Review

Varying virulence: epigenetic control of expression noise and disease processes

Kathryn Miller-Jensen[1,2*], Siddharth S. Dey[3*], David V. Schaffer[2,3,4] and Adam P. Arkin[2,4,5]

[1] Department of Biomedical Engineering, Yale University, New Haven, CT 06511, USA
[2] Institute for Quantitative Biosciences, University of California, Berkeley, CA 94720, USA
[3] Department of Chemical and Biomolecular Engineering and the Helen Wills Neuroscience Institute, University of California, Berkeley, CA 94720, USA
[4] Department of Bioengineering, University of California, Berkeley, CA 94720, USA
[5] Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Gene expression noise is a significant source of phenotypic heterogeneity in otherwise identical populations of cells. Phenotypic heterogeneity can cause reversible drug resistance in diseased cells, and thus a better understanding of its origins might improve treatment strategies. In eukaryotes, data strongly suggest that intrinsic noise arises from transcriptional bursts caused by slow, random transitions between inactive and active gene states that are mediated by chromatin remodeling. In this review, we consider how chromatin modifications might modulate gene expression noise and lead to phenotypic diversity in diseases as varied as viral infection and cancer. Additionally, we argue that this fundamental information can be applied to develop innovative therapies that counteract 'pathogenic noise' and sensitize all diseased cells to therapeutic intervention.

Introduction

Every cell population exhibits differences among individual cells, even when the cells are genetically identical and the environment is carefully controlled. Although non-genetic heterogeneity might arise from varied sources (Box 1), some fraction of it continuously arises from 'biochemical noise': random fluctuations in molecular concentrations and biochemical reactions that affect cellular mechanisms. Biochemical noise is especially apparent in gene expression within individual cells, because genes are usually present in very low numbers (typically 1–2 copies per cell). Gene expression is thus a fundamentally noisy process that can result in non-genetic heterogeneity in prokaryotic and eukaryotic cell populations [1,2].

When stochastic fluctuations are amplified by biological mechanisms such as regulatory circuits, phenotypic heterogeneity that might be advantageous in certain biological contexts can arise [3]. In an illustrative study, two yeast strains were engineered to transition randomly between two phenotypes in response to stochastic fluctuations in gene expression, but each with different switching rates [4]. When environmental conditions were stable, the strain with slower phenotypic transitions dominated the population, but when the environment fluctuated more quickly, faster transitions were beneficial. Similarly, microbial populations use phenotypic heterogeneity as a strategy to respond to unpredictable changes in the environment [5,6]. For example, *Escherichia coli* populations contain a physiologically distinct subset of slow growing 'persister' cells that can survive sudden exposure to antibiotics, caused by reversible drug resistance that is physiological rather than genetic [7]. The presence of such persister cells, possibly arising from non-genetic heterogeneity, could be of clinical importance for other pathogens [8,9]. In another example, stochastic noise underlies the lysis–lysogeny cell fate decision of the bacteriophage λ virus, which permits some viruses to persist in a dormant state [1].

Phenotypic heterogeneity that arises from cell fate decisions driven by stochastic gene expression is also emerging as a persistence mechanism in diverse mammalian diseases. For example, recent evidence suggests that biological noise underlies probabilistic entry into and exit from mammalian viral latency [10], in which a subset of viruses establish silent infections that might permit viruses to evade the host immune system and reactivate later to produce more progeny [11]. In a very different example, cell-to-cell variability in the proteome of cancer cells appears to allow a small population of persister cells to survive chemotherapy [12]. Importantly, in contrast to prokaryotes discussed above, eukaryotic genes are subject to complex mechanisms of chromatin regulation mediated by transcription factors and chromatin modifying enzymes that modulate stochastic fluctuations in gene expression [13]. Thus, chromatin might provide a mechanism for varying the probability of stochastic transitions between phenotypes, and possibly increase the stability of one phenotype versus another.

Here, we review evidence of this in the probabilistic fate decision that underlies latency in human immunodeficiency virus (HIV). We also propose that understanding the mechanisms that underlie stochastic gene expression and

Corresponding author: Arkin, A.P. (aparkin@lbl.gov)
* These authors contributed equally to this work.

  doi:10.1016/j.tibtech.2011.05.004

Box 1. Sources of noise and non-genetic heterogeneity

Gene expression noise is usually categorized as either intrinsic or extrinsic [67,68]. Commonly, intrinsic noise is defined as fluctuations in molecular numbers that arise from the discrete and stochastic nature of the mechanisms of direct interest and under measurement (e.g. expression from a single gene), whereas extrinsic noise includes those factors outside those mechanisms. Experimentally, these two sources of non-genetic heterogeneity have been distinguished by measuring gene expression from two identical promoters in the same cell. Cell-to-cell differences in the expression from both promoters (i.e. extrinsic noise) might result from systematic changes in the availability and activity of general host factors that underlie cellular processes, as well as from variation in the cellular microenvironment, cell size, and noise induced by unequal partitioning of cellular components during division [69,70]. In these cases, the changes in activity are correlated between the two promoters. By contrast, uncorrelated expression in single cells is indicative of intrinsic noise. In yeast, transcriptional noise has primarily been attributed to extrinsic factors [13]. It is more difficult to parse the role of intrinsic and extrinsic noise in mammalian cells, and fewer studies have been completed, but evidence so far points to a larger role for intrinsic noise [16,18]. Mechanistically, the boundary between intrinsic and extrinsic sources of variation can be vague, because coupled systems in cells propagate intrinsic noise (possibly amplifying or filtering it), leading sometimes to changes in cellular decisions that then further increase heterogeneity in the population [71–73]. Extrinsic variation might also be static relative to the time scale of the process being studied. It might then be characterized as altering the initial conditions of the process, such as when a phage infects a cell at different points in the cell cycle, thereby changing the cell volume, availability of host factors, and time to division [70,74]. Extrinsic variation could also alter boundary conditions, such as structural/chemical differences in the microenvironment of the cell that change physiology [75]. In these cases, the resultant phenotypic heterogeneity, although still arising from non-genetic cell-to-cell variability, is attributed to a deterministic, rather than stochastic, mechanism [76].

phenotypic heterogeneity might reveal strategies to counteract mechanisms of disease persistence.

Chromatin remodeling as a source of transcriptional bursts

It is well documented that intrinsically noisy gene expression results, in large part, from bursts of transcript and protein production in a number of cellular systems [2]. In prokaryotes, such noise is primarily attributed to translational bursts that occur when ribosomes generate many proteins from a single transcript [1,13,14]. By contrast, noise in eukaryotic cells primarily arises from transcriptional bursts, which are compatible with a model in which the promoter infrequently transitions between an inactive and an active gene state (Figure 1a and Box 2) [15–17]. Transcriptional bursting has been studied most extensively in yeast, but there is also evidence of such bursting in mammalian cells [16,18]. Furthermore, cell-to-cell variability in transcript and protein levels in human cells is consistent with a stochastic gene state transition model [18–20].

