein amount) and side scatter (SSC-A, cell size) (Fig. 3G). This shows that histone proteins are less correlated with cell size than the average protein in the proteome, *i.e.*, the slope between cell size (SSC-A) and GFP intensity is smaller (Fig. 3H&I). To confirm that histone protein synthesis is indeed size-independent, we quantified the amount of histone synthesized across the cell cycle of single cells by time-lapse fluorescence microscopy and compared it with the increase in size during the same period. Cells that grow more produce proportionally more of a scaling protein (Rpb3), but less histones (Hta2, Htb2 and Htz1) (Fig. 3J&K). Taken together these experiments have identified histones as a rare class of size-independent genes whose transcription, and therefore protein synthesis, is uncoupled from cell size. In this way, histone production is coordinated with genome-content rather than cellular growth.

Both Whi5 and histones are stable proteins synthesized in a size-independent amount during S/G2/M of the cell cycle meaning that their amounts in G1 are determined by inheritance from previous cell cycles. The asymmetric division of budding yeast poses a problem for maintaining size-independent amounts of these proteins because the smaller daughter cells would inherit fewer such proteins if they were partitioned in proportion to cell volume as is generally the case across the proteome. For volume-based partitioning, concentrations are expected to be similar in the mother and daughter cells following division as seen for a freely diffusing mCitrine (Fig. 4A). However, this is not the case for Whi5, which we find is not partitioned evenly by volume so that daughter cells all inherit a more similar number of molecules (Fig. 4A). To perfectly maintain size-independent amounts, a mechanism partitioning equal amounts to the daughter and mother cells is required.

To calculate the impact of amount- and volume-based partitioning modalities on the amounts of inherited protein, we employed a full cell-cycle model that simulates growth and division of a population of cells. This model was parameterized by single-cell microscopy measurements and therefore accounts for cell-to-cell variability and the size-dependence of cell cycle progression (Chandler-Brown et al., 2017). To this model, we added a protein synthesized either at a rate proportional to cell size or independent of size. At division, these proteins were either partitioned in proportion to cell-volume or in the manner empirically determined for Whi5 where a significant fraction is partitioned by amount (Fig. 4B, S6A&B). Our simulations show that the amount of Whi5 inherited in

G1 should be significantly less size-independent if it were partitioned by cell volume rather than by protein amount. This is in part because bud size varies significantly even for mothers with same the volume. Together this suggests that both partitioning by amount and size-independent synthesis ensures the size-independent amount of Whi5 in G1.

One possibility to generate amount-based partitioning of Whi5 would be to harness the equal partitioning of genome segregation. This hypothesis was suggested by the fact that Whi5 binds the DNA-bound SBF transcription factor complex following its dephosphorylation at mitotic exit. To test this, we analyzed the partitioning of a Whi5 variant (Whi5WIQ) that does not bind SBF and is not recruited to the SBF binding sites in the *CLN2* or *SVS1* promoter (Travesa et al., 2013). First, we confirmed that the WIQ mutation reduces Whi5 binding at SBF bound DNA elements across the genome by ChIP-seq (Fig. S6A). Next, we analyzed cells expressing Whi5(WIQ)-mCitrine, which revealed it has a significantly lower bud-to-mother concentration ratio at division than wild-type Whi5. This supports our model that partitioning by amount is indeed mediated by DNA-binding (Fig. 4A). Crucially, Whi5(WIQ) amounts at birth are higher in larger cells and lower in smaller cells when compared to wild-type Whi5, demonstrating that when Whi5 partitioning is disrupted, its size-independence in G1 is also disrupted (Fig. 4C&D). Thus, our data support a model where Whi5 binding to chromatin results in its volume-independent partitioning into the mother and bud at division to ensure daughter cells inherit approximately the same amount of Whi5 regardless of their size.

In conclusion, while most proteins are synthesized so that their amounts are proportional to cell size and their concentrations remain constant, some are not (Fig. 1A&B). We previously identified the cell cycle inhibitor Whi5 and its functional ortholog in human cells, the retinoblastoma protein, as proteins whose amount synthesized per cell is independent of cell size (Schmoller et al., 2015; Zatulovskiy et al., 2020). This results in smaller cells having proportionally higher concentrations of these cell cycle inhibitors to promote more growth in G1 to compensate their initially smaller birth size. Here, we show that in budding yeast this is achieved via a transcriptional mechanism and our transcriptomics experiments have revealed that the G1/S cell cycle inhibitors are not alone. Strikingly, the histone mRNAs constitute the majority of transcripts in this small group of genes whose expression is size-independent. The size-independent expression of histones allows the maintenance of a stoichiometric relationship between histone proteins and the genome without engaging the potentially wasteful feedback mechanisms, which are known to operate when histone expression is artificially elevated (Cross and Smith, 1988; Gunjan and Verreault, 2003; Moran et al., 1990; Norris and Osley, 1987).

For both histones and *WHI5*, size-independent amounts of protein are reflected in the mRNA amounts synthesized during the cell cycle. Our data from cycling cells mirrors the recent finding that in cells arrested in G1 the mRNA concentration of a number of cell cycle activators increases in larger G1 cells (Chen et al., 2020). Such a size-dependent concentration increase was first found in fission

yeast for the cell cycle activator *cdc25* (Keifenheim et al., 2017). However, understanding the mechanisms controlling the various size-dependencies of transcription is limited by the fact we do not understand how cell size controls transcription even in the generic size-dependent case where global transcription rates increase in proportion to cell size (Padovan-Merhar et al., 2015; Sun et al., 2020; Zhurinsky et al., 2010). Nevertheless, despite these limitations, our *WHI5* promoter-swap experiments are informative and suggest that specific promoter elements, with dedicated molecular mechanisms to circumvent the global transcriptional increase, are at least in part responsible for size-independent gene expression.

Here, we found that size-independent gene expression, where a constant number of molecules is made per cell cycle, is insufficient to guarantee that a constant number of molecules are inherited in newborn daughter cells. This problem is most acute when there is a major size asymmetry in cytokinesis such as in the case in budding yeast but also metazoan cells such as *D. melanogaster* neuroblasts or the early embryonic divisions of *C. elegans* (Chia et al., 2008; Sulston et al., 1983). In these cases, if proteins synthesized in constant amounts per cell cycle were simply partitioned in proportion to the relative volumes of the newborn cells, the size-independence would be lost. To maintain size-independent protein expression following division, cells can use chromatin-based partitioning mechanisms that harness the faithful segregation of sister chromosomes to partition near equal amounts to the two newborn cells (Fig. 5). While it has long been appreciated that big and small cells both have the same amount of DNA, we identified here a set of genes that are similarly size-independent. Size-independent synthesis is reflected in the function of many size-independent genes such as the histones, which require stoichiometry with the genome. It is both curious and elegant that the size-independent gene set includes *WHI5* which regulates DNA replication, while the DNA itself is used as a scaffold for the synthesis and maintenance of size-independent gene expression.

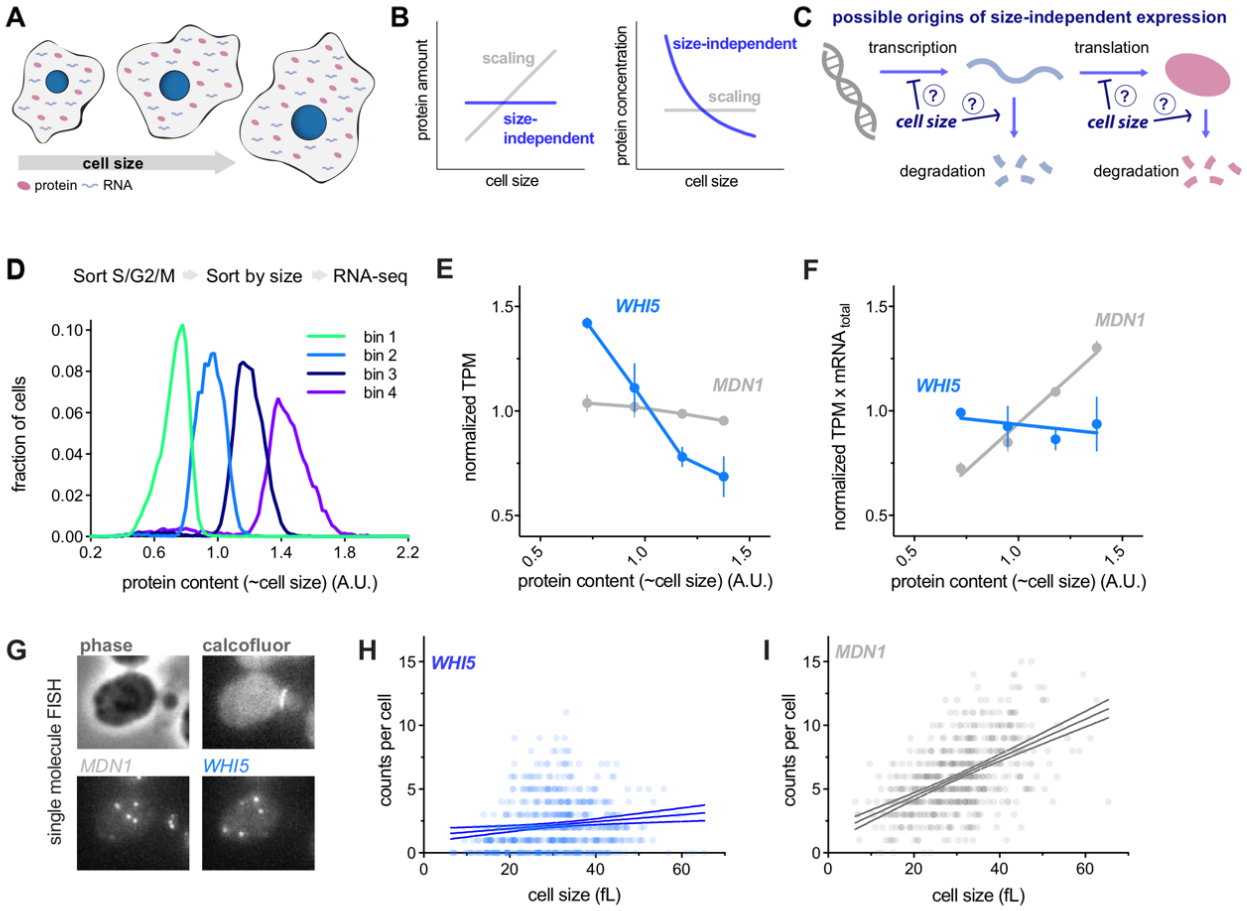


**Figure 1 | *WHI5* mRNA concentrations are uncorrelated with cell size**

(A-C) schematics illustrating gene expression in proportion to or independent of cell size. (A) Total protein and RNA copy numbers per cell generally increase in proportion to cell volume so that their concentrations remain constant during growth. (B) However, some proteins are expressed at constant amounts independent of cell size and are diluted by cell growth, which (C) could result from regulation at any step of gene expression.

(D-F) Cells in S/G2/M were sorted into four bins based on the intensity of total protein dye. See Fig. S1 and Materials and Methods for details. (D) Histogram of total protein content per cell in each bin measured after sorting. (E) Normalized Transcripts Per Million (TPM / mean TPM) for *WHI5* and *MDN1* mRNA in cells of different sizes (total protein content). The mean (+/- range) of two biological replicates is plotted. Changes in TPM are proportional to changes in mRNA concentration. (F) Normalized TPM x total-mRNA for *WHI5* and *MDN1* mRNA in cells of different sizes (total protein content). Mean (+/- range) of two biological replicates is plotted. Changes in TPM x total mRNA are proportional to changes in mRNA amount. Relative total mRNA per cell was determined by the number of reads relative to those from a fixed number of *S. pombe* cells added to the sample.

