nature genetics

Contributions of low molecule number and chromosomal positioning to stochastic gene expression

Attila Becskei[1], Benjamin B Kaufmann[1,2] & Alexander van Oudenaarden[1]

The presence of low-copy-number regulators and switch-like signal propagation in regulatory networks are expected to increase noise in cellular processes. We developed a noise amplifier that detects fluctuations in the level of low-abundance mRNAs in yeast. The observed fluctuations are not due to the low number of molecules expressed from a gene per se but originate in the random, rare events of gene activation. The frequency of these events and the correlation between stochastic expressions of genes in a single cell depend on the positioning of the genes along the chromosomes. Transcriptional regulators produced by such random expression propagate noise to their target genes.

In living cells, fluctuations of molecule copy number are inevitable under certain conditions. On one hand, such random fluctuations may impair signal propagation and hamper the coordination of cellular activities[1]. On the other hand, noise in gene expression introduces phenotypic heterogeneity in an isogenic population, which may facilitate cellular differentiation or may be beneficial in temporally or spatially heterogeneous environments[2,3].

Total noise is typically divided into two components: intrinsic and extrinsic[4,5]. Intrinsic noise, by definition, originates in the randomness associated with discrete, rare biomolecular events (*e.g.*, mRNA synthesis), when few molecules are involved[6,7]. The remaining noise, which measures fluctuations in the regulation of a gene, is lumped together as extrinsic noise[6]. Noise passed on from upstream transcription factors, as well as changes in the global and local cellular environment of a gene, can all contribute to this component[8,9].

When the number of molecules is low, the discrete nature of molecular concentrations becomes pronounced. If the steady-state concentration corresponds to only a few molecules per cell, changing the concentration of a molecular species by even a single molecule is relatively substantial. Protein concentration ranges from 50 to one million copies per cell in yeast[10], and mRNA abundance varies between 0.001 and 100 copies per cell. Fully 75% of yeast genes have steady-state transcript levels of one or fewer copies per cell averaged over a cell population[11,12]. mRNAs encoding transcription factors, including those regulating the cell cycle, are usually expressed at low to very low levels. For example, the steady-state level of mRNAs encoding the cell cycle transcriptional regulators *SWI6*, *SWI4* and *SWI5* have been reported to be 0.005, 0.3 and 1 copy per cell, respectively[11]. Among all expressed genes, *SWI6* belongs to the set of genes expressed at the lowest level. Because the copy number of mRNAs is typically small compared with protein number, intrinsic noise in regulatory genes is believed to be determined mostly by mRNA rather than protein levels[13].

In contrast to intrinsic noise, extrinsic noise does not depend on the number of molecules expressed from a gene but stems from fluctuations in the transcription factor activity that regulates the gene. Such fluctuations might arise from environmental heterogeneity or from global cell-to-cell variations in metabolic or biosynthetic activities[8]. These fluctuations, even if small, can be magnified when they are coamplified along with signals during signal amplification. Amplification is crucial, for example, in linear transcriptional networks where oscillatory signals of the cell cycle are considerably damped during their propagation[14]. Feedforward and feedback loops, and regulatory switches generated by multistep phosphorylation or cooperative binding, are capable of counteracting the dampening of signals[15–19]. In addition to amplifying the signal, transcription factors acting through positive feedback or ultrasensitive switches can generate considerable cell-to-cell variation in gene expression[2,20–22].

Given that a large proportion of mRNAs is expressed at low level, noise due to the low number of molecules is expected to be great. How intrinsic noise affects regulators expressed at very low levels has not been explored. Noise in gene expression is usually measured by the cell-to-cell variation in expression of fluorescent reporters driven by promoters of interest[23,24]. Intrinsic and extrinsic noise in gene expression can be quantified by monitoring the expression of two fluorescent proteins of different colors driven by the same promoter[6]. In this case, uncorrelated fluctuations in reporter gene expression reflect intrinsic noise, because the realization of rare probabilistic events of mRNA production is not correlated between the two genes. Extrinsic noise is usually associated with correlated fluctuations, because fluctuations in the concentration of a transcription factor are transmitted equally to the identical promoters.

[1]Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [2]Division of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA. Correspondence should be addressed to A.v.O. (avano@mit.edu).

Published online 7 August 2005; doi:10.1038/ng1616

When fluorescent proteins are expressed at very low levels, cellular background fluorescence hampers direct noise measurement by the two-color method. To circumvent this limitation, we developed an amplifier of fluctuations in gene expression. We then raised the number of a given mRNA molecule by multiplying the number of gene copies encoding the mRNA species. Increased molecule number is expected to reduce the intrinsic noise. In this work, we calculate the contribution of intrinsic noise to total noise from the ratio of amplified noise intensity of a gene having multiple copies to that of a single-copy gene in the genome.

RESULTS

Design of a genetic circuit to amplify fluctuations

To amplify the signal generated by weak promoters, we designed a synthetic gene network in the yeast *Saccharomyces cerevisiae* (**Fig. 1a**). A yeast promoter of interest drives the expression of a potent transcriptional activator, rtTA, in the input module. When rtTA is bound to the inducer, doxycycline, it drives the expression of yellow fluorescent protein (YFP) in the response module. Thus, the input signal (expression strength of rtTA, x_1) of the yeast promoter is amplified, and the resulting output signal x_2 is measured by the YFP fluorescence intensity. The ratio of x_2 to x_1 can be varied by adjusting the doxycycline concentration. Thus, the input signals transmitted from both weak and strong promoters can be tuned to produce the same mean output signal while the corresponding noise intensities reflect faithfully the difference in the input noise. This method, therefore, offers the advantage to measure different noise levels at same intensity of fluorescence. The output noise η_2 is calculated by dividing the standard deviation of the YFP expression distribution in a cell population by the mean YFP expression (**Fig. 1b**); this is also called the coefficient of variation. To determine the value of the input noise, noise transmitted to the output module must be measured accurately. For this purpose, we explored how the input noise can be amplified.

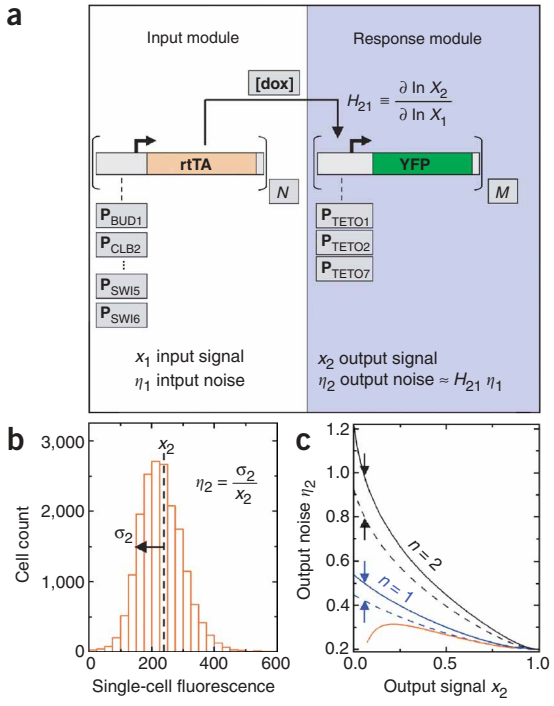

Figure 1 Synthetic gene circuits for noise amplification and transmission. (**a**) The input and output modules are integrated into various chromosomal locations in a diploid cell. Gray boxes denote the experimental variables: the input and response promoters, the copy numbers N and M of the respective modules and the external doxycycline concentration. (**b**) The output noise equals the ratio of the standard deviation σ_2 to the mean x_2 of YFP expression distribution. (**c**) Output noise η_2 of response promoters with a Hill number of $n_H = 1$ (blue lines) or $n_H = 2$ (black lines) was calculated using equation S2 (**Supplementary Note** online), $\eta_g = 0.2$ included. The dashed and full lines stand for input noises of $\eta_1 = 0.4$ and $\eta_1 = 0.5$, respectively. The effect of basal transcription ($b = 0.05$) on η_2 is shown by a solid red line ($n = 1$, $\eta_1 = 0.4$). The difference between the output noises elicited by the two input promoters (arrows) is bigger when the response promoter is cooperative, $n_H = 2$.