The gene state transition model is widely accepted; however, the source of transcriptional bursts is incompletely understood (Box 2). One hypothesis that has gained considerable traction is that stochastic events in nucleosome remodeling cause infrequent transitions between an inactive and active gene state, and thus underlie transcriptional bursting (Figure 1b) [13]. Nucleosomes are the fundamental unit of chromatin, consisting of ~147 base pairs of DNA wrapped around an octamer of the four core histone proteins. Transcription factors must compete with

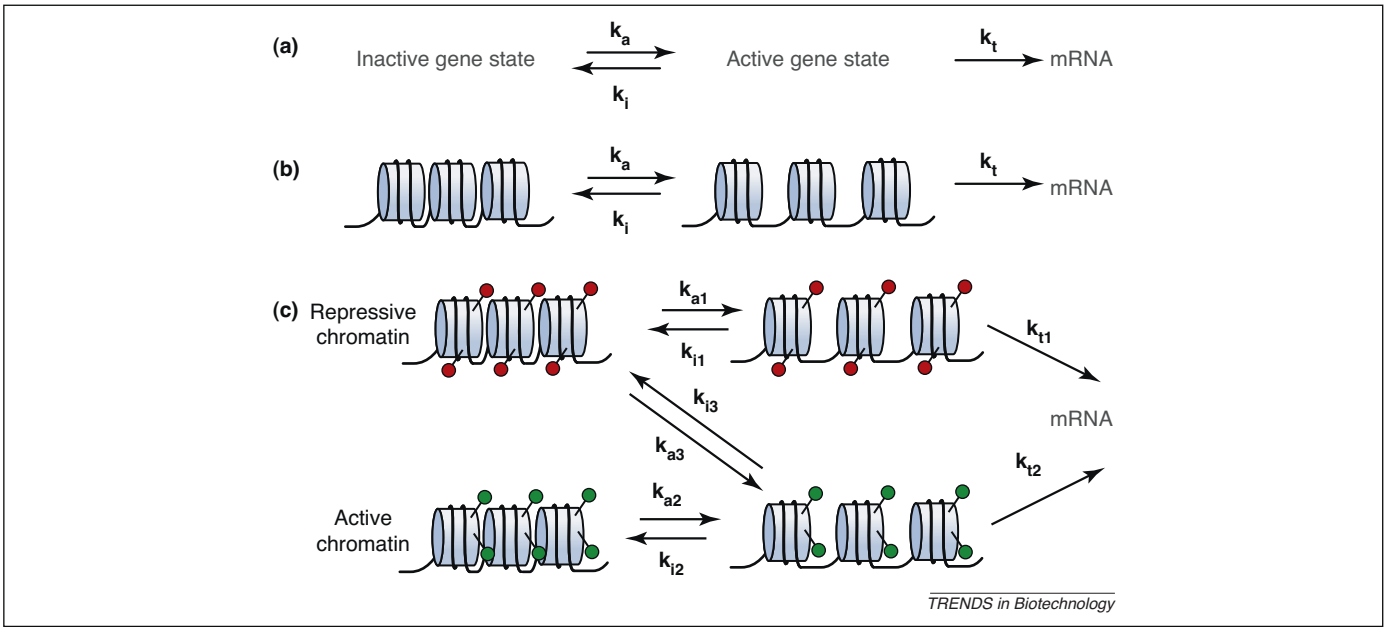
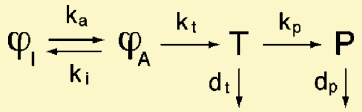


Figure 1. Chromatin remodeling can account for a two-state gene expression model. **(a)** Basic model of gene expression in which the gene transitions from an inactive state with no transcriptional activity to an active state in which it produces transcripts at a rate of k_t. The rates of transition to and from the active state are k_a and k_i, respectively. When k_a and k_i are slow relative to k_t and the transcript degradation rate, transcriptional bursting results. Inactive and active gene states potentially describe many different underlying mechanisms. **(b)** One interpretation of the gene state transition is between a state in which nucleosomes block DNA accessibility (e.g. to transcription factors) and remodeled chromatin in which the nucleosomes have been moved or removed. **(c)** Chromatin modifications directly and indirectly modify the stability of the DNA–nucleosome contacts, which might change the rates of k_a and k_i, and results in different regimes of gene expression noise. Many different chromatin modifications are possible, including both repressive (red) and activating (green), which might have different rates of transition between active and inactive promoter states. Thicker arrows indicate faster rates, with repressive modifications leading to more stable inactive states, and activating modifications leading to more stable active states. Chromatin modifications that are inherited by the next cell generation could lead to short-term 'cell memory' and bistable gene expression. See Box 2 for further discussion.

Box 2. Biological interpretation of the two-state model of gene expression

The two-state model of gene expression assumes that the promoter exists in either an activated state (φ_A) that produces mRNA probabilistically at a fixed rate (k_t), or an inactivated state (φ_I) that is unproductive. Transitions between these states occur at rates k_a and k_i, protein production occurs at a rate k_p, and transcript and protein degradation occur at rates d_t and d_p, respectively (see also Figure 1a).



As discussed in the review, extensive evidence supports the hypothesis that these states represent different configurations of chromatin or nucleosome binding; however, other mechanisms could determine the transitions, including bound and unbound pre-initiation complexes, cell cycle effects, and the presence of transcription factories [77]. This simple stochastic model has been used to reproduce a range of single-gene expression profiles [16–18,68,78]. Importantly, the relative values of the model rate constants, relative to the transcript degradation rate (d_t), determine the regime of gene expression. If the rates of transition are very fast relative to transcript degradation (k_a, $k_i >> d_t$), then gene expression is relatively continuous, and intrinsic noise is low. By contrast, if gene inactivation is much faster than activation ($k_i >> k_a$) and transcript degradation ($k_i >> d_t$), then transcriptional 'bursting' results. In this regime, transcripts are produced in bursts during short-lived transitions to the active promoter state (φ_A). The dynamics of bursting are often described using two parameters: burst size (defined as k_t/k_i, with at least one transcript produced in the active state) and burst frequency (defined as k_a). Intrinsic noise is high in this regime, with lower burst frequency increasing noise even more. Finally, if gene state transitions are extremely slow relative to transcript degradation (k_a, $k_i << d_t$), then each promoter state is relatively stable, with transcripts produced in pulses that result in bimodal protein expression [68,78]. This might be reflective of more stable chromatin states that confer short-term memory on the system (see main text).

nucleosomes for binding DNA, and therefore nucleosomes are considered to be general repressors of transcription [21,22]. ATP-dependent chromatin remodeling enzymes periodically move or disassemble nucleosomes along the DNA, which opens the chromatin to favor transcription factor binding and gene activation [23].

Some of the first experimental studies to measure directly transcriptional bursts have shown that the probability of gene activation varies with chromosomal position, which strongly suggests a link between chromatin remodeling and promoter state transitions [16,24]. Recent genome-wide studies in yeast have further demonstrated that increased variability in gene expression is positively correlated with nucleosome density close to the transcriptional start site [25–27]. Yeast genes with higher expression noise are also more sensitive to perturbation of chromatin regulators, suggesting that noisy genes are subject to chromatin remodeling [25,26].