(G-I) single-molecule Fluorescence In Situ Hybridization (smFISH) analysis of *WHI5* and *MDN1* mRNA. (G) representative smFISH images. (H&I) mRNA counts per cell as a function of cell size for *WHI5* and *MDN1* determined by smFISH, n=567 cells. Linear regression (solid line) and 95% confidence interval (dashed lines) are shown. Data are pooled from two biological replicates. The same data with replicates plotted independently are shown in Fig. S3A&B.

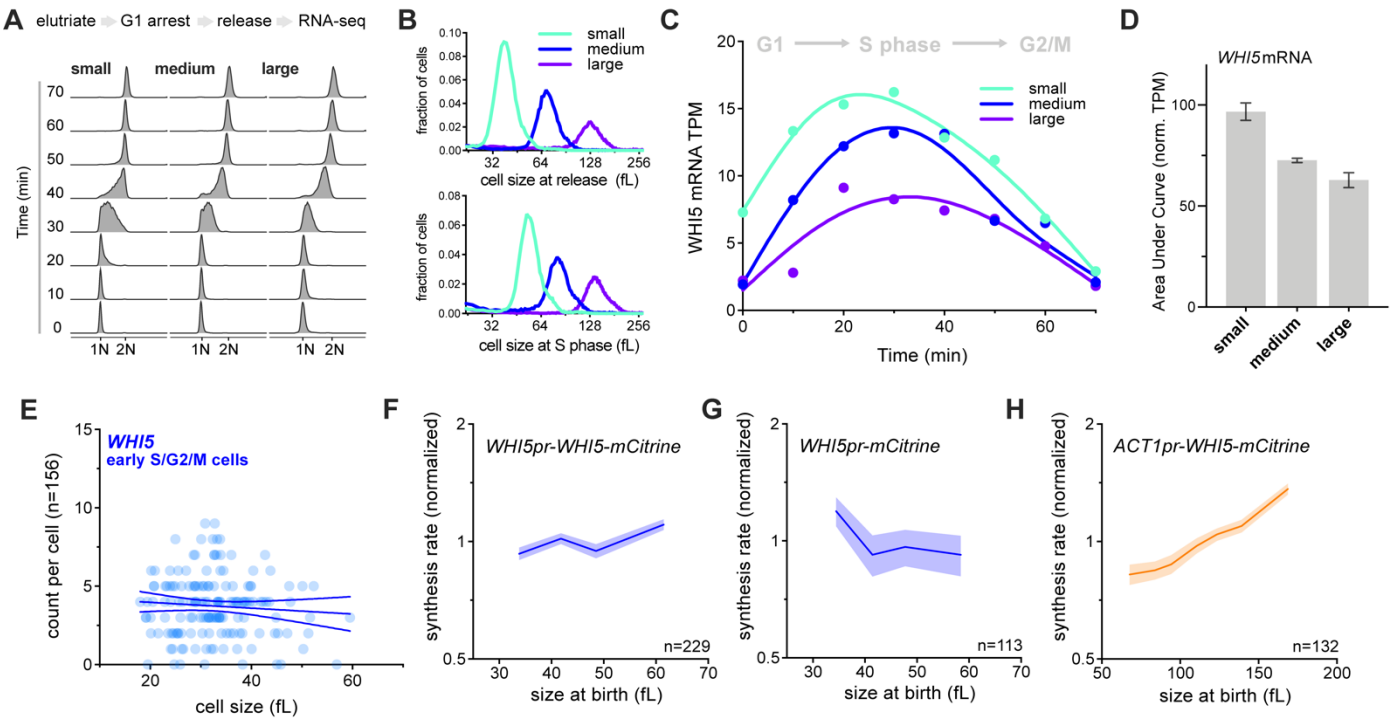


**Figure 2 | *WHI5* transcription is independent of cell size across the cell cycle**

(A-D) G1 cells of different sizes (small, medium or large) were arrested for increasing amounts of time in G1 using a temperature sensitive *cdc28-13* allele at 37°C. Cells were then released from G1 to progress synchronously through the cell cycle and analyzed by RNA-seq across the full cell cycle. See Fig. S2 for details. (A) DNA content analysis determined by flow cytometry of small, medium and large as they trasverse the cell cycle. (B) Size distribution of small, medium or large cells at point of release from G1 arrest (top panel) and at mid S-phase (bottom panel, corresponds to the 40 minute time point). (C) *WHI5* mRNA TPM for small, medium or large cells synchronously progressing through the cell cycle. (D) The Area Under the Curve (AUC) of mean normalized *WHI5* mRNA TPM for small, medium or large cells synchronously progressing through the cell cycle. The AUC mean (+/- range) of two biological replicates is plotted.

(E) mRNA counts per cell for *WHI5* as a function of cell size in early S/G2/M cells determined by smFISH, n=156 cells. Early S/G2/M cells were defined as budded cells with a small ( ≤ 0.2) bud-to-mother volume ratio. Linear regression (solid line) and 95% confidence interval (dashed lines) are shown. Data are pooled from two biological replicates. The same data with replicates plotted independently, including data for *MDN1*, are shown in Fig. S3E&F.

(F-H) Protein synthesis rates (normalized to the mean) as a function of cell volume at budding, measured by time-lapse fluorescence microscopy measuring Whi5-mCitrine expressed from (F) the endogenous *WHI5* promoter or (H) the *ACT1* promoter, and (G) mCitrine expressed alone from the *WHI5* promoter. Synthesis rates were determined as in (Schmoller et al., 2015) for single cells using linear fits of total protein traces for the period between bud emergence and cytokinesis (S/G2/M). Data are binned according to cell size at budding and the mean (+/- SEM) of each bin is plotted. Un-binned single-cell values of the same data are plotted in Fig. S2G-I.

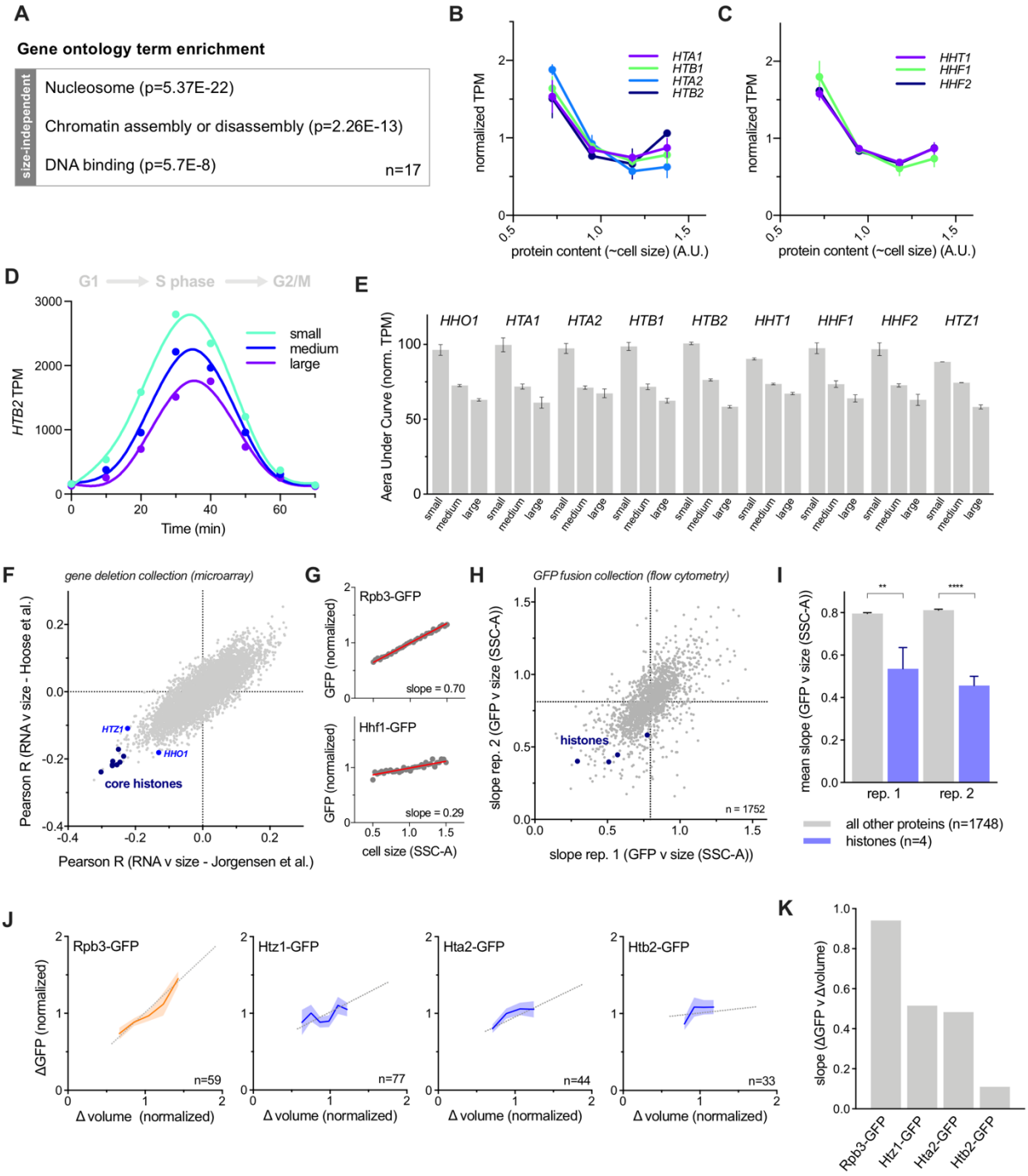


**Figure 3 | Histones are a rare class of size-independent genes**

(A) Gene ontology terms enriched in size-independent genes. 9 of the 17 size-independent genes encode histones and 1 is *WHI5*. See Fig. S4A and Materials and Methods for classification details.

(B&C) Normalized TPM (TPM / mean TPM) for size-independent histone mRNAs in cells of different sizes (total protein content). The mean (+/- range) of two biological replicates is plotted. Changes in TPM are proportional to changes in mRNA concentration. See Fig. S1 for experimental details.

(D) *HTB2* mRNA TPM for small, medium, and large cells synchronously progressing through the cell cycle as in Fig.1J&K. See Fig. S2 for experimental details.
(E) The Area Under the Curve (AUC) of mean normalized size-independent histone mRNA TPM of small, medium, and large cells synchronously progressing through the cell cycle. The AUC mean (+/- range) of two biological replicates is plotted.
(F) Pearson R correlation coefficient for the correlation between mRNA levels relative to wild-type in 1,484 gene deletion strains (Kemmeren et al., 2014; O'Duibhir et al., 2014) and the cell size of the respective gene deletions for two different data sets of size measurements (Jorgensen et al. 2002 (x-axis), Hoose et al. 2012 (y-axis)). Each point represents an individual gene. Histone mRNAs are shown in blue.
(G-I) Flow-cytometry analysis of size-dependent expression in the genome-wide collection of GFP fusion strains. Side scatter (SCC-A) was used as a proxy for cell size. The slope of the linear fit between cell size (SCC-A) and GFP intensity in budded cells was used to estimate the degree of size-dependence for each protein. See Materials and Methods for details.
(G) Plot of example protein-GFP levels (intensities, normalized to the mean intensity) against cell size (SCC-A, normalized to the mean SCC-A) in budded cells. Grey dots are bin means after binning on cell size. Red lines show the linear regression to the un-binned data.
(H) Slope values (for the linear regression between GFP and cell size in budded cells) of 1752 proteins analyzed in two replicates. Slopes closer to 0 correspond to more size-independent behavior. Histone proteins are shown in blue.
(I) Average slope values (for the linear regression between GFP and cell size in budded cells) for histones (blue) and all other proteins (grey). Four histones proteins were present in the 1752 proteins analyzed. Histone proteins have significantly smaller slopes than the average protein (replicate 1: $p=0.0014$, replicate 2: $p<0.0001$).
(J) The amount of Rpb3-GFP (RNA polymerase II subunit) and three histones (Hta2-GFP, Htb2-GFP and Htz1-GFP) synthesized ($\Delta$GFP, normalized to its mean) between birth and division as a function of the amount of growth ($\Delta$volume, normalized to its mean), determined by single cell time-lapse fluorescence microscopy. Data are binned according to $\Delta$volume and the bin means (+/- SEM) are plotted. Robust linear fit to un-binned data are shown (dashed grey line).
(K) Slope of the robust linear fits to $\Delta$GFP against $\Delta$volume as in (J).