Noise in the output module η_2 is directly proportional to the product of noise transmitted from the input module η_1 and the logarithmic gain H_{21} of the promoter when intrinsic noise due to low copy numbers of YFP mRNA or protein is negligible. H_{21} reflects how a relative change in the input signal affects the output signal (**Fig. 1a**)[7]. A theoretical analysis shows that increasing the cooperativity of rtTA binding to the response module should amplify fluctuations. Assuming that a Hill-type function describes the cooperative binding of rtTA to the response promoter leads to a simple direct proportionality between H_{21} and the Hill coefficient n. Thus, a promoter with a Hill coefficient of 2 doubles the output noise η_2 relative to a noncooperative ($n = 1$) but otherwise identical promoter (**Fig. 1c**). η_2 decreases monotonously as the output signal x_2 increases, unless basal transcription $x_{2,bas}$ sets the lower limit of promoter activity (**Fig. 1c**). In the latter case, $|H_{21}|$ reaches a maximum when the output signal $x_2 = \sqrt{x_{2,bas}x_{2,max}}$ and vanishes when the output signal approaches either the basal expression $x_{2,bas}$ or the maximal expression level $x_{2,max}$ (**Fig. 1c** and **Supplementary Note** online).

We used the inherently cooperative binding of transcription factors to multiple binding sites around chromatin to increase cooperativity. The first binding event alters chromatin structure, which enhances the binding of the next transcription factor to the neighboring binding sites[25,26]. This chromatin-mediated cooperativity occurs even in the absence of direct protein-protein interaction between transcription factors. To test the effect of cooperativity, response promoters containing one, two or seven binding sites (P_{TETO1}, P_{TETO2} or P_{TETO7}) for rtTA were incorporated in the synthetic network (**Fig. 1a**).

Cooperativity-based amplification of fluctuations

Varying the concentration of doxycycline generates a response function characteristic of the promoter in the output module (**Fig. 2a**). We used the Hill coefficient as a measure of sensitivity (switch-like nature) of promoter response. When the input module contained the *SWI5* promoter, P_{SWI5}, the Hill numbers were 0.95, 1.1 and 1.45 for P_{TETO1}, P_{TETO2} and P_{TETO7}, respectively. Therefore, the sensitivity of the response increased with but was not proportional to the number of binding sites. In addition to the steeper response, the presence of multiple binding sites substantially lowered the induction threshold of the promoter. Approximately 5 and 45 times lower doxycycline concentrations are needed for P_{TETO2} and P_{TETO7}, respectively, than for P_{TETO1} to reach half-maximal expression. Notably, the basal expression of P_{TETO1} was higher than that of the other constructs. Increased basal expression of promoters with a single binding site has been described for other transcription factors as well[27].

As predicted, η_2 was intensified by cooperative binding and decreased as the output signal (YFP expression) increased (**Fig. 2b** and **Supplementary Fig. 1** online). The maximal value of the output noise was approximately double that of the input noise as measured

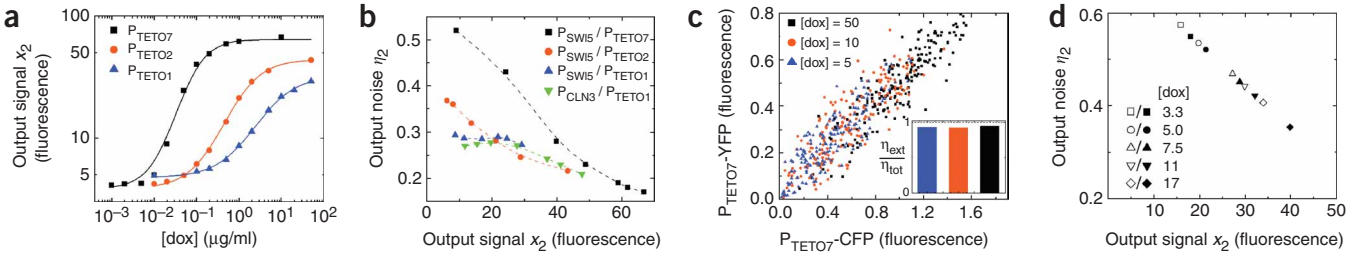

Figure 2 Amplification of noise by transcriptional cooperativity. (**a**,**b**) Output signal x_2 and output noise η_2 of a single copy P_{TETO1}-, P_{TETO2}- or P_{TETO7}-YFP constructs driven by P_{SWI5}-rtTA (ABY0529, ABY0530 and ABY0531). (**a**) The output signal was fitted to a Hill function with basal expression *b*, which is ~0.05 times the maximal expression. (**b**) P_{CLN3}-rtTA P_{TETO1}-YFP (ABY0511a) with the relatively stronger P_{CLN3} was included to show that noise decreases at high transcription rates. (**c**) Extrinsic and intrinsic (uncorrelated) noise in the noise amplifier system. Single-cell two-color fluorescence intensities are shown for the strain ABY0528b (P_{SWI4}-rtTA(S2) $[P_{TETO7}$- CFP$]_1$ / $[P_{TETO7}$-YFP$]_1$) using three different doxycycline concentrations (in μg ml^{-1}). In the inset, the ratio of extrinsic to total noise is shown. (**d**) The output signal and noise of $[P_{TETO7}$-YFP$]_1$ was measured in the absence (empty symbols, ABY0535) and presence of 14 additional $[P_{TETO7}$-CFP$]_2$ (filled symbols, ABY528a) rtTA binding sites. Symbols of the same shape denote induction with the same doxycycline concentration (in μg ml^{-1}).

directly by placing YFP under the control of P_{SWI5}. In the case of P_{TETO1}, the $\eta_2(x_2)$ relation had a plateau. This phenomenon is consistent with the above theoretical analysis of how the basal transcription rate affects noise.

Noise in the activity of cell cycle promoters

Next, we examined how the presence of low-copy-number mRNAs affected the intensity of fluctuations in gene expression. Weak promoters are inaccessible to direct noise measurement by fluorescent proteins. Even P_{SWI5}-YFP produces a fluorescence that is only five times greater than the cellular background fluorescence, although the *SWI5* mRNA has ~100 times higher copy number than the *SWI6* mRNA. Therefore, we calculated the noise intensities η_1 by nonlinear regression of the experimentally determined amplified noise (**Supplementary Note** and **Supplementary Fig. 2** online). The output noise η_2 was determined predominantly by noise transmitted from the input module η_1, because intrinsic noise in the response module was both low and approximately constant in the range of measured YFP fluorescence (**Fig. 2c**).

