www.sciencemag.org/cgi/content/full/science.1250220/DC1

Supplementary Materials for

Screening for noise in gene expression identifies drug synergies

Roy D. Dar, Nina N. Hosmane, Michelle R. Arkin, Robert F. Siliciano, Leor S. Weinberger*

*Corresponding author. E-mail: leor.weinberger@gladstone.ucsf.edu

Published 5 June 2014 on *Science* Express
DOI: 10.1126/science.1250220

This PDF file includes

Other Supplementary Material for this manuscript includes the following:
(available at www.sciencemag.org/cgi/content/full/science.1250220/DC1)

Table S1. The 85 Noise Enhancers with Compound ID, name, ΔFL, and ΔCV^2, for both d2GFP and mCherry reporters and percent reactivation with activators.

Movie S1. Time-lapse fluorescence microscopy movie of LTR-d2GFP polyclonal cells for 20 hours, 24 hours posttreatment with noise enhancer V1. The movie plays on a loop and displays 20 unique x-y locations sampled over time in a 4 by 5 array. Tracked single cells are marked with a red dot for quality control and postprocessing (*2*, *31*).

Supplementary Material

TABLE OF CONTENTS

Methods:

Screen Design and Processing:

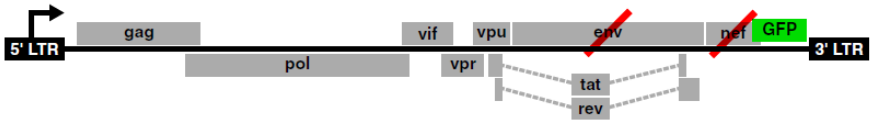
Advanced Studies of Noise Enhancement:

Supporting Figures, Files, and References:

Materials and Methods

Cell Culture and Growth Conditions

A Jurkat isoclone consisting of LTR-d2GFP and LTR-mCherry was used for noise drug screening and chosen from a previously reported isoclone library (*1*, *2*). J-Lat clones 8.6, 9.2, 15.4, and 10.6 were previously described (*3*). J-Lat 8.6 was used in Figures 3, 4A, and 4B. Jurkat cells were cultured in RPMI 1640 supplemented with 10% FBS and 1% penicillin-streptomycin. Cells were grown at a density of ~10^6/mL at 37 ºC, 5% CO_2, under humidified conditions. TNFα (Sigma-Aldrich) was used at concentrations of 10 ng/ml. Prostratin (Sigma) at a concentration of 3 µM, SAHA (Cayman Chemical) at 2.5 µM, Trichostatin A (TSA; Sigma-Aldrich) at 400nM, JQ1 (Cayman Chemical) at 1µM, phorbol ester (PMA) at 200ng/mL, valproic acid (VPA) at 1mM, azadeozyciditine (AZA) at 5µM, and MS-275 (Cayman Chemical) at 10µM.



Schematic of the full-length HIV Jurkat Latency (J-Lat) construct (*3*). GFP is inserted in place of *nef* and *env* is deleted making the viral construct replication incompetent. The following lists information on the integration sites of J-Lat clones used both in the main text and Supplementary Material (personal communication, Eric Verdin):

Clone	band (kb)	Chr	Gene	Orientation	description	Category	Alias
8.6	no						
9.2	1.2	19q13.32	ppp5c	same	Serine/threonine protein phosphatase 5 (EC 3.1.3.16) (PP5) (Protein phosphatase T) (PP-T) (PPT).	cell cycle/transcription	
10.6	1.6	9q34.3	KIAA0310	opposing	Regucalcin gene promoter region-related protein (RGPR)		KIAA1928
15.4	1.1	19q13	UBA2		ubiquitin-activating/SUMO-1activating enzyme subunit 2	ubiquitin	

Generation of Latently Infected Primary Cells

As previously described in detail (*4*), bulk CD4+ T cells were isolated from healthy adult peripheral blood mononuclear cells (PBMCs) using a $CD4^+$ T Cell Isolation Kit II (Miltenyi Biotec). The cells were activated with plate-bound anti-CD3 and soluble anti-CD28 antibodies (BD Biosciences). To coat the anti-CD3 antibodies to the plate, 6-well plates were coated with 1mL of PBS containing 10 µg/ml anti-CD3 antibody per well and incubated at 37 C for 90 minutes. The primary cells were added to the plate and activated by incubation with 1 µg/ml anti-CD28 monoclonal antibody in RPMI 1640 media (Gibco by Life Technologies) enriched with T cell growth factors and 100U/mL IL-2. After 72 hours, the cells were transduced with the EB-FLV lentiviral vector, for constitutive *bcl-2* expression, and spinoculated at 1200 g at room temperature for 2 hours. The cells were then expanded in the presence of 100U/mL IL-2 and T cell growth factor-enriched medium for an additional 72 hours, and subsequently were replaced in RPMI 1640 with 10% FBS and 1% penicillin/streptomycin and cultured for 4 weeks in the absence of

exogenous cytokines. At the end of the culture period, viable cells were collected by Ficoll-Hypaque density gradient centrifugation. The viable *bcl-2* transduced cells were reactivated with anti-CD3 and anti-CD28 antibodies, as previously described, and were expanded in culture by adding 100U/mL IL-2 in RPMI 1640 every other day for approximately a week. Viable cells collected by density gradient centrifugation were activated again with plate-bound anti-CD3 and soluble anti-CD28 antibodies and were infected 72 hours after with a reporter virus NL4-3-Δ6-drEGFP construct that contains a deletion in *env* and mutations in *gag*, *vif*, *vpr*, *vpu*, *nef* (*4*). Cells were infected by spinoculation at 1200 g at room temperature for 2 hours and were cultured in media enriched with T cell growth factors and 100U/mL IL-2 for 3 days, and subsequently cultured in RPMI 1640 with 10% FBS and 1% Penicillin/Streptomycin for 6 weeks without the addition of exogenous cytokines. Viable cells were collected as described previously. To isolate latently infected cells containing the reporter virus, GFP-negative cells were sorted using a MoFlo cell sorter (Beckman Coulter) at the Johns Hopkins Flow Cytometry Core Facilities.

Screening and flow cytometry analysis.

The screen used the Pharmakon1600 library of approved drugs from Microsource Discovery Systems, Inc.. Drug treatments were performed for 24 hours at 10μM final concentration in a 96-well plate format. Automated compound addition to plates was performed using a liquid handling system (Beckman Coulter Biomek FXp Laboratory Automation Workstation) at the UCSF Small Molecule Discovery Center (SMDC). Flow cytometry was performed using a high throughput sampler (HTS) on a BD LSRII cytometer, maintained by the Gladstone Institutes Flow Cytometry Core Facility. Tracking of cytometer performance was performed daily using Cytometer Setup & Tracking beads (CS&T). Plates were kept in a 37 ºC tissue culture incubator under 5% CO_2, and humidified conditions until measurement and wells were mixed before acquiring samples. Treated cells were measured unfixed and live to avoid additional sources of variability from the fixation. 50k live cells were collected from each well for noise measurements. A very conservative gating for a live subset of ~3k cells of similar size, volume, and state, was applied on the FSC vs SSC to reduce extrinsic noise contributions (*5*) (see sample dataset and gating below). Plates included a non-fluorescent naïve cell population to correct for autofluorescent contributions to mean fluorescence and noise. For reactivation experiments, 10k cells were collected in the live scatter gate defined by the untreated sample.

For primary cells, to measure maximal reactivation of latent HIV-1, or percentage of latently infected cells, within a given GFP-negative sorted batch, 50 ng/mL of PMA (Sigma Aldrich) and 1 uM of Ionomycin (Sigma Aldrich) are added to approximately 50,000 cells in a well and expression of GFP is measured 48 hours later by flow cytometry analysis on a BD FACSCalibur (BD Biosciences). The positive control presented in this work was the percentage of latently infected cells maximally activated by PMA and Ionomycin. For the drug combination experiments involving the noise enhancers and suppressors, PMA and Prostratin were used at concentrations of .1 ng/mL and .3 uM, respectively, and were performed in duplicate on 60,000 cells per treatment. Reactivation from each drug combination treatment was analyzed by normalizing the expression of GFP 48 hours later as a percentage of the positive control, or total percent of latently infected cells, within the batch of model cells used.

Bliss Independence Score calculation

The Bliss Independence Score (*6*) is the expected additive value for latent reactivation when combining drugs A and B. Any values significantly above the calculated bliss would be considered synergistic. Bliss was calculated for all 85 noise enhancers used in combination with either TNF or Prostratin in Fig. 3A and an average of all bliss scores was calculated.

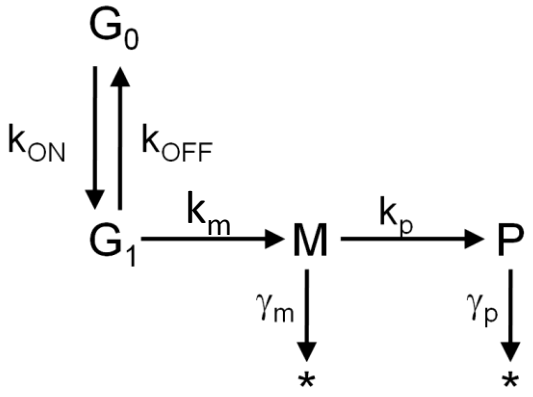
The Bliss Score was calculated using:

$$F_{AB} = F_A + F_B * (1 - F_A)$$

where F_A and F_B represent the fractional response or % reactivation of cells when treated with drugs A or B alone, and F_{AB} for treatment with drugs A and B in combination.

Stochastic simulations of the two-state promoter switching model

The transcriptional bursting model of gene expression used here was previously validated for HIV gene expression (*2*) and specifies switching between discrete high and low transcriptional rates. This minimal model is governed by the following rate parameters (see Figure below): (i) the initiation rate into the transcribing state, k_{on}; (ii) the rate at which the promoter switches off or time dwelled in the ON state, k_{off}; (iii) the transcription rate in the high expression state, k_m; (iv) the translation rate, k_p; and finally (v) the degradation rates of mRNA and protein species (M and P respectively) γ_m and γ_p.



Exact stochastic simulations in Fig. 1C were performed by implementing the Gillespie algorithm (*7*). Physiologically relevant rate parameters were adapted from (2) which quantified the range of burst size and frequency for the LTR promoter in addition to values for the protein and mRNA half-lives.