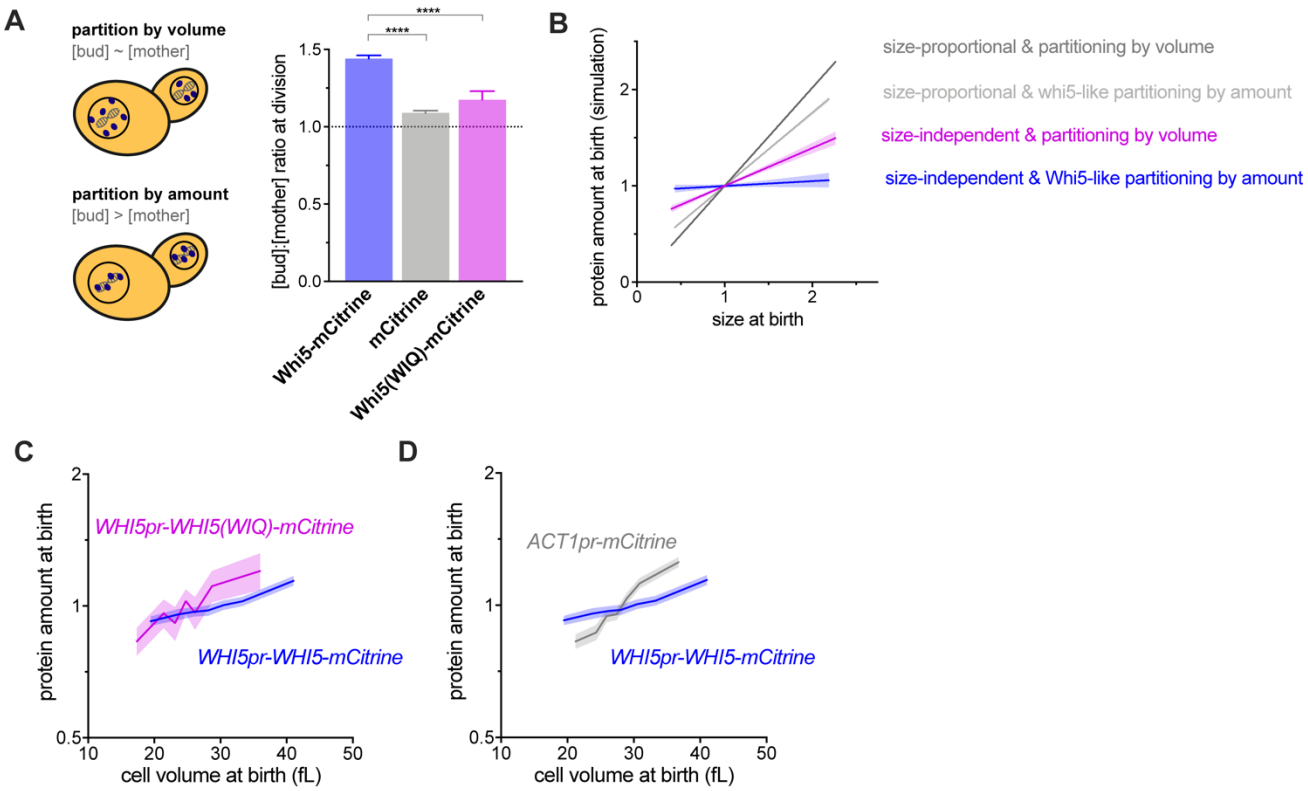


**Figure 4 | Size-independent expression is inherited through asymmetric division due to DNA-mediated partitioning**

(A) The bud-to-mother concentration ratios for Whi5-mCitrine, free mCitrine or Whi5(WIQ)-mCitrine at cytokinesis. Whi5(WIQ)-mCitrine has reduced recruitment to DNA because it does not bind and inhibit the SBF transcription factor (Travesa et al., 2013) (Fig. S6A). A ratio ~ 1 is expected for proteins that are partitioned in proportion to volume. A ratio > 1 is expected for proteins that are partitioned by protein amount (*i.e.*, independent of the mother and bud volume). **** $p < 0.0001$.

(B) Simulation of protein amount at birth as a function of daughter volume at birth. Four different conditions were simulated: proteins expression is either size-dependent or size-independent, and protein partitioning is either by protein amount or by cell volume. Linear fit to simulated cell size and protein amounts at birth are shown. Raw simulated values are plotted in Fig. S6B&C. See Materials and Methods for details of the full stochastic cell cycle simulation. The simulation predicts that partitioning by protein amount is required for full size-independence of protein amounts at birth.

(C&D) Protein amount at birth (normalized to the mean) as a function of cell size at birth for *WHI5pr-WHI5-mCitrine* and (C) *WHI5pr-WHI5(WIQ)-mCitrine* or (D) *ACT1pr-mCitrine* cells. Data are binned according to cell size at birth and the bin means (+/- SEM) are plotted. Un-binned single-cell values of the same data are plotted in Fig. S6D&E.

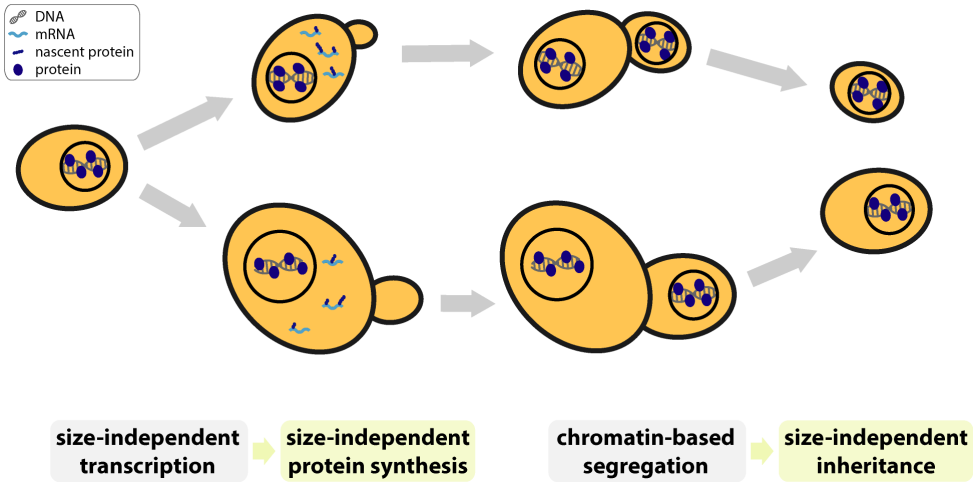


**Figure 5 | Summary schematic**

A small class of genes including the cell cycle inhibitor *WHI5* and histones are transcribed in a size-independent manner. This in turn, results in size-independent protein synthesis during the cell cycle. However, alone this is insufficient because, as is the case for most proteins, the asymmetry of cell division results in inheritance of fewer molecules in the smaller daughter cell. Thus, size-independent proteins must also be partitioned during cell division in a size-independent manner, which is achieved through chromatin-binding based partitioning.

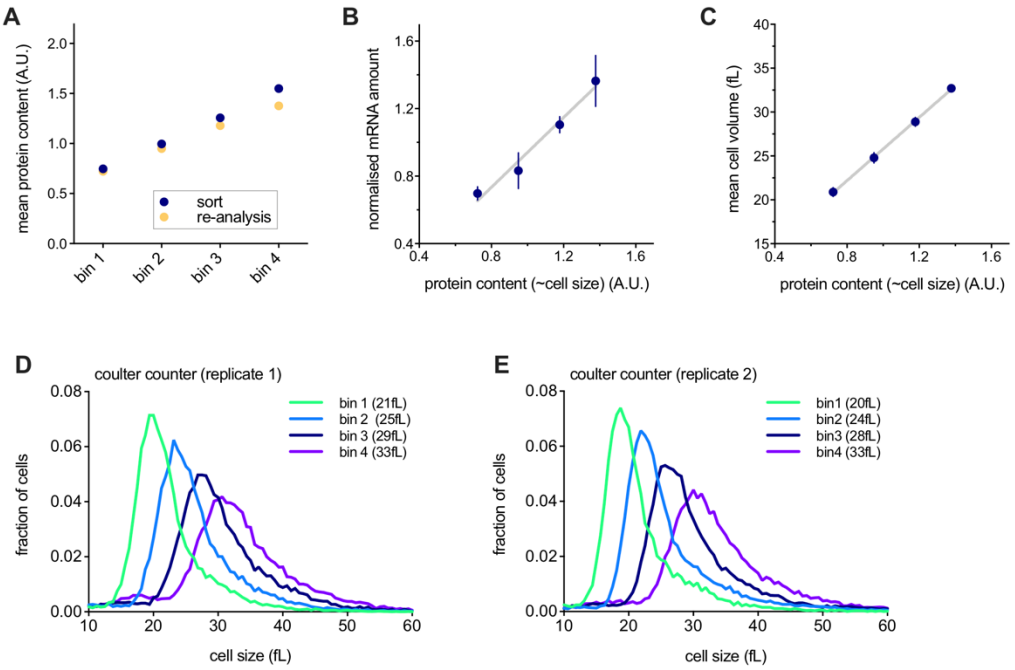


**Figure S1 | Transcriptomic analysis of cells sorted by cell size**

Related to Figures 1D-F.

(A-E) Cells in S/G2/M were sorted into four different bins based on the intensity of a total protein dye that we used as a proxy for cell size. See Materials and Methods for further details.

(A) Protein dye intensity (normalized to the mean) measured during sorting (blue) and re-measured for cells from each bin after the sort (yellow). The re-analysis (yellow) confirms high sort fidelity.

(B) The amount of mRNA (mean +/- SEM) per cell for each bin, determined by the number of reads relative to those from a fixed number of *S. pombe* spike-in cells added to the sample. Total protein amount is well correlated with total cellular mRNA amount.

(C-E) Cell volume of cells after sorting, measured by Coulter counter. Total protein amount is correlated with cell volume.

