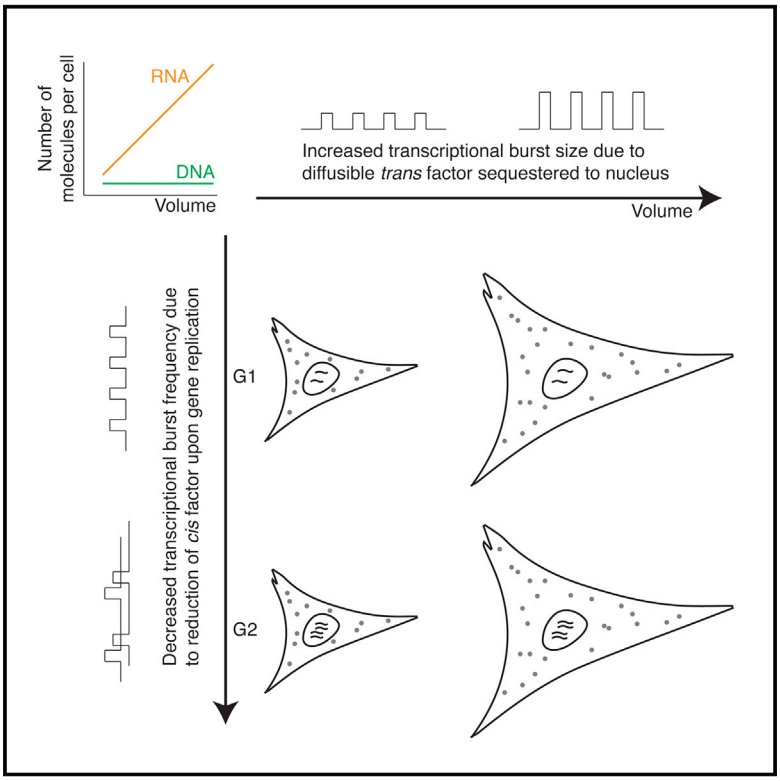
Article

Molecular Cell

Single Mammalian Cells Compensate for Differences in Cellular Volume and DNA Copy Number through Independent Global Transcriptional Mechanisms

Graphical Abstract



Authors

Olivia Padovan-Merhar, Gautham P. Nair, ..., Abhyudai Singh, Arjun Raj

Correspondence

rajlaboratory@gmail.com

In Brief

Padovan-Merhar et al. combine single-molecule transcript counting with computational measurement of cellular volume, showing that single cells maintain transcript abundance despite variability in cell size. Large cells have greater transcriptional burst size, whereas burst frequency halves upon DNA replication.

Highlights

- Transcription scales with cell volume to maintain transcript concentration
- Cell fusion shows that increasing cellular content can increase transcription
- Transcriptional burst size changes with cell volume and burst frequency with cell cycle
- The burst frequency mechanism allows for proper transcription during early S phase

Accession Numbers

GSE66053

Padovan-Merhar et al., 2015, Molecular Cell *58*, 339–352
April 16, 2015
http://dx.doi.org/10.1016/j.molcel.2015.03.005

Molecular Cell
Article

Single Mammalian Cells Compensate for Differences in Cellular Volume and DNA Copy Number through Independent Global Transcriptional Mechanisms

Olivia Padovan-Merhar,[1] Gautham P. Nair,[2] Andrew G. Biaesch,[2] Andreas Mayer,[3] Steven Scarfone,[1] Shawn W. Foley,[4] Angela R. Wu,[5] L. Stirling Churchman,[3] Abhyudai Singh,[6] and Arjun Raj[2,*]
[1]Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA
[2]Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104, USA
[3]Department of Genetics, Harvard Medical School, Boston, MA 02115, USA
[4]Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA
[5]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA
[6]Department of Electrical and Computer Engineering, University of Delaware, Newark, DE 19716, USA
*Correspondence: rajlaboratory@gmail.com
http://dx.doi.org/10.1016/j.molcel.2015.03.005

SUMMARY

Individual mammalian cells exhibit large variability in cellular volume, even with the same absolute DNA content, and so must compensate for differences in DNA concentration in order to maintain constant concentration of gene expression products. Using single-molecule counting and computational image analysis, we show that transcript abundance correlates with cellular volume at the single-cell level due to increased global transcription in larger cells. Cell fusion experiments establish that increased cellular content itself can directly increase transcription. Quantitative analysis shows that this mechanism measures the ratio of cellular volume to DNA content, most likely through sequestration of a transcriptional factor to DNA. Analysis of transcriptional bursts reveals a separate mechanism for gene dosage compensation after DNA replication that enables proper transcriptional output during early and late S phase. Our results provide a framework for quantitatively understanding the relationships among DNA content, cell size, and gene expression variability in single cells.

INTRODUCTION

Within a population, individual mammalian cells can vary greatly in their volume, often independently of their position in the cell cycle (Bryan et al., 2014; Crissman and Steinkamp, 1973; Tzur et al., 2009). Biochemical reaction rates, however, depend on the concentration of reactants and enzymes. Thus, to maintain proper cellular function, most molecules must be present in the same concentration despite these volume variations, meaning that the absolute numbers of molecules would have to scale roughly linearly with cellular volume (see Marguerat and Bähler, 2012 for an excellent review).

One critical molecule whose concentration need not scale with cellular volume, however, is DNA. Most mammalian cells have two or four copies of the genome per cell, and even cells with the same number of genomes can differ widely in size; thus, DNA concentration can vary dramatically from cell to cell. This poses a problem: if two otherwise identical cells with the same DNA content had different volumes, then the larger cell must somehow maintain a higher absolute number of biomolecules despite them being expressed from the same amount of DNA.

Previous efforts to resolve this puzzle have largely focused on analyzing bulk population measurements of size-altering mutants. A number of such studies have shown that the amount of both RNA and protein generally scales with cellular volume (Marguerat and Bähler, 2012; Marguerat et al., 2012; Schmidt and Schibler, 1995; Watanabe et al., 2007; Zhurinsky et al., 2010) and ploidy (Wu et al., 2010), with some further finding that transcription changes in mutants with larger or smaller cell volumes (Fraser and Nurse, 1979; Schmidt and Schibler, 1995; Zhurinsky et al., 2010). Most of these studies utilized yeast, with a few notable exceptions (Miettinen et al., 2014; Schmidt and Schibler, 1995; Watanabe et al., 2007).

These experiments do not, however, establish a causal relationship between cellular volume changes and transcript abundance. Causality could change the interpretation of gene expression measurements because, if cellular volume changes can in and of themselves change global expression levels, observations of changes in global expression levels in response to various perturbations may actually be the indirect consequence of changes to cellular volume rather than resulting from direct global transcriptional responses to the perturbations, per se.

Also unclear is how or even whether these mechanisms manifest in individual cells. Much recent evidence has shown that individual transcripts levels can vary from cell to cell due to stochastic effects in gene expression (Raj and van Oudenaarden, 2008, 2009; Sanchez and Golding, 2013) such as transcriptional bursts (Chubb et al., 2006; Golding et al., 2005; Raj et al., 2006; Suter et al., 2011; Zenklusen et al., 2008). Yet it remains unclear how differences in cellular volume affect the interpretation of such measurements or whether transcriptional measurements can reveal further characteristics of homeostatic mechanisms.

We used single-molecule RNA imaging and computational image analysis to measure transcript abundance and cellular volume simultaneously in individual human cells. Cell fusion experiments showed that cellular size can directly and globally affect gene expression by modulating transcription. Quantitative analysis of these experiments revealed that the mechanism underlying this global regulation does not merely sense cellular volume but rather integrates both DNA content and cellular volume to produce the appropriate amount of RNA for a cell of a given size, consistent with a model in which a factor limiting for transcription is sequestered to the DNA, either through direct titration by DNA or restriction to the nuclear compartment. We also provide a quantitative framework for interpreting gene expression variability in single cells and extended this framework to genome-wide single-cell RNA-sequencing analysis, showing that cell-type-specific genes are more variable than ubiquitously expressed genes.

RESULTS

mRNA Counts Scale with Cellular Volume in Single Mammalian Cells

We first looked at the number of mRNA molecules in individual primary fibroblast cells (human primary foreskin fibroblasts, CRL2097) within a population to see whether mRNA counts scale with cellular volume at the single-cell level. We measured both mRNA abundance and volume simultaneously using single-molecule multicolor mRNA fluorescence in situ hybridization (RNA FISH; Femino et al., 1998; Raj et al., 2008), which allowed us to detect the positions of individual mRNAs in three dimensions as fluorescent spots in the microscope (Figure 1A). We measured the abundance of a particular mRNA (e.g., *TBCB*) labeled with one color and then calculated the volume of the cell using the 3D positions of mRNA from a ''volume guide'' gene labeled with another color to define the cellular boundary (Figure 1B; Experimental Procedures). The cell volumes we measured varied over a 6-fold range, agreeing with other estimates (Bryan et al., 2014; Tzur et al., 2009), and are robust to choice of guide gene and the fixation procedure itself (Figure S1). We ultimately measured the abundance of 25 to 30 different mRNA species in both this primary fibroblast line and a lung cancer line (A549).

For most genes, mRNA counts and volumes in single cells exhibited a strongly positive, linear correlation (e.g., Figure 1C; see Figure S2 for all genes examined). Because larger cells had proportionally more transcripts than did smaller cells, the mRNA concentration remained relatively constant from cell to cell despite considerable variation in absolute mRNA numbers. This scaling property was not confined to high-abundance mRNAs such as *GAPDH* and *EEF2*—genes expressing as few as 10 to 20 mRNA per cell such as *ZNF444* and *KDM5A* scaled similarly, as did rRNA (Figure S2). We also observed the same behavior for short-lived mRNA such as *UBC* and *IER2* mRNA, whose half-lives are 2.9 and 2.2 hr, respectively (Tani et al., 2012).

