CHAPTER TWENTY-SIX

Lentiviral Vectors to Study Stochastic Noise in Gene Expression

Kate Franz, Abhyudai Singh, *and* Leor S. Weinberger

Contents

Abstract

Lentiviral vectors are vehicles for gene delivery that were originally derived from the human immunodeficiency virus type-1 (HIV-1) lentivirus. These vectors are defective for replication, and thus considered relatively safe, but are capable of stably integrating into the genomic DNA of a broad range of dividing and nondividing mammalian cell types. The ability to stably integrate at semi-random genomic positions make lentiviral vectors a unique and ideal tool for studying stochastic variation in gene expression. Here, we describe the experimental and mathematical methods for using lentiviral vectors to study stochastic noise in gene expression.

Abbreviations

cPPT	central polypurine tract
CV	coefficient of variation

Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California, USA

Methods in Enzymology, Volume 497
ISSN 0076-6879, DOI: 10.1016/B978-0-12-385075-1.00026-3

FACS	fluorescence-activated cell-sorting
FSC	forward scatter
HIV-1	human immunodeficiency virus type-1
LTR	long terminal repeat
RRE	Rev responsive element
SIN vectors	self-inactivating vectors
SSC	side scatter
IU	infectious units

1. Introduction

Fluctuations in the levels of gene products are an inevitable consequence of the inherent stochastic nature of biochemical processes that constitute gene expression (Blake *et al.*, 2006; Kaern *et al.*, 2005; Raj and van Oudenaarden, 2008). These fluctuations are referred to as stochastic 'noise' and in a population of cells the noise is often characterized as either due to 'intrinsic' sources or 'extrinsic' sources (Elowitz *et al.*, 2002; Swain *et al.*, 2002). In isogenic populations (i.e., clonal populations where all cells are derived from a single parent), intrinsic noise manifests as cell-to-cell variation that is uncorrelated, while extrinsic noise manifests as variation that is correlated between cells. The origin of the noise in gene products is biochemical: intracellular processes are driven by reactant molecules randomly diffusing and colliding within the cell and are thus inherently stochastic. Specifically, noise in gene expression can arise from the random timing in individual reactions associated with promoter remodeling, transcription, and translation (Blake *et al.*, 2003; Elowitz *et al.*, 2002; Kaern *et al.*, 2005; Swain *et al.*, 2002). Moreover, intercellular differences in the *amount* of cellular components (e.g., RNA polymerase, transcription factors, and ribosomes) also cause variations in expression levels. Measurements in live, single cells have shown that gene expression noise can lead to large statistical fluctuations in protein and mRNA levels in both prokaryotes and eukaryotes (Bar-Even *et al.*, 2006; Golding *et al.*, 2005; Newman *et al.*, 2006; Raj *et al.*, 2006). These fluctuations (i.e., noise) can have significant effects on biological function and phenotype and noise is now recognized to exert significant influence on probabilistic fate decisions in bacteria (Eldar *et al.*, 2009; Suel *et al.*, 2006), viruses (Singh and Weinberger, 2009), and stem cells (Hanna *et al.*, 2009).

A common method to analyze gene-expression noise is to measure the variation in protein levels across isogenic cells as a function of mean protein levels (Bar-Even *et al.*, 2006; Blake *et al.*, 2003). This method is often referred to as CV^2 versus mean analysis, where CV^2 is a dimensionless statistical measure of intercellular variability in protein levels. More

specifically, CV^2 stands for the coefficient of variation (CV) squared and is defined as $CV^2 = \sigma^2/\langle protein \rangle^2$, where σ^2 is the variance in protein abundance and $\langle protein \rangle$ is the average number of protein molecules per cell (Paulsson, 2004).

To study how CV^2 varies as a function of $\langle protein \rangle$, it is essential that one be able to alter mean expression levels by changing the transcriptional rate of the promoter of interest. This modulation can be challenging for many promoters, especially when transcriptional activity cannot be modulated using small-molecule compounds. In such cases, transcriptional efficiency is typically altered through mutations in the promoter sequence (Ozbudak *et al.*, 2002). Studying expression noise using lentiviral vectors provides a unique advantage as it allows one to exploit their known ability to integrate semi-randomly into sites across the human genome (Schroder *et al.*, 2002). Differences in local chromatin microenvironment at each integration site generate differences in promoter strength (Jordan *et al.*, 2001) and this difference provides a natural method to study CV^2 as a function of mean protein levels.

Our laboratory has extensively studied gene-expression noise using lentiviral vectors. We have used lentiviral vectors to study: (i) the role of stochastic noise in viral gene expression (Weinberger and Shenk, 2007; Weinberger *et al.*, 2005), (ii) the influence of stochastic noise upon probabilistic decision making in HIV (Weinberger *et al.*, 2008), and (iii) the molecular source of noise in viral gene expression (Singh *et al.*, 2010). Here, we describe the experimental methods for establishing isogenic populations using lentiviral vectors as well as the quantitative methods used to analyze stochastic noise in these isogenic populations.

The classical method for studying gene-expression noise is to work in an isogenic background—where each cell is grown from a single parent cell—and analyze expression from a specific genetic locus within the isogenic population. Here, we describe how lentiviral-vector technology provides a convenient method for constructing isogenic populations of cells and for minimizing external variation sources in the analysis of gene-expression noise. While usage of isogenic backgrounds does not eliminate all nonexpression sources of cell-to-cell variation (e.g., cell size, cell-cycle state, other extrinsic factors), below we describe how these other sources of variation can be dealt with.

2. The Lentiviral-Vector Approach

The need to deliver foreign genetic information to living cells (i.e., genetically transduce cells with foreign DNA or RNA) has long presented a challenge for molecular and cell biologists. Various methods of gene delivery have been developed including bioballistic 'gene-gun' approaches,

electroporation, chemical methods, and viral methods. Lentiviral vectors are one of the most recent vehicles for gene delivery having been derived originally from the human immunodeficiency virus type-1 (HIV-1) lentivirus (Naldini *et al.*, 1996) and subsequently from related lentiviruses (Mitrophanous *et al.*, 1999; Olsen, 1998; Poeschla *et al.*, 1998). These vectors are defective for replication, and thus considered relatively safe, but are capable of stably transducing a broad range of dividing and nondividing cell types. This provides a significant advantage over other methods of gene transfer, such as gammaretroviral-based vectors (e.g., Murine Leukemia Viruses, MLVs), which can only transduce dividing cells. Lentiviral vectors also carry a number of other recognized benefits over other gene-delivery methods: (i) lentiviral vectors are less susceptible to position-effect variegation as compared to retroviral vectors which are rapidly silenced (Challita and Kohn, 1994); (ii) unlike stable transfections lentiviral vectors do not concatenate or continually change in copy number over multiple cell passages; (iii) lentiviral-integrated constructs are not progressively diluted out with each cell division (as occurs with transient transfections), which makes lentiviral-transduced cell lines easy to maintain. Finally, lentiviral vectors integrate throughout the genome and preferentially integrate, with ~69% preference, in active transcriptional units (Schroder *et al.*, 2002). This semi-random integration allows gene-expression characteristics to be analyzed in various chromatin contexts throughout the genome.