The relationship between nucleosome remodeling and noise has been quantitatively studied in the repressed yeast PHO5 gene, which contains positioned nucleosomes upstream of the transcription start site (TSS) that are remodeled upon activation of the transcription factor Pho4. A simple computational model assuming random assembly and disassembly of the positioned nucleosomes can recreate experimentally measured stochastic properties of PHO5 gene expression, and suggest that nucleosome disassembly is the rate-limiting step in gene expression [28]. In a follow-up study, Mao *et al*. have demonstrated a strong correlation between nucleosome loss at the TSS and the level of gene expression, with higher levels of nucleosome occupancy resulting in lower gene expression but higher intrinsic noise [29]. This study has provided strong experimental evidence for a mechanistic relation between the dynamics of nucleosome remodeling and intrinsic noise (Figure 1b).

Additional experimental evidence for nucleosome remodeling as a source of stochastic bursting has been obtained through studying the endogenous promoter architecture of two cell-cycle-regulated promoters in yeast. An activating transcription factor binding site (TFBS) for the promoters is exposed in a nucleosome-depleted region proximal to the TSS, which results in reliable gene expression in every cell during a brief window of the cell cycle [30]. Bai *et al.* recently have demonstrated that embedding the TFBS into a nucleosome results in highly variable bimodal 'on–off' gene expression, in which the cell cycle genes are expressed in only a fraction of cells in each cycle [30]. Increasing the concentration of the activating transcription factor increases the 'on' fraction, suggesting that the bimodal expression pattern is related to lower accessibility of the activating transcription factor to its TFBS [30]. Finally, the promoter state is partially heritable to the next generation with a half life of one cell cycle, providing direct evidence that nucleosome-induced noise gives rise to cellular heterogeneity.

Chromatin modifications appear to modulate noise and increase heterogeneity

Nucleosome remodeling provides a basis for transcriptional bursting in eukaryotes, but the dynamics of that remodeling are more complex than implied by the two-state model. The histones comprising nucleosomes are subject to extensive post-translational modifications on at least 60 histone residues, with combinatorial possibilities adding even greater diversity [31]. Modified histones primarily serve as binding sites for transcription factors and chromatin remodeling enzymes that affect nucleosome stability and remodeling [32]. These modifications are generally classified as 'activating' and 'repressive', but the effect of histone modifications can also be conceptualized as changing the rate of transcription and/or the rates of transition between discrete promoter states (Figure 1c), which affects gene expression noise. For example, in highly heterochromatic regions of the chromosome with many repressive chromatin marks and high nucleosome density, the probability of gene activation is extremely low and thus might lead to rare, highly stochastic gene expression [24]. A model similar to this has been proposed to explain position–effect variegation in *Drosophila melanogaster* [33].

When the rate of chromatin remodeling decreases relative to the cell cycle, and chromatin modifications are inherited by daughter cells, memory is conferred on intermediate time scales (e.g. multiple generations although perhaps not tens of generations) that can increase cell-to-cell variability [34]. For example, in single *Dictyostelium* cells, the frequency of stochastic transcriptional activation of a gene is inherited from mother to daughter cells and

Box 3. Gene expression noise in an epigenetic landscape

One framework for conceptualizing the heterogeneous phenotypic states that can arise within a cell is a multidimensional space in which stability is plotted as a function of cell state, or the expression levels of the various genes that are being studied [79,80]. Stability can be loosely defined as the probability of finding the cell in a given state. This multidimensional space is commonly referred to as the epigenetic landscape, and states that are at the bottom of a trough represent stable cellular states or phenotypes. Such 'minimum energy' (i.e. high probability) states are called 'attractor states', because other higher energy cellular states in the vicinity of this minimum energy state are attracted to them (Figure I). Gene regulatory networks within a cell can give rise to several attractor states within the epigenetic landscape that correspond to the distinct phenotypes displayed by the cell. However, metastable states arising from further layers of gene regulation might render the epigenetic landscape rugged, such that the landscape now consists of several local minima or 'sub-attractors', in addition to the attractor states (Figure I). Biologically induced noise could potentially drive cells from a stable attractor state towards one of the sub-attractors in its vicinity, and trap such cells within these local minima. Thus, stochastic perturbations could provide a mechanism for a cell to jump from one sub-attractor state to another (i.e. from one distinct phenotype to another) until it is pulled towards a stable attractor state, from which transitions are rare. Chromatin remodeling and the associated modifications that affect remodeling mechanisms could potentially alter the depth (or stability) of a phenotypic state to raise or lower the probability of transitioning away from that state (Figure 1c). Alternatively, chromatin modifications could themselves define a metastable sub-attractor state if those modifications result in a distinct gene expression program maintained over several cell generations (Figure 1c).