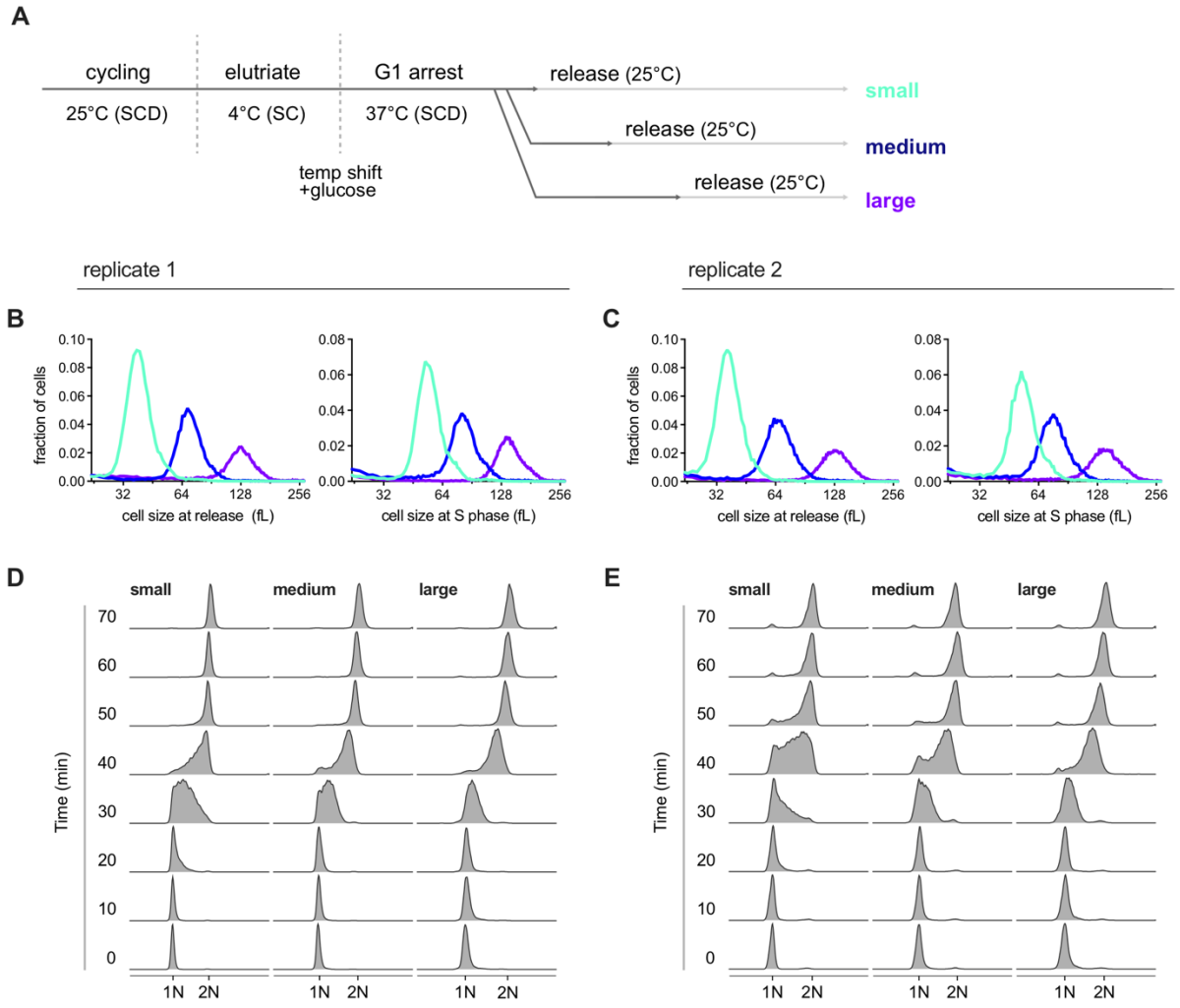


**Figure S2 | Transcriptomic analysis of different sized cells synchronously progressing through the cell cycle**

Related to Figures 1J-L.

(A) Schematic of experimental design. G1 cells harboring a temperature sensitive Cdk1 allele (*cdc28-13*) were recovered by centrifugal elutriation. Cells were then shifted to the restrictive temperature (37°C) to arrest them in G1. After increasing amounts of time during the arrest, cells reached small, medium or large cell sizes. The culture was then shifted to the permissive temperature (25°C) to allow cell cycle entry. Samples for RNA-seq were then collected every ten minutes at the eight time points corresponding to synchronous progression through the cell cycle. The 0 minute time point is designated as that 30 minutes before early S phase as determined by DNA-content analysis using flow cytometry (Fig. S2D&E). See Materials and Methods for further details.

(B&C) Cell size distributions, measured by Coulter counter, for small, medium, and large cells when they are released from the G1 arrest (left panel) or at mid S-phase (the 40 min time point, right panel) for two biological replicates.

329 (D&E) DNA content analysis of small, medium, and large cells synchronously progressing through the cell
330 cycle for two biological replicates.
331 Note that the data in Figures S2B&D is also presented in Figure 2A&B and is shown here for comparison
332 with replicate experiment (S2C&E).

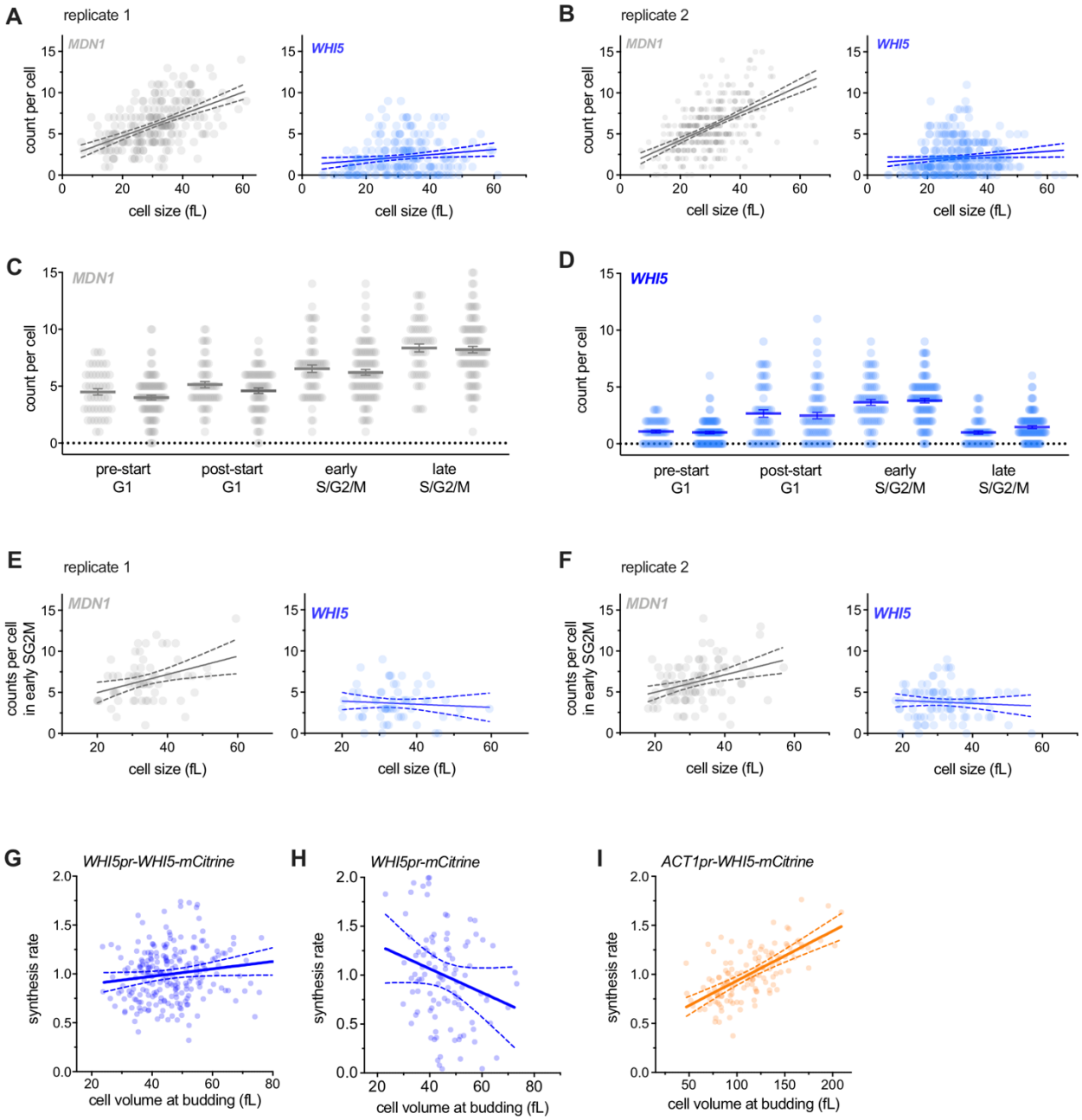


**Figure S3 | Single-cell analysis of *WHI5* mRNA expression**

Related to Figures 1G-I&M-P.

(A&B) mRNA counts per cell, measured by single molecule FISH, for *MDN1* or *WHI5* as a function of cell size in two independent biological replicates. All cells of any cell cycle stage are included. Linear regression (solid line) and 95% confidence interval (dashed lines) are shown. The same data are re-plotted in Fig. H&I with replicates pooled.

(C&D) mRNA counts per cell, measured by single molecule FISH, for *MDN1* or *WHI5* at different cell cycle stages in two independent biological replicates. Mean (solid line) +/- SEM (error bars) are shown. See Materials and methods for details of cell cycle stage classification.

(E&F) mRNA counts per cell, measured by single molecule FISH, for *MDN1* and *WHI5* as a function of cell size during early S/G2/M in two independent biological replicates. Early S/G2/M cells were defined as budded cells with a small ( ≤ 0.2) bud-to-mother volume ratio. Linear regression (solid line) and 95% confidence interval (dashed lines) are shown. The same data are re-plotted in Fig. 2E with replicates pooled.

(G-I) Protein synthesis rates (normalized to the mean) as a function of cell volume at budding, measured by time-lapse microscopy for Whi5-mCitrine fusion proteins expressed from (G) the endogenous *WHI5* promoter or (I) the *ACT1* promoter and (H) mCitrine expressed alone from the *WHI5* promoter. Synthesis rates were determined for single cells for the period between bud emergence and cytokinesis using a linear fit as in (Schmoller et al., 2015). Linear regression (solid line) and 95% confidence interval (dashed lines) are shown. (G) n=229, 2 data points outside the axis limits. (I) n=113, 4 data points outside the axis limits. (H) n=132. The same data, binned by cell volume at budding, are shown in Fig. 1F-H.

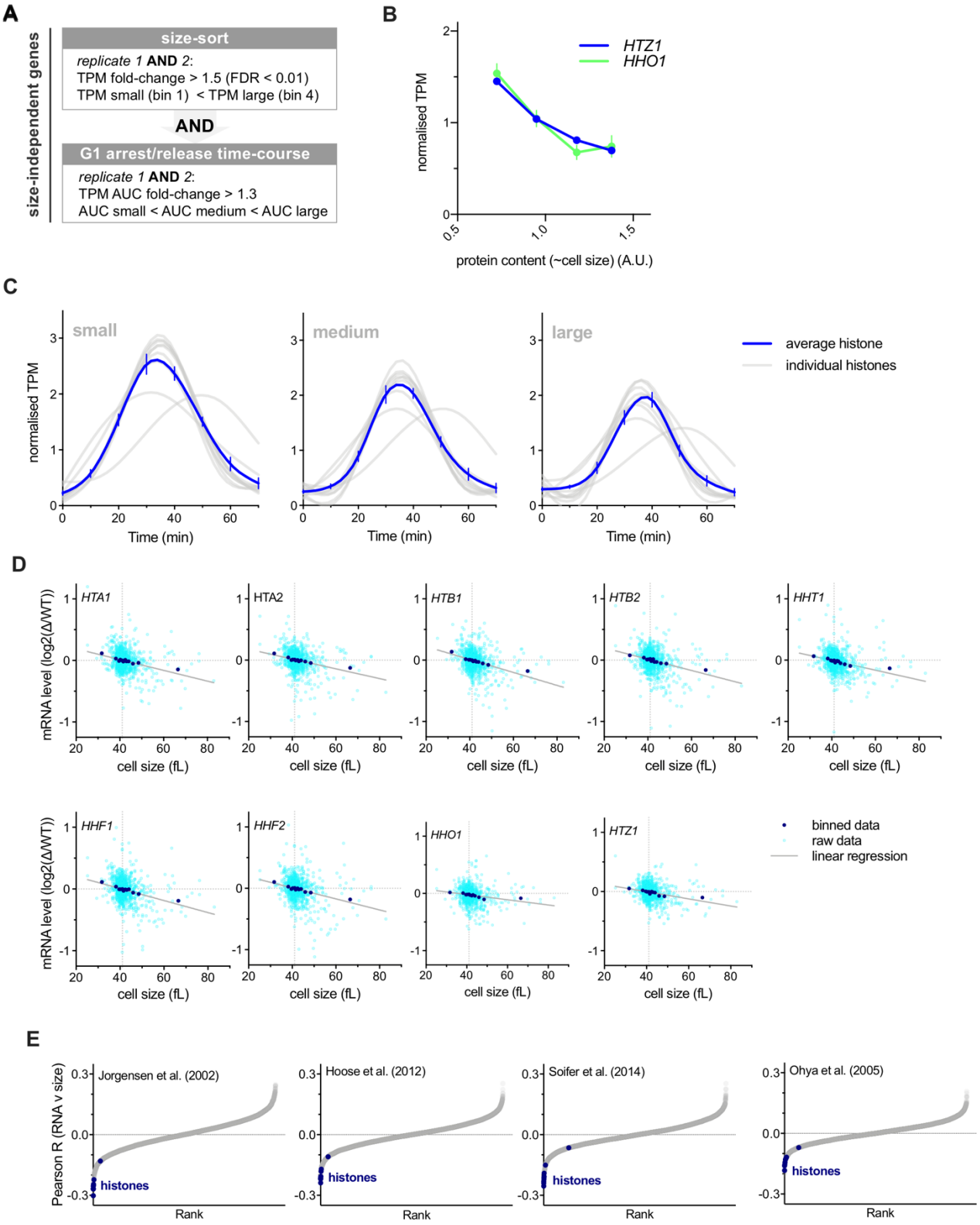


**Figure S4 | Histone gene expression is size-independent**

Related to Figures 2A-F.