Lentiviral-vector systems have been developed from many nonprimate lentiviruses such as equine infectious anemia virus (EIAV) and feline immunodeficiency virus (FIV). However, our discussion will focus on the generation of HIV-1 derived lentiviral vectors as they are the most widely used vector system.

The HIV-1 genome encodes 15 proteins necessary for infection and pathogenesis of the virus. Additionally, the genome contains *cis*-acting elements that are required for genomic RNA packaging (e.g., the Ψ signal), reverse transcription (e.g., the central polypurine tract or cPPT), and viral gene expression (i.e., the 5′ and 3′ LTRs and the Rev responsive element or RRE) (Fig. 26.1). *Trans*-acting elements required for infection and assembly of viral particles are expressed from constructs lacking the above *cis*-acting sequences. The *cis*-acting sequences are found only on the transfer vector containing the promoter/reporter gene cassette. When both cis-acting and trans-acting constructs are transfected into a cell, infectious viral particles will assemble. However, the particles can only encapsidate the transfer vector RNA which encodes the *cis*-acting elements and no *trans*-acting elements thus, thus limiting the lentiviral vectors to a single round of infection.

Either a three-plasmid or four-plasmid approach can be used to generate lentiviral vectors depending on whether the 2nd or 3rd generation lentiviral system is used. In the 2nd-generation system, one plasmid, the packaging construct, encodes all *trans*-acting elements necessary for assembly of the

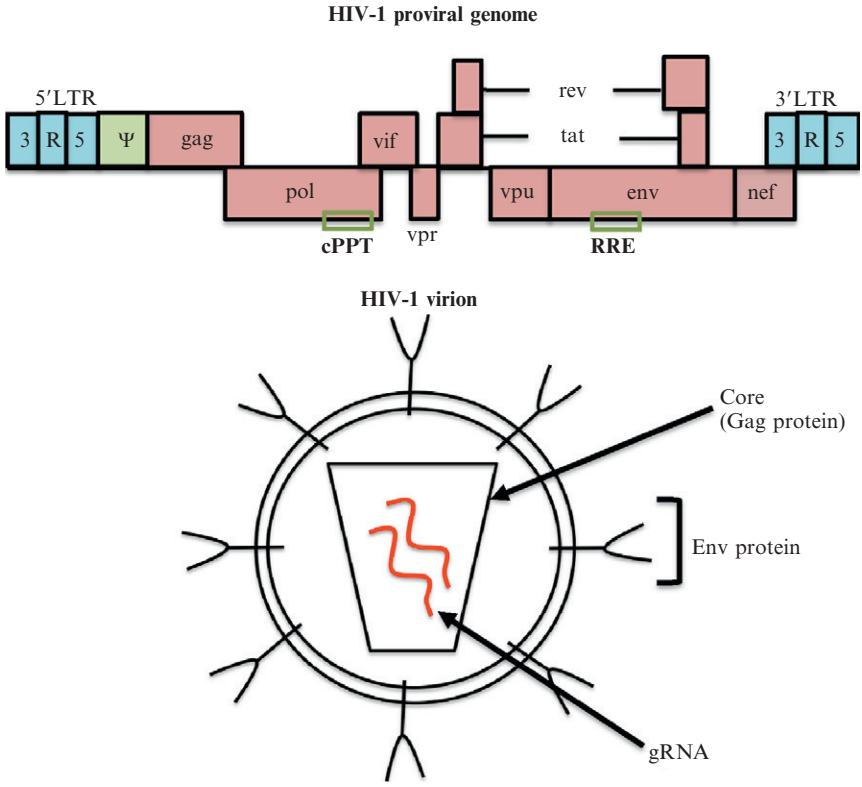

Figure 26.1 HIV-1 genomic organization and virion structure. HIV-1 genome contains nine open reading frames (in red) that code for 15 proteins. *cis*-acting elements important for reverse transcription, gene expression, and packaging are in green. Note that the integrated HIV-1 genome has a full 5′LTR (in blue).

lentiviral core. Accessory genes that act only in the pathogenesis of infection (*nef, vif, vpu*, and *vpr* genes in HIV-1) are eliminated, as well as the transactivator of transcription (*tat*) and *env* gene (Dull *et al.*, 1998; Kim *et al.*, 1998). As explained above, none of the *cis*-acting elements required for encapsidation or transfer into the target cell are present on this construct. This reduces the risk of recombinant, replication-competent viruses arising. The 3rd-generation packaging system encodes the same *trans*-acting elements, but on two separate plasmid constructs to further minimize the risk of recombinant viruses (Fig. 26.2).

The envelope construct expresses a heterologous envelope protein, which serves to expand the cell-targeting tropism of the viral vector.

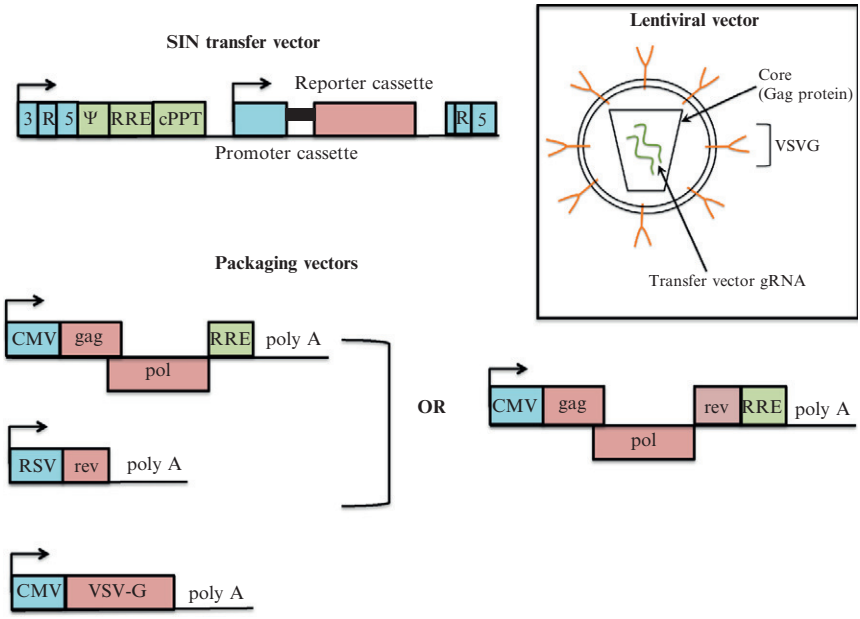

Figure 26.2 Schematic of the 2nd and 3rd generation lentiviral plasmid system. Promoters are in blue, open reading frames are in red, and *cis*-acting elements are in green. The U3 region of the 3′LTR of the transfer vector is mostly deleted. Upon reverse transcription, this deletion will be carried into the 5′LTR of the integrated construct and inactivate it. After transfection of the system into HEK 293FT cells, lentiviral particles are generated that contain transfer vector as the gRNA and VSVG protein for envelope protein (inset).