Simulation Parameters for Noise Enhancement in Fig. 1B

Simulation Parameter	Biological Interpretation	Untreated Simulation	Noise ENHANCEMENT Simulation	ACTIVATOR Simulation	ACTIVATOR + Noise ENHANCEMENT Simulation
k_{ON}	Initiation rate	0.000208333 sec^{-1}	0.000104167 sec^{-1}	0.000625 sec^{-1}	0.000625 sec^{-1}
k_{OFF}	OFF rate	0.002083333 sec^{-1}	0.001041667 sec^{-1}	0.002083 sec^{-1}	0.001041667 sec^{-1}
k_m	Expression rate	0.208333333 sec^{-1}	0.208333333 sec^{-1}	0.208333333 sec^{-1}	0.208333333 sec^{-1}
k_p	Translation rate	0.032346868 sec^{-1}	0.032346868 sec^{-1}	0.032346868 sec^{-1}	0.032346868 sec^{-1}

γ_m	mRNA decay rate	0.000115525 sec^{-1}	0.000115525 sec^{-1}	0.000115525 sec^{-1}	0.000115525 sec^{-1}
γ_p	GFP decay rate	7.70164E-05 sec^{-1}	7.70164E-05 sec^{-1}	7.70164E-05 sec^{-1}	7.70164E-05 sec^{-1}

Simulation Parameters for Noise Suppression in Fig. 4D:

Simulation Parameter	**Biological Interpretation**	**Untreated Simulation**	**Noise SUPPRESSOR Simulation**	**ACTIVATOR Simulation**	**ACTIVATOR + Noise SUPPRESSOR Simulation**
k_{ON}	Initiation rate	0.000208333 sec^{-1}	0.000354167 sec^{-1}	0.000625 sec^{-1}	0.000625 sec^{-1}
k_{OFF}	OFF rate	0.002083333 sec^{-1}	0.003541667 sec^{-1}	0.002083 sec^{-1}	0.00354167 sec^{-1}
k_m	Expression rate	0.208333333 sec^{-1}	0.208333333 sec^{-1}	0.208333333 sec^{-1}	0.208333333 sec^{-1}
k_p	Translation rate	0.032346868 sec^{-1}	0.032346868 sec^{-1}	0.032346868 sec^{-1}	0.032346868 sec^{-1}
γ_m	mRNA decay rate	0.000115525 sec^{-1}	0.000115525 sec^{-1}	0.000115525 sec^{-1}	0.000115525 sec^{-1}
γ_p	GFP decay rate	7.70164E-05 sec^{-1}	7.70164E-05 sec^{-1}	7.70164E-05 sec^{-1}	7.70164E-05 sec^{-1}

All simulations used a steady-state basal expression level of 30k proteins for illustration (i.e. to preclude extinction of expression trajectories).

Analytic approach showing that noise enhancer synergizes with transcriptional activators to increase mean expression level

Applying established theory for the two-state model of episodic transcription, the expression for mean protein abundance is (*8, 9*):

$$\langle P \rangle = O \frac{k_m k_p}{\gamma_m \gamma_p}$$

where k_m, k_p, γ_m, and γ_p, are the mRNA and protein transcription, translation, and degradation rates respectively. The average time in the ON state (or the "On Fraction") can be defined as:

$$O = \frac{k_{ON}}{k_{ON}+k_{OFF}} \ .$$

The Activator increases expression by increasing k_{ON} (or O). The Noise Enhancer conserves abundance and increases noise either by decreasing k_{OFF} while decreasing k_{ON}, by increasing k_m with a decrease in k_{ON}, or by a mixture of k_{OFF} and k_m changes with a decrease in k_{ON} . Enhanced activation requires and assumes that any changes in k_{ON} by the noise enhancer are overly-compensated by the activator.

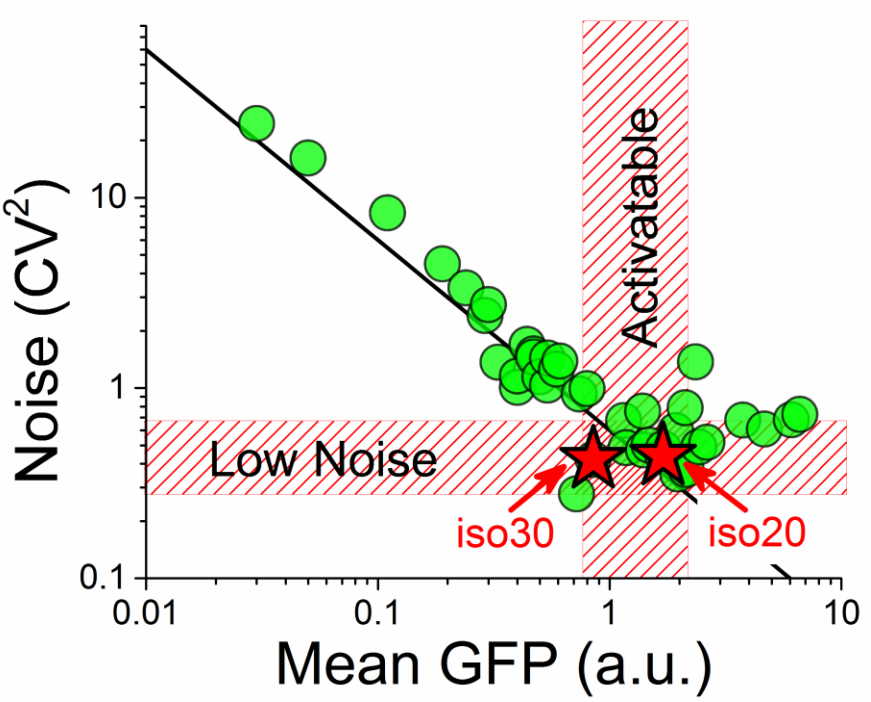
For synergy between noise enhancers and activators:

$$\langle P \rangle_{Untreated} = \langle P \rangle_{Noise\ Enhancer} < \langle P \rangle_{Activator} < \langle P \rangle_{Noise\ Enhancer+Activator}$$

The Noise Enhancer and Activator increase expression by a total increase in k_{ON} coupled with a burst size increase ($\Delta k_m/\Delta k_{OFF}$).

Selection of the LTR-d2GFP isoclone cell line for screening

The screened LTR-d2GFP isoclone (i.e. clonal population of cells with an identical LTR-d2GFP integration site) was selected from a previously characterized isoclone library consisting of 40 clones (*1*, *2*). The clone was selected at a mid-range of gene expression and at a pivotal abundance domain where transcription switches from constant burst size with increasing burst frequency to increasing burst size (red stars below) (*2*). This region in the noise versus mean space has the lowest noise which allows for the largest dynamic range in noise enhancement for the screen. In addition, the mid-abundance region is sufficiently activatable in the mean expression level and allows gene expression movements in any direction in the noise versus mean abundance space. From this region, 2 isoclones named "iso20" and "iso30" were selected and further characterized as potential drug screen candidates. Each of the LTR-d2GFP isoclones harbored an additional LTR driving a stable mCherry reporter (*1*, *2*) which was also used in the screening. For representative raw flow cytometry data and histograms see below.



The above plot shows untreated flow cytometry measurements of a previously reported library of LTR-d2GFP LTR-mCherry isoclones. The black line is a constant burst size model line (*2*). Between the two isoclones considered for the screen (red stars), "iso20" was chosen from the region depicted as lowest in noise and most activatable (highest dynamic range in both noise and mean expression).

Plate Layout for Drug Screening

The 1600 compound FDA-approved library was screened using 20 x 96 well plates. In each plate the left column included the transcriptional activator TNF as a positive control at a concentration of 10 ng/mL. The right column was left untreated, and a single well, A12, of each plate included a non-fluorescent Naïve Jurkat cell population for autofluorescence controls. The remaining 10 columns included 80 unique compounds from the library (depicted here as D1-D80). Isoclone 20 (iso20) is the name of the cell line used in the screen and described in the section above.

	1	2	3	4	5	6	7	8	9	10	11	12
A	iso 20 + TNF	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	Naïve
B	iso 20 + TNF	D11	D12	D13	D14	D15	D16	D17	D18	D19	D20	Untreated iso 20
C	iso 20 + TNF	D21	D22	D23	D24	D25	D26	D27	D28	D29	D30	Untreated iso 20
D	iso 20 + TNF	D31	D32	D33	D34	D35	D36	D37	D38	D39	D40	Untreated iso 20
E	iso 20 + TNF	D41	D42	D43	D44	D45	D46	D47	D48	D49	D50	Untreated iso 20
F	iso 20 + TNF	D51	D52	D53	D54	D55	D56	D57	D58	D59	D60	Untreated iso 20
G	iso 20 + TNF	D61	D62	D63	D64	D65	D66	D67	D68	D69	D70	Untreated iso 20
H	iso 20 + TNF	D71	D72	D73	D74	D75	D76	D77	D78	D79	D80	Untreated iso 20

Z' Calculation for screen

Z* was tested for sensitivity and quality of the screen setup. A plate including the positive TNF activator control and untreated negative control for both iso20 and iso30 cell line candidates were measured for mean GFP and noise on the LSR II fast settings (see next section).

	1	2	3	4	5	6	7	8	9	10	11	12
A	Naiive	iso 20	iso 20	iso 20	iso 20	iso 20	iso 30	iso 30	iso 30	iso 30	iso 30	Naiive + TNF
B	iso 20 + TNF	iso 20	iso 20	iso 20	iso 20	iso 20	iso 30	iso 30	iso 30	iso 30	iso 30	iso 20 + TNF
C	iso 30 + TNF	iso 20	iso 20	iso 20	iso 20	iso 20	iso 30	iso 30	iso 30	iso 30	iso 30	iso 30 + TNF
D	iso 20 + TNF	iso 20	iso 20	iso 20	iso 20	iso 20	iso 30	iso 30	iso 30	iso 30	iso 30	iso 20 + TNF
E	iso 30 + TNF	iso 20	iso 20	iso 20	iso 20	iso 20	iso 30	iso 30	iso 30	iso 30	iso 30	iso 30 + TNF
F	iso 20 + TNF	iso 20	iso 20	iso 20	iso 20	iso 20	iso 30	iso 30	iso 30	iso 30	iso 30	iso 20 + TNF
G	iso 30 + TNF	iso 20	iso 20	iso 20	iso 20	iso 20	iso 30	iso 30	iso 30	iso 30	iso 30	iso 30 + TNF
H	Naiive + TNF	iso 20	iso 20	iso 20	iso 20	iso 20	iso 30	iso 30	iso 30	iso 30	iso 30	Naiive

Z-Prime* was calculated using the following equation:

$$\text{Z-Prime*:} \quad 1 - \left(\frac{3 \times (\text{Standard Deviation of the Maximum Signal Control}) + 3 \times (\text{Standard Deviation of the Minimum Signal Control})}{|(\text{Mean of the Maximum Signal Control} - \text{Mean of the Minimum Signal Control})|} \right.$$

where the maximum signal for the mean d2GFP (or d2GFP CV^2) would be TNF addition, and the minimum signal (or negative control) is the untreated-cell population.