(A) Criteria used to classify size-independent genes. Genes whose expression changed in a manner similar to *WHI5* in both replicates of the size-sort (see Fig. 1A-C) and in the G1 arrest/release time course (see Fig. 2A-D) were classified as size-independent genes. See Materials and Methods for further details.
(B) Normalized TPM (TPM / mean TPM) for *HTZ1* and *HHO1* mRNA in cells of different sizes (total protein content). Mean (+/- range) of two biological replicates is plotted. Changes in TPM are proportional to changes in mRNA concentration. See Figure S1 and Materials and Methods for further details.
(C) Normalized histone mRNA TPM (TPM / mean TPM) of small, medium or large cells at 10-minute time intervals synchronously progressing through the cell cycle. A smoothing spline fitted to the mean normalized TPM from two biological replicates for each histone is plotted in grey. A Spline fitted to mean (+/- SEM) of all histones is shown in blue.
(D) Comparison of histone mRNA levels in 1,484 gene deletions, relative to wild-type (Kemmeren et al., 2014; O'Duibhir et al., 2014) with cell size of the respective gene deletions from Jorgensen et al., 2002. See Materials and Methods for details.

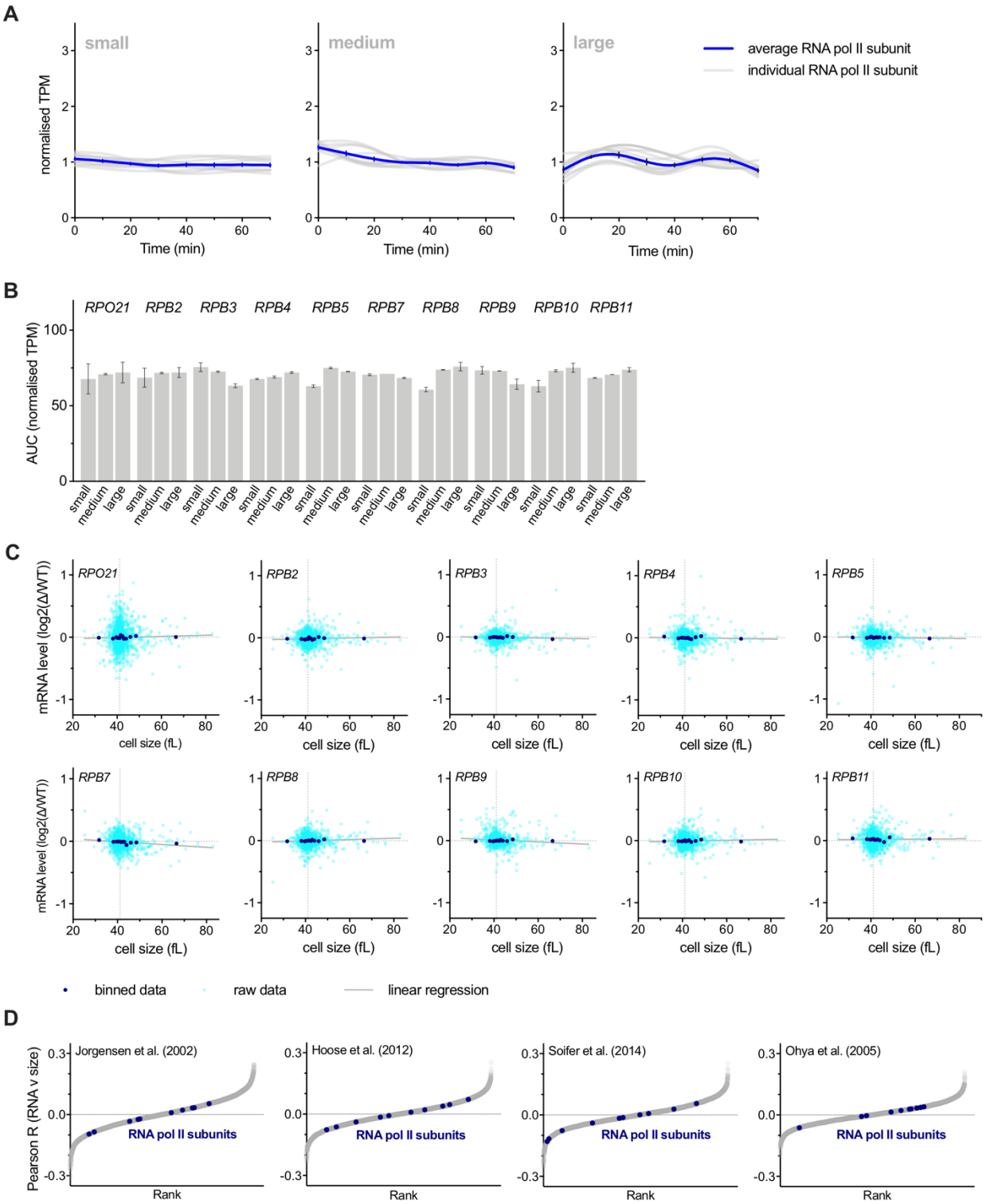


**Figure S5 | RNA polymerase II subunit gene expression is proportional to cell size**

(A) Normalized RNA polymerase II subunit mRNA TPM (TPM / mean TPM) for small, medium, and large cells at 10-minute time intervals synchronously progressing through the cell cycle. A smoothing-spline fitted

to the average normalized TPM from two biological replicates for each subunit is plotted in grey. A Spline fitted to the mean (+/- SEM) of all subunits is shown in blue.
(B) The Area Under the Curve (AUC) of mean normalized RNA polymerase II subunit mRNA TPM of small, medium, and large cells synchronously progressing through the cell cycle. The AUC mean (+/- range) of two biological replicates is plotted.

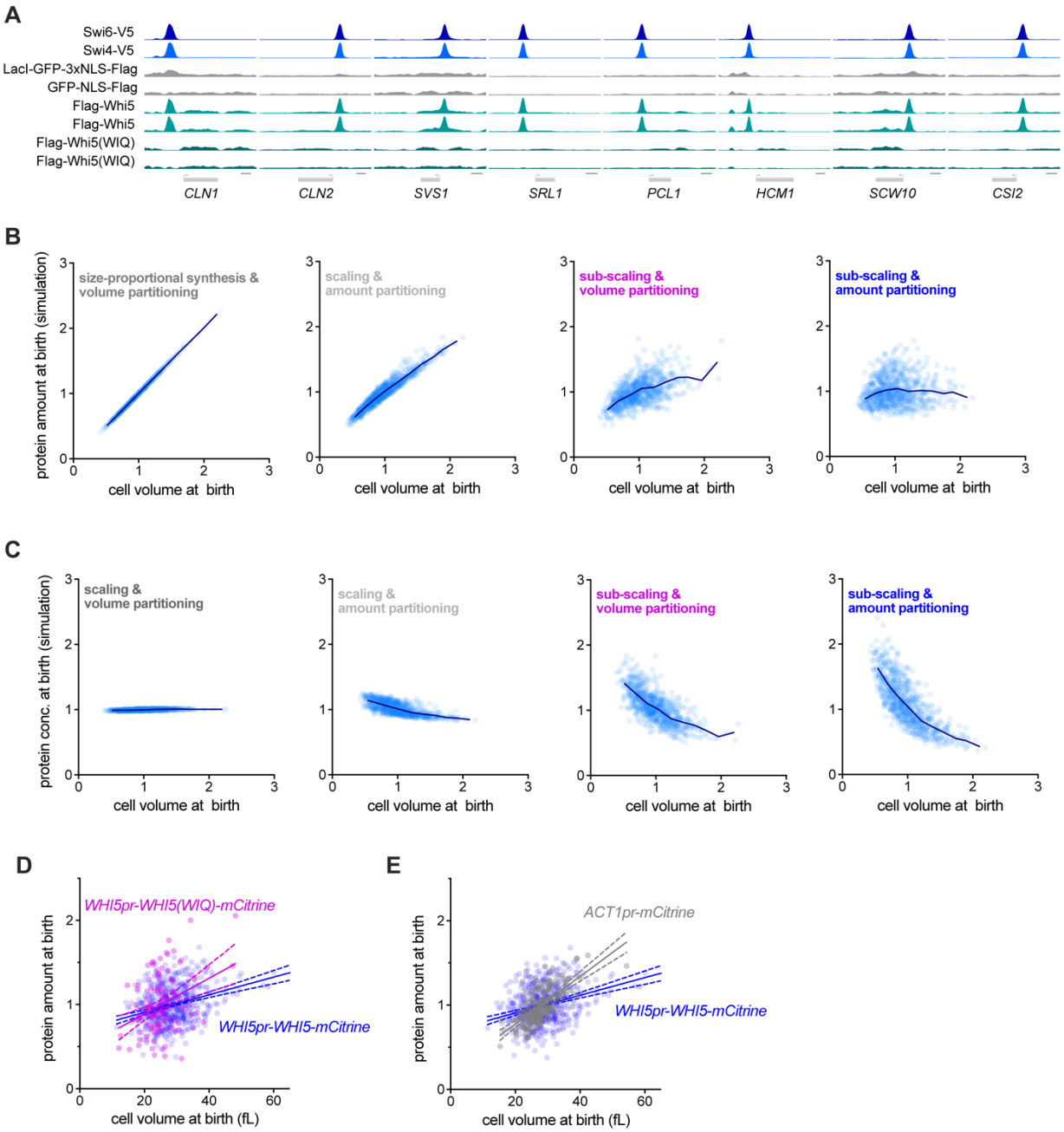


**Figure S6 | Size-independent expression is inherited through asymmetric division due to DNA-mediated partitioning**

Related to Figure 3.

(A) anti-V5 or anti-Flag ChIP-seq experiments were performed using cells expressing the indicated fusion proteins (left hand side). Data are shown for 8 example SBF binding sites near the denoted genes. Cells expressing LacI-GFP-3xNLS-FLAG or GFP-NLS-FLAG were included as controls for non-specific ChIP

signal. Whi5 recruitment to DNA overlaps with SBF (Swi4 and Swi6) binding sites but is lost in the Whi5(WIQ) variant (Travesa et al., 2013).