We checked whether the scaling of mRNA count with volume depended on cell-cycle progression or cell growth. We co-stained cells with cell-cycle markers (Eward et al., 2004; Levesque and Raj, 2013; Robertson et al., 2000; Whitfield et al., 2002) to classify them as being in the G1, S, or G2 phases of the cell cycle (Figure S3). Cell volume varied as much for cells in individual phases of the cell cycle as the population overall, with a shift in the distribution toward G2 cells being larger, and the linear relationship between mRNA count and volume did not depend on cell-cycle phase (Figure 1D), showing that mRNA count did not depend on DNA content of the cell. We also note that the primary fibroblast cells exhibit normal ploidy (Levesque and Raj, 2013), so our results are not simply explained by differences in ploidy. We also found that nuclear size increased somewhat with cellular volume and that nuclear size increased in later stages of the cell cycle (Figure S3). To check whether progression through the cell cycle or cell growth is responsible for maintaining scaling, we grew the primary fibroblasts for 7 days in medium lacking serum, making them quiescent. Despite growth and cell cycle arrest, we found that mRNA count and volume still scaled strongly for all genes examined, showing that neither progression through cell cycle nor continual cell growth is required for mRNA count to scale with cellular volume (Figure 1F). Interestingly, we found that although the mean count of mRNA decreased in quiescent cells as compared with proliferating cells, the cells maintained a similar concentration of *GAPDH* and other mRNA between the two conditions (Figures 1G–1H and S1).

We also checked whether we could observe similar behavior in intact organisms. We looked at both RNA and DNA density in the heads and gonads of adult nematodes, comparing measurements from both wild-type worms and worms with mutations leading to decreased organismal size but with roughly the same number of cells (Figures 2A–2B) (Watanabe et al., 2007). We found that the RNA concentrations were roughly the same between the two strains and that the number of RNA per molecule of DNA decreased by a factor similar to that of the volume differences between the strains (Figures 2C–2D), verifying that our observations can hold in vivo.

It is important to note that although the mRNA abundance is strongly correlated with cellular volume, the y intercept (a) of a line fit to the data (mRNA = $a + bV$) was nonzero, indicating that mRNA count in individual cells has a volume-independent component in addition to the volume-correlated component. Quantifying the relative fractions of mRNA that are volume correlated versus volume independent in a cell of average volume (Figure 1E) [i.e., $a / (a + bV_{avg})$ versus $bV_{avg} / (a + bV_{avg})$] revealed a range of values for different genes (Figure S1), although the volume-dependent fraction was dominant for most genes examined. Thus, the mRNA concentration is actually somewhat greater in smaller cells than in larger ones; for most genes, the smallest cells have an mRNA concentration 1.2 to 3 times greater than that of the largest cells (Figure S2). We later describe a mathematical model providing a potential explanation for this increased concentration based on nuclear volume measurements (see Supplemental Information).

Transcriptional Activity, Not mRNA Degradation, Scales Globally with Cellular Volume

These data show that larger cells have a proportionally higher number of mRNA than smaller cells, even if they have the same absolute number of DNA molecules. To maintain this

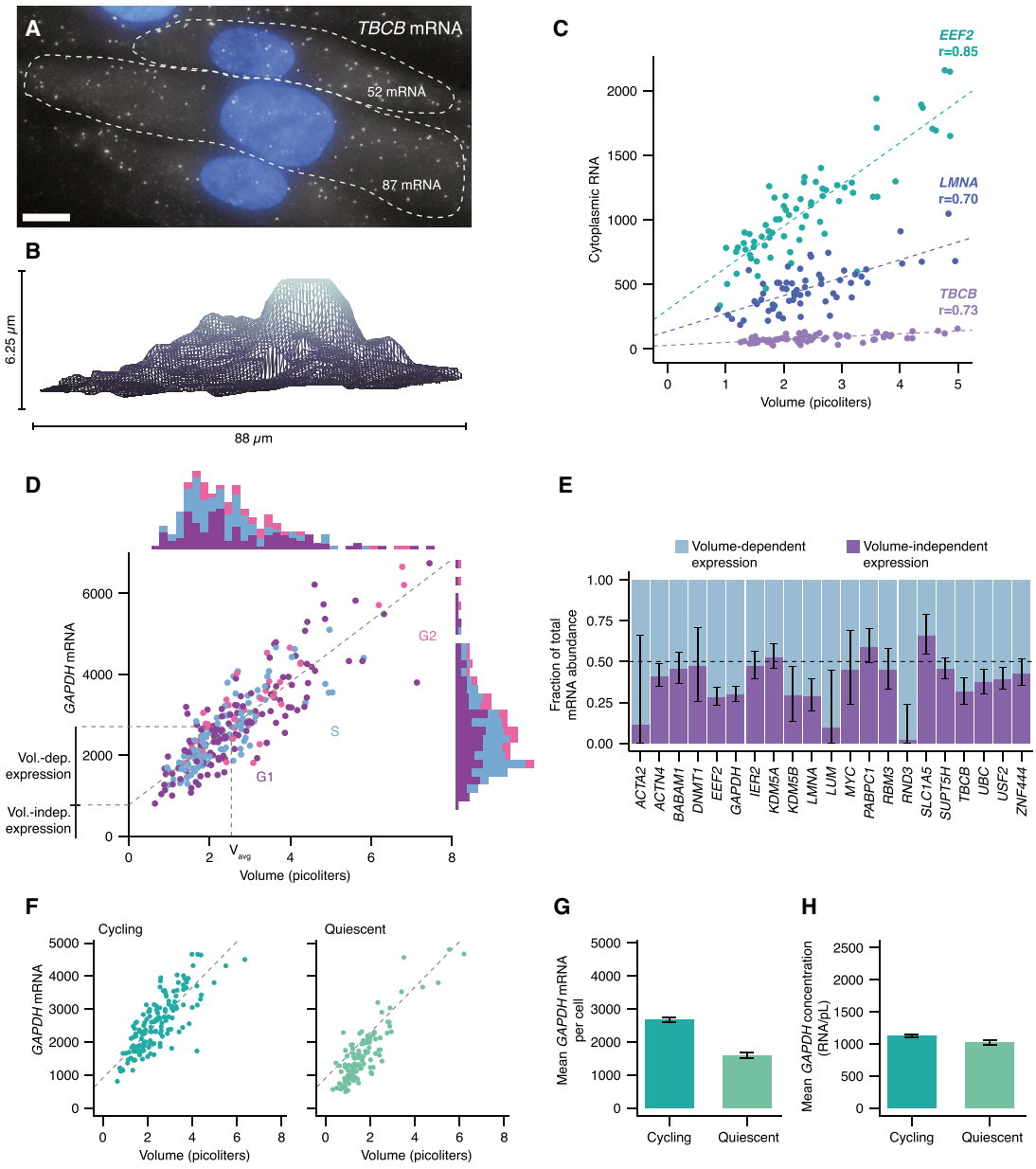


Figure 1. mRNA from Many Genes Scales with Cellular Volume

(A) Single-molecule RNA FISH. The DAPI stain is in blue, and the *TBCB* mRNA FISH probe is in white.

(B) Representative outline of a primary fibroblast cell found using our volume calculation algorithm.

(C) mRNA versus volume for *EEF2*, *LMNA*, and *TBCB*. Each point represents one single-cell measurement. Each dataset is a combination of at least two biological replicates, with at least 30 cells per replicate.

(D) *GAPDH* mRNA and volume in primary fibroblast cells. Marginal histograms show volume and mRNA distributions. Colors indicate cell-cycle stage determined by Cyclin A2 (*CCNA2*) mRNA count. Dashed diagonal line is the best linear fit of RNA versus volume. V_{avg} indicates the average primary fibroblast volume. We determined volume-independent and -dependent transcript levels using the linear fit and V_{avg}. Data are a 15% subset of 1,868 cells spanning >30 biological replicates.

(E) Fraction of volume-independent and -dependent RNA expression from the linear fit of RNA versus volume for 21 genes in primary fibroblast cells (we omitted highly variable genes whose volume-independent fractions were less than zero). Data for each gene are a combination of at least two biological replicates, with at least 30 cells per replicate.

(F) *GAPDH* mRNA versus volume in cycling and quiescent primary fibroblast cells. Dashed lines are best fit line for *GAPDH* in cycling cells. Data are an 8% subset of 1,868 cells spanning >30 biological replicates for cycling cells and 10% subset of 1,105 cells for quiescent. We only analyzed quiescent cells that had less than 20 *CCNA2* mRNA.

(G and H) Mean *GAPDH* mRNA count (G) and concentration (H) in different growth conditions for data from (F).

All error bars represent SEM. See also Figures S1–S3.

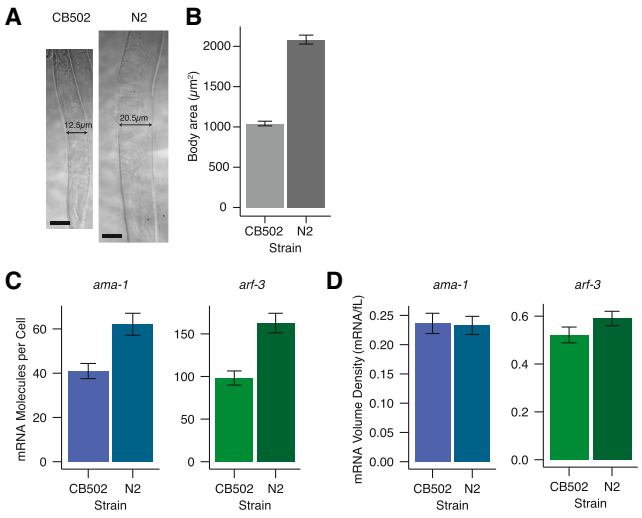


Figure 2. mRNA Scales with Volume In Vivo

(A) Images of the two *C. elegans* strains.

(B) Quantification of the relative sizes of the two strains (n = 24 for N2, n = 20 for CB502).

(C) Number of mRNA molecules per cell in the gonad region for each type of worm for genes *ama-1* and *arf-3*. We estimated the number of cells in each segment by counting nuclei stained with DAPI. Each bar is a compilation of three biological replicates, with >3 worms per replicate.

(D) Concentration of mRNA in the gonad region. All scale bars represent 10 μm. All error bars represent SEM.

proportionality, larger cells must either transcribe more mRNA from the same number of DNA molecules or degrade those mRNA more slowly. To distinguish these possibilities, we determined the rate of mRNA degradation in cells of different sizes by measuring volumes and mRNA counts for *UBC* and *IER2* after a period of 4 hr of transcriptional inhibition (transcriptional inhibition did not affect cell volume; Figure 3D). By measuring mRNA counts before and after transcription inhibition (Figures 3A–3B, inset; see Experimental Procedures for details), we calculated the effective decay constant for each measured cell (Figures 3A–3B). We found that the degradation rate was the same in cells of all volumes, showing that slower degradation is not responsible for the increased number of mRNA in larger cells.