Exchanging the envelope protein for one of another species is termed 'pseudo-typing.' Most lentiviral vectors utilize the vesicular stomatitis virus glycoprotein (VSVG) due to its broad cell-targeting tropism and its ability to withstand ultracentrifugation and freeze-thaw cycles with only minimal loss of infectivity (Akkina *et al.*, 1996; Reiser *et al.*, 1996). Many other envelope proteins may be used (Cronin *et al.*, 2005), and it is vital to determine whether a given choice of envelope can effectively target the cell type being studied.

All relevant *cis*-acting elements for infection and are located on the transfer vector (a.k.a. the lentiviral "backbone" vector). The transfer vector includes the LTR promoter, Ψ signal, RRE, and cPPT followed by the promoter and reporter cassettes to be studied. The Ψ signal enables the dimerization and encapsidation of only the transfer vector and the cPPT functions in reverse transcription. Since nuclear export is blocked for unspliced mRNAs in eukaryotic cells, inclusion of the RRE is important

for efficient expression of the transcribed construct. Rev—coded for within the packaging plasmid—binds the RRE and utilizes the CRM1 nuclear export pathway to transport unspliced RNAs into the cytoplasm (Fischer *et al.*, 1995; Neville *et al.*, 1997).

One advance in transfer-vector design that greatly increased the biosafety of lentiviral vectors has been the creation of *s*elf-*in*activating vectors, or SIN vectors. SIN vectors carry a near complete deletion of the promoter-enhancer sequences in the U3 region in the 3′LTR. Since reverse transcriptase uses the 3′LTR as the template for the proviral copy of the 5′LTR, upon reverse transcription this deletion is transferred to the 5′LTR, rendering the 5′LTR transcriptionally inactive. This deletion prevents future mobilization of the construct, decreases risk of endogenous oncogene activation, and eliminates any interference in gene expression due to transcription from the lentiviral LTR (Miyoshi *et al.*, 1998).

Design of a transfer vector depends significantly on the aims of the study. Often the transferred construct consists of the promoter of interest driving the expression of a fluorescent reporter gene. Fluorescent reporters can be fused (either transcriptionally or translationally) to another gene product. Flanking the promoter and/or reporter gene with restriction enzymes allows for quick swapping of elements within the transfer vector and generation of new vectors that vary with respect to their promoter and reporter gene. Design of transfer vectors can be simplified by using lentivirus expression vectors like those available from Invitrogen (Carlsbad, CA), which use Gateway and TOPO cloning to insert the construct of interest into a transfer-vector backbone. Transfer vectors can also be generated using basic cloning techniques to ligate your construct into a previously described transfer vector.

For example, we have studied gene-expression noise in HIV and its consequences, by utilizing lentiviral vectors expressing fluorescent reporter genes (e.g., green fluorescent protein, GFP) from the HIV-1 LTR promoter. These lentiviral vectors allowed us transduce Jurkat T cells (ATCC # TIB-152), isolate single cells using fluorescence-activated cell-sorting (FACS), and expand these cells into isogenic populations of cells for flow cytometry analysis to obtain data on the CV and mean fluorescence. Below, we describe the methods used to construct these lentiviral vectors and analyze these isogenic populations for stochastic noise (Singh *et al.*, 2010).

3. Production of Lentiviral Vectors and Transduced Cell Lines

All steps should be performed using BSL 2+ containment procedures. Overview of protocol found in Fig. 26.3.

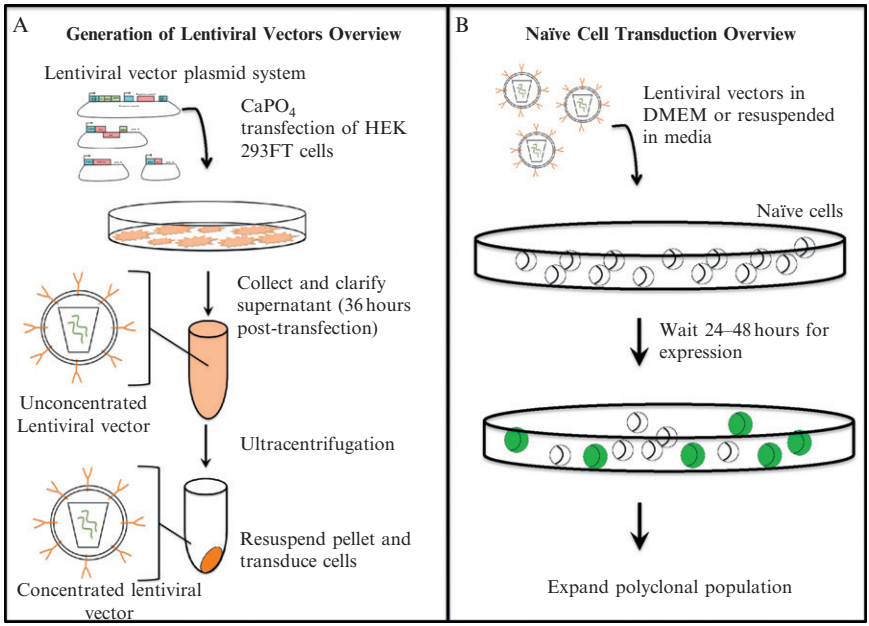

Figure 26.3 Overview of the protocol for production of lentiviral particles (A) and transduction of naïve cells (B).