(B&C) Simulation of (B) protein amount or (C) protein concentration at birth as a function of daughter volume at birth. Four different conditions were simulated where protein expression was either in proportion to cell size or independent of cell size, and protein partitioning is either by amount or in proportion to cell volume. Individual simulated cells (light blue) as well as bin means (dark blue) are plotted. Note binned values for (B) are also plotted in Figure 3B.

(D&E) Protein amount at birth for *WHI5pr-WHI5-mCitrine* and (D) *WHI5pr-WHI5(WIQ)-mCitrine* or (E) *ACT1pr-mCitrine* cells. Note, Fig. 4C&D shows bin means of the same data binned by cell volume at birth.

401 **Table S1 – list of *S. cerevisiae* strain used in this study and full genotypes**

| | | |
|---|---|---|
| A17896 | *Amon lab* | *W303; MATa, ADE2, GAL, cdc28-13* |
| DCB99 | *This study* | *W303; MATa, ADE2, WHI5-mCitrine::URA3* |
| MK551-1 | *This study* | *W303;* MATa, bar1Δ, HISG, whi5Δ::LEU2, cln3Δ::HphMX |
| *HTA1-GFP* | *GFP collection (Huh et al., 2003)* | *BY4741; HTA1-GFP* |
| *HTB2-GFP* | *GFP collection (Huh et al., 2003)* | *BY4741; HTB2-GFP* |
| *HTZ1-GFP* | *GFP collection (Huh et al., 2003)* | *BY4741; HTZ1-GFP* |
| *RPB3-GFP* | *GFP collection (Huh et al., 2003)* | *BY4741; RPB3-GFP* |
| DCB72 | *This study* | *W303; MATa, ADE2, ura3:: WHI5pr(1kb)-mCitrine-CYC1term* |
| KSY108-1 | *Lab collection* | *W303; ADE2, WHI5-mCitrine-HIS3* |
| KSY190-2 | *This study* | *W303; ADE2, whi5Δ::KanMX URA3::WHI5pr(1kb)-WHI5-WIQ-mCitrine-ADH1term* |
| KSY160-2 | *This study* | *W303; ADE, URA3::ACT1pr(1kb) -Whi5-mCitrine-CYC1term* |
| KSY158-1 | *Lab collection* | *W303; ADE2, URA3:: ACT1pr(1kb)-mCitrine-CYC1term* |
| *MS534* | *This study* | *W303; bar1Δ::HisG GFP-3xNLS-FLAG* |
| *MS535* | *This study* | *W303; bar1Δ::HisG LacI-GFP-3xNLS-FLAG* |
| *MS536* | *This study* | *W303; bar1Δ::HisG; whi5Δ::LEU2; URA3::WHI5pr(1kb)-3xFLAG-WHI5-CYC1term* |
| *MS537* | *This study* | *W303; bar1Δ::HisG; URA3::WHI5pr(1kb)-3xFLAG-WHI5(WIQ)-CYC1term* |
| *MK653-1* | *This study* | *W303; bar1Δ; SWI4-V5::hphMX6* |
| *MK645-1* | *This study* | *W303; bar1Δ; SWI6-V5::hphMX6* |

402

403 **REFERENCES**

404

405 Chandler-Brown, D., Schmoller, K.M., Winetraub, Y., and Skotheim, J.M. (2017). The Adder
406 Phenomenon Emerges from Independent Control of Pre- and Post-Start Phases of the Budding
407 Yeast Cell Cycle. Curr Biol *27*, 2774-2783 e2773.
408 Chen, Y., Zhao, G., Zahumensky, J., Honey, S., and Futcher, B. (2020). Differential Scaling of Gene
409 Expression with Cell Size May Explain Size Control in Budding Yeast. Mol Cell *78*, 359-370 e356.
410 Chia, W., Somers, W.G., and Wang, H. (2008). Drosophila neuroblast asymmetric divisions: cell
411 cycle regulators, asymmetric protein localization, and tumorigenesis. J Cell Biol *180*, 267-272.
412 Creanor, J., and Mitchison, J.M. (1982). Patterns of protein synthesis during the cell cycle of the
413 fission yeast Schizosaccharomyces pombe. J Cell Sci *58*, 263-285.
414 Cross, S.L., and Smith, M.M. (1988). Comparison of the structure and cell cycle expression of
415 mRNAs encoded by two histone H3-H4 loci in Saccharomyces cerevisiae. Mol Cell Biol *8*, 945-954.
416 Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M.,
417 and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.
418 Doncic, A., Eser, U., Atay, O., and Skotheim, J.M. (2013). An algorithm to automate yeast
419 segmentation and tracking. PLoS One *8*, e57970.
420 Doncic, A., Falleur-Fettig, M., and Skotheim, J.M. (2011). Distinct interactions select and maintain a
421 specific cell fate. Mol Cell *43*, 528-539.
422 Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and
423 visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics *10*, 48.
424 Elliott, S.G. (1983). Coordination of growth with cell division: regulation of synthesis of RNA during
425 the cell cycle of the fission yeast Schizosaccharomyces pombe. Mol Gen Genet *192*, 204-211.
426 Elliott, S.G., and McLaughlin, C.S. (1979). Regulation of RNA synthesis in yeast. III. Synthesis
427 during the cell cycle. Mol Gen Genet *169*, 237-243.
428 Elliott, S.G., Warner, J.R., and McLaughlin, C.S. (1979). Synthesis of ribosomal proteins during the
429 cell cycle of the yeast Saccharomyces cerevisiae. J Bacteriol *137*, 1048-1050.
430 Feng, J., Liu, T., Qin, B., Zhang, Y., and Liu, X.S. (2012). Identifying ChIP-seq enrichment using
431 MACS. Nat Protoc *7*, 1728-1740.
432 Fraser, R.S., and Nurse, P. (1978). Novel cell cycle control of RNA synthesis in yeast. Nature *271*,
433 726-730.
434 Fraser, R.S., and Nurse, P. (1979). Altered patterns of ribonucleic acid synthesis during the cell
435 cycle: a mechanism compensating for variation in gene concentration. J Cell Sci *35*, 25-40.
436 Gunjan, A., and Verreault, A. (2003). A Rad53 kinase-dependent surveillance mechanism that
437 regulates histone protein levels in S. cerevisiae. Cell *115*, 537-549.
438 Hoose, S.A., Rawlings, J.A., Kelly, M.M., Leitch, M.C., Ababneh, Q.O., Robles, J.P., Taylor, D.,
439 Hoover, E.M., Hailu, B., McEnery, K.A.*, et al.* (2012). A systematic analysis of cell cycle regulators in
440 yeast reveals that most factors act independently of cell size to control initiation of division. PLoS
441 Genet *8*, e1002590.
442 Hu, B., Petela, N., Kurze, A., Chan, K.L., Chapard, C., and Nasmyth, K. (2015). Biological
443 chromodynamics: a general method for measuring protein occupancy across the genome by
444 calibrating ChIP-seq. Nucleic Acids Res *43*, e132.
445 Huh, W.K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O'Shea, E.K.
446 (2003). Global analysis of protein localization in budding yeast. Nature *425*, 686-691.
447 Jorgensen, P., Edgington, N.P., Schneider, B.L., Rupes, I., Tyers, M., and Futcher, B. (2007). The
448 size of the nucleus increases as yeast cells grow. Mol Biol Cell *18*, 3523-3532.
449 Jorgensen, P., Nishikawa, J.L., Breitkreutz, B.J., and Tyers, M. (2002). Systematic identification of
450 pathways that couple cell growth and division in yeast. Science *297*, 395-400.
451 Keifenheim, D., Sun, X.M., D'Souza, E., Ohira, M.J., Magner, M., Mayhew, M.B., Marguerat, S., and
452 Rhind, N. (2017). Size-Dependent Expression of the Mitotic Activator Cdc25 Suggests a Mechanism
453 of Size Control in Fission Yeast. Curr Biol *27*, 1491-1497 e1494.

Kemmeren, P., Sameith, K., van de Pasch, L.A., Benschop, J.J., Lenstra, T.L., Margaritis, T., O'Duibhir, E., Apweiler, E., van Wageningen, S., Ko, C.W., *et al.* (2014). Large-scale genetic perturbations reveal regulatory networks and an abundance of gene-specific repressors. Cell *157*, 740-752.

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.

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.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol *15*, 550.

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

Marguerat, S., Schmidt, A., Codlin, S., Chen, W., Aebersold, R., and Bahler, J. (2012). Quantitative analysis of fission yeast transcriptomes and proteomes in proliferating and quiescent cells. Cell *151*, 671-683.

Moran, L., Norris, D., and Osley, M.A. (1990). A yeast H2A-H2B promoter can be regulated by changes in histone gene copy number. Genes Dev *4*, 752-763.

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.

Norris, D., and Osley, M.A. (1987). The two gene pairs encoding H2A and H2B play different roles in the Saccharomyces cerevisiae life cycle. Mol Cell Biol *7*, 3473-3481.

O'Duibhir, E., Lijnzaad, P., Benschop, J.J., Lenstra, T.L., van Leenen, D., Groot Koerkamp, M.J., Margaritis, T., Brok, M.O., Kemmeren, P., and Holstege, F.C. (2014). Cell cycle population effects in perturbation studies. Mol Syst Biol *10*, 732.

Ohya, Y., Sese, J., Yukawa, M., Sano, F., Nakatani, Y., Saito, T.L., Saka, A., Fukuda, T., Ishihara, S., Oka, S., *et al.* (2005). High-dimensional and large-scale phenotyping of yeast mutants. Proc Natl Acad Sci U S A *102*, 19015-19020.

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.

Parts, L., Liu, Y.C., Tekkedil, M.M., Steinmetz, L.M., Caudy, A.A., Fraser, A.G., Boone, C., Andrews, B.J., and Rosebrock, A.P. (2014). Heritability and genetic basis of protein level variation in an outbred population. Genome Res *24*, 1363-1370.

Pramila, T., Wu, W., Miles, S., Noble, W.S., and Breeden, L.L. (2006). The Forkhead transcription factor Hcm1 regulates chromosome segregation genes and fills the S-phase gap in the transcriptional circuitry of the cell cycle. Genes Dev *20*, 2266-2278.

Qu, Y., Jiang, J., Liu, X., Wei, P., Yang, X., and Tang, C. (2019). Cell Cycle Inhibitor Whi5 Records Environmental Information to Coordinate Growth and Division in Yeast. Cell Rep *29*, 987-994 e985.

Raj, A., and Tyagi, S. (2010). Detection of individual endogenous RNA transcripts in situ using multiple singly labeled probes. Methods Enzymol *472*, 365-386.

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

Schmoller, K.M., Turner, J.J., Koivomagi, M., and Skotheim, J.M. (2015). Dilution of the cell cycle inhibitor Whi5 controls budding-yeast cell size. Nature *526*, 268-272.

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.

Soifer, I., and Barkai, N. (2014). Systematic identification of cell size regulators in budding yeast. Mol Syst Biol *10*, 761.

Sulston, J.E., Schierenberg, E., White, J.G., and Thomson, J.N. (1983). The embryonic cell lineage of the nematode Caenorhabditis elegans. Dev Biol *100*, 64-119.

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.

Travesa, A., Kalashnikova, T.I., de Bruin, R.A., Cass, S.R., Chahwan, C., Lee, D.E., Lowndes, N.F., and Wittenberg, C. (2013). Repression of G1/S transcription is mediated via interaction of the GTB motifs of Nrm1 and Whi5 with Swi6. Mol Cell Biol *33*, 1476-1486.

Trcek, T., Chao, J.A., Larson, D.R., Park, H.Y., Zenklusen, D., Shenoy, S.M., and Singer, R.H. (2012). Single-mRNA counting using fluorescent in situ hybridization in budding yeast. Nat Protoc *7*, 408-419.