The above formula is described in (*10*) and yielded the following results for the drug screen candidates:

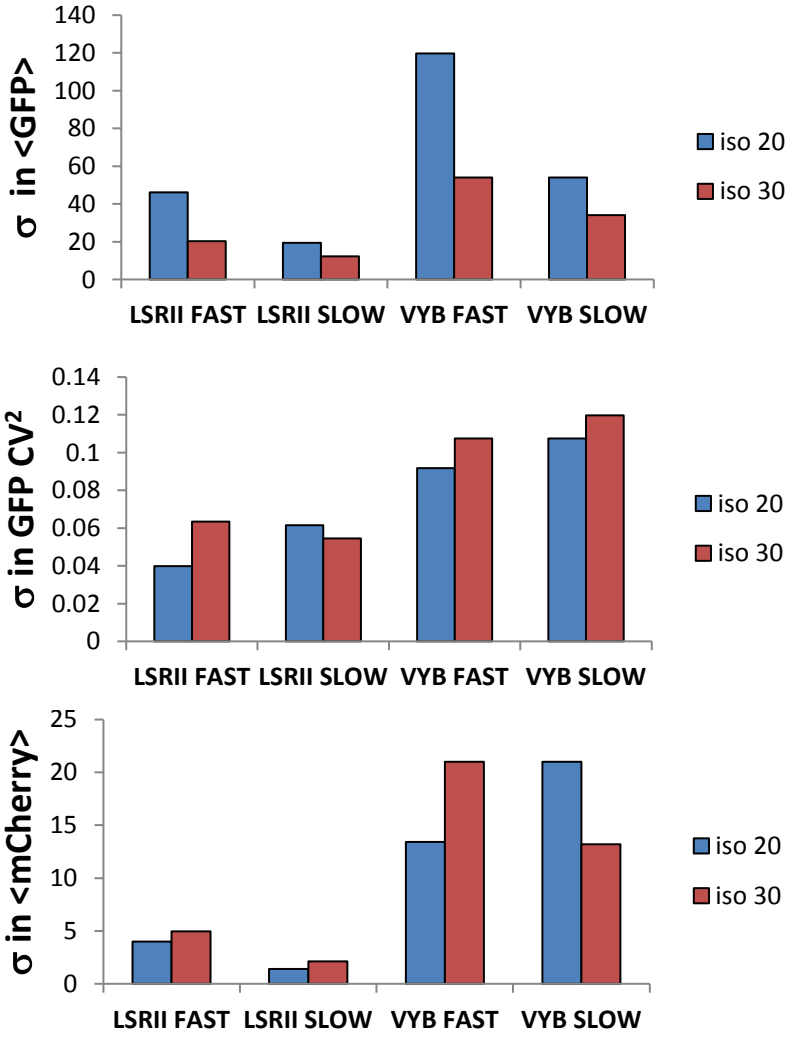
iso 20	<d2GFP>	CV^2 d2GFP	<mCherry>
Z'	0.959476	0.811440702	0.825170772

iso 30	<d2GFP>	CV^2 d2GFP	<mCherry>
Z'	0.922826	0.697299996	0.921690533

The Z' for GFP CV^2 of iso20 was found to be higher than iso30 further validating its selection as the screening cell line.

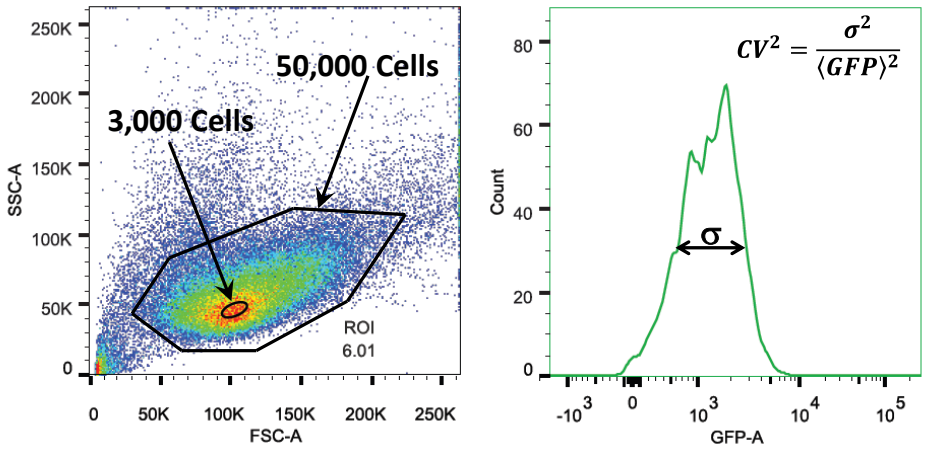
Selection of a high-throughput flow cytometer and sample inject speed based on σ of CV^2

Control plates consisting of 2 LTR-d2GFP candidates for the screen (iso 20 and iso 30) were tested on two flow cytometers (BD LSRII and Miltenyi MACSQuant VYB) at two different sample injection speeds (3 and 0.5 uL/sec). Isoclone 20 was selected for screening on the BD LSRII with fast injection speed, which had the lowest standard deviation in GFP noise.



Sample of raw flow cytometry data and conservative gating for the LTR-d2GFP population

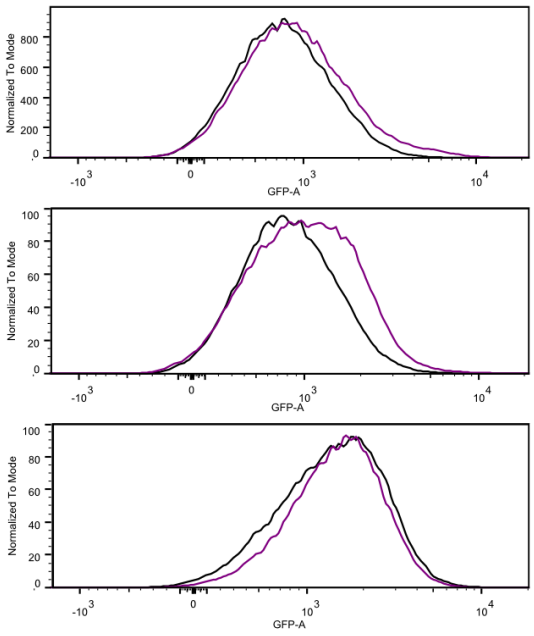
Raw data for 50k cells collected from the LTR-d2GFP cell line used in the screen. Forward versus side-scatter of cells (left) are displayed along with the conservative gating used for noise and mean quantifications (black oval). The conservative gating used ~3k of the ~50k Live gated cells. Noise and mean GFP (and mCherry) are calculated using the histogram from the conservative gating (right). The conservative gating approach is a previously validated method to gate out extrinsic noise (*1*, *5*, *11*).



Here the variance and mean GFP (or mCherry) have been corrected for autofluorescence by the naïve population as previously described (*5*).

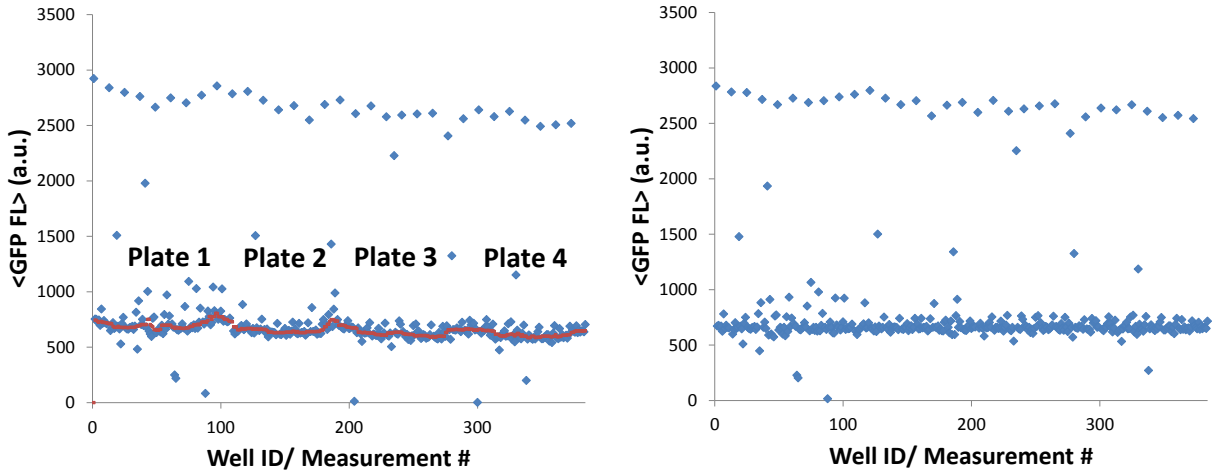
LTR-d2GFP Fluorescence Histograms for Treatments with Noise Enhancers or Suppressors

Below are representative examples of 2 noise enhancer distributions in purple compared to untreated distributions in black (upper 2 panels) and one example of a noise suppressor in purple (bottom panel).



Removal of mean fluorescence and σ general trends for each plate set

The screen was run over five days, with 4 × 96-well plates run per day. Each of the 5 sets displayed a global dip in fluorescence in the middle of each plate. In addition, a general slight decrease in mean GFP was observed (left panel below). This effect was attributed to system and sample drift in measuring live cells over a period of 4-6 hours in a given day. To correct for this non-drug related artifact a median trendline was calculated (red in left panel below) and removed from each individual measurement (right panel below). Similarly, the same process was performed for the standard deviation across all measurements in each plate set for noise calculations.



Raw and median-corrected measurements of mean GFP for a set of 4 plates. (left) raw GFP fluorescence data for 4 x 96-well plates measured, (right) detrended measurement after subtraction of the median trendline.