We next checked whether larger cells transcribe more than smaller cells (as observed in bulk populations; Fraser and Nurse, 1979; Schmidt and Schibler, 1995; Zhurinsky et al., 2010). We inferred global transcription rate by incorporating a labeled uridine into all newly synthesized RNA produced during a 60-min time window (Figure 3C), which we then rendered fluorescent via click chemistry (Jao and Salic, 2008). The intensity of fluorescence is equivalent to the global transcription rate. We found that transcription rate is linearly proportional to volume, thus showing that individual cells vary in their overall transcription (das Neves et al., 2010), and these variations correlate strongly with volume. We conclude that larger cells maintain proportionally higher levels of RNA by increased transcription rather than decreased degradation as compared to smaller cells. Also, quantification of fluorescence from probes targeting the internal transcribed spacer of the rRNA (the "intronic" sequence of rRNA) showed that transcription of rRNA also scales linearly with volume (Figure S2), indicating that RNA polymerase I transcription is also volume dependent.

The scaling of transcription with cellular volume could be due to global factors regulating transcription of all genes in a volume-correlated manner, or it could be that gene regulatory networks sense deviations in each particular gene's protein concentration and modulate transcription to restore concentration. In the latter case, reducing protein concentration of any one gene would result in increased transcription to compensate, whereas in the global scenario, reducing the concentration of any one gene would not appreciably affect the cellular volume, thus leaving the gene's transcription unchanged. We tested this by reducing the level of lamin A/C mRNA and protein in the cell via small interfering RNA (siRNA) (Figures 3E and 3F); we chose lamin A/C because its expression scales strongly with volume (Figure 1C) and is thought to be tightly regulated (Swift et al., 2013). To measure the transcriptional response, we took advantage of the fact that transcription occurs in bursts (Chubb et al., 2006; Dar et al., 2012; Golding et al., 2005; Suter et al., 2011; Vargas et al., 2005; Zenklusen et al., 2008), and genes that are actively undergoing a transcriptional burst have bright accumulations of nascent RNA at the site of transcription itself (Levesque and Raj, 2013; Levsky et al., 2002; Raj et al., 2006; Zenklusen et al., 2008) (note that siRNA does not affect nuclear RNA; Maamar et al., 2013). We measured both the average number of active lamin A/C transcription sites per cell and the intensity of those transcription sites, finding both metrics unchanged upon reduction of lamin A/C protein levels (Figure 3G). We conclude that increased mRNA counts in larger cells result from a global difference in transcription rather than the activity of a particular gene network that regulates the concentration of lamin A/C. There may be other situations in which mRNA levels are regulated by specific networks.

A Diffusible *trans* Factor Sensing DNA Content and Volume Links Cellular Volume and Transcription

What links volume and transcription? One possibility is that the total cellular content itself exerts a global influence on transcription, thus making transcription scale with cellular volume; alternatively, transcription may affect cellular volume. To distinguish these possibilities, we fused a small human melanoma cell that expressed GFP mRNA (WM983b-GFP-NLS) at constant density to the larger fibroblasts that do not express GFP (Figure 4A) to form heterokaryons (Pomerantz et al., 2009). We found that absolute GFP mRNA counts increased in fused cells as compared to the original small cells (Figure 4B), showing that increasing total cellular content is by itself sufficient to increase absolute mRNA abundance. Moreover, the GFP mRNA counts scaled with heterokaryon volume (Figure 4C), suggesting that the rate of GFP transcription scaled with the ultimate volume of the fused cell. The fact that the nucleus from the WM983b-GFP-NLS cell could change its overall transcriptional activity showed that the modulation occurs via the activity of a diffusible *trans* factor.

How might such a factor transmit volume information to the GFP gene in order to increase its transcription concordantly with the increase in cellular volume? There are two broad categories of mechanism: (1) the factor acts as a "volume sensor" and does not know about the amount of DNA in the cell. An

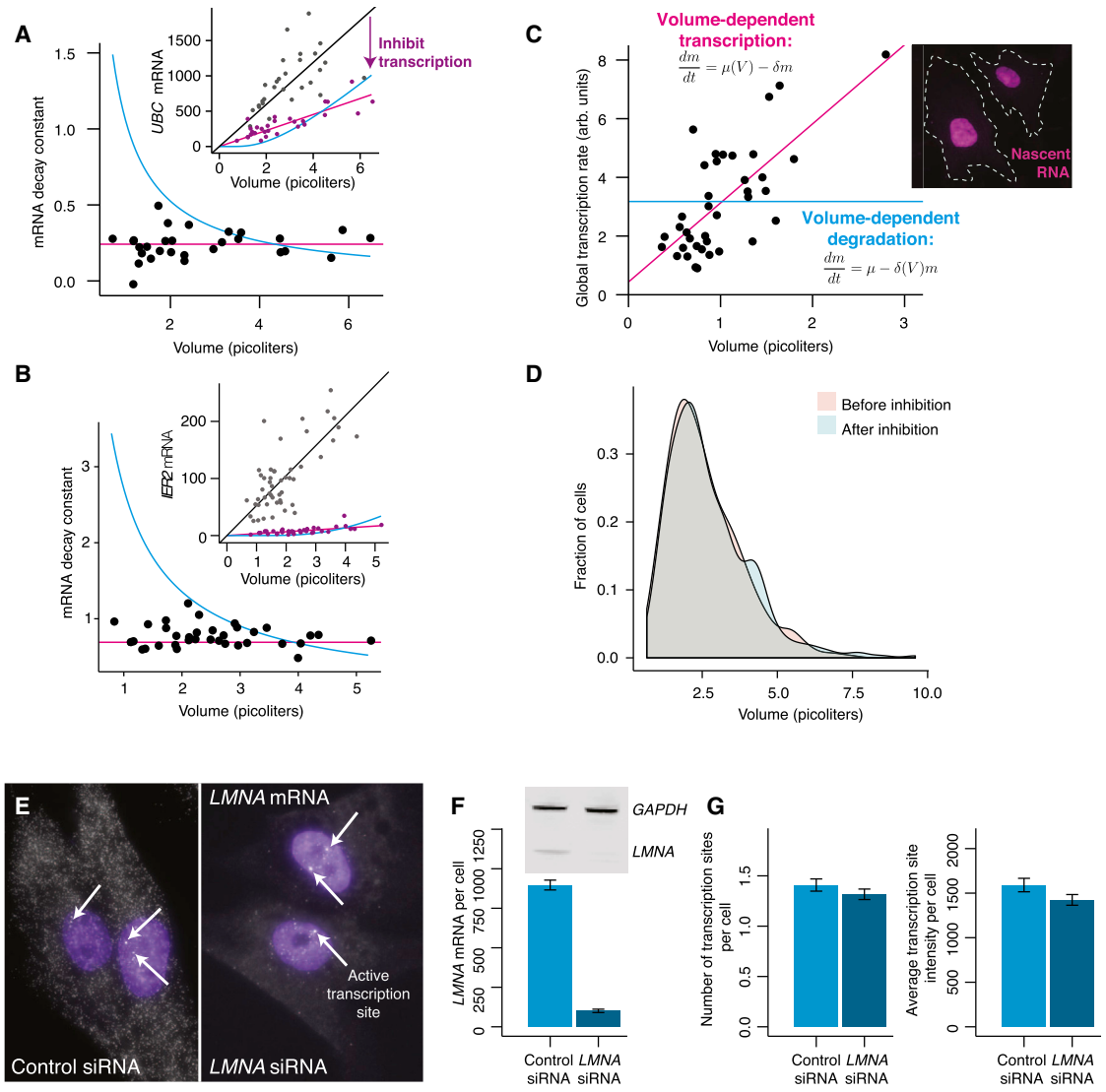


Figure 3. Cells Exhibit Global Volume-Dependent Transcriptional Control over mRNA Abundance

(A and B) We inhibited transcription in primary fibroblast cells using actinomycin D for 4 hr and allowed *UBC* (A) and *IER2* (B) mRNA to degrade. Inset shows mRNA before and after inhibition. Each point represents a single-cell measurement. We calculated the decay constant for each cell using the best-fit line before inhibition (see Experimental Procedures). Blue line shows fit if degradation were volume-dependent; red line shows fit if transcription were volume-dependent. Data represent one of two biological replicates.

(C) We fluorescently labeled nascent RNA produced in 1 hr using the Click-iT eU assay in primary fibroblast cells and quantified the total fluorescence intensity by imaging the nuclei of single cells. Inset shows raw micrograph data. Blue line shows fit for volume-dependent degradation; red line shows fit for volume-dependent transcription. Data shown are from quiescent cells and are one of three biological replicates.

(D) Distribution of cell volumes before and after transcription inhibition.

(E) We performed siRNA treatment for 72 hr in primary fibroblast cells using either a control siRNA (left) or an siRNA targeting *LMNA* mRNA (right). DAPI stain is shown in purple, and *LMNA* mRNA FISH probe is shown in white. White arrows indicate active transcription sites.

(F) Quantification of cytoplasmic *LMNA* mRNA knockdown by RNA FISH. Inset shows protein knockdown.