Tutucci, E., Vera, M., Biswas, J., Garcia, J., Parker, R., and Singer, R.H. (2018). An improved MS2 system for accurate reporting of the mRNA life cycle. Nat Methods *15*, 81-89.

Youk, H., Raj, A., and van Oudenaarden, A. (2010). Imaging single mRNA molecules in yeast. Methods Enzymol *470*, 429-446.

Zatulovskiy, E., Zhang, S., Berenson, D.F., Topacio, B.R., and Skotheim, J.M. (2020). Cell growth dilutes the cell cycle inhibitor Rb to trigger cell division. Science *369*, 466-471.

Zenklusen, D., Larson, D.R., and Singer, R.H. (2008). Single-RNA counting reveals alternative modes of gene expression in yeast. Nat Struct Mol Biol *15*, 1263-1271.

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.

**Acknowledgments.** We would like to thank Bruce Futcher and Yuping Chen for invaluable advice on centrifugal elutriation, members of the Skotheim laboratory for constructive feedback, Gabriel Neurohr for sending strain A17896, Chris You for assistance with the ChIP-seq experiments and Jon Turner for help in optimizing the smFISH protocol. 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. KMS was supported by the Human Frontier Science Program (Postdoctoral Fellowship and Career Development Award).

**Author contributions**. JMS supervised the work. MPS and JMS wrote the manuscript. MPS performed and analyzed all experiments except for live cell microscopy experiments performed and analyzed by DCB & KMS and smFISH experiments performed and analyzed by ML. Sequencing data analysis was performed by MPS and GM, supervised by WG & AK.

**Conflict of interests.** The authors declare no conflicts of interest.

## MATERIALS AND METHODS

### Yeast genetics

Standard procedures were used for *Saccharomyces cerevisiae* strain construction. Full genotypes of all strains used in this study are listed in Table S1.

### Transcriptomic analysis of different sized cells synchronously progressing through the cell cycle

To determine transcript levels in cells of different sizes, S/G2/M cells were sorted according to total protein content by fluorescence activated cell sorting (FACS) before RNA extraction and sequencing. 500ml *S. cerevisiae* (*HTB2-GFP*) was grown (synthetic complete media + 2 % glucose at 30°C), and fixed at O.D. ~ 0.3 by addition of 500ml 80% methanol 20mM TRIS (-20°C) and then incubated at -20°C for 30 minutes. Cells were fixed to prevent gene expression changes during the course of the cell sorting, which requires multiple generations equivalents of time to complete. Cells were pelleted (13krpm, 3 minutes) and washed 3x in PBS, before gentle sonication and then addition of 5µg/ml total protein dye (Alexa Fluor™ 647 NHS Ester dye; ThermoFisher Scientific; A20006) and incubation (4°C, 30 minutes). Cells were again pelleted (13krpm, 3 minutes) and washed 3x in PBS to remove excess dye before again being sonicated. Cells from four different size fractions were sorted on a FACSAria II sorter (BD Biosciences) according to the following strategy. First singlets were gated based on scatter (FSC and SSC), then S/G2/M cells were identified using an Htb2-GFP signal, and then finally, four bins of different total protein content cells were sorted based on the total protein intensity (bin 1 = lowest signal, bin 4 = highest, Fig. S1A). The fidelity of the total protein dye sort was confirmed by re-analyzing 10,000 cells on the same sorter (Fig. 1D & S1A). Furthermore, total protein content was validated as a proxy for size by measuring the cell volume of the different protein dye sorted cells using a Coulter counter (Fig. S1C-E).

Two biological replicates were performed. Within each biological replicate, two technical replicates were performed for each size bin so that four replicates were performed per bin in total. For the different bins within the same replicate set, a constant number of *S. pombe* cells, fixed as above, were added as a spike-in to measure total RNA content per *S. cerevisiae* cell. The number of *S. cerevisiae* cells and *S. pombe* cells (972 h-) per sample was constant within a set of replicates but varied slightly between each set of replicates (5-10 million *S. cerevisiae* cells and 5-10 million *S. pombe* cells). To maximize the number of reads from the experimental *S. cerevisiae* samples, *S. pombe* cells were nitrogen starved (grown in EMM before media switch to EMM - $NH_4Cl$ for 24 hr at 30°C) because this reduces their mRNA copy number per cell (and their cell size) (Marguerat et al., 2012). Cells were then pelleted (4krpm, 15 minutes), and their RNA extracted and sequenced as

described below. For each set of replicates, the *S. pombe* spike-in was added independently and RNA was extracted independently.

To estimate the relative amount of mRNA per cell in each size bin, the number of *S. cerevisiae* reads per *S. pombe* read was calculated (see RNA-seq data processing below) and then normalized to the mean value within a set of replicates. The normalized total mRNA per sample was then averaged between the four replicates (Fig. S1B).

**Transcriptomic analysis of different sized cells synchronously progressing through the cell cycle**

To determine transcript levels during the cell cycle in cells of different sizes, cells were elutriated and arrested in G1 for different amounts of time. Samples were then collected during the synchronous progression from G1 through S, G2 and M phases of the cell cycle for RNA extraction and sequencing (See Fig. S2 for schematic of experimental design). Specifically, 4L *S. cerevisiae* (A17896: *W303 cdc28-13*) were grown in synthetic complete (SC) media with 2% glucose at 25°C to OD ~0.75 and then collected on a filter membrane and resuspend in ice-cold SC media (no carbon source). Cells were then sonicated (3x 20s, 3 minutes on ice 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 (2.4krpm, 4°C). The elutriation chambers were pre-equilibrated and run with SC media (4°C, no carbon source). The pump speed was gradually increased until G1 cells with minimal debris were being collected. G1 fractions were then collected on a filter and resuspended in 37°C conditioned SC media + 2% glucose in a 37°C shaking water-bath (OD ~ 0.1). The G1 arrest was maintained at 37°C until cells reached either 36-39fL (small), 67-69fL (medium) or 129-131fL (large) as determined by Coulter counter (Fig. S2B&C). When they reached our target size, cells were released from the G1 arrest. To do this, cells were collected on a filter membrane and resuspended in 25°C SC media + 2 % glucose (OD ~0.35). Samples for size measurement by Coulter counter, DNA-content analysis, and RNA-extraction were taken at 10 minute intervals after release with the 0 minute time point being designated as the time point 30 minutes before the onset of DNA replication (Fig. S2D&E). For small cells, the 0 minute time point was collected 40-50 minutes after the shift to the permissive temperature, for medium cells the 0 minutes time point was collected 10 minutes after the shift to the permissive temperature and for large cells the 0 minute time point was collected 0 minutes after shift to the permissive temperature. Two biological replicates were performed.

For DNA-content analysis, 0.4ml culture was added to 1ml 100% 4°C ethanol and stored at 4°C. Cells were pelleted (13krpm, 2 minutes), washed, and resuspended in 50mM Sodium Citrate (pH = 7.2), incubated with 0.2mg/ml RNAse A (overnight, 37°C) and then 0.4mg/ml proteinase K (1 hour, 50°C) before addition of 25μM Sytox Green (ThermoFisher Scientific). Cells were then sonicated and DNA-content was analyzed for >10000 events on a FACScan Analyzer (BD Biosciences). For RNA-

extraction 1.5ml cells were pelleted (13krpm, 30 seconds) and snap frozen in liquid $N_2$. Samples were then thawed in TRI Reagent (Zymo Rearch) and RNA was extracted as described below (RNA extraction and sequencing).

**RNA extraction and sequencing**

To extract RNA, cell pellets were lysed in 300µl TRI Reagent (Zymo Research) by bead beating using a Fastprep 24 (4°C, settings: 5.0 m/s, 1x30s). Cell debris was pelleted (13 krpm, 5 minutes) and the supernatant recovered. RNA was then extracted using the direct-zol RNA microprep kit (Zymo Research). mRNA was enriched using the NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB, E7490) and NEBNext Ultra II RNA Library Prep Kit for Illumina® (NEB, #E7775) was then used to prepare libraries for paired-end (2x150bp) Illumina sequencing (Novogene). More than 20 million reads were sequenced per sample.

**RNA-seq data processing**

Because some samples analyzed in this study contained *S. cerevisiae* as well as reference spike-in *S. pombe* RNA, a combined *S. cerevisiae* and *S. pombe* genome file was created using the sacCer3 and ASM294v2 versions of the respective genomes and a combined transcriptome annotation was created using the *S. pombe* gene models available from PomBase (Lock et al., 2019) and an *S. cerevisiae* set of gene models updated using transcript-end mapping data as previously described (Shipony et al., 2020). 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 (Langmead et al. 2009; version 1.0.1; settings: -e 200 -a -X 1000). Alignments were then quantified using eXpress (version 1.5.1) (Roberts and Pachter, 2013) before effective read count values and TPM (Transcripts Per Million transcripts) were then separated for each genome and renormalized TPMs were calculated with respect to the total reads for *S. cerevisiae*.

Differential expression analysis by DESeq2 was performed using technical replicates to compare RNA-seq data from different size bins in the experiment shown in Figures 1D-F and S1 (Love et al., 2014). To calculate the total amount of transcription during the G1 arrest/release RNA-seq time course experiment (Figures 1J-L and S2), TPM values were normalized to the mean for each

experiment and the Area Under the Curve (AUC) for TPM as a function of time was calculated for each time course using the R function auc(*type* = *"spline"*) from the *MESS* package.

**Classification of size-independent transcripts**

To classify transcripts whose expression was independent of cell size, we analyzed data from two experiments: (1) the RNA-seq experiment on size-sorted populations of cells (Fig. 1D-F and S1) and (2) the G1 arrest/release RNA-seq time course experiment (Fig. 2A-D and S2). Two biological replicates of each experiment were performed. Sub-scaling genes were classified as genes that passed the following criteria in both biological replicates of each experiment:

(1) At least one pair-wise comparison between the four-size bins has a false-discovery rate (FDR) adjusted p-value <0.01, a TPM fold-change > 1.5, and bin 1 TPM > bin 4 TPM.

(2) Small cells' TPM Area Under Curve (AUC) > medium cells' TPM AUC > large cells' TPM AUC and TPM AUC fold-change > 1.3. See above (RNA-seq data processing) for details of the AUC calculation.

**GO term enrichment**

GO term enrichment (Fig. 3A) was performed using the GOrilla GO analysis tool (Eden et al., 2009). Enrichment of size-independent gene transcripts was performed versus a background set of all genes that had a TPM value > 0 in all RNA-seq samples.

**Single Molecule Fluorescence In Situ Hybridization (smFISH)**

smFISH was used to image *WHI5* and *MDN1* mRNAs in single cells. A Whi5-mCitrine tagged strain (DCB099) was used to discriminate pre- and post-*Start* G1 based on Whi5-mCitrine nuclear localization (Doncic et al., 2011). Early S/G2/M cells were defined as budded cells with a small (≤ 0.2) bud-to-mother volume ratio. Cells were grown at 30°C in synthetic complete (SC) media + 2 % glycerol + 1 % ethanol. Two biological replicates were performed. Each biological replicate contained two technical replicates (*i.e.*, two independent hybridizations to cells from the same culture). In addition, two negative controls were performed regularly where (i) FISH probes were omitted and (ii) a *whi5*Δ (MK551-1) strain was analyzed.