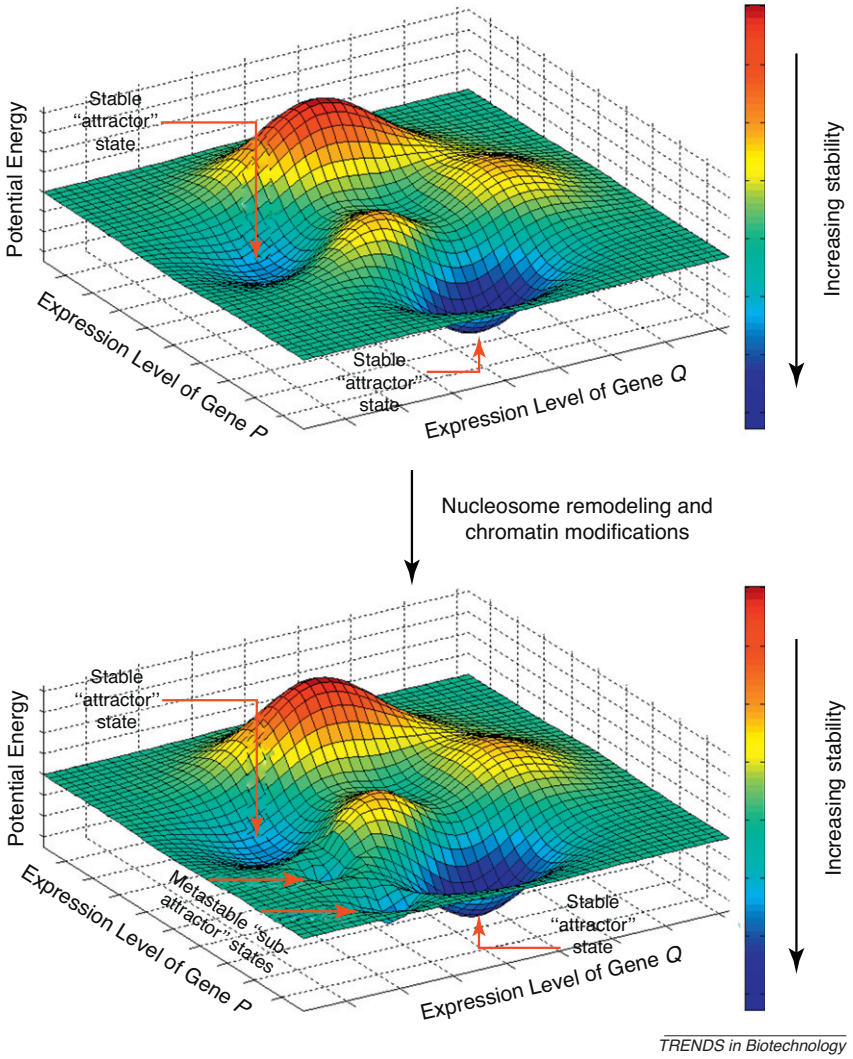


Figure I. Epigenetic landscapes: theoretical framework for conceptualizing the emergence of heterogeneous phenotypic states in genetically identical cells. The figure shows a hypothetical 3D epigenetic landscape in which the state of the system is described by the expression levels of two genes, *P* and *Q*, along the x and y axis, respectively. The z axis shows the stability or potential energy associated with different expression level combinations for genes *P* and *Q*. The upper figure shows a hypothetical model in which the two troughs represent two stable or attractor phenotypic states. Gene expression noise – generated by inactive–active promoter transitions in genes regulating those states – leads to the probabilistic transition between states. Chromatin modifications might change the topography of the landscape, which might change the probability of transitioning between two phenotypes. Global chromatin modifications might further lead to metastable or sub-attractor states, represented by local minimas in the lower figure, in which cells may be briefly trapped, giving rise to transient heterogeneity within a cell population. These landscapes are hypothetical and were generated using Matlab®.

persists for more than one cell cycle [35]. Importantly, chromatin modification (methylation of histone 3 at lysine 4) is required for this memory because it stabilizes the frequency of transcription between generations. In another example, stochastic bimodal expression of the cytokine interleukin (IL)-4 gene in T helper lymphocytes is explained by short-term memory conferred through chromatin remodeling [36]. Mariani *et al.* have demonstrated that a computational two-state model describing the opening of the chromatin at the IL-4 promoter could account for experimental measurements of stochastic IL-4 expression. The model predicts that the rate of inactivation of the promoter is slow (on the order of a cell cycle), which increases the probability of subsequent IL-4 stimulation in the same cells. Thus, slow chromatin dynamics result in distinct subpopulations with different levels of activation that persist over several generations (Figure 1c).

Modulation of noise by chromatin and its modifications becomes even more pronounced when coupled with the complex regulatory mechanisms that exist in multicellular organisms. Thus, chromatin modifications might affect the stability of phenotypic states and the probability of transitioning from one phenotypic state to another; concepts that can be visualized with an epigenetic landscape (Box 3). This chromatin complexity presents some limitations for using nucleosome remodeling and gene expression noise in yeast as a proxy for higher eukaryotes such as mammals. Although the yeast genome does contain some heterochromatin (or heterochromatin-like) structures, most yeast genes are transcribed, in contrast to multicellular organisms in which gene expression attenuation and silencing mediated by repressive chromatin modifications are broadly used to maintain distinct cell type functionalities [23]. Thus, it is possible that the range of promoter dynamics is narrower in yeast, and the time scale of heritability is shorter. However, there is currently no experimental evidence to support these concepts directly.

A challenge going forward is to quantify experimentally the relation between the dynamics of chromatin modification, stochastic fluctuations in transcription, and phenotypic heterogeneity. Here, we argue that HIV-1 might provide a strong experimental system to study gene expression noise in the context of more complex eukaryotic chromatin environments.

Role for gene expression noise and chromatin in viral latency

The choice between HIV replication and latency, a decision with substantial consequences for human health, is an example of a heterogeneous fate decision that might result from stochastic gene expression [10]. Following infection and integration in $CD4^+$ T lymphocytes, HIV usually actively replicates in the cell, but on rare occasions, it fails to establish a productive infection and enters a latent state [37]. Latent HIV proviruses are highly stable and persist even in patients on highly active antiretroviral therapy [38]. Upon activation of the host T cells, latent virus can reactivate and reseed viremia, and for this reason, patients must continuously take antiviral therapy. Consequently, HIV latency is the most significant barrier to curing viral infection [39].

The virally encoded transcriptional activator Tat is essential for establishing a productive infection and for reactivating latent virus. Tat is transcribed early during

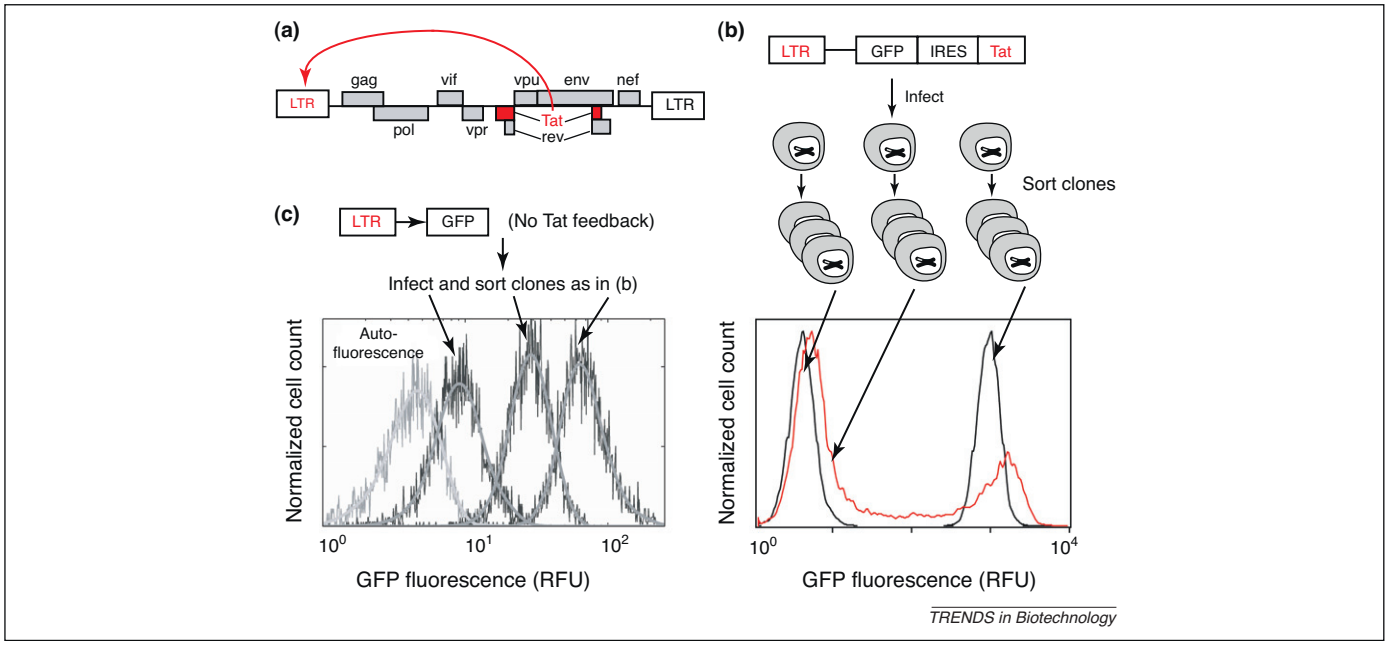


Figure 2. Noise-driven HIV gene expression circuits might underlie the replication versus latency decision. **(a)** The HIV-1 protein Tat strongly activates gene expression from the viral LTR promoter by setting up a positive feedback loop. **(b)** Jurkat cells were infected with a model HIV with the LTR promoter driving expression of a GFP reporter gene and Tat. Jurkat clones containing a single viral integration were expanded and analyzed by flow cytometry. GFP distributions for clones with three different viral integrations are presented, demonstrating three infected cell fates. No Tat expression or high Tat expression results in unproductive or active infection, respectively (black histograms). Alternatively, when levels of Tat are low, stochastic fluctuations can result in low and high gene expression in the same clone (red histogram) [40]. Low gene expression might result in a delay in transcriptional activation, leading to a latent state. **(c)** Jurkat cells were infected and analyzed as in **(b)** but the model HIV did not contain Tat feedback, to examine noise in basal LTR expression. GFP distributions for clones with three different viral integrations are presented, demonstrating how mean expression levels vary with genomic environment. The large gene expression variation is consistent with a bursting model of gene expression. For histograms in **(b)** and **(c)**: total cell counts were normalized to the maximum number of cells found at a given expression level.