(G) Comparison of the number of *LMNA* transcription sites and transcription site intensity in siRNA control and *LMNA* knockdown conditions. We detected transcription sites through intron/exon colocalization using RNA FISH. All error bars represent SEM. Data in (D) and (E) are a combination of two biological replicates, n = 323 cells for control siRNA and 284 cells for *LMNA* siRNA.

example could be a modifiable global transcription factor protein whose degree of modification/activity is proportional to cellular size (Figure 4D, left). Or (2) the factor acts as a "volume/DNA sensor" whose activity depends on both cellular volume and DNA content. One such mechanism is the existence of a general transcription factor of limiting abundance relative to the number of binding sites in the DNA (limiting factor). Here, the DNA "counts" how big the cell is by binding all available factor molecules (Figure 4D, right), thus increasing transcription in bigger cells as more factor binds to DNA. Another possibility is

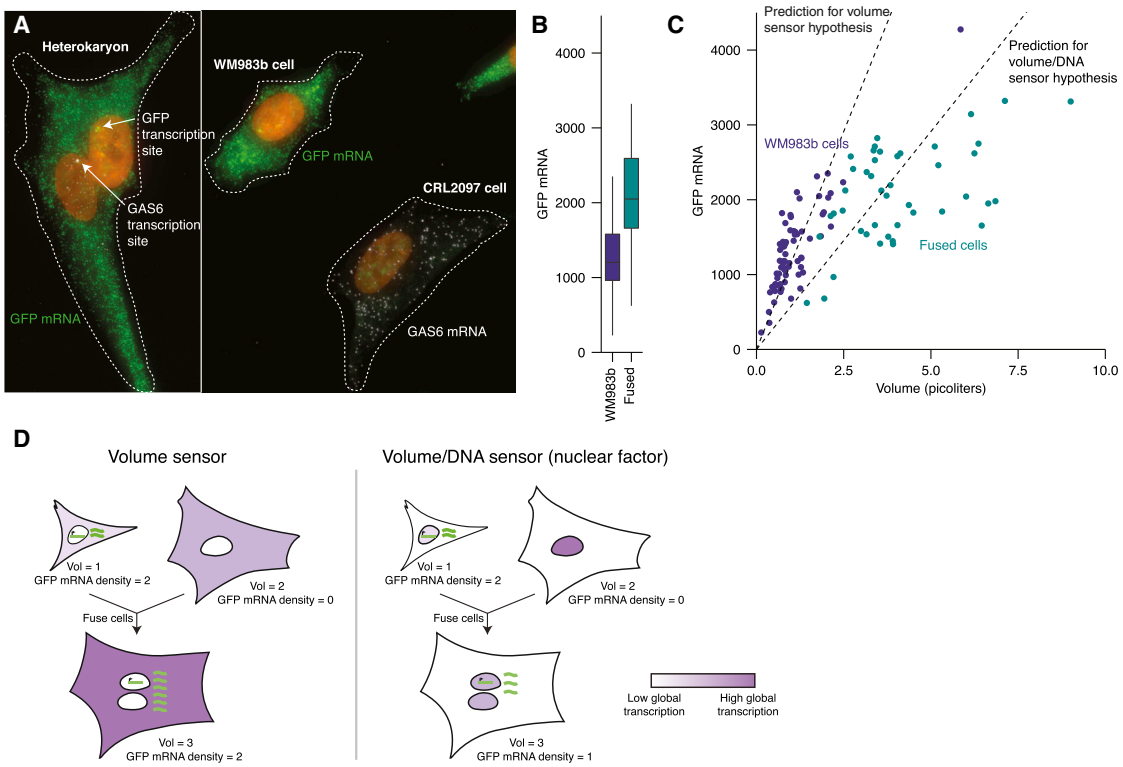


Figure 4. A *trans*-Acting Limiting Factor Links Gene Expression to Volume

(A) Representative image of fused cells (heterokaryon, left) and unfused cells (WM983b, primary fibroblast, right). DAPI stain is in orange, GFP mRNA is in green, and *GAS6* mRNA is in white. White arrows indicate transcription sites.

(B) Quantification of GFP mRNA in unfused and fused cells. Box extends to first and third quartiles, and whiskers extend to the maximum-distance points within 1.5 interquartile ranges of the box. Data are a combination of two biological replicates.

(C) GFP versus volume for fused and unfused cells. Upper dashed line represents fit for unfused cells. Lower dashed line has a slope that is half of the upper fit line.

(D) Schematic of transcriptional output of fused cells if the scaling of expression with volume were mediated by a volume sensor or a volume/DNA sensor. All error bars represent SEM.

sequestration of the factor to the nucleus, which can also achieve the same volume/DNA-sensing behavior if nuclear volume is only weakly dependent on cellular volume (see Supplemental Information).

We distinguished these two alternatives by comparing the concentration of GFP mRNA in the fused and unfused cells (Figure 4C). In the volume sensor scenario, the fusion cell will have the same concentration of GFP mRNA as the original small cells because the factor transmits the volume information to the GFP gene independent of the number of nuclei in the cell. In the volume/DNA sensor scenario, the fusion cell will have half the concentration of GFP mRNA because the factor senses both the increased volume and the two nuclei (for example, a limiting factor would be diluted between the two nuclei in the fused cell). We found that the concentration of GFP mRNA in the fused cells is strictly less than and very close to half the concentration in unfused cells. We conclude that the factor responsible for increased transcription in smaller cells is not a volume sensor but responds to both the size and DNA content. These results also suggest that perturbations that change cell size will indirectly change global transcript counts per cell through this generic mechanism.

Transcriptional Burst Size Increases in Larger Cells

We next sought to characterize this mechanism further by examining the relationship between volume and transcription of individual genes. mRNA is produced in bursts, marked by bright accumulations of nascent mRNA at the site of transcription. We characterize this bursting behavior through burst size (how much RNA is produced during a single burst) and burst fraction (how often a gene is actively transcribing, which is related to burst frequency; Levesque and Raj, 2013). To quantify burst size, we measured the intensity of transcription sites for four genes in our fibroblasts and found higher intensity transcription sites in larger cells (Figure 5A). We verified that the intensity of transcription sites reflected the degree of transcriptional activity by treating primary fibroblast cells with 100 nM triptolide, which targets RNA polymerase II for degradation (Bensaude, 2011), reducing its levels (Figure 5D). After 1 hr, we saw a reduction of bright transcription sites for two different genes, showing that transcription site intensity depends directly on the amount of transcriptional machinery available. This intensity is often proportional to burst size (Levesque and Raj, 2013; Senecal et al., 2014). Interestingly, transcription site intensity did not depend on cell-cycle stage (Figure 5B). We concluded

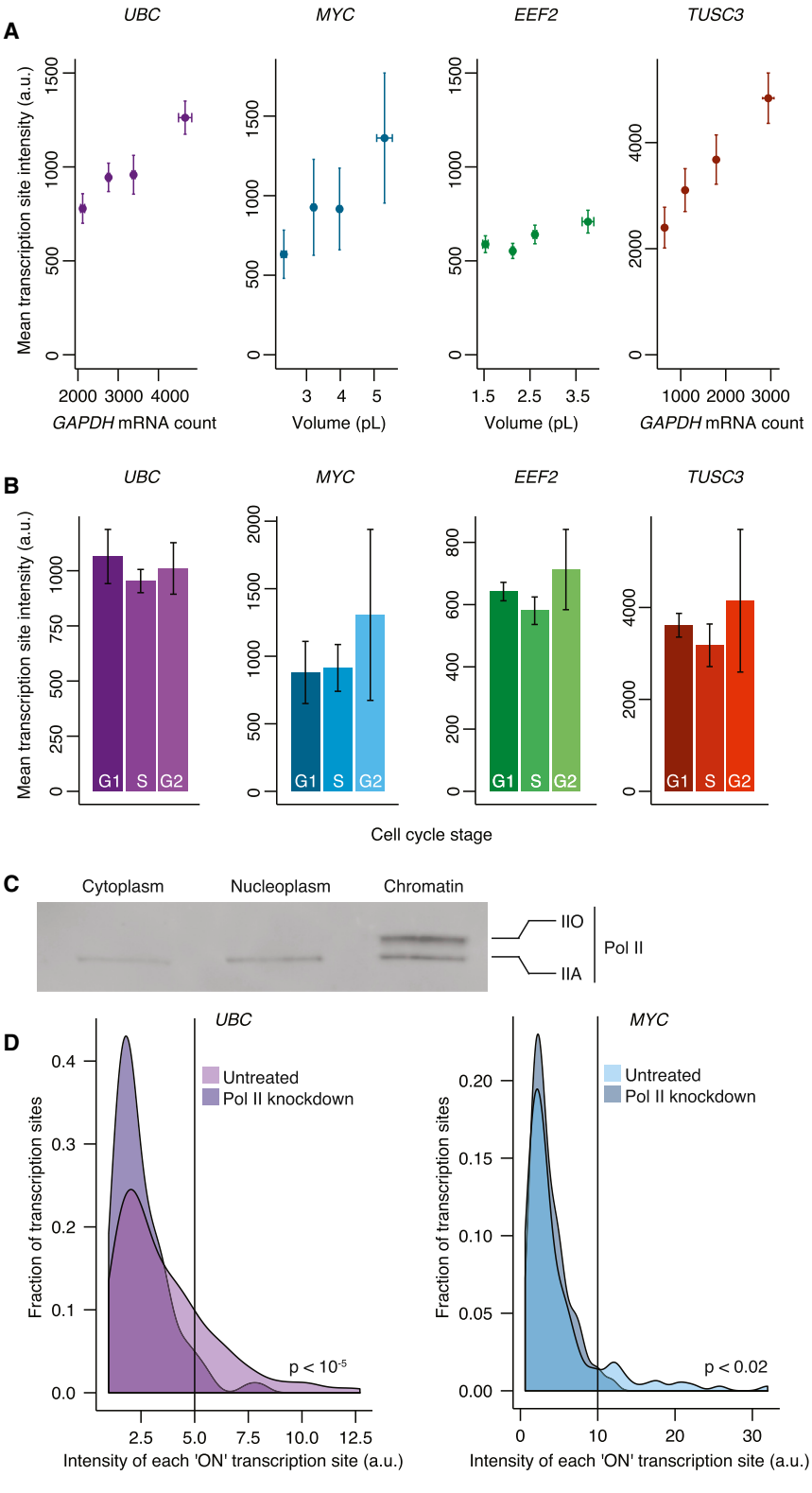


Figure 5. Transcriptional Burst Size Increases in Larger Cells

(A) Transcription site intensity and volume in primary fibroblast cells for genes *UBC*, *MYC*, *EEF2*, and *TUSC3*. Each data point represents the mean transcription site intensity per cell for a quartile of cells classified by volume or *GAPDH*. We detected transcription sites through intron/exon colocalization using RNA FISH. We calculated volume for *EEF2* data using *EEF2* as a guide and volume for *MYC* data using *GAPDH*. We use *GAPDH* as a proxy for volume for *UBC* and *TUSC3*.

(B) Transcription site intensity and cell-cycle stage in primary fibroblast cells. We determined cell-cycle stage by Cyclin A2 and the histone 1H4E mRNA counts (see Experimental Procedures and Figure S3). For intensity measurements, data for *UBC*, *MYC*, and *EEF2* are from one of two biological replicates (*EEF2*: n = 190, *UBC*: n = 202, and *MYC*: n = 103 transcription sites). Data for *TUSC3* are combined from two biological replicates (n = 255 transcription sites).

(C) Western blot analysis reveals that >99% of the C-terminal domain hyperphosphorylated form of RNA polymerase II (IIO) is present in the chromatin fraction. The hypophosphorylated form of Pol II (IIA) is captured in all cellular fractions. We generated subcellular lysates from the same batch of primary fibroblast cells and probed with the F-12 antibody (Santa Cruz Biotechnology) that is directed against the N-terminal region of RPB1, the largest subunit of RNA polymerase II. We adjusted sample volumes so that western blot signals of the subcellular fractions are comparable.