We used a triple integration of P_{TETO7}-YFP to enhance the output signal further, because the largest absolute differential amplification of noise was observed at low intensities of the output signal. The introduction of 14 additional rtTA binding sites reduced the active pool of rtTA only slightly. This reduction in the number of active rtTA did not increase the intensity of output noise, even when rtTA was expressed from a weak promoter (**Fig. 2d**). Therefore, YFP expression induced by rtTA reflected noise in the input module faithfully; no additional noise was introduced by fluctuations in the rtTA activity itself.

We tuned doxycycline concentrations to obtain equal levels of active rtTA for a broad range of promoter strength (**Fig. 3a**). Among the examined promoters, P_{MYO2} had the lowest noise level (**Fig. 3b**,**c**). P_{MYO2} drives constitutive expression of a cytoskeletal motor protein. In contrast, P_{SWI6} had the highest noise level, roughly two times higher than that of P_{MYO2}. Other cell cycle promoters, P_{SWI4}, P_{SWI5}, P_{CLN2} and P_{CLB2}, had noise levels between those of P_{MYO2} and P_{SWI6}. Because expression at P_{SWI6} is very low, it is expected to be affected substantially by intrinsic noise.

Fully correlated fluctuations at low expression level

To differentiate between intrinsic and extrinsic components of *SWI6* noise, we examined how raising the input gene copy number affected η_2. If noise in gene expression is purely intrinsic due to the low copy number of transcribed mRNA, gene duplication should decrease the output noise by a factor of $\sqrt{2} \approx 1.41$, because intrinsic noise is not correlated between different copies of a gene[6]. But when five copies of P_{SWI6}-rtTA were integrated in tandem at the *ade2* locus, the input noise η_1 was not reduced (**Fig. 4a**). This suggests the absence of intrinsic, uncorrelated noise due to low copy number of expressed molecules. The five-copy construct had

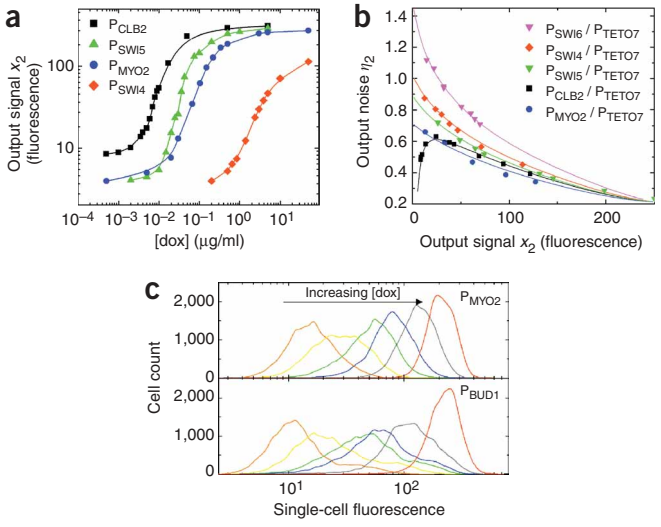

Figure 3 Input noise measurement using the noise amplifier system. (**a**) Output signal of $[P_{TETO7}$-YFP$]_3$ with different input modules. (**b**) The output noise η_2 for P_{CLB2}-, P_{MYO2}-, P_{SWI4}- or P_{SWI6}-rtTA and $[P_{TETO7}$-YFP$]_3$ constructs. Solid lines denote fits by equation S2 (**Supplementary Note** online). P_{CLB2}-rtTA induces transcription even in the absence of doxycycline, which accounts for the observed plateau in $\eta_2(x_2)$. For P_{SWI6}-rtTA no data were obtained at higher output mean values, because of the effect of doxycycline on cell growth when its concentration is higher than 50 μg ml^{-1}. (**c**) Distribution of single-cell $[P_{TETO7}$-YFP$]_3$ fluorescence intensity, using the P_{MYO2}-rtTA (ABY0520) or P_{BUD1}-rtTA (ABY0548c) input modules at different doxycycline concentrations.

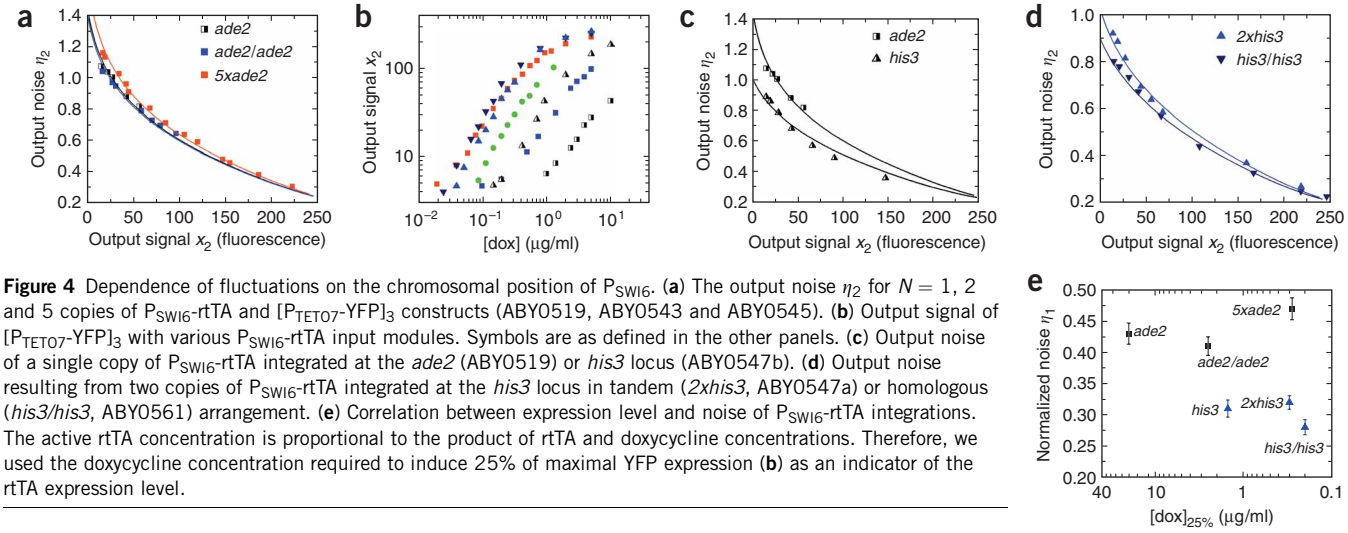

Figure 4 Dependence of fluctuations on the chromosomal position of P_{SWI6}. (**a**) The output noise η_2 for $N = 1$, 2 and 5 copies of P_{SWI6}-rtTA and $[P_{TETO7}\text{-YFP}]_3$ constructs (ABY0519, ABY0543 and ABY0545). (**b**) Output signal of $[P_{TETO7}\text{-YFP}]_3$ with various P_{SWI6}-rtTA input modules. Symbols are as defined in the other panels. (**c**) Output noise of a single copy of P_{SWI6}-rtTA integrated at the *ade2* (ABY0519) or *his3* locus (ABY0547b). (**d**) Output noise resulting from two copies of P_{SWI6}-rtTA integrated at the *his3* locus in tandem (*2xhis3*, ABY0547a) or homologous (*his3/his3*, ABY0561) arrangement. (**e**) Correlation between expression level and noise of P_{SWI6}-rtTA integrations. The active rtTA concentration is proportional to the product of rtTA and doxycycline concentrations. Therefore, we used the doxycycline concentration required to induce 25% of maximal YFP expression (**b**) as an indicator of the rtTA expression level.

a slightly higher output noise than the single-copy construct, which may reflect the fact that longer chromosomal insertions with multiple tandem gene copies have slower association rates with the transcriptional machinery.