HIV infection and significantly amplifies expression from the HIV long terminal repeat (LTR) promoter in a strong positive feedback loop (Figure 2a). However, when Tat protein levels are low, such as just after infection or before reactivation, stochastic fluctuations in Tat gene expression can lead to delays before activation of the Tat-mediated positive feedback loop, resulting in subpopulations of low and high gene expression in a genetically identical population of cells (Figure 2b) [40]. In this case, stochastic gene expression noise, coupled with a strong positive feedback loop, operates as a genetic 'switch' that regulates entry and exit from latency [41]. If viral replication is sufficiently delayed, other cellular factors mediate chromatin changes that further suppress viral transcription and maintain (and further stabilize) the latent state [42,43].

If stochastic fluctuations in gene expression underlie the HIV replication-versus-latency decision, what is the source of the noise? Consistent with studies of stochastic fluctuations in eukaryotic gene expression discussed previously, our group and others recently have demonstrated that a two-state model of transcriptional bursting could account for HIV gene expression variance in the absence of Tat feedback (Figure 2c) [18,44]. Both studies used flow cytometry to analyze gene expression in Jurkat T cell clones infected with single integrations of a model HIV-1 virus that contained the full-length LTR that drives expression of a GFP reporter, but without Tat or other viral genes. The mean GFP expression depended on the genomic environment at the site of integration, but GFP expression for all clones was highly variable (Figure 2c). A two-state stochastic model of gene expression was fitted to the experimentally measured GFP distribution to determine values of burst size (defined as k_t/k_i) and burst frequency (defined as k_a) for each clone (Figure 1a). Both studies concluded that burst size and burst frequency vary across integration sites, and this could account for 'noisy' HIV gene expression [18,44].

Mean expression levels from model HIV-1 vectors correlate with specific chromatin features at their integrations [45], and therefore it is likely that similar chromatin features also largely determine the genomic environment that contributes to expression variation. Thus, our analysis suggests that chromatin features at the site of integration strongly influence transcriptional burst size, which varies from a few to tens of transcripts (Figure 3a) [18]. By contrast, burst frequency values do not vary systematically with integration position, and the variation in burst frequency among clones is less, suggesting that it depends on more global factors (Figure 3b) [18]. Importantly, however, because the time scale of bursting is shorter than protein degradation, a model based on protein data cannot resolve whether integration position primarily affects the dynamics of promoter inactivation (k_i), or transcriptional productivity in the active state (k_t). Additional experiments are necessary to distinguish between these mechanisms; however, given the overlapping effects of many molecular regulatory mechanisms on transcriptional dynamics, these experiments are not trivial.

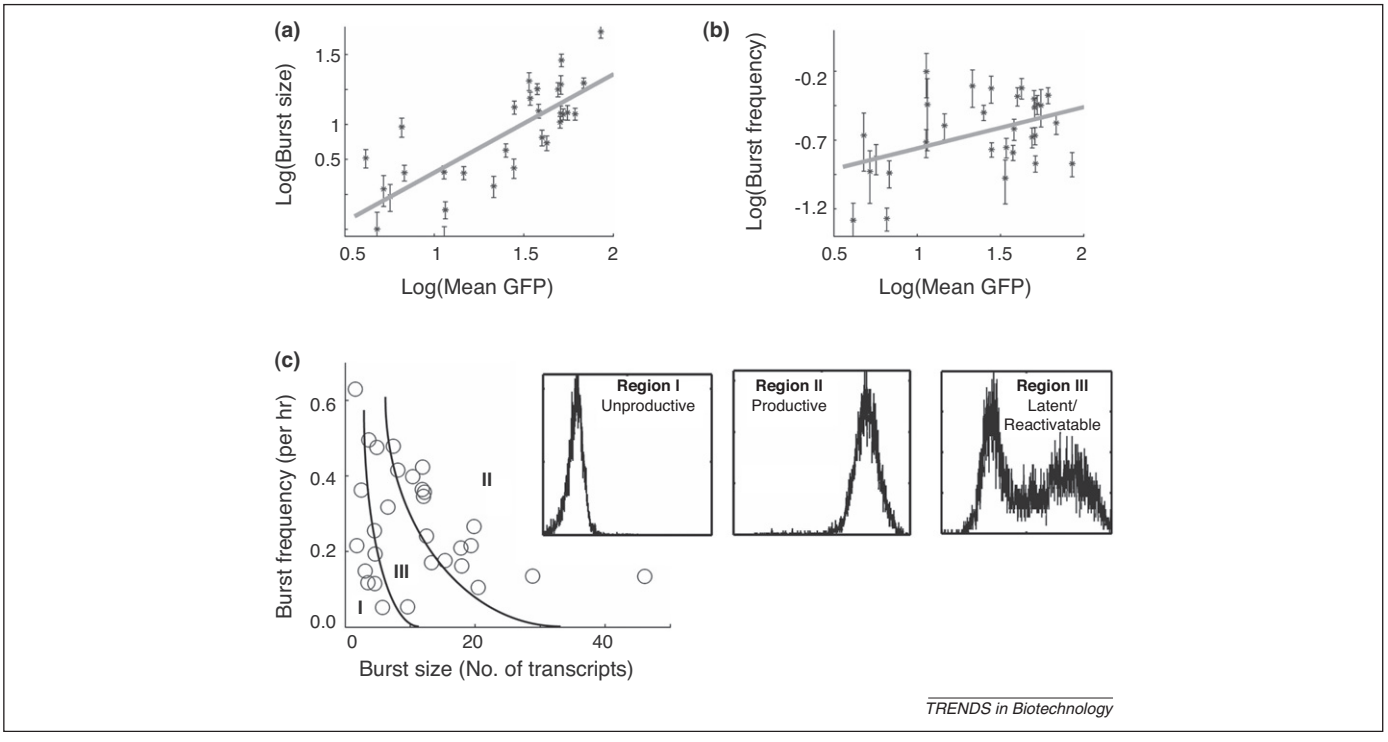


Figure 3. Chromatin environment at the site of HIV integration results in transcriptional bursts and high gene expression noise. GFP distribution from 31 HIV-infected clones were systematically fit to the two-state gene expression model described in Figure 1a and Box 2, to define values for the key parameters of burst size (k_t/k_i; Figure 1a) and burst frequency (k_a; Figure 1a) for each clone. **(a)** Burst size correlates with the chromatin environment at the site of integration (as characterized by the mean level of GFP expression). **(b)** Burst frequency is not strongly correlated with integration position. **(c)** Basal promoter fluctuations of the HIV transactivator Tat, which activates a strong positive feedback loop depicted in Figure 2, might lead to distinct infected-cell fates (described in Figure 2b): Region I: low burst size and burst frequency values never lead to stable Tat expression (unproductive infection); Region II: high burst size and burst frequency values always lead to stable Tat expression (active replication); Region III: large fluctuations in basal transcription infrequently lead to high levels of Tat expression (latent state). Histograms demonstrate representative expression patterns for single-integration clones of a similar vector that includes Tat (Figure 2b). Bursting parameters plotted are based on 31 clonal fits but region boundaries are hypothetical. Adapted with permission from [18].