(D) Quantification of transcription site intensity before and after treatment with 100 nM triptolide for 1 hr. p value represents the probability of randomly finding the distributions of bright transcription sites (values to the right of the black line) in each condition.

All error bars represent SEM. See also Figures S3 and S4.

that the factor connecting gene expression and cellular volume affected mRNA abundance through modulation of transcriptional burst size.

What might this factor be? In order to formalize the possibilities, we developed a mathematical model for how this factor may act (see Supplemental Information), assuming that the factor is almost purely nuclear and that the total amount of factor in the cell is proportional to cellular volume. Our model of perfect linear scaling of transcription with volume requires that either (1) the volume of the nucleus remains fixed between cells irrespective of cytoplasmic volume or (2) the factor has such a high affinity for DNA that it is essentially titrated by DNA (or both). In both scenarios, either the nucleus or the DNA "counts" the amount of factor in the cell to modify transcription, thus making transcription proportional to the total amount, rather than concentration, of the factor. Further, in scenario 1, it predicts that if nuclear volume changes with cell size, it will manifest as a

deviation from pure linear scaling of transcription with cellular volume. Our data cannot distinguish these two possibilities, but our analysis did predict that in scenario 1, the slight increase in nuclear volume we saw in larger cells (Figure S3) could lead to a slight decrease in mRNA concentration, manifesting as a nonzero intercept when fitting RNA abundance to volume, as observed in Figure 1D. Indeed, the quantitative magnitude of the increased concentration at low cell volumes matches what our model predicts based on our measured relationship between nuclear volume and cellular volume (see Supplemental Information), perhaps favoring the model in which transcriptional scaling results from sequestration of this factor to the nucleus over a pure titration mechanism. It is possible that the factor is some component of the general transcriptional machinery, which is 94% nuclear based on fractionation (Figure 5C) (A.M. and L.S.C., unpublished data).

A DNA-Linked *cis*-Acting Factor Reduces Transcription Frequency, Not Burst Size, Immediately after DNA Replication

We have shown that cells in the G1 and G2 phases of the cell cycle can have the same volume and same mRNA density, although cells in G2 have twice the DNA content as those in G1. How, then, do two cells of the same size produce the same amount of mRNA, despite having different amounts of DNA?

We already found that transcriptional burst size does not change dramatically between phases of the cell cycle (Figure 5B), so we now measured burst fraction. To measure transcriptional burst fraction, we counted active transcription sites in each cell and divided by the total number of gene copies for each stage of the cell cycle (two copies in G1, four copies in G2). For all of the genes we measured, the number of active sites per gene copy in G2 was approximately half of that in G1 (Figure 6A). This is not due to repression of replicated copies of DNA, as we observed G2 cells with four transcription sites (Figure S4). This showed that the cell has a mechanism to precisely reduce transcription fraction in G2 to keep overall transcription constant across the G1 and G2 phases of the cell cycle. Burst fraction did not depend on cell volume (Figure 6B), showing that the mechanism is distinct from the volume-compensating mechanism described above.

We were surprised that the mechanism for cell cycle was different from the one compensating for volume. In principle, limiting factor models that may be responsible for volume conservation would also compensate for increased gene copy number due to DNA replication. However, these models predict an inappropriate boost in transcription for genes that replicate early in S phase because a limiting factor would distribute itself over all the DNA, essentially ''double counting'' the small percentage of genes that replicate considerably earlier than the majority of DNA (Figure 6C). To see whether this was the case, we measured transcriptional burst fraction for early-replicating genes in S phase (*EEF2*, *MYC*, *UBC*; see Figure S4 for replication timing). These genes all showed the same number of active transcription sites per gene copy in S and G2, ruling out the possibility that a factor simply gets diluted between copies of replicated DNA.

This leaves two alternatives for burst fraction reduction between G1 and G2. Transcription fraction could universally decrease by a factor of two upon the initiation of S phase, but this would potentially cause the opposite problem—genes that replicate late in S phase would be *under*-transcribed for the majority of S phase. Alternatively, transcription fraction could decrease on a gene-by-gene basis (i.e., in *cis*) immediately upon DNA replication. To test between these alternatives, we imaged transcription of a gene that replicates very late in S phase (*TUSC3*; Figure S4). If transcription fraction was universally reduced by a factor of two at the beginning of S phase, this gene would have the same transcription fraction in S and G2. However, we found that this late-replicating gene maintains G1 levels of transcription through S phase and does not reduce transcription fraction until G2. Therefore, we conclude that there exists a mechanism whereby transcription fraction is reduced by a factor of two immediately upon replication of that gene, with different timing for different genes. Candidates for a mechanism include the partitioning or modification of DNA-linked factors, such as histones with particular modifications, upon DNA replication, resulting in half the transcriptional burst fraction as before replication.

Together, our data demonstrate the existence of two separate transcriptional mechanisms that allow cells to maintain RNA concentration homeostasis despite changes in DNA content and cellular volume. Cells modulate transcriptional burst size through a *trans* mechanism to allow larger cells to produce more mRNA from the same amount of DNA and modulate burst frequency over the cell cycle in *cis* to maintain RNA concentration despite changes in DNA content.

Single-Cell RNA Sequencing Reveals that Cell-Type-Specific Genes Are "Noisier" Than Ubiquitously Expressed Genes

Our RNA FISH data revealed that, although the expression of most genes was consistent with a volume-dependent transcription rate, many of the genes also showed strong variability in transcript concentration from cell to cell (Li and Xie, 2011; Raj and van Oudenaarden, 2009; Raj et al., 2008; Sanchez and Golding, 2013).

To accurately quantify gene expression noise while accounting for cell volume, we developed a metric we call volume-corrected noise measure (Nm; Figure 7A), defined as (Bowsher and Swain, 2012; Johnston et al., 2012; Swain et al., 2002) follows:

$$Nm = CV_m^2 - \left(\frac{b\langle V\rangle}{a + b\langle V\rangle}\right)\left(\frac{Cov(m,V)}{\langle m\rangle\langle V\rangle}\right),$$

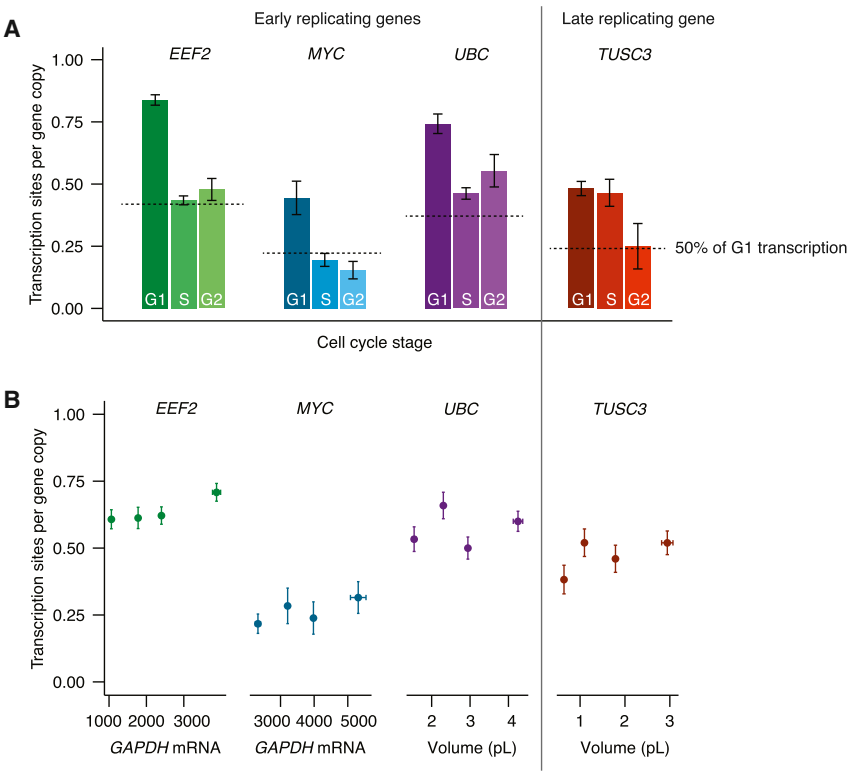
where

$$CV_m = \frac{\sigma_m}{\langle m\rangle},$$

and a and b indicate the intercept and slope of a best-fit line for mRNA and volume,

$$\langle m\rangle = a + b\langle V\rangle.$$

Volume-corrected noise measure is in principle similar to the squared coefficient of variation of mRNA concentration but accounts for volume-independent transcription (Figure S5).



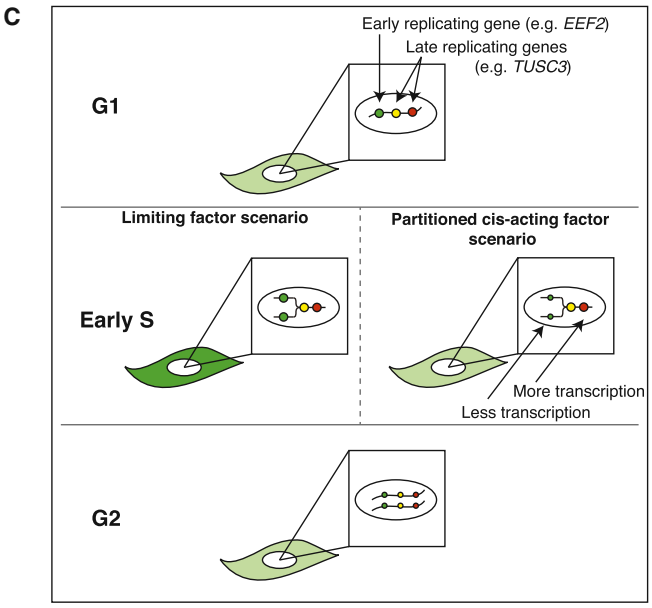


Figure 6. A *cis*-Acting Factor Decreases Transcription Frequency Immediately upon DNA Replication

(A) Number of transcription sites by cell-cycle stage in primary fibroblast cells. We determined cell-cycle stage by Cyclin A2 and histone 1H4E mRNA counts. Dashed lines represent half the number of transcription sites in G1. We normalize all G1 data to two gene copies and all G2 data to four gene copies. For *EEF2*, *MYC*, and *UBC* (early replicators), we normalize S-phase data to four gene copies. For *TUSC3* (late replicator), we normalize S-phase data to two gene copies.