The two-reporter method to filter for transcriptional noise enhancement

The two-color method using differential stability reporters is based upon a derived theory (*9*, *12*). The theory shows that the transcriptional bursting component of the total noise is negligible for stable long-lived reporters:

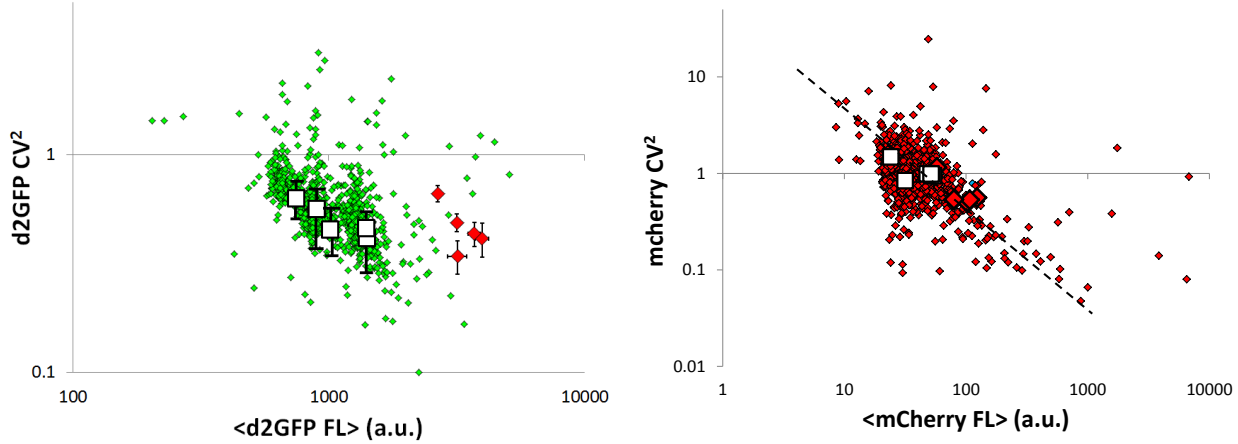
$$CV^2_{TOTAL} = CV^2_{POISSON} + CV^2_{TRANSCRIPTIONAL\ BURSTING} + CV^2_{EXTRINSIC}$$

$$CV^2 = \frac{1}{\langle p \rangle} b + C_k \frac{(1-O)}{O} + C_{ext}$$

$$\text{and} \quad \langle p \rangle = \frac{k_m \text{Ob}}{\gamma_p} \quad ; \quad O = \frac{k_{ON}}{k_{ON}+k_{OFF}}$$

Where b is the translational burst rate (or the average number of proteins translated from each mRNA), O is the fraction of time spent in the ON state (referred to as the "on fraction"), and C_k is a scaling factor that approaches 0 for fast bursting ($k_{ON} + k_{OFF} >> \gamma_p$, e.g. long-lived mCherry) and 1 for slow bursting ($\gamma_p >> k_{ON} + k_{OFF}$, e.g. d2GFP) relative to the protein reporter stability used.

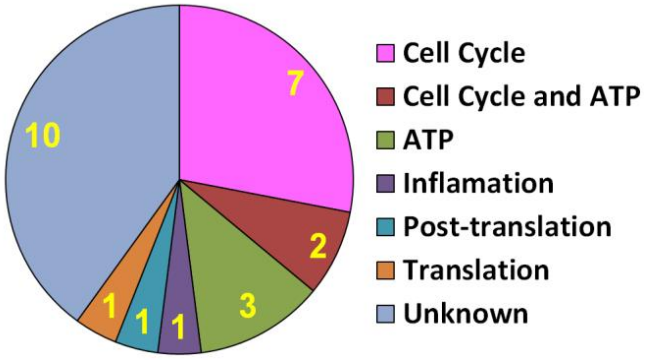
In addition to the GFP reporter in the LTR-d2GFP cell line, an LTR driving a long half-life mCherry reporter was present in the cell line. This differential-stability 2-reporter system enabled the differentiation between drugs that were primarily extrinsic (global cellular resources) and post-transcriptional variability modifiers in which the noise magnitude changed significantly in both reporters. To remove compounds that altered CV^2 post transcriptionally, we removed compounds that affected the CV^2 long-lived red reporter while conserving its mean mCherry level. This method continues development of the original and decade-old 2-reporter system for intrinsic versus extrinsic noise measurements in bacteria (*13*).



Noise versus mean fluorescence for all compounds screened on both the d2GFP and mCherry channels. White squares represent the untreated plate sets, red diamonds their corresponding TNF controls. Simultaneous changes in both d2GFP and mCherry noise enable to identify compounds causing non-transcriptional perturbations. From the 5 plate sets screened above, 25 compounds enhanced noise in both reporters by more than 2-σ and are plotted below.

The "filtered out" post-transcriptional noise-enhancer compounds correlate with post-transcriptional MOAs

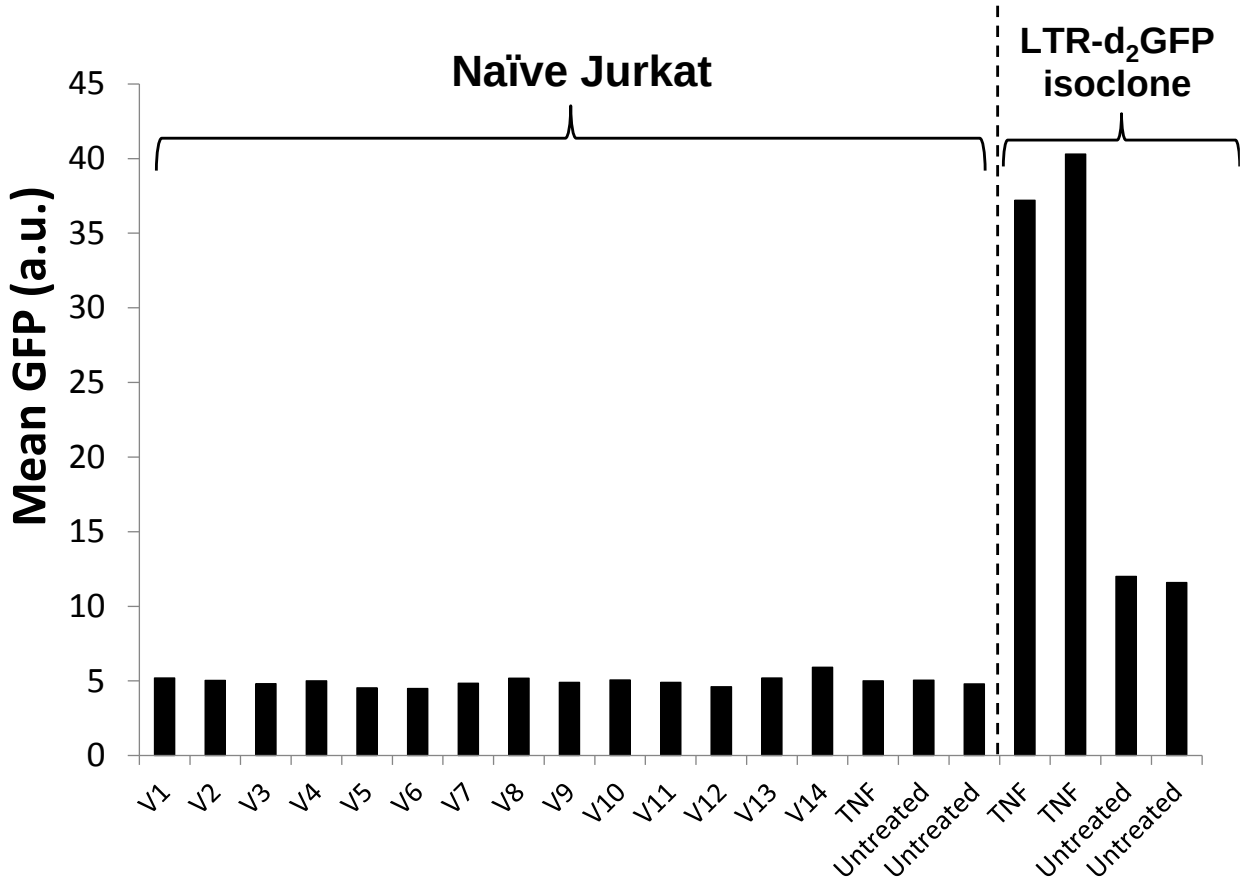
MOA	Compound Name
CC	arginine hydrochloride
CC	clofarabine
CC	fludarabine phosphate
CC	melphalan
CC	methotrexate(+/-)
CC	mitomycin c
CC	Taxol
ATP	carnitine hydrochloride
ATP	digitoxin
ATP	lonidamine
ATP CC	taurine
CC ATP	albendazole
INFL TR	rifaximin
PTL	tetramizole hydrochloride
TL	thiostrepton
	aminacrine
	Arecoline hydrobromide
	dobutamine hydrochloride
	doxapram hydrochloride
	editol
	erythrosine sodium
	nikethamide
	oxfendazole
	rutin
	zalcitabine



Compound list and classification of the 25 post-transcriptional compounds filtered from the detected noise enhancer hits (CC = Cell Cycle, PTL = Post-translation, TL = Translation, INFL = Inflammation).

Control for autofluorescence of noise-enhancing compounds

Naïve Jurkat cells were treated with a representative set of 14 noise enhancers for 24 h at 10μM. No enhanced autofluorescence was detected compared to untreated naïve Jurkats, TNF activated Jurkats, or LTR-d2GFP iso 20 cells either treated with TNF or untreated.



Vertical "Noise Streams" are evidence of transcriptional modulation

Vertical noise streams were observed from independently screened plate sets measured on different days (Fig. 2A). Since it seemed unlikely that compounds would randomly generate almost perfectly vertical increases in CV^2, we examined two potential explanations based on the underlying two-state model.