When a pair of P_{SWI6}-rtTA constructs was integrated into homologous chromosomal sites of a diploid cell, the input noise η_1 was reduced by a factor of 1.03 ± 0.01 relative to the single-copy integration (**Fig. 4a**). This reflects a strong but not full correlation ($\rho = 0.88 \pm 0.04$) between two copies of a gene when integrated into separate chromosomes. A correlation coefficient of $\rho = 1$ indicates a full correlation between individual events of promoter activation at different copies of a given gene. The full correlation of fluctuations in expression of tandem gene copies indicates that noise is not due to the low number of rtTA mRNA molecules but is generated in an event upstream of mRNA expression.

The effect of the frequency of promoter activation on noise

The absence of intrinsic noise due to the low copy number of expressed molecules can be explained by a stochastic model of gene expression. In such a model, if a promoter is rarely activated and the rate of promoter activation and inactivation is slower than the mRNA production rate at an activated gene, then noise originates in the rare instances of transcription. The number of mRNAs transcribed at a single event of promoter activation has a negligible effect on noise in this case. Thus, if the relatively high noise of *SWI6* is not due to the low copy number of the expressed mRNA, then it might originate in the low frequency of random gene activation. This model predicts that increasing the frequency of promoter activation leads to reduction of noise.

Chromosomal positioning changes expression level and noise

P_{SWI6} has no known transcriptional regulation, because it lacks a TATA box or binding site for any known transcriptional regulator[28,29]. The expression rate of several genes depends strongly on the chromosomal or intranuclear positioning of the gene[30]. Therefore, we changed the chromosomal position of P_{SWI6}-rtTA to examine whether this leads to a change in the rate of gene expression. When a single copy of P_{SWI6}-rtTA was integrated at the *his3* locus, its expression rate doubled and was comparable to that of the two-copy integration at the *ade2* locus (**Fig. 4b**). Noise at *his3* was reduced substantially, by a factor of 1.37 ± 0.04 relative to the *ade2* locus (**Fig. 4c**,**e**), which is close to the maximal potential reduction (factor of 1.41). This is the expected factor of noise reduction when a twofold increase in gene expression stems from the increased frequency of gene activation, which indicates that noise at *SWI6* originates mostly in the rare frequency of gene activation.

Chromosomal positioning affects correlation of fluctuations

We observed a small reduction of noise when two copies of P_{SWI6}-rtTA were integrated at homologous *ade2* loci (**Fig. 4a**). To explore further the effect of chromosomal positioning of genes on fluctuations, two copies of P_{SWI6}-rtTA were integrated either in tandem at a *his3* locus or one copy was integrated at each homologous *his3* locus. The resulting strains had very similar mean expression rates (**Fig. 4b**), but the noise associated with the homologous pair was lower by a factor of 1.09 ± 0.02 than that of the single-copy integration (**Fig. 4d**). In contrast, the tandem integration did not decrease the noise. These data indicate that the events of promoter activation are fully correlated at genes integrated in tandem at a given chromosomal locus. When copies of a gene were positioned on separate homologous chromosomes, the correlation was reduced to $\rho = 0.68 \pm 0.06$, which is still significant. We calculated the uncorrelated noise from the correlation coefficient. Values of η_{UNCOR} were 0.14 ± 0.05 and 0.16 ± 0.05 for the homologous *ade2* and *his3* loci, respectively, which reflects the fact that the noise originating in the uncorrelated fluctuations between homologous chromosomal loci were similar for both the *ade2* and *his3* loci.

Plotting the calculated (input) noise of various P_{SWI6}-rtTA integrations as a function of rtTA expression level showed that noise was independent of rtTA mRNA level (**Fig. 4e**). Noise was determined primarily by the chromosomal-position dependent frequency of gene activation and to a lesser degree by fluctuations uncorrelated between distinct chromosomes.

Propagation of fluctuations in single cells

These results shed light on the origins of fluctuations in concentration of regulators. Next, we studied how regulator noise is propagated to target genes. Although cooperative binding of transcription factors increased the sensitivity in response, it also made the response noisier (**Fig. 2b**). To explore how fluctuations weaken the correlation between the input and output signals, we measured the expression of an rtTA–cyan fluorescent protein (CFP) fusion protein driven by the stronger P_{SWI5} promoter (**Fig. 3a**) and that of the P_{TETO7}-YFP construct in single cells (**Fig. 5a**). We plotted the output fluorescence (YFP) as a