(B) Number of transcription sites per gene copy classified by volume in primary fibroblast cells. Each data point represents the mean number of transcription sites for a quartile of cells classified by volume. We calculated volume for *EEF2* data using *EEF2* as a guide and volume for *MYC* data using *GAPDH*. We use *GAPDH* as a proxy for volume for *UBC* and *TUSC3*. For frequency measurements, data for *EEF2*, *UBC*, and *TUSC3* are a combination of two biological replicates (*EEF2*: n = 516, *UBC*: n = 332, and *TUSC3*: n = 255 transcription sites). Data for *MYC* is from one of two biological replicates (*MYC*: n = 103 transcription sites).

(C) Schematic depicting different potential mechanisms for changing gene expression with cell cycle. All error bars represent SEM. See also Figures S3 and S4.

We evaluated Nm for all of the genes we measured by RNA FISH. A standard measure of gene noise is the coefficient of variation (*CV*; standard deviation divided by mean), which takes into account only the spread and the mean of the expression data. With such a measurement, most of the genes we measured by RNA FISH would be deemed "noisy," or far from Poisson noise levels (Figure S5). However, when we instead measured Nm for each of our RNA FISH genes, we saw that many of them displayed levels of variability near to or indistinguishable from Poisson (Figure 7B) once we accounted for measurement error (upper bound of approximately 15%) (Raj et al., 2006).

To quantify this variability genome-wide, we performed single-cell RNA sequencing (Brennecke et al., 2013; Grün et al., 2014; Shalek et al., 2013) on human foreskin fibroblast cells, calibrating the sequencing data using our RNA FISH results (Figures 7C and S6) and adding synthetic RNA at known concentrations to estimate cell volume. Briefly, we used the ratio of RNA-sequencing reads mapping to the transcriptome to the reads mapping to spiked-in RNA from the External RNA Controls Consortium (ERCCs) (Devonshire et al., 2010) to estimate the total relative amount of RNA in the cell (Marinov et al., 2014; Wu et al., 2014), dropping a small number of cells that were qualitatively inconsistent with the rest (Figure S7). We used the measured relationship between *GAPDH* mRNA count and cellular volume from RNA FISH to convert total RNA to relative cellular volume, matching this to our measured volume distribution to obtain cellular volumes. We then used the correlation between FPKM (fragments per kilobase per million fragments mapped) and RNA FISH counts for our gene panel to provide estimates of absolute RNA counts

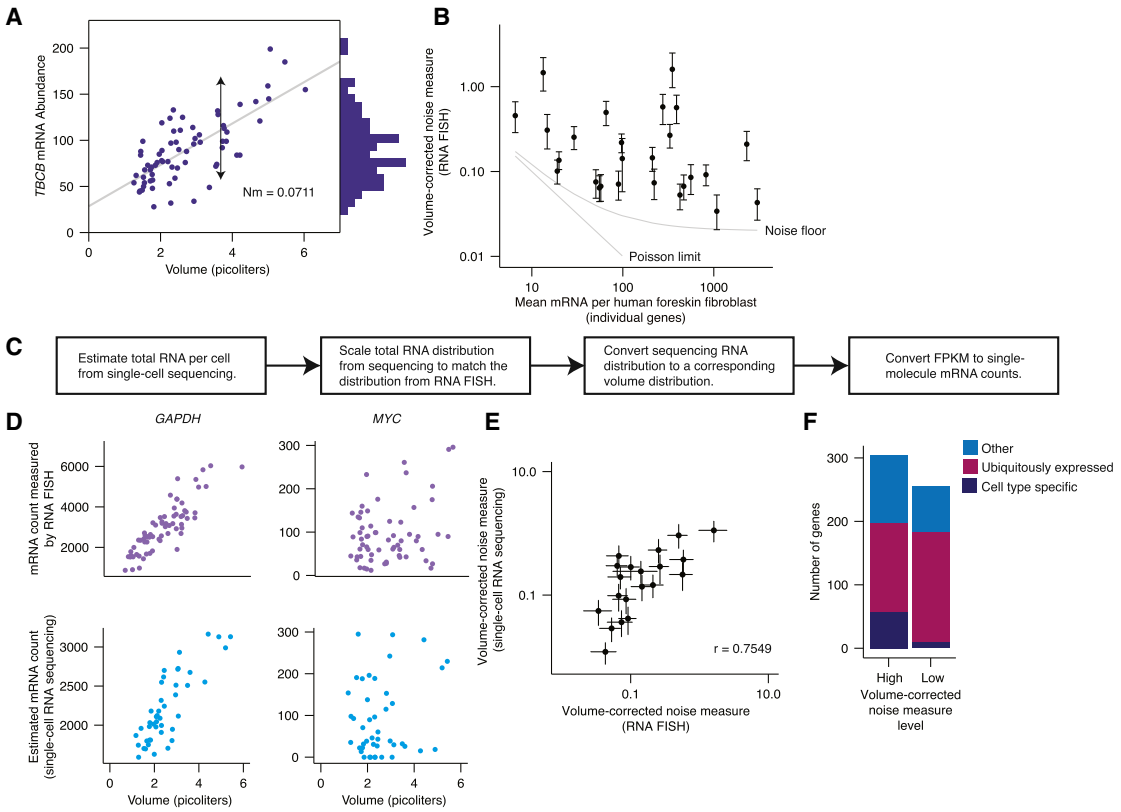


Figure 7. Connection between Single-Cell RNA Sequencing and RNA FISH Reveals that Cell-Type-Specific Genes Exhibit Higher Noise Levels

(A) RNA FISH data. *TBCB* mRNA abundance and volume in primary fibroblast cells. Each point represents a single-cell measurement. Histogram indicates mRNA distribution. Arrow indicates volume-corrected noise measure. Gray line is best linear fit.

(B) Volume-corrected noise measure values for different genes in primary fibroblast cells. Each data point represents a collection of single-cell measurements for one gene. The straight gray line represents the Poisson limit. The curved gray line is the Poisson limit plus our experimental noise limit, a combination of the Poisson limit and a 15% measurement error. Data for each gene is a combination of at least two biological replicates, with at least 30 cells per replicate.

(C) Pipeline for converting FPKM from single-cell sequencing to RNA-FISH-equivalent counts and cellular volume in picoliters.

(D) Qualitative comparison of count versus volume from RNA FISH and single-cell RNA sequencing. Example low-Nm (*GAPDH*) and high-Nm genes (*MYC*).

(E) Comparison between Nm calculated from RNA FISH data and single-cell RNA-seq data. Each point represents a single gene. Nm is calculated by bootstrapping.

(F) Breakdown of high- and low-noise genes into ubiquitously expressing genes and genes that express in a cell-type-dependent manner.

All error bars represent 95% confidence intervals as calculated by bootstrapping. See also Figures S5–S7.

for all genes in individual cells. (It is important to note, however, that there is substantial variance in this relationship and that the relationship is nonlinear.) We found that both RNA FISH and single-cell mRNA sequencing yielded similar results for noisy and non-noisy genes (Figures 7D and 7E).

Upon quantifying Nm for the genes in our RNA FISH experiments in both human foreskin fibroblast and lung cancer cells, we noticed that three of the four genes (*ICAM1*, *LUM*, and *ACTA2*) with the strongest degree of cell-type specificity were also the three genes with the highest noise measure of all the genes in our study (Figure S7). To see whether this trend held more generally, we used our single-cell RNA-sequencing data to explore noise measure across all genes. We selected genes with high or low noise measures and looked for enrichment in genes exhibiting cell-type-specific expression between human lung cancer cell and foreskin fibroblast cell data. We found that the set of high noise measure genes contained a significantly higher proportion of cell-type-specific genes than did the set of low noise genes (Figure 7F; see Figure S7 for classification details). Such findings mirror those showing that more ubiquitously expressed "housekeeping" genes typically exhibit lower levels of noise than other types of genes, although the notion of cell-type specificity is more difficult to relate to studies performed in single-celled organisms (Bar-Even et al., 2006; Newman et al., 2006; Taniguchi et al., 2010).

DISCUSSION

We have shown that individual cells globally control transcription to compensate for variability in the ratio of DNA to cellular content. Our results point to two independent mechanisms: one that compensates for cell size fluctuations and one that

compensates for DNA content changes during the cell cycle. These compensatory mechanisms help to maintain the concentration of mRNA in the cell, which is presumably useful from the perspective of the cell because the rates of most chemical reactions in the cell depend on concentration rather than absolute number. Importantly, the fact that these mechanisms seem to be global in nature, and not gene-specific, means that other specific forms of transcriptional regulation, for instance, during development or in response to particular cues, can still function properly in both large and small cells without the need for a complex interplay between the specific regulation and the mechanism governing concentrational homeostasis; indeed, the expression of most genes is likely not specifically compensated to maintain concentration (Springer et al., 2010). This is not to say that the concentration of most gene expression products is arbitrary and unregulated. Rather, these global mechanisms provide a means by which any such regulation may operate to achieve said concentration without having to take into account differences in DNA content due to cellular volume or DNA content differences. This is important in a number of biological contexts such as development and embryogenesis, in which rapid cell divisions lead to an exponential decrease in individual cell volume, but the organism must maintain the concentration of most proteins while still enabling dynamic transcriptional programs to occur (Nair et al., 2013).

Our work also highlights the importance of taking cellular volume into account when interpreting gene expression data and points to the significance of global factors in studying single cell expression in general (Elowitz et al., 2002; das Neves et al., 2010; Volfson et al., 2006). In particular, our cell fusion experiments show that changing the amount of cellular content in and of itself can lead to changes in total RNA abundance, whereas previous experiments largely relied on cell-size mutants that make inferences of cause and effect more difficult (Fraser and Nurse, 1979; Miettinen et al., 2014; Schmidt and Schibler, 1995; Zhurinsky et al., 2010). These cell fusion experiments directly establish that any perturbation that changes overall cellular volume may result in global changes in overall transcript abundance as a secondary rather than primary effect per the generic mechanism of a diffusible *trans* factor that senses an increased ratio of volume to DNA. Thus, we believe one must take care in interpreting experiments showing global changes in transcript abundance (Lin et al., 2012; Nie et al., 2012), both from the perspective of establishing causal relationships, given that cellular volume/content can by itself change transcription rates, and in the interpretation of the functional significance, given that the concentration of many transcripts will remain roughly the same despite these overall changes. Our study does not, however, address the question of why cells have different volumes and how expression plays a role in that heterogeneity—such questions necessarily involve the examination of mechanisms regulating cell growth and proliferation. Rather, our results show how cells may globally cope with such changes to maintain biomolecule concentration.