A two-state bursting model of HIV gene expression is consistent with the long-standing knowledge that nucleosomes are positioned at the HIV TSS, and the observation that, in the absence of Tat, the HIV promoter binds repressive factors that maintain an inactive chromatin configuration [46]. The HIV LTR promoter also contains binding sites for activating factors; therefore, binding competition with repressive factors could lead to an infrequent all-or-none binding of activating factors that directly remodel promoter-bound nucleosomes to establish a short-lived transcriptionally active chromatin configuration [47]. In fact, it has been suggested that tumor necrosis factor (TNF)-α, which activates the LTR via the binding of transcription factor nuclear factor-κB, does so primarily by increasing burst frequency [44], although it is challenging to use a steady-state gene expression model to elucidate transcriptional activation mechanisms that are dynamic in nature following cytokine or other stimulation. The development of new experimental methods to investigate HIV's dynamic gene expression properties might yield general insights into mammalian transcriptional regulation, since features of the HIV promoter that account for its heterogeneous expression pattern are probably present in other mammalian promoters.

Together these data suggest that by sampling different chromatin environments, HIV establishes a range of noisy gene expression distributions, which might act to specify distinct infected cell fates when coupled to Tat positive feedback (Figure 2a,b). In particular, we might speculate that, because productive viral replication depends on robust expression of the HIV protein Tat, integration with high basal gene expression (large burst size) generates sufficient viral Tat protein to replicate, whereas integration with very low burst size results in unproductive infection (Figure 3c). By contrast, those integrations with small or intermediate basal burst sizes might infrequently (stochastically) generate sufficient Tat for positive feedback activation, which favors latency [18]. Therefore, we suggest that nucleosome remodeling and features of the chromatin environment lead to HIV phenotypic diversity that might facilitate viral persistence through the establishment of latency.

Herpesviruses also persist as latent infections for the lifetime of the host, but can reactivate and establish new infections under certain environmental conditions [48]. The alpha herpesvirus herpes simplex virus (HSV) establishes latency in infected neurons, where the viral genome exists episomally and the promoter of VP16 – a transactivating protein that positively regulates viral replication – is incorporated into heterochromatin [49]. Recent evidence suggests that exiting viral latency depends on stochastic expression of VP16, which must reach an adequate level, in combination with the necessary host factors, to start viral replication [50], somewhat analogous to HIV. Similar evidence of stochastic gene expression has been reported for latent cytomegalovirus infection in the lung [51]. In another example, the latent gamma herpesvirus Kaposi's sarcoma-associated herpesvirus (KSHV) is also associated with chromatin modifications that silence expression of the viral transactivating protein RTA, a protein required for exiting the latent state [52,53]. Despite chromatin silencing, it has been speculated that low ongoing transcriptional fluctuations in the RTA activating protein might lead to spontaneous reactivation of KSHV [54]. It was recently discovered that KSHV expresses a miRNA during latency, miRK9, which appears to suppress these low fluctuations [54].

Deciphering the underlying mechanisms of viral latency also provides insights into novel antiviral strategies. The capacity for HIV-1 to generate genetic diversity that leads to drug resistance has long been a problem for antiviral therapy [55]. However, non-genetic heterogeneity in the form of viral latency is arguably an even more challenging problem [39]. If all viral infections could be driven towards replication, then all viruses would be susceptible to antiviral drugs. As an alternative strategy, if infections could be forced into a permanent latent state, then the viral genomes could persist without phenotypic consequences. Thus, effective treatment strategies for HIV-1 and herpesvirus infections can benefit from eliminating non-genetic heterogeneity.

One option being actively pursued is to purge the latent viral pool by simultaneously inhibiting enzymes that maintain repressive chromatin marks (e.g. histone deacetylase inhibitors) and activating transcription factors to induce viral activation [56–58]. A similar idea has recently been proposed for HSV therapy [49]. An inverse strategy, proposed by Weinberger *et al*., is to reduce the strength of the Tat-positive feedback loop by inhibiting Tat function, such that the proportion of infected cells entering latency is increased [10,41,59]. In both cases, the therapeutic intervention perturbs gene expression such that the contributions of Tat fluctuations to the viral cell fate decision are suppressed.

It remains to be quantified *in vivo* how the contributions of stochastic gene expression noise couple to other known modulators of latency to regulate the entry and exit from latency in HIV-1 and herpesvirus infections. However, HIV-1 and herpesviruses provide examples of how chromatin environment, coupled with the level of host and viral transcription factors in the cell, might result in a situation in which stochastic gene expression significantly influences an important cell fate decision.

Non-genetic variability in cancer

Stochastic gene expression might play a role in other diseases. For example, cancer cells can rapidly develop resistance to a variety of drugs, a phenomenon not readily explained by the mutation theory because it would require having the correct mutation within the original tumor in all instances. Therefore, alternative theories have arisen to explain such phenomena, including non-genetic heterogeneity.

It has been proposed that heterogeneity in key protein levels that are sustained over a few cell generations could confer a fitness advantage to a subpopulation of genetically identical cancer cells, especially when exposed to drugs. For example, when protein levels were tracked in individual human H1299 lung carcinoma cells treated with camptothecin, cells with higher levels of certain proteins were observed to escape the chemotherapy drug, whereas cells with low levels of these proteins die [60]. Similarly, Spencer *et al.* have shown that clonal cell populations treated

with TNF-related apoptosis-inducing ligand (TRAIL) exhibit significant temporal differences in the induction of apoptosis due to cell-to-cell variability in the levels of endogenous proteins activated by TRAIL [61]. In yet another recent example, Singh *et al.* have used high-content imaging to show that patterns of basal signaling heterogeneity correlate well with drug sensitivities [62]. Such proteomic heterogeneity could possibly function in concert with the genetic mutation theory by providing sufficient time for acquiring further genetic mutations that eventually manifest as malignant cancers.