Reagents and equipment

Human Embryonic Kidney (HEK) 293FT cells (Invitrogen cat. no. R700-07)
Dulbecco's modified eagle medium (DMEM) (Cellgro cat. no. 10-013-CV) supplemented with 10% fetal bovine serum (FBS) (Cellgro 35-016-CV) and 1% penicillin/streptomycin (PS) (Cellgro cat. no. 30-009-CI)
1× phosphate buffered saline (PBS) (Cellgro cat. no. 21-030-CV)
0.05% trypsin (Cellgro cat. no. 25-051-CI)
Plasmids: lentiviral transfer vector, lentiviral packaging vectors (Invitrogen cat. no. K4975-00)
Molecular biology grade water (Cellgro cat. no. 46-000-CM)
2× HEPES-buffered saline (HeBS) (SIGMA cat. no. 51558), filter-sterilized
2.5 *M* $CaCl_2$, autoclaved
10-cm tissue culture dishes
15-mL conical tube
0.45-μm syringe filters
30-mL syringe
Beckman ultracentrifuge tube (Beckman cat. no. 344059)
Beckman SW41 rotor

Beckman ultracentrifuge
Fluorescence microscope
Humidified tissue-culture incubator at 5% CO_2

1. HEK 293FT cell culture
   - **1.1** Seed low passage-number (<20 passages) HEK 293FT cells on 10-cm dishes in a total of 10 mL DMEM + 10% FBS + 1% PS.
     - **1.1.1** If seeding from a 10-cm dish of confluent cells, aspirate the media and wash the cells with 4 mL of PBS. Add 0.5 mL of 0.05% trypsin and incubate dish at 37 °C in 5% CO_2 for 5 min. Resuspend cells in 4.5 mL of DMEM + 10% FBS + 1% penicillin and streptomycin. Aliquot resuspended cells in 10-cm dishes. Incubate HEK 293FT cells at 37 °C in 5% CO_2. HEK 293FT cells express the SV40 T antigen which allows plasmids containing the SV40 origin of replication to be retained and replicated during cell division and enhances expression of lentiviral plasmids.
   - **1.2** Observe the seeded dishes. Cells should be in exponential growth. When cells are 60–80% confluent, they are ready for transfection.

   Important note: Never use 293FT cells that are growing slowly. Switch to lower passage-cells if growth is impaired.
2. Transfection of plasmids by $CaPO_4$ precipitation
   - **2.1** Prepare one aliquot of DNA/$CaCl_2$ solution per dish of 293FT cells to transfect. In a 15-mL conical tube, add 3.5 μg of pVSVG, 1.5 μg pRev, 5 μg pMDL, and 10 μg transfer vector to 120 μL 2.5 *M* $CaCl_2$. Bring solution to 1 mL with ultrapure water. Mix.
   - **2.2** Perform the following one aliquot at a time. Add 1 mL 2× HeBS. Pipette entire volume four times. Let solution incubate 1.5 min. Gently add HeBS/DNA/$CaCl_2$ solution drop-wise to HEK 293FT cells. Swirl the dish very gently to evenly distribute the transfection mixture being careful not to detach cells.

   Important note: $CaPO_4$ transfection efficiency is pH-dependent. Titrate 2× HeBS to pH 6.95 before using.

   (alternatively, we use Fugene™ transfection with the 3-plasmid packaging system in the following ratios: 5.6 μg Δ8.9, 2.8 μg of VSVG, and 2.2 μg of transfer vector in 30 μL of Fugene and 300 μL of serum-free DMEM)
   - **2.3** Incubate the cells overnight at 37 °C, 5% CO_2 in a humidified environment.
   - **2.4** Visually inspect Observe 12-18 hours later; transfection efficiency can be assessed visually through expression of fluorescent reporter proteins. Transfection efficiency should be >90%. Aspirate the

media and carefully add 10 mL of fresh DMEM + 10% FBS + 1% PS. Incubate 24-28 hours at 37 °C, 5% CO_2 in a humidified environment.

3. Harvest the supernatant
   - **3.1** Collect the supernatant from the dish. *Use BSL 2+ containment procedures* as all supernatants and cell culture contain infectious vector. If a second harvest of lentiviral vector is desired, carefully add 10 mL of fresh DMEM + 10% FBS + 1% PS to the dish. Incubate cells 24-28 hours at 37 °C, 5% CO_2. Spray exteriors of all tubes containing lentivirus vector with 70% ethanol before removing from hood. Store the collected supernatant in a secondary container at 4 °C.
   - **3.2** The next morning, collect the second harvest of supernatant from the cells. Pool supernatants from both harvests and centrifuge at 500×*g* for 5 min to pellet cell debris. Clarify the supernatant by filtering through a 0.45-μm filter attached to a 30-mL syringe. Clarified supernatants can be stored up to a year at −80 °C. At this point, the titer of the supernatant can be estimated to be $\sim 10^6$ IU/mL (and up to twofold lower for vectors with transfer constructs over 4kb). If desired, store 1 mL of this filtrate at −80 °C until ready to titer. Supernatants can now be used for transduction, but ultracentrifugation is recommended to concentrate the viral prep and increase transduction efficiency.
4. Ultracentrifugation of clarified lentiviral prep

   For ultracentrifugation, we use a Beckman SW41 swinging-bucket rotor and a Beckman ultracentrifuge, however equivalent equipment will work as well.
   - **4.1** Pipette filtrate into Beckman ultracentrifuge tubes. It is critical that the tubes opposite from one another in the rotor are precisely balanced. Place tubes into the appropriate rotor buckets, secure the lids on the buckets and weigh opposing tubes to ensure they are balanced.
   - **4.2** Centrifuge the viral prep at 20,000×*g* for 90 min at 4 °C. Aspirate the supernatant, tilting the tube slightly horizontally to avoid aspirating the viral pellet. The pellet will most likely not be visible, but may appear as a glassy spot on the bottom of the tube.
   - **4.3** Resuspend the pellet in the amount of PBS or media that allows you to transduce your cells at your target multiplicity of infection (MOI) and at MOIs 1–2 logs above and below target. A good target MOI for noise analysis is MOI = 0.1, which limits double integrations to 10–15% of your transduced population. We usually resuspend the pellets in 500 μL of media. Viral prep can be used directly from this step or aliquoted and stored at −80 °C for up to a year.
5. Transduction of lentiviral vector into cell line of interest
   - **5.1** Maintain a low passage-number cell line of interest. Use only freshly passaged cells for transduction.

5.1.1 *Suspension cells*: Count cells and resuspend 500,000 cells for each transduction in 1–2 mL of media (depending on your cell line's preferred growth conditions) in a 12-well plate.

5.1.2 *Adherent cells*: Count cells and resuspend 500,000 cells for each transduction in 1.5 mL of media in a 6-well plate. Plating confluency may need to be adjusted to accommodate cell size, growth rate, and optimal culture confluency.