The smFISH protocol (detailed below) was optimized based on protocols from multiple prior studies (Raj and Tyagi, 2010; Trcek et al., 2012; Tutucci et al., 2018; Youk et al., 2010; Zenklusen et al., 2008). 45ml cells (OD600 ~0.2) were fixed with 5 mL 37 % formaldehyde and incubated (45 minutes, room temperature, rotating). Cells were then pelleted (1600 g, 5 minutes) and washed twice in 1 mL of ice-cold fixation buffer (pH 7.5, 218 mg/mL sorbitol, 84 mM potassium phosphate dibasic, 16 mM potassium phosphate monobasic, dissolved in water for RNA work (Thermo Fisher Scientific, BP561-1)). Cells were again pelleted and resuspended in 900 µL fixation buffer and gently sonicated before 100 µL of 200 mM RNase inhibitor vanadyl ribonucleoside complex was added (New England

BioLabs, S1402S). Cells were then digested by adding 3.5-5 µL zymolyase stock (5 mg/mL 100T, MP Biomedicals, 0832093) and incubated (70-80 minutes, 30° C, rotating). Cells were then pelleted (400 g, 6 minutes) and washed twice in 1 mL of ice-cold fixation buffer to stop digestion and finally permeabilized by resuspension in 1ml 70 % ethanol. Permeabilized cells were kept at 4°C for 1 to 3 days. 300 µL of the permeabilized cells per hybridization sample were then pelleted (400 g, 7 minutes) and washed in 500 µL of wash buffer A (Biosearch Technologies, SMF-WA1-60: prepared fresh on the day of use according to manufacturer's instructions always using a fresh aliquot of deionised formamide (EMD Millipore, S4117, stored at -20° C)). Permeabilized cells were then resuspended in 100 µL hybridisation solution (Biosearch Technologies, SMF-HB1-10) containing 1-3x of standard probe concentrations for *WHI5* and *MDN1* probes, 10 mM VRC and 0.5 mg/mL smFISH probe competitor E. coli tRNA (Roche, TRNAMRE- RO). Note VRC and probe competitor were omitted for half the cells analyzed in replicate 1. For *MDN1* mRNAs two probe sets, totaling 86 probes (38 + 48), coupled to the FAM (fluorescein amidite) dye (Biosearch Technologies, Stellaris Custom Probes) were used. The sequences of these probes were taken from Tutucci et al., 2018 (*MDN1*-3'ORF and *MDN1*-ORF). For *WHI5*, a set of 46 probes coupled to the Quasar570 dye (Biosearch Technologies, Stellaris Custom Probes) were used. *WHI5-mCitrine* Probe sequences were designed using Stellaris Probe Designer applied to the *WHI5-mCitrine* mRNA sequence. Probes were hybridised in the dark (30°C, overnight, with end-over-end rotation). 100 µL of wash buffer A was then added before cells were pelleted (400 g, 8 min) and supernatant was aspirated. Cells were resuspended in 1 mL wash buffer A, incubated in the dark (30° C, 30 min), pelleted (400 g, 6 min), resuspended in 1 mL wash buffer A +350 µg/mL calcofluor white (Sigma, F3543), and again incubated in the dark (30°C, 30 min). Cells were then resuspended in 1 mL wash buffer B (Biosearch Technologies, SMF-WB1-20), incubated for 2- 5 minutes at room temperature, pelleted (400 g, 6 minutes) and resuspended in 2 to 3 drops (~75 µL) of Vectashield Antifade Mounting Medium (Vector Laboratories, H-1000). This suspension was then mixed thoroughly by pipetting to separate clumped cells. 1.5 µL of this solution was mounted on an acid washed slide and imaged on a wide-field epifluorescence Zeiss Observer Z1 microscope (63X/1.4NA oil immersion objective and a Colibri LED module). 30-step *z*-stacks (step size = 200 nm) were imaged. Cell outlines were identified using phase contrast images. Quasar570 probes (*WHI5*) were imaged in the orange channel (white LED module for 555 nm wavelength, 100 % light power, 5 sec exposure per stack image). *Whi5-mCitrine* protein and FAM probes (*MDN1*) were imaged in the yellow channel (505 nm LED module, 100 % light power, 3.5 seconds exposure time). FAM FISH probes alone were imaged in the green channel (470 nm LED module, 75 % light power, 5 seconds exposure time). Calcofluor white stain was imaged in the blue channel (365 nm LED module, 25 % light power, 20 ms exposure). Under these conditions, no significant photobleaching was observed after taking multiple images of the same cells.

723 smFISH Image analysis was performed manually using ImageJ (version 2.0.0). Single cells
724 were manually selected in each image. A cell was only selected if its morphology was sufficiently intact
725 (following zymolyase treatment) and if the absence/presence of a bud and the nuclear/cytoplasmic
726 localization of Whi5-mCitrine protein could be assigned. Cell size was measured by drawing cell
727 outlines in the phase *z*-plane with the largest cell area, fitting a two-dimensional ellipse, and then
728 rotating the ellipse along its major axis to obtain a volume estimate. Separately calculated volumes for
729 mothers and buds were added together. Absolute counts of *WHI5-mCitrine* and *MDN1* mRNAs in
730 single cells were obtained by manual counting and single dots were counted as one mRNA (*i.e.*, we
731 did not quantify single dot intensities to try to discern multiple overlapping mRNAs). Linear regression
732 was performed in Prism 8 (default settings).
733
734 **Live Cell Microscopy**
735 Cells were grown to early log phase in synthetic complete (SC) media + 2 % glycerol + 1 % ethanol
736 and gently sonicated before being loaded into a CellASIC Y04C microfluidics plate (Milipore SIGMA)
737 under continuous media flow at 2 psi. Imaging and pedigree tracking was performed as previously
738 described (Doncic et al., 2013; Schmoller et al., 2015). Cells expressing GFP proteins were exposed
739 for 50 ms and cells expressing mCitrine were exposed for 400 ms. Background subtraction for variation
740 in background fluorescence in each frame of the movie was performed as previously describe
741 (Chandler-Brown et al., 2017). Briefly, in each frame cell and non-cell area was defined. A 4-pixel
742 average filter was then applied and the background was taken to be the median filtered pixel value of
743 the non-cell area. Differences in background fluorescence due to cell volume dependent
744 autofluorescence were accounted for as previously described (Schmoller et al., 2015). Note that
745 analysis of *WHI5pr-WHI5-mCitrine* (Fig. 2F, S3G, 3C&D) include cells previously imaged, analyzed
746 and reported in Schmoller et al., 2015.
747
748 **Gene deletion collection screen by microarray**
749 We analyzed the correlation between RNA levels and cell size in 1,484 gene deletion stains using
750 published microarray data and cell size measurements of these strains (Hoose et al., 2012; Jorgensen
751 et al., 2002; Kemmeren et al., 2014; O'Duibhir et al., 2014; Ohya et al., 2005; Soifer and Barkai, 2014).
752 Gene expression changes relative to wild-type were from (O'Duibhir et al., 2014), where we used the
753 dataset transformed to correct for effects of slow growth. The same trends were observed in the
754 uncorrected dataset. For each gene we then analyzed the correlation between then relative fold-
755 change its expression in a given deletion with the size of that deletion strain across all the deletion
756 strains for which both data were available. Pearson correlation coefficients were calculated using the
757 R function *cor* (Fig. 3F).
758

**GFP fusion collection screen by flow cytometry**

To examine the size-dependence of an individual protein's expression, we analyzed a genome-wide dataset of flow cytometry-based GFP intensity measurements (Parts et al., 2014), where each measurement is from a single well containing two strains both expressing the same protein C-terminally fused to GFP. One strain is in the BY4741 background (replicate 1) and the other in the RM11 background (replicate 2). Cells were grown in low fluorescence media containing 2% glucose and measured using an BD LSRII flow cytometer as described in Parts et al., 2014. Cells were separated into budded and unbudded populations on the basis of the side scatter width (SSC-W). Co-cultured strains of each background were separated on the basis of HTB2-mCherry intensity (RM low, BY high). Size was defined by the area of the side scatter signal (SSC-A). We used the lowest expressed gene from each plate as a background for that plate, thereby controlling for plate-to-plate variation in measurements. To calculate the background we fitted a linear function to SSC-A and total GFP fluorescence for these low-expressing cells (python function *polyfit*(matplotlib)). We then subtracted the fit for these lowest expressing cells from the GFP intensity for all other cells. Strains with noisy signals (*i.e.*, their mean expression is less than the standard deviation) and cells with saturated signals (mean expression is greater than 200000) were excluded. SSC-A and background subtracted GFP intensity were then normalized to the mean and a linear function was then fitted (python function *polyfit*(matplotlib)). The slope of this function was used as a measurement for a protein's size-dependence.

**ChIP-seq experiments**

Cells expressing Swi4-V5, Swi6-V5, 3xFLAG-WHI5, 3xFLAG-WHI5, GFP-NLS-5xFLAG or LacI-GFP-NLS-5xFLAG were grown in SC media with 2% glycerol 1% ethanol. 500ml of cells at OD ~0.5 were fixed with 1% formaldehyde (30 minutes) and quenched with 0.125 M glycine (5 minutes). Fixed cells were washed twice in cold PBS, pelleted, snap-frozen and stored at −80°C. Cell lysis and ChIP reactions were performed as previously described (Hu et al., 2015) with minor modifications. Pellets were 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 using 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 15 minutes (10 seconds on, 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. For each ChIP reaction, 30 µL Protein G Dynabeads (Invitrogen) were blocked (PBS + 0.5% BSA), prebound with 10 µL anti-V5 antibody (SV5-Pk1, BioRad Cat# MCA1360G) or 10 µL anti-FLAG antibody (M2, SIGMA Cat# F1804) and washed once with PBS before incubation with supernatant (4°C, overnight). Dynabeads were then washed (5 minutes per wash) twice in FA lysis buffer, twice 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), twice in ChIP wash buffer (10 mM TrisHCl pH 7.5, 0.25 M LiCl, 0.5% NP-40, 0.5% sodium deoxycholate, 1 mM EDTA, 1 mM PMSF) and once in TE wash buffer (10 mM TrisHCl pH 7.5, 1 mM EDTA, 50 mM NaCl). DNA was eluted in ChIP elution buffer (50 mM TrisHCl pH 7.5, 10 mM EDTA, 1% SDS) at 65°C for 15-20 minutes. Eluted DNA was 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 purified using the ChIP DNA Clean & Concentrator kit (Zymo Research). Indexed sequencing libraries were generated using the NEBNext Ultra II DNA Library Prep kit (NEB, # E7645), pooled and sequenced on an Illumina HiSeq instrument as paired end 150bp reads (Genewiz, NJ).

**ChIP-seq analysis**

Demultipexed fastq files were mapped to the sacCer3 assembly of the *S. cerevisiae* genome 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. Duplicate reads were removed using picard-tools (v.1.99). Peaks were called using MACS2 (v.2.1.0) (Feng et al., 2012) with the following settings: -g 12000000-f BAMPE. RPM (Reads Per Million) normalized read coverage genome browser tracks were generated using custom-written python scripts.

**Cell Cycle and Protein Partitioning Modeling**

The cell cycle was modeled as reported in (Chandler-Brown et al., 2017). We simulated the entire cell cycle, where cells grew and divided according to measured growth and cell cycle transition rates. This accounts for cell-to-cell variability. To examine the role of protein partitioning in the overall scaling of protein expression, we simulated the synthesis of a constitutively expressed protein ($p$) in each cell. Within the model, protein synthesis and partitioning properties were varied. Protein synthesis was modelled as either size-dependent ($\frac{dp}{dt} = kV$) or size-independent ($\frac{dp}{dt} = k$) and protein partitioning was modeled as either volume-proportional partitioning at cytokinesis (($p_{mother} = p_{total}\frac{V_{mother}}{V_{mother}+V_{daughter}}$) and ($p_{daughter} = p_{total}\frac{V_{daughter}}{V_{mother}+V_{daughter}}$)) or partitioned in the manner empirically measured for Whi5 manner where a significant fraction is partitioned by amount ($\frac{p_{mother}}{V_{mother}} = 1.441\frac{p_{daughter}}{V_{daughter}}$). Cells were simulated until a steady-state distribution was achieved and all cells at the last time-point were plotted.