These examples of non-genetic heterogeneity probably arise in part because of stochastic gene expression, but the underlying mechanisms are not clear. However, evidence suggests that this non-genetic heterogeneity is, in some cases, mediated by global chromatin changes in the cell. For example, it has recently been shown in a non-small cell lung cancer cell line that reversible drug tolerance in a small fraction of cells is mediated by insulin-like growth factor-1 receptor signaling, and elevated levels of a histone-modifying enzyme, which results in an alternative, transient chromatin state in a set of genes [63]. Importantly, drug sensitivity could be restored in these persister cells by inhibiting the histone-modifying enzyme. In another study, Roesch *et al.* have shown that the proliferative capacity of a heterogeneous population of melanoma cells also correlates with the levels of a histone-modifying enzyme, and the dynamic regulation of this enzyme results in individual tumor cells switching between states [64]. In these examples, the chromatin modifications produce a distinct but transient phenotypic state, in addition to underlying transitions to and from that state. Given the dramatically altered and distorted epigenetic landscape of cancer cells [65], it is plausible that transitions between different chromatin states confer phenotypic heterogeneity to a clonal population of cancer cells, providing it with the opportunity to adapt to varying selective pressures (Box 3). Therefore, further mechanistic understanding of the role of epigenetic marks and modifications are important for developing drug cocktails to target effectively all heterogeneous subpopulation of cells within a tumor, especially persister cells [63].

Conclusions

Transient phenotypes generated from non-genetic heterogeneity with functionally important consequences have been studied across multiple biological contexts, including development, immune activation, viral infections, and drug resistance in cancer [7,36,40,60,66]. Non-genetic heterogeneity provides a newly appreciated layer of selectable traits that can confer rapid adaptability to a varying environment in addition to the classical mutation-driven Darwinian evolution. Although the sources of non-genetic heterogeneity are probably variable, stochastic gene expression, driven by complex epigenetic regulation, is emerging as a key driving force behind such phenomena in eukaryotic systems. A thorough understanding of the myriad epigenetic modifications that might control the establishment of and reactivation from latency in HIV and herperviruses is crucial for identifying drugs that reverse these modifications. Similarly, identifying the sources of non-genetic heterogeneity in cancer cells might be important for increasing the susceptibility of the entire population to chemotherapy drugs.

Acknowledgments

We thank Dr. Ron Skupsky and Priya Shah for comments on the manuscript. This work was supported by NIH 2R01GM073058 to A.P.A and D.V.S.

References

1 McAdams, H.H. and Arkin, A. (1997) Stochastic mechanisms in gene expression. *Proc. Natl. Acad. Sci. U.S.A.* 94, 814–819
2 Eldar, A. and Elowitz, M.B. (2010) Functional roles for noise in genetic circuits. *Nature* 467, 167–173
3 Losick, R. and Desplan, C. (2008) Stochasticity and cell fate. *Science* 320, 65–68
4 Acar, M. *et al.* (2008) Stochastic switching as a survival strategy in fluctuating environments. *Nat. Genet.* 40, 471–475
5 Wolf, D.M. *et al.* (2005) A microbial modified prisoner's dilemma game: how frequency-dependent selection can lead to random phase variation. *J. Theor. Biol.* 234, 255–262
6 Lidstrom, M.E. and Konopka, M.C. (2010) The role of physiological heterogeneity in microbial population behavior. *Nat. Chem. Biol.* 6, 705–712
7 Balaban, N.Q. *et al.* (2004) Bacterial persistence as a phenotypic switch. *Science* 305, 1622–1625
8 Lewis, K. (2010) Persister cells. *Annu. Rev. Microbiol.* 64, 357–372
9 Gefen, O. and Balaban, N.Q. (2009) The importance of being persistent: heterogeneity of bacterial populations under antibiotic stress. *FEMS Microbiol. Rev.* 33, 704–717
10 Singh, A. and Weinberger, L.S. (2009) Stochastic gene expression as a molecular switch for viral latency. *Curr. Opin. Microbiol.* 12, 460–466
11 Stumpf, M.P. *et al.* (2002) Herpes viruses hedge their bets. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15234–15237
12 Niepel, M. *et al.* (2009) Non-genetic cell-to-cell variability and the consequences for pharmacology. *Curr. Opin. Chem. Biol.* 13, 556–561
13 Raj, A. and van Oudenaarden, A. (2008) Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* 135, 216–226
14 Thattai, M. and van Oudenaarden, A. (2001) Intrinsic noise in gene regulatory networks. *Proc. Natl. Acad. Sci. U.S.A.* 98, 8614–8619
15 Newman, J.R. *et al.* (2006) Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* 441, 840–846
16 Raj, A. *et al.* (2006) Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* 4, e309
17 Zenklusen, D. *et al.* (2008) Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat. Struct. Mol. Biol.* 15, 1263–1271
18 Skupsky, R. *et al.* (2010) HIV promoter integration site primarily modulates transcriptional burst size rather than frequency. *PLoS Comput. Biol.* 6, e1000592
19 Sigal, A. *et al.* (2006) Variability and memory of protein levels in human cells. *Nature* 444, 643–646
20 Cohen, A.A. *et al.* (2009) Protein dynamics in individual human cells: experiment and theory. *PLoS ONE* 4, e4901
21 Segal, E. and Widom, J. (2009) From DNA sequence to transcriptional behaviour: a quantitative approach. *Nat. Rev. Genet.* 10, 443–456
22 Mao, C. *et al.* (2011) Occlusion of regulatory sequences by promoter nucleosomes *in vivo*. *PLoS ONE* 6, e17521
23 Kundu, S. and Peterson, C.L. (2009) Role of chromatin states in transcriptional memory. *Biochim. Biophys. Acta* 1790, 445–455
24 Becskei, A. *et al.* (2005) Contributions of low molecule number and chromosomal positioning to stochastic gene expression. *Nat. Genet.* 37, 937–944
25 Tirosh, I. and Barkai, N. (2008) Two strategies for gene regulation by promoter nucleosomes. *Genome Res.* 18, 1084–1091
26 Choi, J.K. and Kim, Y.J. (2008) Epigenetic regulation and the variability of gene expression. *Nat. Genet.* 40, 141–147
27 Field, Y. *et al.* (2008) Distinct modes of regulation by chromatin encoded through nucleosome positioning signals. *PLoS Comput. Biol.* 4, e1000216
28 Boeger, H. *et al.* (2008) Nucleosome retention and the stochastic nature of promoter chromatin remodeling for transcription. *Cell* 133, 716–726