These explanations arise from the following equations described in (*1*, *2*):

$$CV^2 = (b/\langle P\rangle)*(1+BS) \qquad \text{where} \qquad b = k_p/\gamma_m \quad \text{and} \quad BS = k_m/k_{OFF},$$

$$\text{and} \qquad \langle P\rangle = BF*BS*k_p/(\gamma_m*\gamma_p) \qquad \text{where} \qquad BF = k_{ON}$$

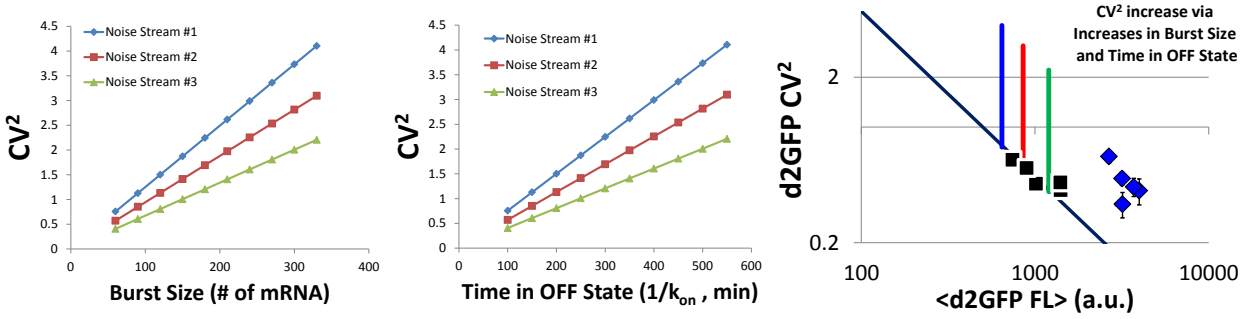
Here, *b* is the translational burst rate, BF the burst frequency, BS the burst size, k_{ON}, k_{OFF}, and k_m were previously defined in the two-state model as rate of initiation, transition rate to the OFF state, and transcription rate, respectively. γ_m and γ_p are the mRNA and protein rates respectively, and k_p is the translation rate.

We consider the case of increasing GFP noise with constant protein abundance.

Case 1: Noise increase is purely transcriptional

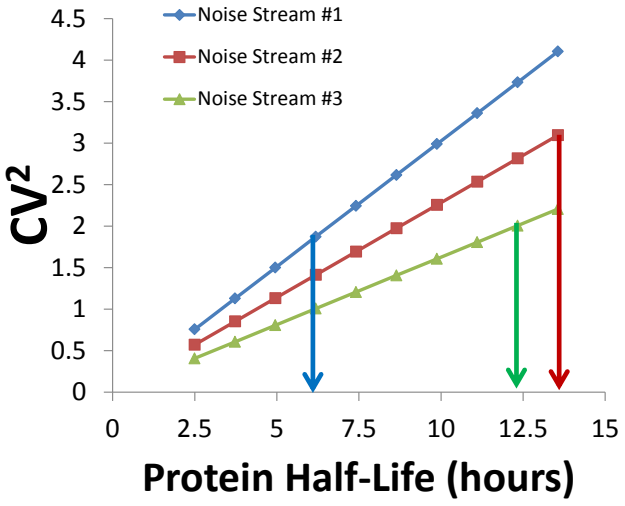
If increases in translational burst rate, *b*, can be ruled out due to no changes in the long-lived mCherry noise for the noise enhancers of d2GFP, then increases in burst size must be counter-balanced by decreases in burst frequency to conserve the mean protein abundance. I.e. for constant <P>, assuming k_p , γ_m, and γ_p remain the same, k_{on} must decrease.

Using relevant transcriptional burst parameters from (*2*) the model for this case is plotted below. The right panel shows the vertical model lines at each noise stream abundance for the parameters in the two left panels. The model lines match the abundance position of the noise streams in Figure 2A of the main text. It shows that increases in burst size and $1/k_{ON}$ can explain the vertical noise streams. The range of k_{ON} needed to reach the top of the streams is consistent with values previously reported (*2*). In addition, the model suggests that equivalent noise enhancement requires much larger burst sizes and refractory periods in the OFF state for noise streams at higher abundances (where k_{ON} is high, green trend) than at lower abundances (blue trend).



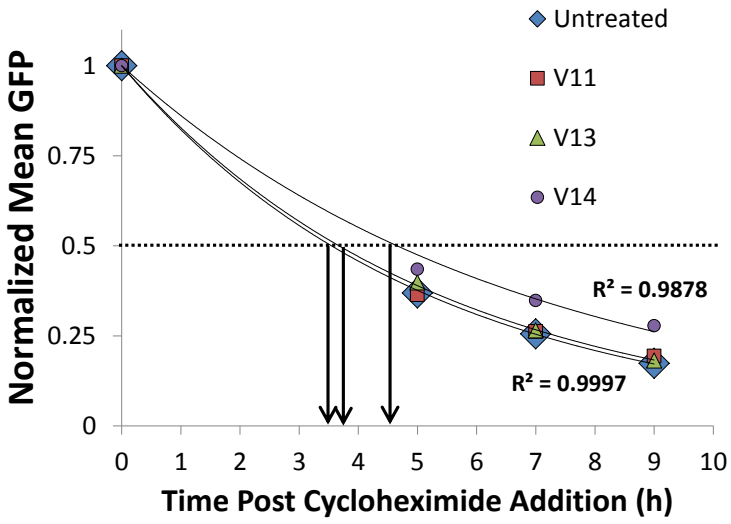
Case 2: Noise increase is purely translational

For the case of noise changes that are purely translational, the BF*BS term governing the mean protein abundance is constant and the protein decay must compensate increases in translational burst rate (b). For the same range of noise enhancement of the noise streams in "Case 1", the model yields that the d2GFP reporter half-life must increase from ~3 hours to 6-13 hours depending on the noise stream (see below, #1 would be the left most stream in Figure 2A). To test this possibility, the GFP half-lives was measured in presence and absence of noise enhancer treatments (below). The data shows that d2GFP half-life is not substantially extended indicating that translation modulation alone cannot explain vertical noise streams.

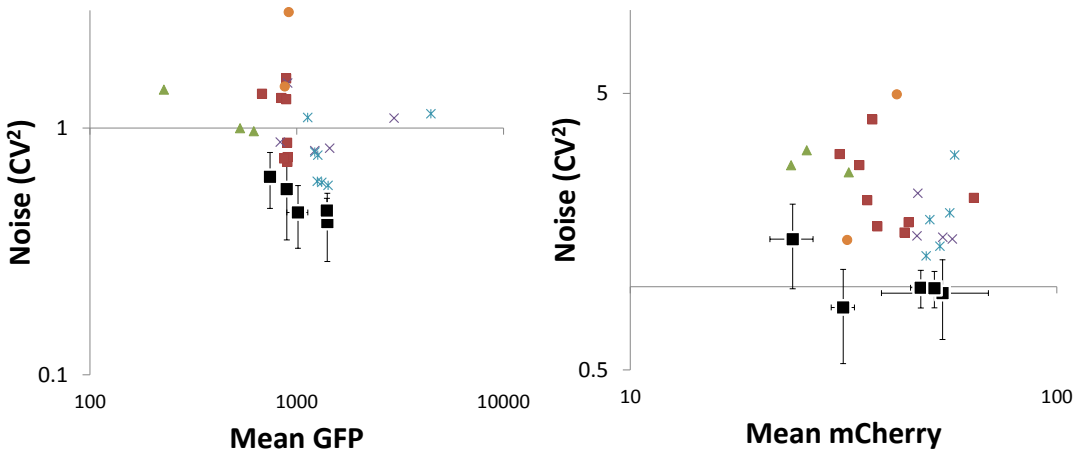


GFP half-life shows little change with noise enhancer treatment

The LTR-d2GFP isoclone from the drug screen was pretreated with 3 different noise enhancers for 24 h and then exposed to cycloheximide (30µg/mL) for 5, 7, and 9 hour durations. An exponential fit of the decay of mean GFP measured by flow cytometry yielded similar half-lives (3-5 hrs) for all compounds.



mCherry Noise Enhancement in the "noise stream":



In filtering out the 25 noise enhancers that increase both GFP and mCherry noise by more than 2σ using the two-reporter method, a subset of these post-transcriptional noise enhancers (~5) land up in the GFP noise stream (red squared and orange circles in left panel above). Based on the change in protein half-life needed in "Case 2", we estimate that these compounds cause noise enhancement dominated by either a mixture of translation and extrinsic noise, or extrinsic noise alone.

Bioactivities of detected noise enhancers ("NEs") and their normalized representation in the screened compound library

Bioactivity	# of NEs	% of Total NEs	% by Library Composition
analgesic	2	3%	0.09%
anthelmintic	3	5%	0.09%
antiarrhythmic	2	3%	0.02%
antibacterial	5	8%	0.98%
antiemetic	2	3%	0.02%
antihyperlipidemic	3	5%	0.04%
antihypertensive	4	6%	0.20%
antiinflammatory	2	3%	0.13%
antineoplastic	11	17%	0.91%
antiviral	4	6%	0.10%
bronchodilator	2	3%	0.02%
Vitamin B2 and B5	2	3%	0.00%
Other	25	38%	24.24%

Mechanisms of action for noise enhancers (NEs) and known effects on LTR transcription

The listed noise enhancers below are those found to have known targets related to regulation and transcription of the HIV LTR. The cis binding site of the LTR promoter included in parentheses describes a reported interaction with the listed drug target. S1 is the noise suppressor shown in Fig. S11.

Target	# of NEs	NE ID#	Reference to LTR Transcription
p53 (SP1)	2	V64, V70	Duan et al, JVI (1994); Bargonetti et al, CMB (1997)
NFKB and IL-6	3	V71, V18, V95	Duh et al, PNAS (1989); Poli et al, JEM (1990)
FOS Activators (AP-1)	3	V28, V87, V100	Van Lint et al, J. Virol (1997); Kharroubi and Verdin, JBC (1994)
ESR1 (SP1)	4	V2, V3, V10, V68	Katagiri et al, Intnl. Immunopharmacology (2006); Asin et al, AIDS Res. and Human Retroviruses (2008)
Dopamine Receptor (p53 and Ca^{2+})	2	V12, V68	Fernandez et al, PNAS (2005); Fernandez et al, Protein Sci. (2008); Liu et al, Mol Pharmacol. (2008);
Methylation Inhibitor	1	V17	Bednarik et al, EMBO (1990); Blazkova et al, PLoS Pathog. (2009)
CCL2/CCL5/CCL11	3	V20, V13, V11	Handen and Rosenberg, FEBS Lett (1997)
Purine nucleotide synthesis	1	V9	
Calcium Channel Blocker	1	S1	Chan et al, PLoS ONE (2013)

Cited in table: p53 (*14*, *15*); NFKB and IL-6 (*16*, *17*); FOS (*18*, *19*); ESR1 (*20*, *21*); Dopamine Receptor (p53 and Ca^{2+}):(*22-24*)); Methylation Inhibition: (*25*, *26*); Calcium Flux and NFKB: (*27*).