5.2 Add the appropriate amount of viral prep suspended in PBS or media to each well. We resuspend the viral pellet in 500 μL of media and add 100 μL, 10 μL, or 1 μL of a 1:10 dillution of the viral stock to 500,000 cells to obtain target MOIs of 10, 1, 0.1, and 0.01, respectively. Incubate the cells normally. If viral prep is suspended in the same culture media of the cells or PBS, a media change after incubation is not necessary. If viral prep is in a different media, let the cells and virus incubate for 1–4 h. Centrifuge the cells at $300\times g$ for 10 min and resuspend in the appropriate media.

5.3 Visually inspect the cells under the microscope the next day. Fluorescent reporters can usually be seen after 24 h. After 48 h, quantify transduction efficiency using flow cytometry. If necessary, small-molecule transcriptional activators can be added to your cells the day before flow cytometry to increase the expression of the fluorescent protein and activate silent promoters. Any activation treatment should also be performed on naive cells to control for any changes in autofluorescence induced by the small-molecule activator treatment. Cells can be analyzed for expression of the transgene using flow cytometry.

5.4 Expand cultures using standard tissue culture techniques and freeze down cells for storage. Cells can usually be frozen in a cryovial at concentration of 10^6 cells/mL in media supplemented with 10% DMSO and stored at −80 °C.

3.1. Sorting isoclonal populations

Isoclonal (isogenic) populations are populations that have been derived from a single parent cell. Single cells can be isolated through the use of fluorescence-activated cell-sorting (FACS). Polyclonal populations used for sorting should be between 5% and 10% positive for fluorescent reporter expression; this ensures that the frequency of double integration is low, as calculated using a Poisson distribution.

Reagents and equipment

Live-cell FACS sorting capabilities
Control (naïve) cells
Lentiviral vector-transduced population, expanded

PBS supplemented with 10% FBS
96-well tissue culture plates
24-well tissue culture plates
Culture media supplemented with 10–20% FBS and 1% penicillin/streptomycin
5 mL FACS polypropylene tubes with cell-strainer caps (sterile)
Mylar plate sealers (Thermo Scientific cat. no. 5701ROC)
Humidified tissue-culture incubator at 5% CO_2

1. Preparing polyclonal lentiviral vector-transduced cells for sorting
   - **1.1** Using standard tissue-culture techniques, passage polyclonal transduced cell population and a naïve untransduced population of the same cell type. Maintain growth in the exponential phase. If necessary, activators can be added to the transduced population to maximize gene expression prior to sorting so that an increased number of transduced cells can be sorted. All activator treatments should also be performed on the naïve population.
   - **1.2** Two hours before sorting, count the cells in each population and centrifuge 1–2 × 10^6 cells of each population in 15-mL conical tubes at 300×*g* for 5 min. Aspirate the media and resuspend the cells in 2 mL of PBS supplemented with 10% FBS.
   - **1.3** Pipette the population through a sterile cell strainer (to remove debris), into a sterile FACS tube. Place cells on ice.
2. Prepare 96-well plates for collection of sorted cells
   - **2.1** Each well of the plate will collect one sorted cell, so prepare as many wells as cells you intend to collect. It is a good assumption that only 5–15% of the single-cell sorts will survive. For this reason, we generally collect between 200 and 300 single cells of each population to be sorted to recover 20–30 isoclonal populations.
   - **2.2** In each well of the plate add 150 μL of culture media supplemented with 10–20% FBS and 1% penicillin and streptomycin. Some cells types are more sensitive to the sorting procedure and may require higher percentages of FBS in the collection media.
3. Gate cells for sorting and collect cells
   - **3.1** The FACS operator will need the naïve cells in order to distinguish the autofluorescence background of the cells from the fluorescent reporter signal. Gate the live cells according to the forward and side scattering measurements (see Fig. 26.4). From this gate, observe the distribution of fluorescence. We gate three regions of fluorescence for sorting, a DIM region encompassing the lowest third of fluorescence signal, a MID region which consists of the middle third of the fluorescence signal, and a BRIGHT region which consists of the highest signal expressed (see Fig. 26.4).

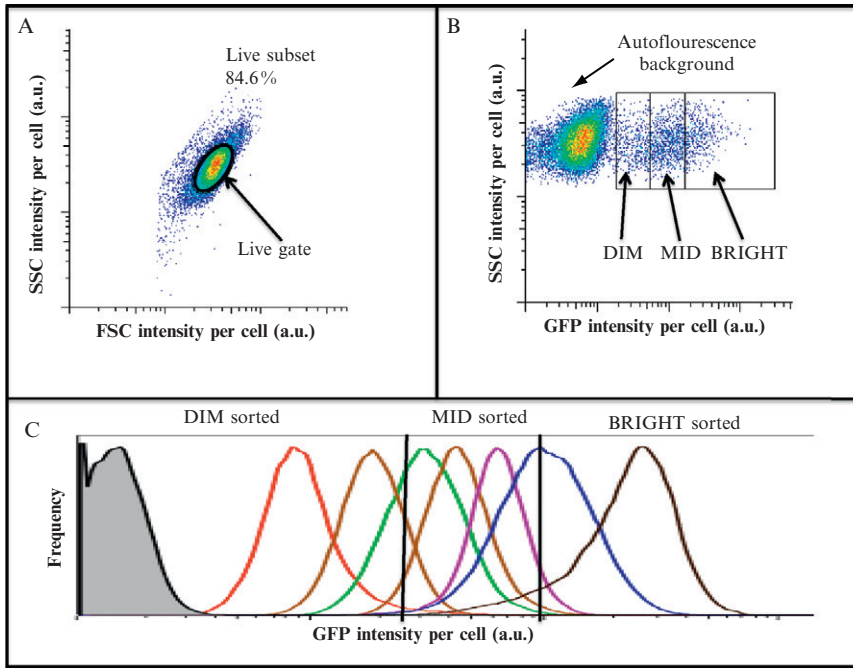

Figure 26.4 (A) Example of live gating off forward-scatter versus side-scatter plots of Jurkat T-cells. (B) Plot shows side scatter versus GFP fluorescence. Gates are drawn to show example sorting gates of DIM, MID, and BRIGHT populations. (C) GFP fluorescence histograms of sorted isoclonal populations.

3.2 Collect 200–300 cells from each gate. After sorting, verify under a microscope that each well has only one cell. Incubate cells under normal conditions. To prevent evaporation of media from the plate, remove the lid and cover the wells with a gas-permeable mylar plate sealer.