29 Mao, C. *et al.* (2010) Quantitative analysis of the transcription control mechanism. *Mol. Syst. Biol.* 6, 431
30 Bai, L. *et al.* (2010) Nucleosome-depleted regions in cell-cycle-regulated promoters ensure reliable gene expression in every cell cycle. *Dev. Cell* 18, 544–555
31 Kouzarides, T. (2007) Chromatin modifications and their function. *Cell* 128, 693–705
32 Henikoff, S. (2008) Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat. Rev. Genet.* 9, 15–26
33 Ahmad, K. and Henikoff, S. (2001) Modulation of a transcription factor counteracts heterochromatic gene silencing in *Drosophila*. *Cell* 104, 839–847
34 Rando, O.J. and Verstrepen, K.J. (2007) Timescales of genetic and epigenetic inheritance. *Cell* 128, 655–668
35 Muramoto, T. *et al.* (2010) Methylation of H3K4 is required for inheritance of active transcriptional states. *Curr. Biol.* 20, 397–406
36 Mariani, L. *et al.* (2010) Short-term memory in gene induction reveals the regulatory principle behind stochastic IL-4 expression. *Mol. Syst. Biol.* 6, 359
37 Han, Y. *et al.* (2007) Experimental approaches to the study of HIV-1 latency. *Nat. Rev. Microbiol.* 5, 95–106
38 Joos, B. *et al.* (2008) HIV rebounds from latently infected cells, rather than from continuing low-level replication. *Proc. Natl. Acad. Sci. U.S.A.* 105, 16725–16730
39 Richman, D.D. *et al.* (2009) The challenge of finding a cure for HIV infection. *Science* 323, 1304–1307
40 Weinberger, L.S. *et al.* (2005) Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. *Cell* 122, 169–182
41 Weinberger, L.S. *et al.* (2008) Transient-mediated fate determination in a transcriptional circuit of HIV. *Nat. Genet.* 40, 466–470
42 Blazkova, J. *et al.* (2009) CpG methylation controls reactivation of HIV from latency. *PLoS Pathog.* 5, e1000554
43 Margolis, D.M. (2010) Mechanisms of HIV latency: an emerging picture of complexity. *Curr. HIV/AIDS Rep.* 7, 37–43
44 Singh, A. *et al.* (2010) Transcriptional bursting from the HIV-1 promoter is a significant source of stochastic noise in HIV-1 gene expression. *Biophys. J.* 98, L32–34
45 Jordan, A. *et al.* (2001) The site of HIV-1 integration in the human genome determines basal transcriptional activity and response to Tat transactivation. *EMBO J.* 20, 1726–1738
46 Sadowski, I. *et al.* (2008) Factors controlling chromatin organization and nucleosome positioning for establishment and maintenance of HIV latency. *Curr. HIV Res.* 6, 286–295
47 Burnett, J.C. *et al.* (2009) Control of stochastic gene expression by host factors at the HIV promoter. *PLoS Pathog.* 5, e1000260
48 Speck, S.H. and Ganem, D. (2010) Viral latency and its regulation: lessons from the gamma-herpesviruses. *Cell Host Microbe* 8, 100–115
49 Knipe, D.M. and Cliffe, A. (2008) Chromatin control of herpes simplex virus lytic and latent infection. *Nat. Rev. Microbiol.* 6, 211–221
50 Thompson, R.L. *et al.* (2009) *De novo* synthesis of VP16 coordinates the exit from HSV latency *in vivo*. *PLoS Pathog.* 5, e1000352
51 Grzimek, N.K. *et al.* (2001) Random, asynchronous, and asymmetric transcriptional activity of enhancer-flanking major immediate-early genes ie1/3 and ie2 during murine cytomegalovirus latency in the lungs. *J. Virol.* 75, 2692–2705
52 Gunther, T. and Grundhoff, A. (2010) The epigenetic landscape of latent Kaposi sarcoma-associated herpesvirus genomes. *PLoS Pathog.* 6, e1000935
53 Toth, Z. *et al.* (2010) Epigenetic analysis of KSHV latent and lytic genomes. *PLoS Pathog.* 6, e1001013
54 Bellare, P. and Ganem, D. (2009) Regulation of KSHV lytic switch protein expression by a virus-encoded microRNA: an evolutionary adaptation that fine-tunes lytic reactivation. *Cell Host Microbe* 6, 570–575
55 Johnson, V.A. *et al.* (2010) Update of the drug resistance mutations in HIV-1: December 2010. *Top HIV Med.* 18, 156–163
56 Burnett, J.C. *et al.* (2010) Combinatorial latency reactivation for HIV-1 subtypes and variants. *J. Virol.* 84, 5958–5974
57 Kauder, S.E. *et al.* (2009) Epigenetic regulation of HIV-1 latency by cytosine methylation. *PLoS Pathog.* 5, e1000495
58 Lehrman, G. *et al.* (2005) Depletion of latent HIV-1 infection *in vivo*: a proof-of-concept study. *Lancet* 366, 549–555
59 Weinberger, L.S. *et al.* (2003) Theoretical design of a gene therapy to prevent AIDS but not human immunodeficiency virus type 1 infection. *J. Virol.* 77, 10028–10036
60 Cohen, A.A. *et al.* (2008) Dynamic proteomics of individual cancer cells in response to a drug. *Science* 322, 1511–1516
61 Spencer, S.L. *et al.* (2009) Non-genetic origins of cell-to-cell variability in TRAIL-induced apoptosis. *Nature* 459, 428–432
62 Singh, D.K. *et al.* (2010) Patterns of basal signaling heterogeneity can distinguish cellular populations with different drug sensitivities. *Mol. Syst. Biol.* 6, 369
63 Sharma, S.V. *et al.* (2010) A chromatin-mediated reversible drug-tolerant state in cancer cell subpopulations. *Cell* 141, 69–80
64 Roesch, A. *et al.* (2010) A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. *Cell* 141, 583–594
65 Portela, A. and Esteller, M. (2010) Epigenetic modifications and human disease. *Nat. Biotechnol.* 28, 1057–1068
66 Chang, H.H. *et al.* (2008) Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature* 453, 544–547
67 Elowitz, M.B. *et al.* (2002) Stochastic gene expression in a single cell. *Science* 297, 1183–1186
68 Raser, J.M. and O'Shea, E.K. (2004) Control of stochasticity in eukaryotic gene expression. *Science* 304, 1811–1814
69 Huh, D. and Paulsson, J. (2011) Non-genetic heterogeneity from stochastic partitioning at cell division. *Nat. Genet.* 43, 95–100
70 Arkin, A. *et al.* (1998) Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected *Escherichia coli* cells. *Genetics* 149, 1633–1648
71 Hooshangi, S. *et al.* (2005) Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc. Natl. Acad. Sci. U.S.A.* 102, 3581–3586
72 Samoilov, M.S. *et al.* (2006) From fluctuations to phenotypes: the physiology of noise. *Sci. STKE* 2006, re17
73 Ajo-Franklin, C.M. *et al.* (2007) Rational design of memory in eukaryotic cells. *Genes Dev.* 21, 2271–2276
74 St-Pierre, F. and Endy, D. (2008) Determination of cell fate selection during phage lambda infection. *Proc. Natl. Acad. Sci. U.S.A.* 105, 20705–20710
75 Snijder, B. *et al.* (2009) Population context determines cell-to-cell variability in endocytosis and virus infection. *Nature* 461, 520–523
76 Snijder, B. and Pelkmans, L. (2011) Origins of regulated cell-to-cell variability. *Nat. Rev. Mol. Cell Biol.* 12, 119–125
77 Chubb, J.R. and Liverpool, T.B. (2010) Bursts and pulses: insights from single cell studies into transcriptional mechanisms. *Curr. Opin. Genet. Dev.* 20, 478–484
78 Blake, W.J. *et al.* (2003) Noise in eukaryotic gene expression. *Nature* 422, 633–637
79 Huang, S. and Ingber, D.E. (2006) A non-genetic basis for cancer progression and metastasis: self-organizing attractors in cell regulatory networks. *Breast Dis.* 26, 27–54
80 Brock, A. *et al.* (2009) Non-genetic heterogeneity—a mutation-independent driving force for the somatic evolution of tumours. *Nat. Rev. Genet.* 10, 336–342