Using a literature search and the known drug targets of the noise enhancing compounds, a subset of noise enhancers clustered to target factors responsible for binding cis regulatory binding sites found in the HIV

LTR promoter and implicated in regulation of LTR transcription. Noise enhancers included activators of p53 (known to bind CDK9 and stall transcriptional elongation) (*14, 15*), modulators of TNF and NF-KB, methylation inhibitors, and modulators of JUN-B and c-FOS (regulators of transcription at AP-1 sites in the LTR) (*18, 19*). Estrogen receptor (ESR1) agonists appeared in 4 different noise enhancers and are immediately suspect of modifying LTR transcription since ESR1 binds SP1 which the LTR has 3 binding sites for. ESR1 also binds p300 which is involved in Tat transactivation of the LTR promoter (*28*). Estrogen has been reported to directly affect the efficiency of SP1 binding to the LTR in a Tat independent manner (*20*) and estradiol regulates HIV replication in peripheral blood mononuclear cells (PBMCs) through transcriptional mechanisms (*21*). V9, a nucleotide synthesis inhibitor alters transcription. The antihistamines (V11 and V13) inhibit CCL11 and CCL5. CCL11 is a ligand for CCR2, CCR3, and CCR5 (CCR2 binds Tat). CCL5 or RANTES is a beta-chemokine known to suppress HIV LTR transcription and is produced by CD4+ and CD8+ T-cells (*29*).

Figure S1

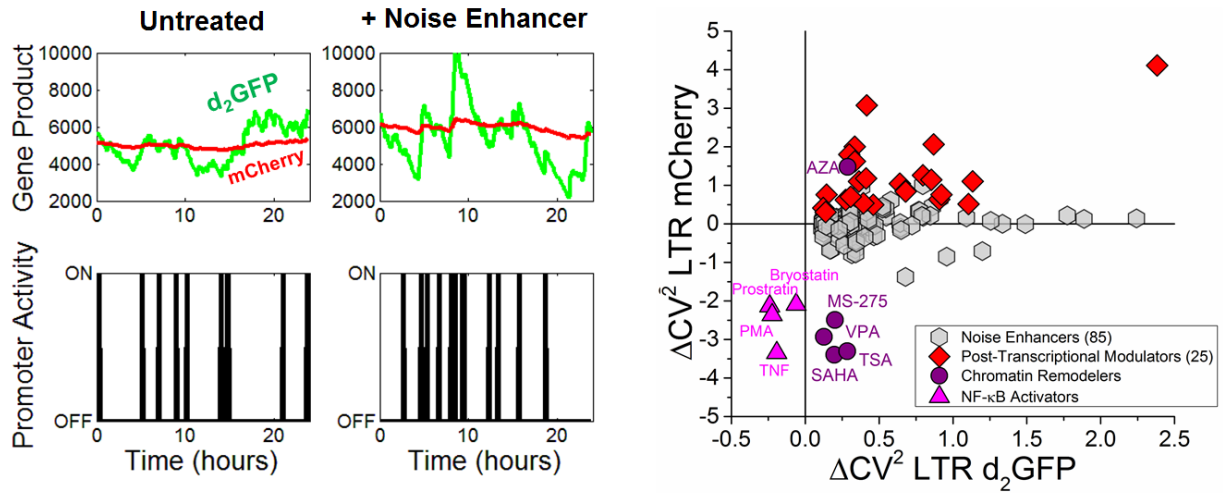


Figure S1: Identification of 25 post-transcriptional modulators using the two-reporter system. (left) Stochastic simulations showing the long-lived mCherry signal averaging out changes in underlying promoter activity with and without a noise-enhancing compound. Conversely, the d2GFP signal is capable of tracking modulations of the underlying episodic transcription. (right) Noise deviations of the two reporters from the untreated origin show minimal mCherry noise deviations for the 85 selected noise enhancers (gray hexagons). 25 compounds with deviations of >2σ in both d2GFP and mCherry noise (red diamonds) were excluded from the latency reactivation assay and are made up of post-transcriptional modifiers (see section above). Treatment of the two-reporter system with known transcriptional activators (HDACi, PKC agonists, AZA) all increased mean mCherry resulting in lower mCherry noise (expected from theory and mCherry scatter sections above). As expected the activators of NF-Kb decrease both mCherry and d2GFP noise (pink triangles) while chromatin remodelers decrease mCherry noise and increase d2GFP noise (purple circles) in agreement with Fig. 2C. The increased d2GFP noise can be directly attributed to modulation of transcriptional bursts.

Figure S2

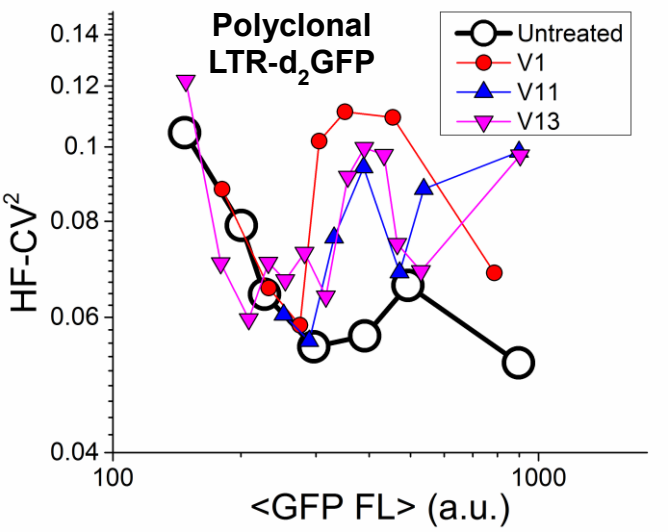


Figure S2: Noise enhancers increase expression noise across 100's of different HIV integration sites. Demonstration that noise enhancers increase noise across hundreds of integration sites for three different noise-enhancer compounds. An LTR d2GFP polyclonal population was imaged by single-cell time-lapse fluorescence microscopy (*2*, *30*). Each data point represents a subcluster of ~90 unique single cells. Each cell is tracked and quantified for GFP fluorescence for 12 hours after a 24 hour pretreatment with the compound. Fluorescent trajectories are high-frequency noise processed by detrending their general deterministic behavior (based on all cells of the population) in addition to mean suppression of individual cell trajectories to focus on the intrinsic noise fluctuations (*2*, *31*). The three noise enhancers in the above experiment display increased high-frequency noise magnitude (HF-CV^2) across ~500 integration sites compared to the untreated polyclonal population (black circles) showing that the detected noise enhancer hits are not specific to the isoclone cell line used in the drug screen. See Supporting Movie S1.

Figure S3

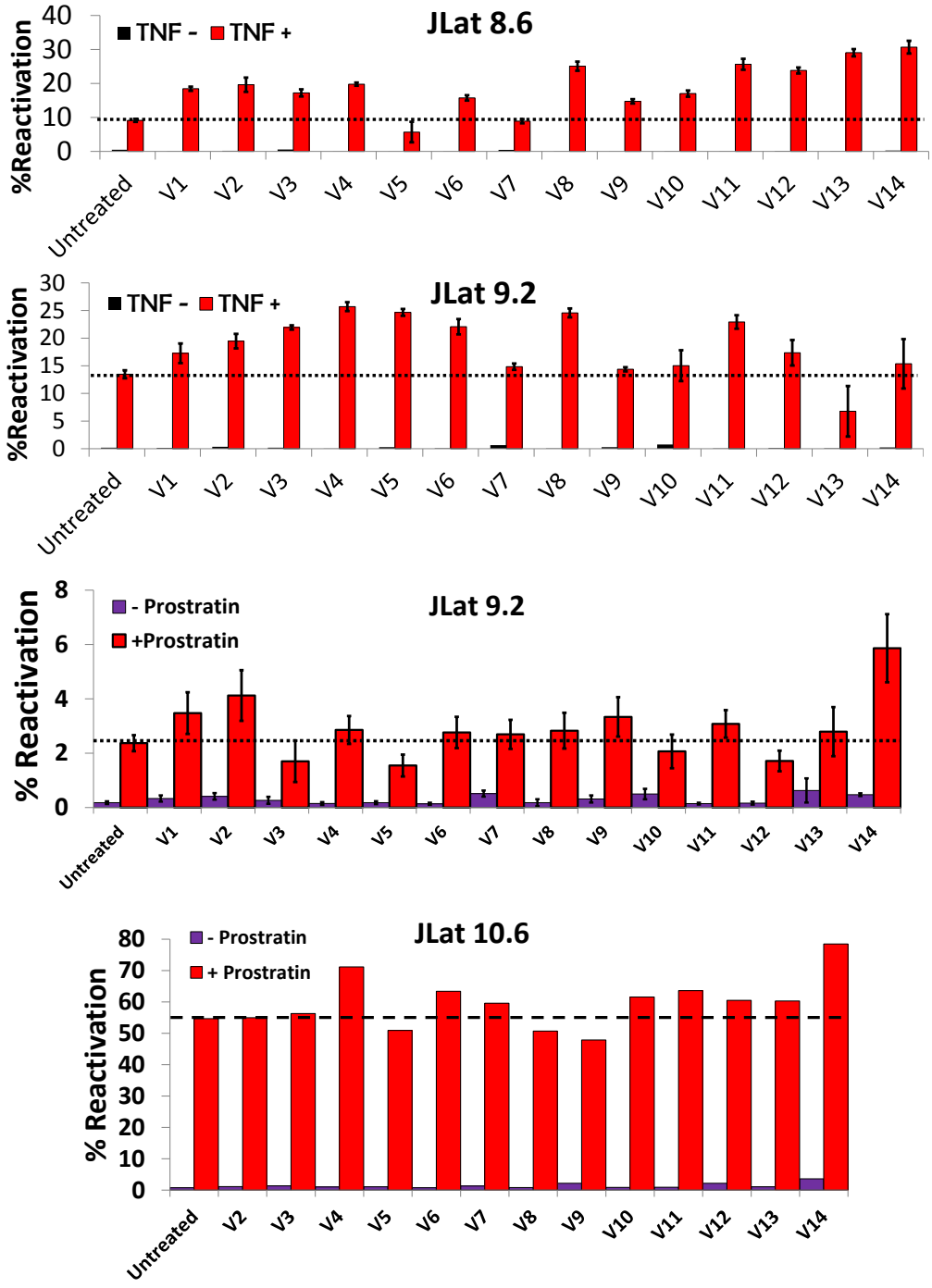


Figure S3: Noise enhancers synergize reactivation in alternate JLat cell lines. A subset of 14 noise enhancers were added with TNF or Prostratin to JLat 9.2 and 10.6 for 24 hours and yielded similar synergies as JLat 8.6 in the main text. Upper three panels quantify the mean and standard error for 4-6 repeated measurements for each drug combination. The bottom panel is a single measurement of a highly activatable JLat cell line.