We do not yet know the factor (or factors) linking the ratio of cellular volume to DNA content to the amount of transcription. One potential candidate is some element of the general transcription machinery, such as the RNA polymerase II holoenzyme. We have shown that RNA polymerase II is almost completely nuclear, and it is required for transcription; indeed, its reduction changes burst size, much as reduced volume does. Its transcription also scales with volume (Figure S4). Other studies (Marguerat and Bähler, 2012; Zhurinsky et al., 2010) have speculated that RNA polymerase II holoenzyme may act as a limiting factor titrated by DNA, but various studies (Kimura et al., 1999, 2002) show that RNA polymerase II is directly associated with DNA only for short periods of time. Thus, we believe that the scenario in which the factor remains in the nucleus and nuclear volume scales only weakly with cell size is the most plausible given our current evidence (see Supplemental Information). Indeed, our model shows that the weak scaling of nuclear volume with cell size we observe can also explain why we observed higher mRNA concentrations in smaller cells, sometimes by a factor of two or more, further supporting this view, although more experiments will be required to establish this model.

Regardless of the origin of the effect, it is clear that mRNA concentration is typically higher in smaller cells. We do not know the consequences of this effect, in particular on cell growth, or how it may vary in different cell types and contexts. One practical consequence of this finding is that the time-honored practice of normalizing transcript levels to *GAPDH* mRNA abundance, although largely sound, does not fully account for differences in mRNA concentration between small and large cells. This suggests that new strategies may be required for measuring cellular volume when interpreting single-cell RNA-sequencing experiments.

We found it striking that the volume compensation mechanism is distinct from the one that compensates for changes in DNA content as the cell cycle progresses. We found that the burst frequency appears to decrease upon DNA replication for each gene rather than at a particular time in the cell cycle for all genes. One plausible explanation for this feature is to ensure proper transcriptional output regardless of whether a gene is replicated early or late in S phase, which can proceed for many hours. The molecular underpinnings of this mechanism remain unclear, although our results demonstrate that it must be a factor that remains bound to DNA and changes in character during DNA replication. A likely candidate may be a DNA or histone modification that completely coats the DNA during G1 but that is diluted by a factor of two upon DNA replication in S phase.

Together, these findings provide a deeper quantitative understanding of single-cell gene expression and its role in maintaining cellular homeostasis. Further work may elucidate how these homeostatic mechanisms for maintaining biomolecule concentration manifest themselves in biological contexts and whether they are an important point of dysregulation in disease processes.

EXPERIMENTAL PROCEDURES

Cell Culture

We grew primary human foreskin fibroblast cells (CCD-1079Sk, ATCC CRL-2097) and A549 cells (human lung carcinoma, A549, ATCC CCL-185) in Dulbecco's Modified Eagle Medium supplemented with 10% FBS and 50 U/mL penicillin and streptomycin (Pen/Strep). To create quiescent cells, we grew primary fibroblast cells in DMEM with Pen/Strep, without FBS for 7 days. We cultured WM983b-GFP-NLS cells (melanoma cell line from the

lab of Meenhard Herlyn) in Tu 2% media. The WM983b-GFP-NLS contains EGFP fused to a nuclear localization signal driven by a cytomegalovirus promoter that we stably transfected into the parental cell line.

RNA Fluorescence In Situ Hybridization and Imaging

We performed single-molecule RNA FISH on the samples as described previously (Femino et al., 1998; Raj and Tyagi, 2010; Raj et al., 2008). We co-stained the actin cytoskeleton with Phalloidin-Alexa 488 (Life Technologies) to detect cell boundaries.

We used Cyclin A2 mRNA to specifically label cells in S, G2, and M phases (Eward et al., 2004). To distinguish cells in S phase from those in G2 (Robertson et al., 2000; Whitfield et al., 2002), we labeled histone H4 mRNA (see Figure S3). Table S1 lists all sequences of oligonucleotide probes.

We imaged the cells with a Nikon Ti-E equipped with appropriate filter sets. We took a series of optical z sections, each 0.2–0.35 microns high, that spanned the vertical extent of the cell.

Image Analysis and Quantification

We manually identified cell boundaries and counted and localized RNA spots using custom software written in MATLAB (Raj and Tyagi, 2010; Raj et al., 2008). We estimate the technical error in our RNA count determination to be at most 15%.

To compute the volume of a cell, we used the 3D positions of a highly abundant mRNA found by RNA FISH to define the outline of the cell, applying corrections for bias in volume estimation. The volume computation did not depend on the number of spots identified nor on the choice of mRNA (Figure S1). We limited ourselves to the cytoplasmic volume by removing a vertical cylinder corresponding to the nuclear outline.

We identified transcription sites through intron/exon probe colocalization. We manually annotated transcription sites by visually inspecting images of intron and exon probes to determine instances of colocalized signal and computationally determined their intensity.

RNA Degradation

We measured *UBC* and *IER2* mRNA degradation by inhibiting transcription for 4 hr by applying actinomycin D at 1 ug/ml. We interpreted the data using models of volume-dependent or -independent degradation.

LMNA siRNA Knockdown

We incubated primary fibroblast cells with an siRNA targeting *LMNA* for 72 hr, verifying protein knockdown via western blot.

Heterokaryon Formation

We created heterokaryons by pelleting cultures of primary fibroblast cells and WM983b-GFP-NLS cells and resuspending in PEG for 2 min. We added media over the course of 5 min to allow cells to fuse, and then we plated the cells onto two-well chambered coverglasses (Lab-Tek) and fixed the cells after 12 hr.

Fractionation and RNA Polymerase II Western Blot

We performed cell fractionation and blotting as described in (Bhatt et al., 2012) and based on (Wuarin and Schibler, 1994) with modifications.

Triptolide

We degraded RNA polymerase II in primary fibroblast cells by incubating cells in 100 nM triptolide for 1 hr and then fixed cells in methanol.

Repli-seq Analysis

We accessed Repli-seq data from Hansen et al. (2010) using the UW Repli-seq track on the UCSC Genome Browser.

Bulk RNA Sequencing

We sequenced total RNA from primary fibroblast cells. We used the NEB Next Ultimate Library Preparation Kit for Illumina and the Ribo-Zero Magnetic Gold Kit. We used 50 b single-end reads and sequenced each of two replicates at a depth of 10–15 M reads. We aligned reads to hg19 using STAR's included annotation. We quantified reads per gene using HTSeq and a RefSeq hg19 annotation. We calculated FPKM for each gene using R. All sequencing data are available at GEO accession number GSE66053.

Single-Cell RNA Sequencing

We prepared 96 cells for RNA sequencing on a Fluidigm C1 Single-Cell Auto Prep System using a large-size chip. We added ERCC (External RNA Controls Consortium) RNA controls, Mix 1 (Ambion 4456740) at a concentration of 1:10,000 before loading and prepared cDNA libraries as per the Fluidigm instructions. We obtained 75b paired-end reads to a depth of ~1–2 M reads per sample. We quantified reads per gene using HTSeq and a RefSeq hg19 annotation and transformed the data to obtain an estimate of molecule count per cell. All sequencing data are available at GEO accession number GSE66053.

C. elegans Growth and Imaging

We grew N2 (wild-type) and CB502 (*sma-2* mutant) *C. elegans* at 20°C under standard conditions. We performed RNA FISH and analyzed the data as previously described (Raj et al., 2008), determining the volume of each analyzed region computationally.

ACCESSION NUMBERS

The GEO accession number for the single-cell RNA sequencing (primary fibroblast cell line) and the bulk RNA sequencing (primary fibroblast and lung cancer cell lines) data reported in this paper is GSE66053.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, seven figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.molcel.2015.03.005.

ACKNOWLEDGMENTS

We thank members of the Raj lab for many helpful suggestions. We thank Jeff Carey for suggesting the *LMNA* knockdown experiment and Sydney Shaffer for the WM983b-GFP-NLS cell line. We thank Jan Skotheim for useful discussions about mechanism and John I. Murray for pointing out that the simple *trans* factor model would lead to overtranscription in early S phase. We thank Hyun Youk and Uschi Symmons for a careful reading of the manuscript. We thank the Herlyn lab for providing the WM983b cell line. A.R. acknowledges support from an NSF CAREER award, NIH New Innovator Award 1DP2OD008514, and a Burroughs Wellcome Fund Career Award at the Scientific Interface. G.P.N. is a Howard Hughes Medical Institute Fellow of the Life Sciences Research Foundation (http://www.lsrf.org). A.M. was supported by Long-Term Postdoctoral Fellowships of the Human Frontier Science Program (LT000314/2013-L) and EMBO (ALTF858-2012). L.S.C. acknowledges support from an NIH grant (NHGRI R01HG007173), a Damon Runyon Cancer Research Foundation Frey Award, and a Burroughs Wellcome Fund Career Award at the Scientific Interface.

Received: September 22, 2014
Revised: January 16, 2015
Accepted: March 4, 2015
Published: April 9, 2015

REFERENCES

Bar-Even, A., Paulsson, J., Maheshri, N., Carmi, M., O'Shea, E., Pilpel, Y., and Barkai, N. (2006). Noise in protein expression scales with natural protein abundance. Nat. Genet. *38*, 636–643.

Bensaude, O. (2011). Inhibiting eukaryotic transcription: Which compound to choose? How to evaluate its activity? Transcription *2*, 103–108.

Bhatt, D.M., Pandya-Jones, A., Tong, A.-J., Barozzi, I., Lissner, M.M., Natoli, G., Black, D.L., and Smale, S.T. (2012). Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. Cell *150*, 279–290.

Bowsher, C.G., and Swain, P.S. (2012). Identifying sources of variation and the flow of information in biochemical networks. Proc. Natl. Acad. Sci. USA *109*, E1320–E1328.

Brennecke, P., Anders, S., Kim, J.K., Kołodziejczyk, A.A., Zhang, X., Proserpio, V., Baying, B., Benes, V., Teichmann, S.A., Marioni, J.C., and Heisler, M.G. (2013). Accounting for technical noise in single-cell RNA-seq experiments. Nat. Methods *10*, 1093–1095.

Bryan, A.K., Hecht, V.C., Shen, W., Payer, K., Grover, W.H., and Manalis, S.R. (2014). Measuring single cell mass, volume, and density with dual suspended microchannel resonators. Lab Chip *14*, 569–576.