4. Expand isoclonal populations.

4.1 Depending on cell type, isoclonal populations may take 1–2 weeks to become confluent in the 96-well plate. During this time, check that the media in the plate has not evaporated below 1/3 the well height and replace media as needed. Keep 100–150 μL of media in each well. Every 3–4 days, check the confluency of the cells in each well under a microscope. Do not check cells more often than every 2–3 days since incandescent light can be toxic to cells when they are at low concentration. Alternatively, if the culture media has a pH indicator, wells can be screened visually by looking for a change in media color in the well indicating that the population is reaching a high confluency.

4.2 Once an isoclonal population has reached 80–90% confluency in the well, transfer the population to 0.5 mL of media in a 24-well

tissue culture plate. Continue expanding the population to maintain cells at a healthy confluency using standard tissue-culture techniques. Store aliquots of the expanded populations at −80 °C. Isoclonal populations can now be analyzed by flow cytometry (Fig. 26.4C).

4. Procedure for Constructing a CV^2 Versus Mean Plot

To perform a CV^2 versus mean analysis, the first step is to create a library of isogenic populations each carrying a single integrated copy of the promoter of interest driving a reporter gene, as described previously. These clonal populations will exhibit considerable differences in mean expression levels as each clone corresponds to a different integration site of the lentiviral vector (Fig. 26.4C). Reporter expression in each clone is measured by flow cytometry with data from at least 100,000 single-cells collected per clone. To quantify the cell-to-cell reporter variation within each clonal population, flow cytometry data is analyzed using standard software packages like FlowJo™ (Treestar Inc., Ashland, Oregon). Before quantifying this variation it is important that one minimizes differences in protein levels due to heterogeneity in cell size, cell shape, and cell-cycle state (i.e., extrinsic noise). A standard approach to reduce extrinsic noise is to draw a small gate around the forward scatter (FSC) and side scatter (SSC) medians that contains at least 30,000 cells (Newman *et al.*, 2006). CV and mean protein levels for different clones is computed from this gated population using the statistics toolbox in FlowJo. For computing CV^2, it should be kept in mind that many software packages, like FlowJo, report CV as a percentage. Next, mean protein levels, $\langle \text{protein} \rangle$, which are quantified in terms of fluorescence intensities, are converted into absolute protein molecular counts. For example, EGFP Calibration Beads™ (BD Biosciences, Clontech, San Jose, CA) can be used to convert GFP fluorescence intensities into GFP molecular equivalents of solubilized flourophores (MESF), a measure of GFP molecular abundance. Once CV^2 and $\langle \text{protein} \rangle$ have been quantified for all clones, the final step is to look at correlations between them by making a scatter plot with CV^2 on the y-axis and $\langle \text{protein} \rangle$ on the x-axis.

5. Inferring Promoter Regulatory Architecture from CV^2 Versus Mean Analysis

To understand how the relationship between CV^2 and $\langle \text{protein} \rangle$ can inform upon promoter architecture, we review mathematical predictions from two different gene-expression models below.

5.1. Constitutive promoter architecture

Constitutive promoter models (Fig. 26.5A), where mRNAs are continuously created from the promoter at exponentially distributed time intervals generate the prediction:

$$\mathrm{CV}^2 = \frac{C}{\langle \text{protein} \rangle} \tag{26.1}$$

where C is a proportionality constant (Paulsson, 2004). Equation (26.1) shows that for constitutive gene expression, increasing the mean protein count will decrease CV^2 such that the product $\mathrm{CV}^2 \times \langle \text{protein} \rangle$ remains

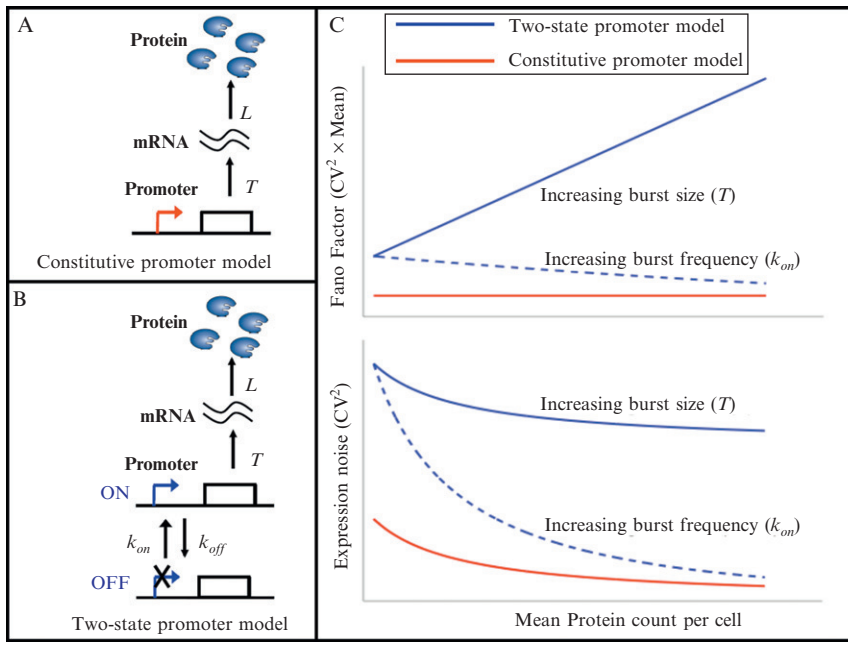

Figure 26.5 (A) Constitutive promoter model where mRNAs are produced continuously from the promoter one at a time. (B) Two-state promoter model where mRNAs are only produced when the promoter transitions to an ON state before returning to an OFF state. (C) Predictions for the scaling of gene-expression noise (CV^2) and Fano factor ($\mathrm{CV}^2 \times$ Mean) as a function of mean protein count for different promoter architectures. Two-state promoter architecture is predicted to generate elevated levels of expression noise compared to a constitutive promoter architecture. In a constitutive promoter, CV^2 decreases with mean such that the Fano factor remains fixed (red line). In a two-state promoter model, CV^2 also decreases with mean but the Fano factor can either increase or decrease with mean depending on the mode of transcriptional activation (soild and dashed blue lines).