Figure S4

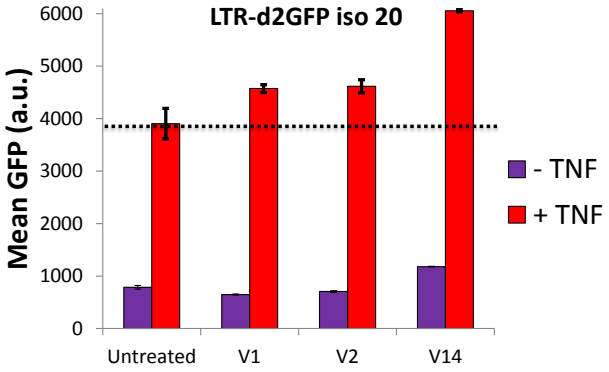


Figure S4: Mean expression levels increase in LTR-d2GFP expression when treated with noise enhancers and activators in combination. Mean GFP levels are significantly increased when noise enhancer compounds are added with TNF to LTR-d2GFP isoclone 20 cells, in agreement with theory (Fig. 2B).

Figure S5

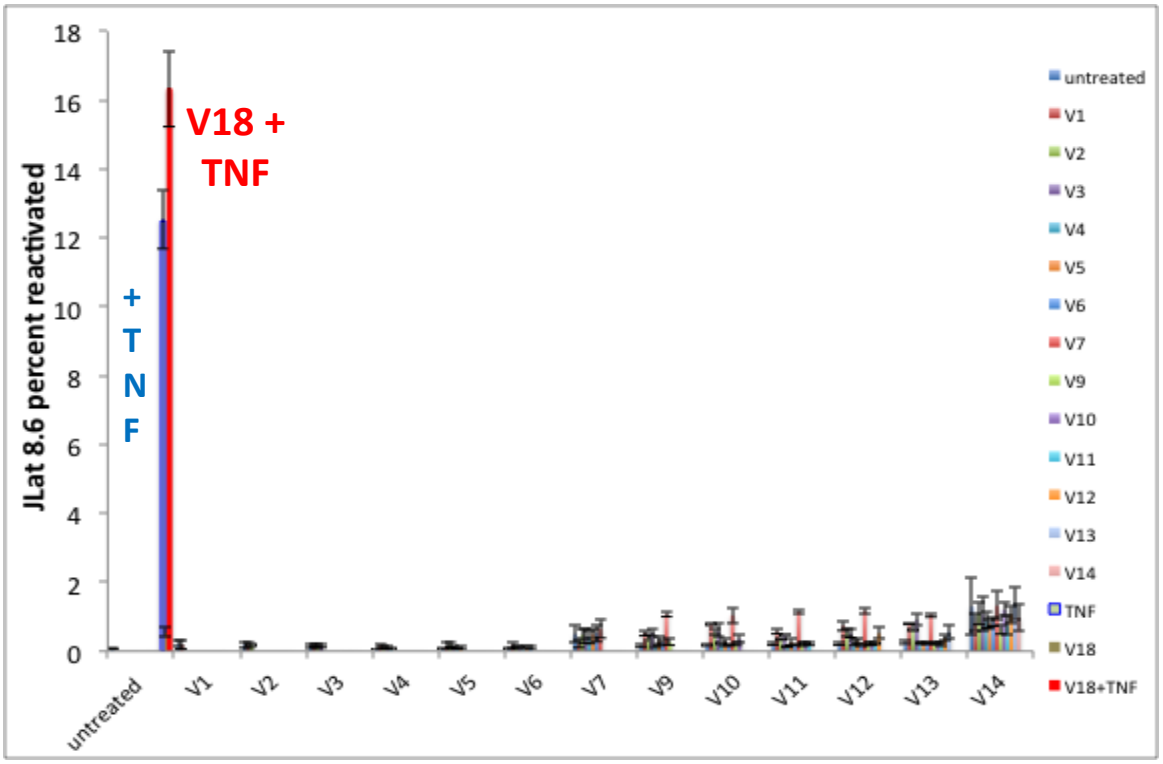


Figure S5: Combination treatments of two noise enhancers have no effect on latent reactivation. The JLat 8.6 latent cell line was subject to combinations of a subset of noise enhancer compounds for 24 h. For comparison a TNF control and TNF + V18 which are shown to synergize in Figure 3A are shown to the left. All measurments performed in triplicate.

Figure S6

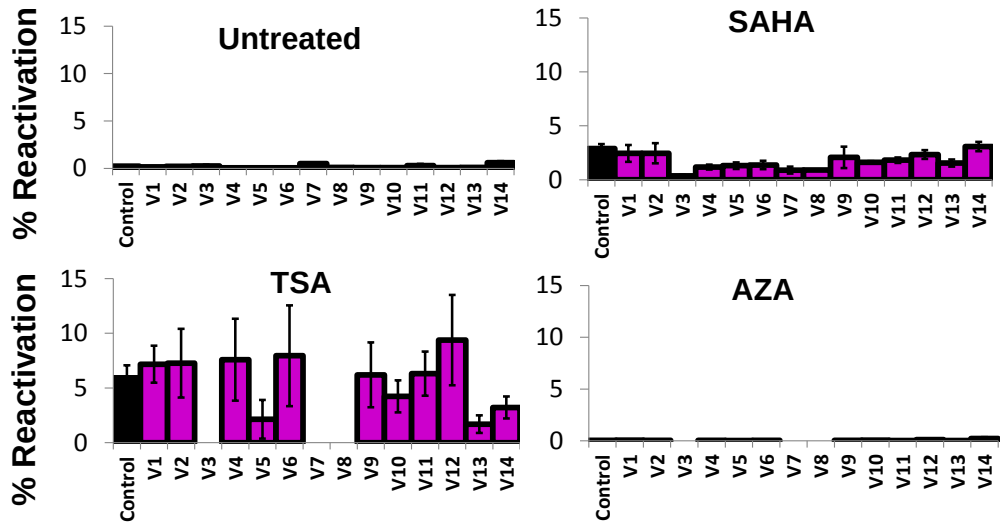


Figure S6: Noise enhancers show no significant reactivation in combination with conventional drugs that are also noise enhancers. A subset of 14 noise enhancers were tested in combination with a variety of known transcriptional modulators shown to be noise enhancers in the main text figure 2. Consistent with the proposed theory the noise enhancers do not increase reactivation with SAHA, TSA, or AZA, which are mixtures of partial activator with noise enhancement.

Figure S7

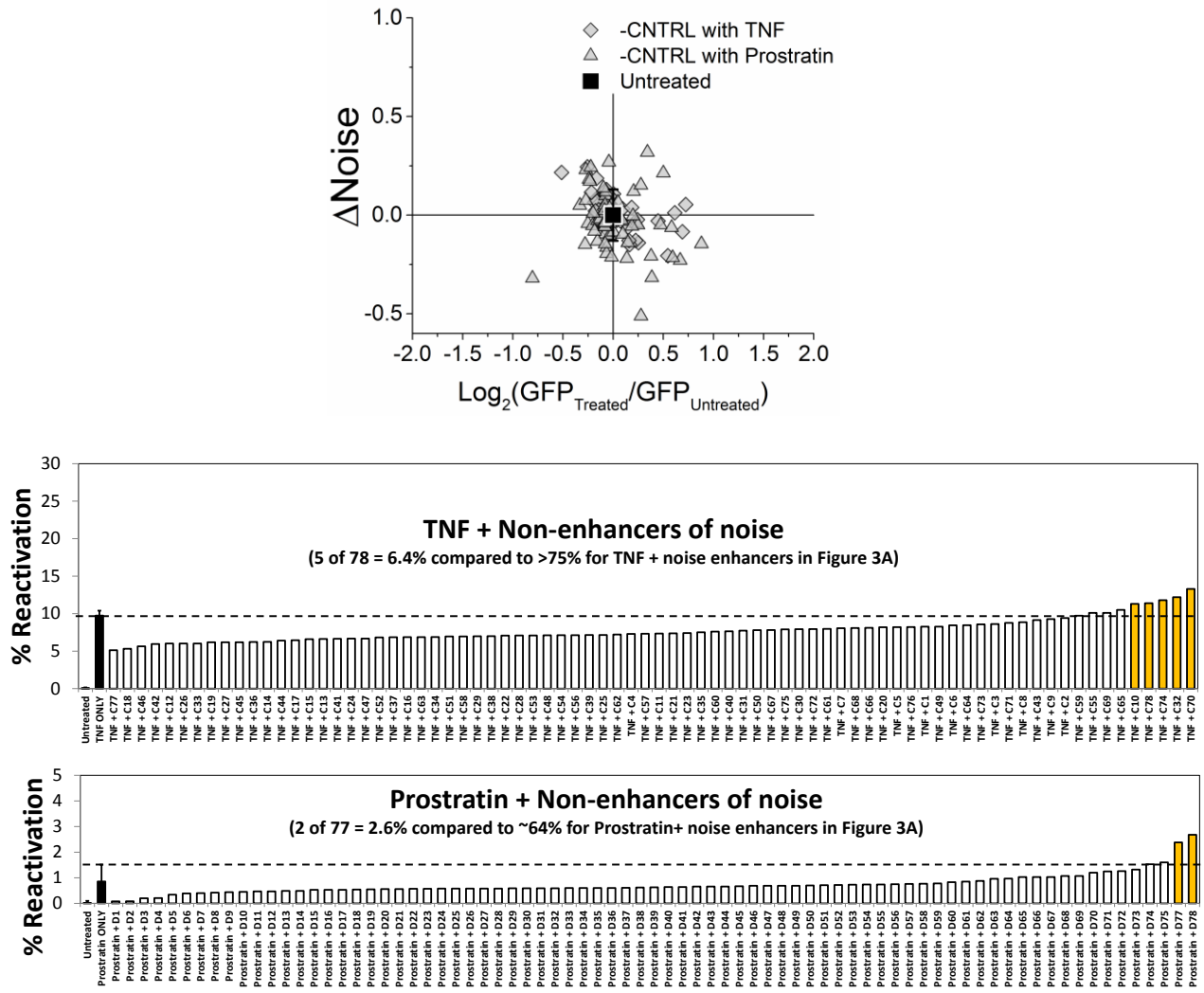


Figure S7: TNF and Prostratin combined with non-enhancers of noise do not increase reactivation. (Upper) Noise vectors for 160 diverse compounds that do not fall within the noise enhancement region of +2σ in the screen (i.e. "non-enhancers of noise"). Two plates containing full-length HIV J-Lat 8.6 were prepared containing 80 compounds combined with either TNF or Prostratin along with positive and untreated controls. (Lower) Results post 24 hours for JLat 8.6 treated with non-enhancers of noise combined with either TNF or Prostratin. For both activators, both the % of compounds that synergize reactivation (yellow) and the level of increased reactivation are significantly lower for the non-enhancers of noise compared to the noise enhancer combinations in Fig. 3A (P-Value < 0.00001 for TNF).