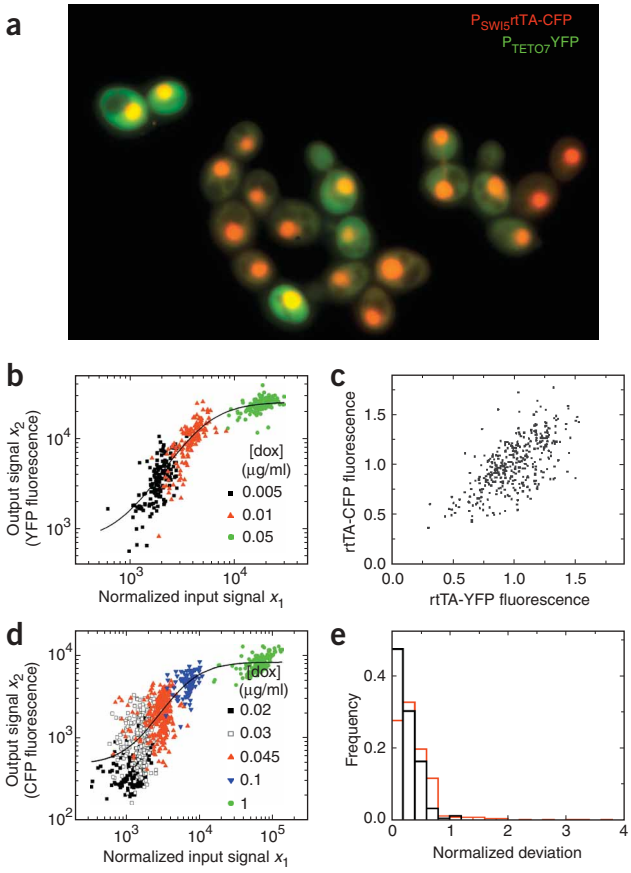

Figure 5 Correlation of fluctuations between the input and output modules. (**a**) Two-color fluorescence imaging of rtTA-CFP (red) and P_{TETO7}-YFP (green) in ABY0557 cells. (**b**) Correlation of input-output fluorescence intensities in single cells (ABY0557). The input fluorescence values (rtTA-CFP) were multiplied by the ratio of the actual to the lowest doxycycline concentration used (0.005 μg ml^{-1}). The basal output signal (P_{TETO7}-YFP) is 752, $n_H = 2.15$ when fitted to single-cell fluorescence data. Fitting the Hill function to the mean output signal when the doxycycline concentration is varied yields $n_H = 2.08$ (data not shown). (**c**) Fluorescence of P_{SWI5}-rtTA-YFP and P_{SWI5}-rtTA-CFP constructs (ABY0559), integrated at homologous *ade2* loci, each having three copies. The total and uncorrelated noises were 0.26 ± 0.03 and 0.15 ± 0.02, respectively. (**d**) The output signal in ABY0549 is the fluorescence intensity of P_{TETO7}-CFP. The input signal corresponds to P_{SWI5}-YFP fluorescence, normalized by doxycycline concentration as in **b**. $n_H = 1.8$ when fitted to single-cell fluorescence data. (**e**) Normalized deviations for ABY0557 (black columns) and ABY0549 (red columns). Doxycycline concentrations of 0.01 and 0.045 μg ml^{-1} were used for respective strains to obtain similar input and output noise values for the two strains. $<x_{2dev}> = 0.25$, $\eta_1 = 0.262$, $\eta_2 = 0.443$ for ABY0557; $<x_{2dev}> = 0.43$, $\eta_1 = 0.292$, $\eta_2 = 0.46$ for ABY0549.

function of CFP fluorescence normalized by the applied doxycycline concentration (**Fig. 5b**). Thus, the concentration of active rtTA can be correlated with the output signal in single cells, assuming a linear relation between doxycycline and active rtTA concentration[14]. We obtained the single-cell response obtained by fitting a Hill-type function to the normalized input and response fluorescence data obtained at three different doxycycline concentrations. The Hill coefficient, obtained from single-cell measurements, was similar to that obtained from population-level measurements. Single-cell measurements showed that even though the output signal was affected by large fluctuations, the input and output signals were significantly correlated in individual cells (**Fig. 5b**). Cells with lower rtTA expression had lower output signals in the sigmoid region of promoter response and vice versa. To quantify the deviation from fully correlated response, we calculated the average difference between single-cell output signals and the fitted response based on the Hill-type function, $<x_{2dev}>$. x_{2dev} equals the absolute value of $(x_2 - x_{2cal})/x_{2cal}$, where x_2 denotes the output signal in a single cell and x_{2cal} denotes the expected output value based on the fitted Hill-type function. We found that $<x_{2dev}> = 0.25$ for the above strain. In comparison, we found a value of 0.43 for a control strain, where the input signal was not measured directly by an rtTA-fluorescent protein fusion but by P_{SWI5}-YFP integrated at homologous chromosomal location with respect to P_{SWI5}-rtTA (**Fig. 5d**). Stochastic expression from the P_{SWI5}-YFP and P_{SWI5}-YFP constructs was not fully correlated because they were positioned on physically distinct chromosomes, even though they were driven by the same promoter (**Fig. 5c**). This accounts for the larger deviation between the input and output signals. In summary, we found a significant correlation between fluctuating concentrations of the transcription regulator and its target gene. In principle, a strong correlation may be absent in some cases of eukaryotic gene activation. A lower correlation is expected, for example, if independent stochastic transitions between accessible and inaccessible states of a promoter influence the efficiency of transcription factor binding.

Transmission of *SWI6* fluctuations to target genes

These results suggest that fluctuations in the concentration of regulatory proteins have a dominant role in determining noise strength in target genes, which is further amplified by the sensitivity in the promoter response. Next, we examined whether fluctuations in the concentration of Swi6p protein, inherited from the *SWI6* promoter, were transmitted to target genes. Promoters that contain multiple binding sites for Swi6p are the most likely to be affected by transmitted noise, because chromatin-mediated cooperativity enhances the propagation of fluctuations. The promoter of the gene *BUD1*, which regulates bud site selection, contains three closely located Swi4p-Swi6p binding sites[28]. Noise associated with P_{BUD1} had a high intensity, comparable to that of P_{SWI6} (**Figs. 3c** and **6a**). To link *SWI6* and *BUD1* fluctuations directly, we placed the expression of *SWI6* under the control of P_{SWI5}, which has roughly half the noise level of P_{SWI6} (**Fig. 3b**). This is expected to reduce the fluctuations in Swi6p concentration. Noise in P_{BUD1} activity was reduced substantially, by a factor of ~2 (**Fig. 6b**). These observations indicate that noise at P_{BUD1} was predominantly transmitted from P_{SWI6}.

P_{CLN2} contains binding sites for Swi6p and for several additional transcriptional factors. The relatively lower noise level at P_{CLN2} might be a result of regulation from multiple transcriptional factors (**Fig. 6a**).

DISCUSSION

Fluctuations in transcription factor activity were amplified through chromatin-mediated cooperativity (**Fig. 2b**), which could affect functioning of cellular regulatory networks because eukaryotic promoters usually contain multiple transcription-factor binding sites. Even though the response signal was more broadly distributed over a cell population, our findings indicate that the input and response signals of cooperative promoters were correlated in individual cells (**Fig. 5b**). In principle, signal amplification through cooperative binding enables an efficient propagation of oscillations in individual cells, but the coamplified extrinsic noise could desynchronize cell-cycle oscillations

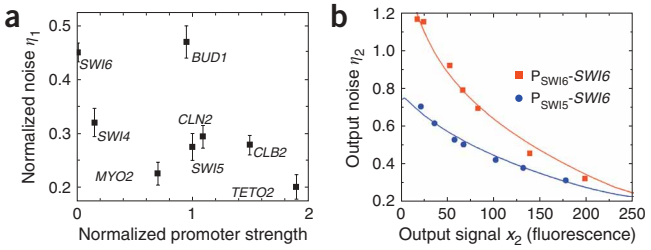

Figure 6 Noise intensities of cell-cycle promoters. (**a**) Relation of noise to promoter strength. Promoter strength was determined using promoter-*lacZ*[14] or promoter-GFP constructs and was normalized by the promoter strength of P_{SWI5} (compare with ~1 copy per cell for *SWI5* mRNA[11]). The input noise was determined as in **Figure 4e**. For TETO2, noise was measured directly by inducing the P_{CLN3}-rtTA:: $[P_{TETO2}$-YFP$]_1$ construct with high concentration of doxycycline (50 μg ml^{-1}) to represent noise at saturated promoter activity. Noise at P_{CLN2} is reduced in *SWI4*/swi4Δ heterozygote strains (from an intensity of 0.294 ± 0.021 in wild-type to 0.265 ± 0.019), which indicates that part of the noise at P_{CLN2} is extrinsic due to regulation by Swi4p-Swi6p. (**b**) Output noise of P_{BUD1} is reduced when *SWI6* is expressed under the control of P_{SWI5} (blue circles, ABY0551a, $\eta_1 = 0.24$) in comparison to a wild-type strain (red squares, ABY0548c, $\eta_1 = 0.47$).