unchanged. A useful way to detect this inverse correlation between CV^2 and mean protein levels is by constructing a secondary scatter plot of the Fano factor on the *y*-axis versus the ⟨protein⟩ on the *x*-axis, where Fano factor is defined as $CV^2 \times \langle \text{protein} \rangle$. The main rationale behind this plot is that since Fano factor is predicted to be a constant for constitutive promoter architecture (Eq. (26.1)), it will appear to be completely uncorrelated with mean protein levels on this plot (Fig. 26.5C). If Fano factor is indeed uncorrelated with mean protein levels, then for constitutive gene expression, the average Fano factor across different clonal populations should be equal to the proportionality constant C given by:

$$C = \frac{L}{\mathrm{d}m + \mathrm{d}p} = \frac{\langle \text{protein} \rangle}{\langle \text{mRNA} \rangle} \frac{\mathrm{d}p}{\mathrm{d}m + \mathrm{d}p} \tag{26.2}$$

where L is mRNA translation rate, ⟨mRNA⟩ is the mean mRNA count per cell and dm and dp represent mRNA and protein degradation rate, respectively (Paulsson, 2004). In essence, to confirm constitutive promoter architecture one needs to independently compute C for the reporter gene in consideration and match it with the experimentally obtained Fano factor. Assuming that the mRNA half-life, protein half-life, and either the mRNA translation rate L or the ratio of protein and mRNA abundance are known for the reporter gene, C can be directly computed from Eq. (26.2). Genome-wide gene-expression studies in yeast have put the average value of C across different genes to be $\sim$1000 molecules; however, this value can vary considerably for different genes (Bar-Even *et al.*, 2006). In summary, an inverse correlation between CV^2 and ⟨protein⟩ such that the Fano factor is uncorrelated with ⟨protein⟩ and equal to the proportionality constant C is a signature of a constitutive promoter architecture. These predictions from stochastic gene-expression models have been useful for analyzing expression noise in different *Escherichia coli* and *Saccharomyces cerevisiae* genes, and have confirmed that many essential genes in these organisms encode constitutive promoters (Bar-Even *et al.*, 2006; Newman *et al.*, 2006; Ozbudak *et al.*, 2002).

5.2. Two-state promoter architecture

Many genes exhibit elevated levels of expression noise that are inconsistent with constitutive gene expression models (Raj *et al.*, 2006). CV^2 or Fano factor values that are much higher than that predicted by Eq. (26.1) are indicative of a two-state promoter architecture, where the promoter fluctuates between an inactive and active state with rates k_{on}, k_{off} and transcription only occurs from the active state at a rate T (Fig. 26.5B and C). In such 'two-state' models, mRNAs are created in bursts during promoter transitions from inactive to active state, with k_{on} and T/k_{off} denoting the burst

frequency and the average size of the transcriptional bursts, respectively (Kepler and Elston, 2001; Simpson *et al.*, 2004). Stochastic analysis of two-state promoter models shows a complex relationship between expression noise and mean protein levels. This is both parameter-dependent and also dependent on whether promoter transcriptional efficiency is increased by increasing the burst frequency or burst size (Kaern *et al.*, 2005). At a qualitative level, CV^2 is always predicted to decrease with mean protein levels in a two-state promoter model; however, the product $CV^2 \times \langle \text{protein} \rangle$ or Fano factor can increase (Case I), decrease (Case II), or stay independent of the mean protein levels (Case III). We next discuss how these different noise profiles inform upon the two-state promoter architecture.

Case I. Fano factor increasing with mean protein counts indicates a two-state promoter architecture where higher promoter transcriptional efficiency is obtained by increasing the transcription rate T, that is, higher transcriptional burst size T/k_{off}.

Case II. Fano factor decreasing with mean protein counts indicates a two-state promoter architecture, where promoter infrequently transitions between stable active and inactive states, and higher promoter transcriptional efficiency is obtained by increasing the burst frequency k_{on}.

Case III. Fano factor uncorrelated with mean protein counts indicates a two-state promoter architecture, where promoter infrequently transitions to an unstable active state that rapidly transitions back to the inactive state. Furthermore, higher prompter transcriptional efficiency is obtained by increasing the burst frequency k_{on}.

A recent study of expression noise in HIV-1 LTR promoter using lentiviral vectors shows noise profiles that are a combination of Cases I and III (Singh *et al.*, 2010). More specifically, across some integration sites CV^2 decreases with mean protein counts such that the Fano factor is invariant. However, at many other integration sites Fano factor increases with mean protein counts. Collectively, this data suggests that HIV-1 encodes a two-state promoter architecture, which infrequently transitions to an unstable active promoter state that is rapidly switched OFF after making a burst of few mRNAs. Moreover, the site of integration uses both burst frequency and burst size to modulate viral gene expression.

In summary, CV^2 versus mean analysis has been instrumental in studying expression noise in both prokaryotic and eukaryotic genes. It has shown that while many essential genes encode constitutive promoters to minimize noise levels, other promoters such as HIV-1 LTR use two-state promoter architectures to increase expression noise. The unique property of lentiviral vectors to stably integrate a promoter or genetic circuit at a semi-random

position in the human genome makes them an ideal tool for studying gene-expression noise and they will likely find an increasing use in inferring and comparing the regulatory architecture of human promoters.

6. Conclusion

Lentiviral vectors provide several unique experimental advantages for studying gene-expression noise in mammalian systems including the ability to easily introduce transgenes that remain stable as single integrations in host genomic DNA, and the ability to generate isogenic populations of cells where the transgene is genetically stable and often transcriptionally stable. In addition, the semi-random integration pattern of lentiviruses and the natural variation in the chromatin microenvironment inherent at the lentiviral integration site generates differences in lentiviral-encoded promoter strength and this difference provides a natural method to study CV^2 as a function of mean protein levels.

Above, we have focused exclusively on the use of lentiviral vectors as probes to measure magnitude of noise, but we have not touched upon the analysis of temporal correlations in noise or frequency-domain analysis of noise (Austin *et al.*, 2006; Cox *et al.*, 2008). Such frequency-domain analysis of noise can be highly informative about underlying biology and lentiviral vectors are also very effective tools for analyzing the frequency and temporal correlations of noise. We have used lentiviral vectors to study temporal correlations in HIV gene-expression noise and to characterize HIV gene-expression circuitry (Weinberger *et al.*, 2008). Thus, the utility of lentiviral vectors for the study of noise is not limited to measurements of noise magnitude (i.e., CV) and lentiviral vectors are effective tools to probe biological noise in multiple dimensions.

ACKNOWLEDGMENTS

This work was supported by the NIH (GM083395), the California HIV/AIDS Research Fund (ID08-SD-01), and UCSD Center for AIDS Research (NIAID 5 P30 AI36214).

REFERENCES

Akkina, R. K., Walton, R. M., Chen, M. L., Li, Q. X., Planelles, V., and Chen, I. S. (1996). High-efficiency gene transfer into CD34+ cells with a human immunodeficiency virus type 1-based retroviral vector pseudotyped with vesicular stomatitis virus envelope glycoprotein G. *J. Virol.* **70,** 2581–2585.