Chubb, J.R., Trcek, T., Shenoy, S.M., and Singer, R.H. (2006). Transcriptional pulsing of a developmental gene. Curr. Biol. *16*, 1018–1025.

Crissman, H.A., and Steinkamp, J.A. (1973). Rapid, simultaneous measurement of DNA, protein, and cell volume in single cells from large mammalian cell populations. J. Cell Biol. *59*, 766–771.

Dar, R.D., Razooky, B.S., Singh, A., Trimeloni, T.V., McCollum, J.M., Cox, C.D., Simpson, M.L., and Weinberger, L.S. (2012). Transcriptional burst frequency and burst size are equally modulated across the human genome. Proc. Natl. Acad. Sci. USA *109*, 17454–17459.

das Neves, R.P., Jones, N.S., Andreu, L., Gupta, R., Enver, T., and Iborra, F.J. (2010). Connecting variability in global transcription rate to mitochondrial variability. PLoS Biol. *8*, e1000560.

Devonshire, A.S., Elaswarapu, R., and Foy, C.A. (2010). Evaluation of external RNA controls for the standardisation of gene expression biomarker measurements. BMC Genomics *11*, 662.

Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic gene expression in a single cell. Science *297*, 1183–1186.

Eward, K.L., Van Ert, M.N., Thornton, M., and Helmstetter, C.E. (2004). Cyclin mRNA stability does not vary during the cell cycle. Cell Cycle *3*, 1057–1061.

Femino, A.M., Fay, F.S., Fogarty, K., and Singer, R.H. (1998). Visualization of single RNA transcripts in situ. Science *280*, 585–590.

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.

Golding, I., Paulsson, J., Zawilski, S.M., and Cox, E.C. (2005). Real-time kinetics of gene activity in individual bacteria. Cell *123*, 1025–1036.

Grün, D., Kester, L., and van Oudenaarden, A. (2014). Validation of noise models for single-cell transcriptomics. Nat. Methods *11*, 637–640.

Hansen, R.S., Thomas, S., Sandstrom, R., Canfield, T.K., Thurman, R.E., Weaver, M., Dorschner, M.O., Gartler, S.M., and Stamatoyannopoulos, J.A. (2010). Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. Proc. Natl. Acad. Sci. USA *107*, 139–144.

Jao, C.Y., and Salic, A. (2008). Exploring RNA transcription and turnover in vivo by using click chemistry. Proc. Natl. Acad. Sci. USA *105*, 15779–15784.

Johnston, I.G., Gaal, B., Neves, R.P., Enver, T., Iborra, F.J., and Jones, N.S. (2012). Mitochondrial variability as a source of extrinsic cellular noise. PLoS Comput. Biol. *8*, e1002416.

Kimura, H., Tao, Y., Roeder, R.G., and Cook, P.R. (1999). Quantitation of RNA polymerase II and its transcription factors in an HeLa cell: little soluble holoenzyme but significant amounts of polymerases attached to the nuclear substructure. Mol. Cell. Biol. *19*, 5383–5392.

Kimura, H., Sugaya, K., and Cook, P.R. (2002). The transcription cycle of RNA polymerase II in living cells. J. Cell Biol. *159*, 777–782.

Levesque, M.J., and Raj, A. (2013). Single-chromosome transcriptional profiling reveals chromosomal gene expression regulation. Nat. Methods *10*, 246–248.

Levsky, J.M., Shenoy, S.M., Pezo, R.C., and Singer, R.H. (2002). Single-cell gene expression profiling. Science *297*, 836–840.

Li, G.-W., and Xie, X.S. (2011). Central dogma at the single-molecule level in living cells. Nature *475*, 308–315.

Lin, C.Y., Lovén, J., Rahl, P.B., Paranal, R.M., Burge, C.B., Bradner, J.E., Lee, T.I., and Young, R.A. (2012). Transcriptional amplification in tumor cells with elevated c-Myc. Cell *151*, 56–67.

Maamar, H., Cabili, M.N., Rinn, J., and Raj, A. (2013). linc-HOXA1 is a noncoding RNA that represses Hoxa1 transcription in cis. Genes Dev. *27*, 1260–1271.

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

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

Marinov, G.K., Williams, B.A., McCue, K., Schroth, G.P., Gertz, J., Myers, R.M., and Wold, B.J. (2014). From single-cell to cell-pool transcriptomes: stochasticity in gene expression and RNA splicing. Genome Res. *24*, 496–510.

Miettinen, T.P., Pessa, H.K.J., Caldez, M.J., Fuhrer, T., Diril, M.K., Sauer, U., Kaldis, P., and Björklund, M. (2014). Identification of transcriptional and metabolic programs related to mammalian cell size. Curr. Biol. *24*, 598–608.

Nair, G., Walton, T., Murray, J.I., and Raj, A. (2013). Gene transcription is coordinated with, but not dependent on, cell divisions during C. elegans embryonic fate specification. Development *140*, 3385–3394.

Newman, J.R.S., Ghaemmaghami, S., Ihmels, J., Breslow, D.K., Noble, M., DeRisi, J.L., and Weissman, J.S. (2006). Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. Nature *441*, 840–846.

Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., Wang, R., Green, D.R., Tessarollo, L., Casellas, R., et al. (2012). c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. Cell *151*, 68–79.

Pomerantz, J.H., Mukherjee, S., Palermo, A.T., and Blau, H.M. (2009). Reprogramming to a muscle fate by fusion recapitulates differentiation. J. Cell Sci. *122*, 1045–1053.

Raj, A., and van Oudenaarden, A. (2008). Nature, nurture, or chance: stochastic gene expression and its consequences. Cell *135*, 216–226.

Raj, A., and van Oudenaarden, A. (2009). Single-molecule approaches to stochastic gene expression. Annu. Rev. Biophys. *38*, 255–270.

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

Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., and Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells. PLoS Biol. *4*, e309.

Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. Nat. Methods *5*, 877–879.

Robertson, K.D., Keyomarsi, K., Gonzales, F.A., Velicescu, M., and Jones, P.A. (2000). Differential mRNA expression of the human DNA methyltransferases (DNMTs) 1, 3a and 3b during the G(0)/G(1) to S phase transition in normal and tumor cells. Nucleic Acids Res. *28*, 2108–2113.

Sanchez, A., and Golding, I. (2013). Genetic determinants and cellular constraints in noisy gene expression. Science *342*, 1188–1193.

Schmidt, E.E., and Schibler, U. (1995). Cell size regulation, a mechanism that controls cellular RNA accumulation: consequences on regulation of the ubiquitous transcription factors Oct1 and NF-Y and the liver-enriched transcription factor DBP. J. Cell Biol. *128*, 467–483.

Senecal, A., Munsky, B., Proux, F., Ly, N., Braye, F.E., Zimmer, C., Mueller, F., and Darzacq, X. (2014). Transcription factors modulate c-Fos transcriptional bursts. Cell Rep. *8*, 75–83.

Shalek, A.K., Satija, R., Adiconis, X., Gertner, R.S., Gaublomme, J.T., Raychowdhury, R., Schwartz, S., Yosef, N., Malboeuf, C., Lu, D., et al. (2013). Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature *498*, 236–240.

Springer, M., Weissman, J.S., and Kirschner, M.W. (2010). A general lack of compensation for gene dosage in yeast. Mol. Syst. Biol. *6*, 368.

Suter, D.M., Molina, N., Gatfield, D., Schneider, K., Schibler, U., and Naef, F. (2011). Mammalian genes are transcribed with widely different bursting kinetics. Science *332*, 472–474.

Swain, P.S., Elowitz, M.B., and Siggia, E.D. (2002). Intrinsic and extrinsic contributions to stochasticity in gene expression. Proc. Natl. Acad. Sci. USA *99*, 12795–12800.

Swift, J., Ivanovska, I.L., Buxboim, A., Harada, T., Dingal, P.C.D.P., Pinter, J., Pajerowski, J.D., Spinler, K.R., Shin, J.-W., Tewari, M., et al. (2013). Nuclear lamin-A scales with tissue stiffness and enhances matrix-directed differentiation. Science *341*, 1240104.

Tani, H., Mizutani, R., Salam, K.A., Tano, K., Ijiri, K., Wakamatsu, A., Isogai, T., Suzuki, Y., and Akimitsu, N. (2012). Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. Genome Res. *22*, 947–956.

Taniguchi, Y., Choi, P.J., Li, G.-W., Chen, H., Babu, M., Hearn, J., Emili, A., and Xie, X.S. (2010). Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. Science *329*, 533–538.

Tzur, A., Kafri, R., LeBleu, V.S., Lahav, G., and Kirschner, M.W. (2009). Cell growth and size homeostasis in proliferating animal cells. Science *325*, 167–171.

Vargas, D.Y., Raj, A., Marras, S.A.E., Kramer, F.R., and Tyagi, S. (2005). Mechanism of mRNA transport in the nucleus. Proc. Natl. Acad. Sci. USA *102*, 17008–17013.

Volfson, D., Marciniak, J., Blake, W.J., Ostroff, N., Tsimring, L.S., and Hasty, J. (2006). Origins of extrinsic variability in eukaryotic gene expression. Nature *439*, 861–864.

Watanabe, N., Ishihara, T., and Ohshima, Y. (2007). Mutants carrying two sma mutations are super small in the nematode C. elegans. Genes Cells *12*, 603–609.

Whitfield, M.L., Sherlock, G., Saldanha, A.J., Murray, J.I., Ball, C.A., Alexander, K.E., Matese, J.C., Perou, C.M., Hurt, M.M., Brown, P.O., and Botstein, D. (2002). Identification of genes periodically expressed in the human cell cycle and their expression in tumors. Mol. Biol. Cell *13*, 1977–2000.

Wu, C.-Y., Rolfe, P.A., Gifford, D.K., and Fink, G.R. (2010). Control of transcription by cell size. PLoS Biol. *8*, e1000523.

Wu, A.R., Neff, N.F., Kalisky, T., Dalerba, P., Treutlein, B., Rothenberg, M.E., Mburu, F.M., Mantalas, G.L., Sim, S., Clarke, M.F., and Quake, S.R. (2014). Quantitative assessment of single-cell RNA-sequencing methods. Nat. Methods *11*, 41–46.

Wuarin, J., and Schibler, U. (1994). Physical isolation of nascent RNA chains transcribed by RNA polymerase II: evidence for cotranscriptional splicing. Mol. Cell. Biol. *14*, 7219–7225.

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., Bähler, J., and Nurse, P. (2010). A coordinated global control over cellular transcription. Curr. Biol. *20*, 2010–2015.