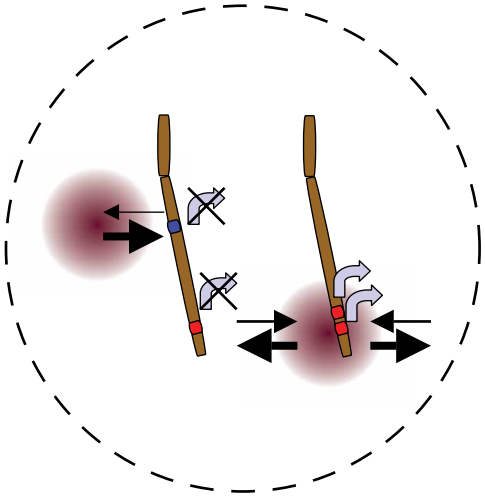

Figure 7 Model of noise generation. Gene activity can be strongly influenced by chromosomal position (blue and red circles stand for different positions; *e.g.*, the *ade2* and *his3* loci). Noise is primarily determined by the association rate of the gene with the transcriptional machinery (purple foci). The width of the arrows denotes the chromosomal position–dependent rate of gene activation and deactivation. The two adjacent circles on a chromosome represent tandem gene integration. Expression of genes at homologous sites is not fully simultaneous. Noise in regulator expression, generated by the above mechanisms (*e.g.*, *SWI6*), is transmitted to target genes (*e.g.*, *BUD1*).

in an initially synchronized cell population. A related phenomenon was observed in the circadian clock of zebrafish cells. Population-level measurements showed a fast decay of biorhythmic oscillations after the cells were placed in a dark environment. Individual cells oscillated regularly, however, but with distributed phases and fluctuating period[31]. Stochastic effects may also be buffered by the robust functioning of the cell cycle[32,33]. Perturbations of the cell-cycle network can weaken noise buffering or further intensify noise and consequently derail the coordination of cell-cycle activities[34]. For example, impairing the positive feedback in cyclin-dependent kinase activation leads to stochastic, quantized cycle times in fission yeast[35].

We started the search for intrinsic noise by correlating noise to promoter strength, expecting that as the promoter became weaker the noise would become correspondingly stronger. This showed that genes with very low expression (*SWI6* and, to some extent, *SWI4*) had high noise level, whereas *BUD1* was affected by transmitted fluctuations (**Fig. 6b**). Notably, genes reported to have copy numbers of ~1 mRNA (*e.g.*, *SWI5*) per cell had only moderate noise, slightly above the noise strength associated with saturated promoter activity (**Fig. 6a**).

The full correlation of gene expression from multiple-copy P_{SWI6} constructs in single cells is notable, given the very low transcript abundance and the fact that P_{SWI6} has no known transcriptional regulators[28]. These observations show that the large cell-to-cell variation in P_{SWI6} activity is not intrinsic noise due to the low number of expressed mRNAs. Large fluctuations may instead originate in the rare, random instances of promoter activation leading to mRNA production[36]. In principle, promoter activation may be initiated by fluctuations in concentration of general transcription factors or by changes in the accessibility of the promoter in a given chromosomal region to transcription factors. For example, transcriptional activity correlates with the spatial positioning of chromosomal regions in the yeast nucleus, and coregulated genes may be clustered in the nucleus[37–39]. Additionally, active genes associate with spatially dispersed foci of RNA polymerase II in mammalian cells[40].

The above experimental observations are consistent with a model in which fluctuations are determined by the random association of chromosomal regions with dispersed foci of the transcriptional machinery (**Fig. 7**). Some chromosomal regions have slower association rates with the foci (such as *ade2*), and others have faster rates (such as *his3*). Less frequent stochastic association results in larger fluctuations. This model also implies that fluctuations are fully correlated when genes are inserted in tandem arrangement, because they associate with the foci simultaneously. The association of homologous chromosomal regions with these foci occurs at equal frequency (hence the equal noise) but not completely simultaneously. As a result, the correlation in fluctuations between the two loci is diminished.

In spatially extended systems, not all uncorrelated fluctuations equal intrinsic noise. Even spatially homogeneous extrinsic fluctuations can result in inhomogeneous accumulations of reaction components in some simple chemical diffusion reaction systems[5]. In principle, such inhomogeneously dispersed regulatory components can generate uncorrelated extrinsic fluctuations. In single cells, spatially inhomogeneous distribution of regulators can generate uncorrelated fluctuations in expression of two identical genes situated in different positions in the nucleus. Therefore, the uncorrelated fluctuations of homologous chromosomal regions could be considered local extrinsic fluctuations.

Our results link stochastic behavior of transcriptional networks to two primary sources: the random, rare instances of gene (promoter) activation determined primarily by chromosomal positioning of a gene (**Fig. 4e**) and noise transmission through transcriptional regulators (**Figs. 5** and **6b**). In higher eukaryotes, positioning of genes along the chromosomes and in the nucleus probably exerts a strong influence on stochastic gene activation for the following two reasons.

With increasing complexity of eukaryotic organisms, transcription of the genome is more and more repressed by multiple layers of inhibitory mechanisms[41,42]. Therefore, events of gene activation by overcoming the repression may occur rarely and randomly for some genes. Such events may cause the large cellular variation in gene expression when gene reporters are positioned close to repressed

regions of chromosomes[43]. In addition, individual genes and components of the transcriptional and post-transcriptional processing machineries have specific intranuclear spatial arrangement[44,45]. Therefore, expression of genes may be, to some extent, temporally uncorrelated, owing to their distinct spatial regulation. The above phenomena may contribute to the stochastic behavior of gene expression observed in higher eukaryotes[46,47].

METHODS

Construction of plasmids and strains. We inserted *Kpn*I-promoter-*Bam*HI, *Bam*HI-rtTA-*Eco*RI, *Bam*HI-YFP-*Eco*RI or *Bam*HI-CFP-*Eco*RI sequences into pRS402, pRS306 or pRS303 vectors unless otherwise specified[14]. The P_{BUD1}, P_{CLN3}, P_{SWI5} and P_{MYO2} promoter sequences correspond to 600-, 1,200-, 892- and 677-bp regions upstream of the start codon of the respective genes. We obtained P_{TETO1} and P_{TETO7} by replacing the tet operator region in P_{TETO2} upstream of the CYC1-TATA region[14]. rtTA corresponds to the rtTA(S2) variant. The rtTA-CFP and rtTA-YFP fusion proteins are constructed from sequences encoding nuclear localization signal–tagged rtTA, rtTA-SV40NLS-*Bgl*II, fused to a C-terminally positioned *Bam*HI-CFP-*Eco*RI. We integrated the P_{SWI5}-*SWI6*(1–500 + stop codon) cassette into the *SWI6* locus by linearizing the sequence with *Bse*RI. All yeast strains (**Supplementary Table 1** online) are cogenic with a diploid W303 strain (*MATa/α ade2-1 leu2-3 ura3 trp1-1 his3-11,15 can1-100*). We determined the number of integrations by Southern blotting by digesting genomic DNA with restriction enzymes that cut the integrated construct at a single site.

Growth conditions and flow cytometric and microscopic analyses. We grew yeast cells in minimal medium (synthetic dextrose supplemented with the appropriate drop-out solution) at 30 °C. Exponentially growing cells were induced with doxycycline for 5 h unless otherwise indicated. Cell division time is ~90–100 min in this regime. We stopped expression by transferring the culture on ice. Cell density was $A_{600\ nm}$ of 0.2–0.5. We evaluated fluorescence intensities using a narrow gating for forward and side scattering so that the most represented population was selected (central region of the scatter distribution). We measured at least 10,000 cells. The fluorescence distributions of promoter-GFP and -YFP constructs were identical.