Austin, D. W., Allen, M. S., McCollum, J. M., Dar, R. D., Wilgus, J. R., Sayler, G. S., Samatova, N. F., Cox, C. D., and Simpson, M. L. (2006). Gene network shaping of inherent noise spectra. *Nature* **439,** 608–611.

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

Blake, W. J., Kaern, M., Cantor, C. R., and Collins, J. J. (2003). Noise in eukaryotic gene expression. *Nature* **422,** 633–637.

Blake, W. J., Balazsi, G., Kohanski, M. A., Isaacs, F. J., Murphy, K. F., Kuang, Y., Cantor, C. R., Walt, D. R., and Collins, J. J. (2006). Phenotypic consequences of promoter-mediated transcriptional noise. *Mol. Cell* **24,** 853–865.

Challita, P. M., and Kohn, D. B. (1994). Lack of expression from a retroviral vector after transduction of murine hematopoietic stem cells is associated with methylation in vivo. *Proc. Natl. Acad. Sci. USA* **91,** 2567–2571.

Cox, C. D., McCollum, J. M., Allen, M. S., Dar, R. D., and Simpson, M. L. (2008). Using noise to probe and characterize gene circuits. *Proc. Natl. Acad. Sci. USA* **105,** 10809–10814.

Cronin, J., Zhang, X. Y., and Reiser, J. (2005). Altering the tropism of lentiviral vectors through pseudotyping. *Curr. Gene Ther.* **5,** 387–398.

Dull, T., Zufferey, R., Kelly, M., Mandel, R. J., Nguyen, M., Trono, D., and Naldini, L. (1998). A third-generation lentivirus vector with a conditional packaging system. *J. Virol.* **72,** 8463–8471.

Eldar, A., Chary, V. K., Xenopoulos, P., Fontes, M. E., Loson, O. C., Dworkin, J., Piggot, P. J., and Elowitz, M. B. (2009). Partial penetrance facilitates developmental evolution in bacteria. *Nature* **460,** 510–514.

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

Fischer, U., Huber, J., Boelens, W. C., Mattaj, I. W., and Luhrmann, R. (1995). The HIV-1 Rev activation domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* **82,** 475–483.

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

Hanna, J., Saha, K., Pando, B., van Zon, J., Lengner, C. J., Creyghton, M. P., van Oudenaarden, A., and Jaenisch, R. (2009). Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature* **462,** 595–601.

Jordan, A., Defechereux, P., and Verdin, E. (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.

Kaern, M., Elston, T. C., Blake, W. J., and Collins, J. J. (2005). Stochasticity in gene expression: From theories to phenotypes. *Nat. Rev. Genet.* **6,** 451–464.

Kepler, T. B., and Elston, T. C. (2001). Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations. *Biophys. J.* **81,** 3116–3136.

Kim, V. N., Mitrophanous, K., Kingsman, S. M., and Kingsman, A. J. (1998). Minimal requirement for a lentivirus vector based on human immunodeficiency virus type 1. *J. Virol.* **72,** 811–816.

Mitrophanous, K., Yoon, S., Rohll, J., Patil, D., Wilkes, F., Kim, V., Kingsman, S., Kingsman, A., and Mazarakis, N. (1999). Stable gene transfer to the nervous system using a non-primate lentiviral vector. *Gene Ther.* **6,** 1808–1818.

Miyoshi, H., Blomer, U., Takahashi, M., Gage, F. H., and Verma, I. M. (1998). Development of a self-inactivating lentivirus vector. *J. Virol.* **72,** 8150–8157.

Naldini, L., Blomer, U., Gallay, P., Ory, D., Mulligan, R., Gage, F. H., Verma, I. M., and Trono, D. (1996). In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. *Science* **272,** 263–267.

Neville, M., Stutz, F., Lee, L., Davis, L. I., and Rosbash, M. (1997). The importin-beta family member Crm1p bridges the interaction between Rev and the nuclear pore complex during nuclear export. *Curr. Biol.* **7,** 767–775.

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

Olsen, J. C. (1998). Gene transfer vectors derived from equine infectious anemia virus. *Gene Ther.* **5,** 1481–1487.

Ozbudak, E. M., Thattai, M., Kurtser, I., Grossman, A. D., and van Oudenaarden, A. (2002). Regulation of noise in the expression of a single gene. *Nat. Genet.* **31,** 69–73.

Paulsson, J. (2004). Summing up the noise in gene networks. *Nature* **427,** 415–418.

Poeschla, E. M., Wong-Staal, F., and Looney, D. J. (1998). Efficient transduction of nondividing human cells by feline immunodeficiency virus lentiviral vectors. *Nat. Med.* **4,** 354–357.

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

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

Reiser, J., Harmison, G., Kluepfel-Stahl, S., Brady, R. O., Karlsson, S., and Schubert, M. (1996). Transduction of nondividing cells using pseudotyped defective high-titer HIV type 1 particles. *Proc. Natl. Acad. Sci. USA* **93,** 15266–15271.

Schroder, A. R., Shinn, P., Chen, H., Berry, C., Ecker, J. R., and Bushman, F. (2002). HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **110,** 521–529.

Simpson, M. L., Cox, C. D., and Sayler, G. S. (2004). Frequency domain chemical Langevin analysis of stochasticity in gene transcriptional regulation. *J. Theor. Biol.* **229,** 383–394.

Singh, A., and Weinberger, L. S. (2009). Stochastic gene expression as a molecular switch for viral latency. *Curr. Opin. Microbiol.* **12,** 460–466.

Singh, A., Razooky, B., Cox, C. D., Simpson, M. L., and Weinberger, L. S. (2010). Transcriptional bursting from the HIV-1 promoter is a significant source of stochastic noise in HIV-1 gene expression. *Biophys. J.* **98,** L32–L34.

Suel, G. M., Garcia-Ojalvo, J., Liberman, L. M., and Elowitz, M. B. (2006). An excitable gene regulatory circuit induces transient cellular differentiation. *Nature* **440,** 545–550.

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

Weinberger, L. S., and Shenk, T. (2007). An HIV feedback resistor: Auto-regulatory circuit deactivator and noise buffer. *PLoS Biol.* **5,** e9.

Weinberger, L. S., Burnett, J. C., Toettcher, J. E., Arkin, A. P., and Schaffer, D. V. (2005). Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. *Cell* **122,** 169–182.

Weinberger, L. S., Dar, R. D., and Simpson, M. L. (2008). Transient-mediated fate determination in a transcriptional circuit of HIV. *Nat. Genet.* **40,** 466–470.