For microscopic analysis, we obtained data on a Nikon TE2000 inverted microscope using Metamorph (Universal Imaging) data acquisition software. For each doxycycline concentration, we measured 150–300 cells. Cells were outlined automatically in MATLAB using only the phase-contrast image. Rare cases where the outline did not follow the cell wall were manually discarded. We then mapped cell regions to the YFP and CFP fluorescent images from which the total and average fluorescence were determined. We obtained average cell fluorescence by dividing the total cell fluorescence by cell area and subtracted a constant camera background from each frame. The autofluorescence of yeast cells was ~7 times higher using a CFP filter cube than using a YFP filter cube.

Data analysis. For presentation of output noise data, measurements are shown for mean fluorescence intensities higher than two times the cellular background fluorescence, $x_{2bg} = 3.6 \pm 0.3$, $\eta = 0.36 \pm 0.35$. We obtained the $\eta_2(x_2)$ relation for parameter fitting by subtracting the residual global noise at maximal induction, $\eta_g = 0.2$, from the measured output noise: $\eta_2(x_2)^2 = \eta_2(x_2)^2_{measured} - \eta_g^2$. We obtained the input noise η_1 by fitting equation S2 (**Supplementary Note** online) to the $\eta_2(x_2)$ relation, with $n = 1.45$ and $K = 30$ (**Supplementary Note** online). We obtained normalized values of η_1 by dividing η_1 by 2.0. This normalization factor corresponds to the ratio of input noise extrapolated from measurement of the amplification method ($\eta_1 = 0.55$; **Fig. 3b**) to the noise measured directly, $\eta = 0.27$, using the P_{SWI5}-YFP construct, which had a mean fluorescence of 29.

Correlation between fluctuations in gene activation. We modeled the production of mRNA by probabilistic transitions between active (A) and inactive (I) states of a gene[48]. The rates of promoter activation (I→A) and inactivation (A→I) are denoted by λ and μ, respectively. The mRNA production at an active promoter (A→A + mRNA) proceeds at a rate of ν, and mRNA degradation (mRNA→∅) at a rate of δ. In this case, the mRNA noise is given by the following equation:

$$\eta_1^2 = \frac{\delta}{\lambda}\left(\frac{\lambda+\mu}{\nu}+\frac{\mu}{\lambda+\mu+\delta}\right).$$

The rate of transcription at an activated promoter is $\nu \approx 10\ min^{-1}$ (ref. 49). In comparison, δ and μ have smaller values (0.05–0.5 min^{-1}). Assuming that the instances of gene activation are very rare events at low level of transcription, $\lambda < 0.01\ min^{-1}$. In this case, $\nu > \delta$ and $\mu > \lambda$, which simplifies the above equation to

$$\eta_1^2 \approx \frac{1}{\lambda}\left(\frac{\ln 2}{t_{1/2}\mathrm{mRNA} + t_{1/2}\mathrm{gene}}\right),$$

where $t_{1/2}$mRNA and $t_{1/2}$gene denote the half-lives of the mRNA and the activated gene, respectively. Assuming further that $\delta > \mu$, noise is given by a binomial distribution of active gene states: $\eta_1^2 = \mu/\lambda$. In this region of parameter values, η_1^2 is inversely proportional to the frequency of gene activation and is independent of mRNA production rate (*e.g.*, as the copy number of the integrated gene is varied).

Mechanisms of transcription have been proposed that are capable of producing very-low-copy-number mRNAs, in such a way that the timing of mRNA bursts are fully correlated even between multiple nonidentical promoters[50]. A hypothetical transcription machinery composed of the RNA polymerase linked to the DNA polymerase could produce RNA in a fully deterministic way. A single mRNA molecule will be transcribed from each gene during the passage of the above hypothetical machinery through the chromosome at a given cell division cycle.

To describe how the random instances of gene activation are correlated between two copies of a gene, we introduced the stochastic variables P_1 (m_1, $\eta_{1,1}$) and P_2 (m_2, $\eta_{1,2}$) with mean and noise given in parenthesis for each gene. If a pair of gene copies is integrated in tandem or at homologous sites ($m_1 = m_2$ and $\eta_{1,1} = \eta_{1,2} = \eta_{1,S}$), then the resulting stochastic process $P_{1+2}(m_1 + m_2, \eta_{1,D})$ depends on the correlation coefficient ρ between P_1 and P_2.

Then, the correlation between the transcriptional bursts at the two gene copies is calculated from the ratio of noise of two gene copies to that of a single gene copy:

$$\rho = 2\left(\frac{\eta_{1,D}}{\eta_{1,S}}\right)^2 - 1.$$

ρ reflects the proportion of correlated noise with respect to total noise:

$$\rho = \frac{\eta_{COR}^2}{\eta_{COR}^2 + \eta_{UNCOR}^2}.$$

Therefore, the uncorrelated noise was calculated as $\eta_{UNCOR} = \eta_{1,S}\sqrt{1-\rho}$.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank J. Pedraza, W. Tansey and M. Thattai for discussions. A.B. is a Long Term Fellow of the Human Frontier Science Program. This work was supported by a grant from the US National Institutes of Health and a US National Science Foundation CAREER grant.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 11 February; accepted 21 June 2005
Published online at http://www.nature.com/naturegenetics/

1. Rao, C.V., Wolf, D.M. & Arkin, A.P. Control, exploitation and tolerance of intracellular noise. *Nature* **420**, 231–237 (2002).
2. Becskei, A., Seraphin, B. & Serrano, L. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. *EMBO J.* **20**, 2528–2535 (2001).
3. Thattai, M. & van Oudenaarden, A. Stochastic gene expression in fluctuating environments. *Genetics* **167**, 523–530 (2004).
4. Erdi, P. & Toth, J. *Mathematical Models of Chemical Reactions: Theory and Applications of Deterministic and Stochastic Models* (Manchester University Press, Manchester, UK, 1989).

5. Horsthemke, W., Doering, C.R., Ray, T.S. & Burschka, M.A. Fluctuations and correlations in a diffusion-reaction system - unified description of internal fluctuations and external noise. *Phys. Rev. A.* **45**, 5492–5503 (1992).
6. Elowitz, M.B., Levine, A.J., Siggia, E.D. & Swain, P.S. Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002).
7. Paulsson, J. Summing up the noise in gene networks. *Nature* **427**, 415–418 (2004).
8. Pedraza, J.M. & van Oudenaarden, A. Noise propagation in gene networks. *Science* **307**, 1965–1969 (2005).
9. Rosenfeld, N., Young, J.W., Alon, U., Swain, P.S. & Elowitz, M.B. Gene regulation at the single-cell level. *Science* **307**, 1962–1965 (2005).
10. Ghaemmaghami, S. *et al.* Global analysis of protein expression in yeast. *Nature* **425**, 737–741 (2003).
11. Holland, M.J. Transcript abundance in yeast varies over six orders of magnitude. *J. Biol. Chem.* **277**, 14363–14366 (2002).
12. Velculescu, V.E. *et al.* Characterization of the yeast transcriptome. *Cell* **88**, 243–251 (1997).
13. Swain, P.S. Efficient attenuation of stochasticity in gene expression through post-transcriptional control. *J. Mol. Biol.* **344**, 965–976 (2004).
14. Becskei, A., Boselli, M.G. & van Oudenaarden, A. Amplitude control of cell-cycle waves by nuclear import. *Nat. Cell Biol.* **6**, 451–457 (2004).
15. Hasty, J., Dolnik, M., Rottschafer, V. & Collins, J.J. Synthetic gene network for entraining and amplifying cellular oscillations. *Phys. Rev. Lett.* **88**, 148101 (2002).
16. Isaacs, F.J., Hasty, J., Cantor, C.R. & Collins, J.J. Prediction and measurement of an autoregulatory genetic module. *Proc. Natl. Acad. Sci. USA* **100**, 7714–7719 (2003).
17. Basu, S., Mehreja, R., Thiberge, S., Chen, M.T. & Weiss, R. Spatiotemporal control of gene expression with pulse-generating networks. *Proc. Natl. Acad. Sci. USA* **101**, 6355–6360 (2004).
18. Xiong, W. & Ferrell, J.E., Jr. A positive-feedback-based bistable 'memory module' that governs a cell fate decision. *Nature* **426**, 460–465 (2003).
19. Pramila, T., Miles, S., GuhaThakurta, D., Jemiolo, D. & Breeden, L.L. Conserved homeodomain proteins interact with MADS box protein Mcm1 to restrict ECB-dependent transcription to the M/G1 phase of the cell cycle. *Genes Dev.* **16**, 3034–3045 (2002).
20. Acar, M., Becskei, A. & van Oudenaarden, A. Enhancement of cellular memory by reducing stochastic transitions. *Nature* **435**, 228–232 (2005).
21. Shibata, T. & Fujimoto, K. Noisy signal amplification in ultrasensitive signal transduction. *Proc. Natl. Acad. Sci. USA* **102**, 331–336 (2005).
22. Hooshangi, S., Thiberge, S. & Weiss, R. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc. Natl. Acad. Sci. USA* **102**, 3581–3586 (2005).
23. Blake, W.J., Kaern, M., Cantor, C.R. & Collins, J.J. Noise in eukaryotic gene expression. *Nature* **422**, 633–637 (2003).
24. Raser, J.M. & O'Shea, E.K. Control of stochasticity in eukaryotic gene expression. *Science* **304**, 1811–1814 (2004).
25. Li, G. & Widom, J. Nucleosomes facilitate their own invasion. *Nat. Struct. Mol. Biol.* **11**, 763–769 (2004).
26. Vashee, S., Melcher, K., Ding, W.V., Johnston, S.A. & Kodadek, T. Evidence for two modes of cooperative DNA binding *in vivo* that do not involve direct protein-protein interactions. *Curr. Biol.* **8**, 452–458 (1998).
27. Melcher, K. & Xu, H.E. Gal80-Gal80 interaction on adjacent Gal4p binding sites is required for complete GAL gene repression. *EMBO J.* **20**, 841–851 (2001).
28. Harbison, C.T. *et al.* Transcriptional regulatory code of a eukaryotic genome. *Nature* **431**, 99–104 (2004).
29. Basehoar, A.D., Zanton, S.J. & Pugh, B.F. Identification and distinct regulation of yeast TATA box-containing genes. *Cell* **116**, 699–709 (2004).
30. Maillet, L. *et al.* Evidence for silencing compartments within the yeast nucleus: a role for telomere proximity and Sir protein concentration in silencer-mediated repression. *Genes Dev.* **10**, 1796–1811 (1996).
31. Carr, A.J. & Whitmore, D. Imaging of single light-responsive clock cells reveals fluctuating free-running periods. *Nat. Cell Biol.* **7**, 319–321 (2005).
32. Lengronne, A. & Schwob, E. The yeast CDK inhibitor Sic1 prevents genomic instability by promoting replication origin licensing in late G(1). *Mol. Cell* **9**, 1067–1078 (2002).
33. Thornton, B.R., Chen, K.C., Cross, F.R., Tyson, J.J. & Toczyski, D.P. Cycling without the cyclosome: modeling a yeast strain lacking the APC. *Cell Cycle* **3**, 629–633 (2004).
34. Magee, J.A., Abdulkadir, S.A. & Milbrandt, J. Haploinsufficiency at the Nkx3.1 locus. A paradigm for stochastic, dosage-sensitive gene regulation during tumor initiation. *Cancer Cell* **3**, 273–283 (2003).
35. Sveiczer, A., Csikasz-Nagy, A., Gyorffy, B., Tyson, J.J. & Novak, B. Modeling the fission yeast cell cycle: quantized cycle times in wee1–cdc25Delta mutant cells. *Proc. Natl. Acad. Sci. USA* **97**, 7865–7870 (2000).
36. Pirone, J.R. & Elston, T.C. Fluctuations in transcription factor binding can explain the graded and binary responses observed in inducible gene expression. *J. Theor. Biol.* **226**, 111–121 (2004).
37. Menon, B.B. *et al.* Reverse recruitment: The Nup84 nuclear pore subcomplex mediates Rap1/Gcr1/Gcr2 transcriptional activation. *Proc. Natl. Acad. Sci. USA* **102**, 5749–5754 (2005).
38. Casolari, J.M. *et al.* Genome-wide localization of the nuclear transport machinery couples transcriptional status and nuclear organization. *Cell* **117**, 427–439 (2004).
39. Hurst, L.D., Pal, C. & Lercher, M.J. The evolutionary dynamics of eukaryotic gene order. *Nat. Rev. Genet.* **5**, 299–310 (2004).
40. Osborne, C.S. *et al.* Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat. Genet.* **36**, 1065–1071 (2004).
41. Struhl, K. Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. *Cell* **98**, 1–4 (1999).
42. Martin, D.I. Transcriptional enhancers–on/off gene regulation as an adaptation to silencing in higher eukaryotic nuclei. *Trends Genet.* **17**, 444–448 (2001).
43. Ahmad, K. & Henikoff, S. Modulation of a transcription factor counteracts heterochromatic gene silencing in *Drosophila*. *Cell* **104**, 839–847 (2001).
44. Misteli, T. Concepts in nuclear architecture. *Bioessays* **27**, 477–487 (2005).
45. Roix, J.J., McQueen, P.G., Munson, P.J., Parada, L.A. & Misteli, T. Spatial proximity of translocation-prone gene loci in human lymphomas. *Nat. Genet.* **34**, 287–291 (2003).
46. Dernburg, A.F. *et al.* Perturbation of nuclear architecture by long-distance chromosome interactions. *Cell* **85**, 745–759 (1996).
47. Nutt, S.L. *et al.* Independent regulation of the two Pax5 alleles during B-cell development. *Nat. Genet.* **21**, 390–395 (1999).
48. Peccoud, J. & Ycard, B. Markovian modelling of gene product synthesis. *Theor. Popul. Biol.* **48**, 222–234 (1995).
49. Iyer, V. & Struhl, K. Absolute mRNA levels and transcriptional initiation rates in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **93**, 5208–5212 (1996).
50. Guptasarma, P. Does replication-induced transcription regulate synthesis of the myriad low copy number proteins of *Escherichia coli?*. *Bioessays* **17**, 987–997 (1995).