Figure S8

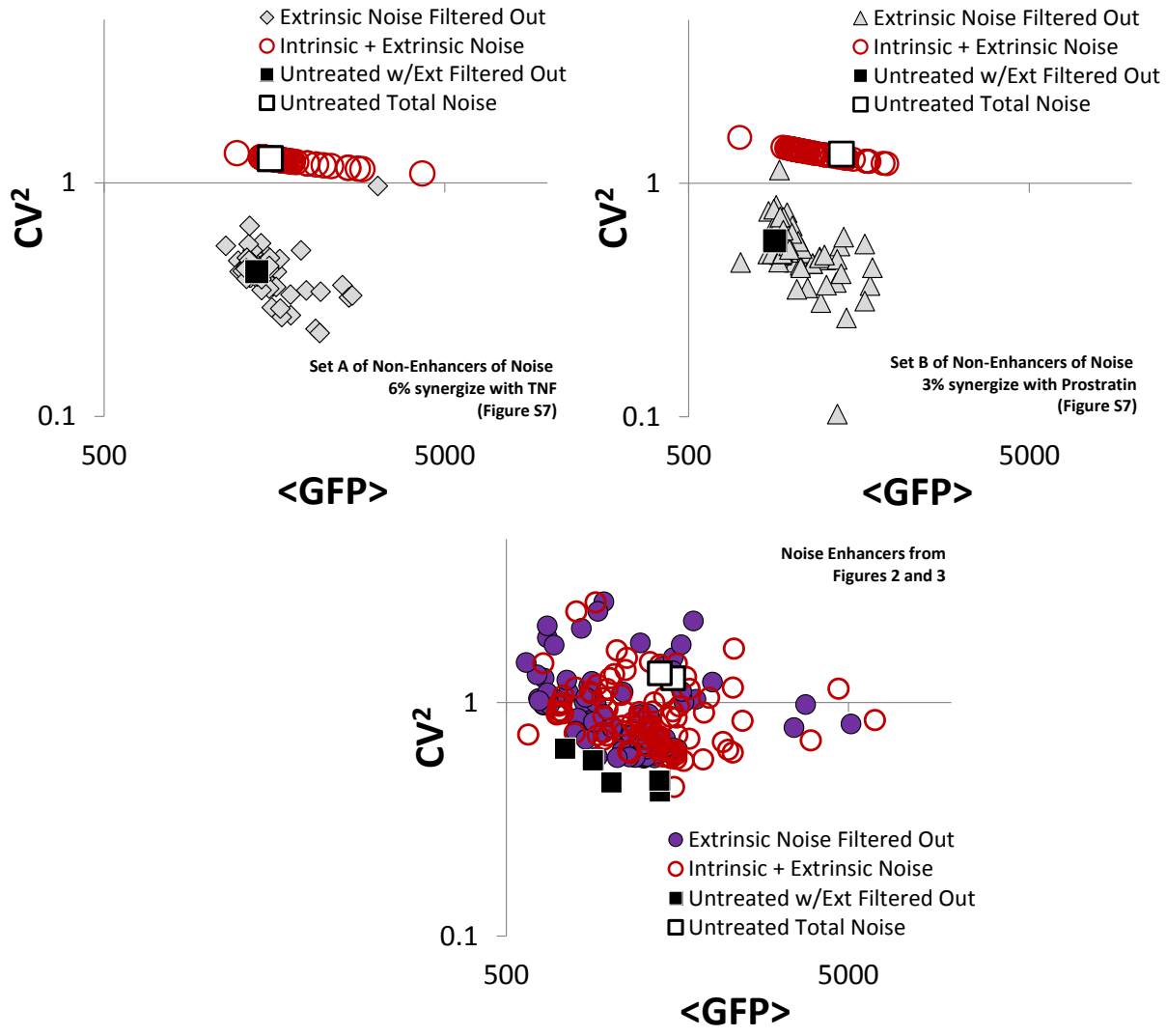


Figure S8: Extrinsic variability cannot predict synergistic effects. (Upper) Noise versus mean plots for the non-enhancers of noise in Figure S7. Noise was processed both with (grey diamonds) and without (red circles) conservative gating on the single-cell forward versus side scatter data. The conservative gating (see Methods above) clearly filters extrinsic cellular noise that would otherwise dominate the noise trends at a constant value (as seen by the red circles). The untreated control follows the same shift in total noise when the conservative gating is relieved (squares). (Lower) Noise verses mean GFP for the noise enhancers identified and assayed in figures 2 and 3. Extrinsic noise filtering was used in the original processing (purple circles) and non-filtered noise accounting for the whole live cell population reveals a similar scatter suggesting that intrinsic noise contributions are large enough that total noise is not dominated by extrinsic noise contributions (red circles).

Figure S9

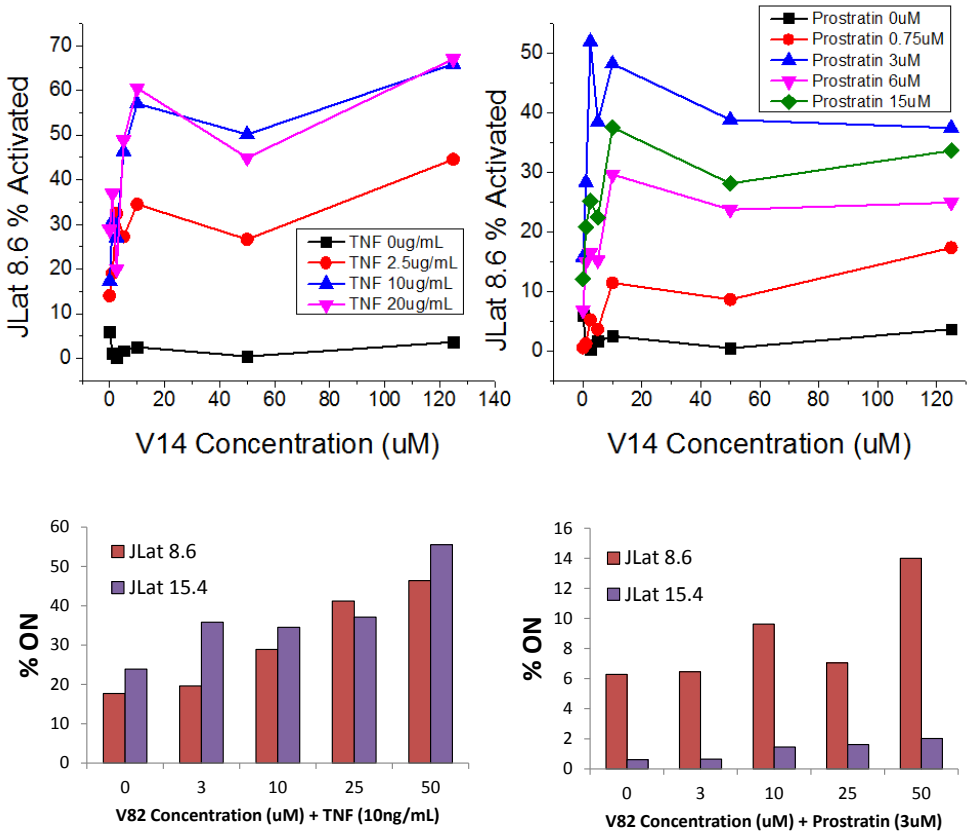


Figure S9: V14 and V82 dose response curves with TNF and Prostratin. (upper) V14 dose response curves with TNF and Prostratin at 24 and 48 hours. Dose response curves show peak activation at the original V14 concentration of 10uM treatment used. (lower) V82 treatement for 48 hours of two JLat cell lines shows maximal reactivation at 50uM.

Figure S10

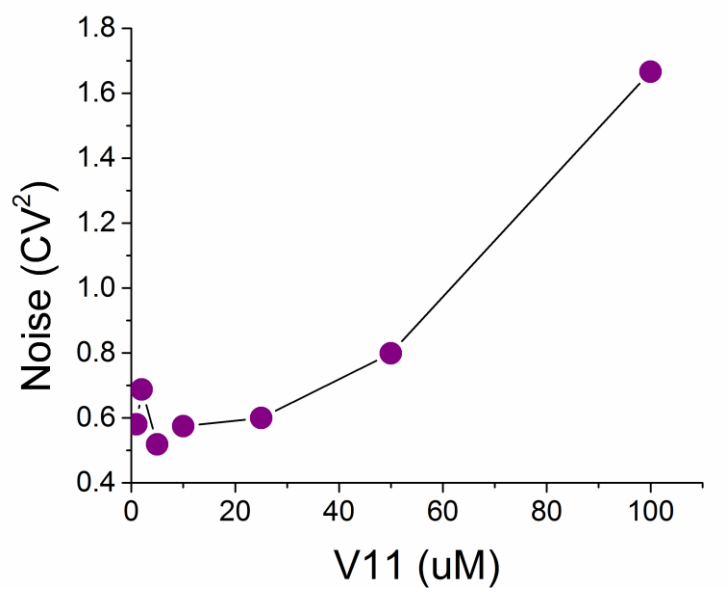


Figure S10: Noise enhancement of the LTR promoter increases with V11 concentration. Treatment of the drug screen LTR-d2GFP cell line with increasing V11 concentration yields an increasing trend in gene-expression noise. This result is consistent with Figs. 3 and 4, which show a correlation between noise enhancement and reactivation of latency.

Figure S11

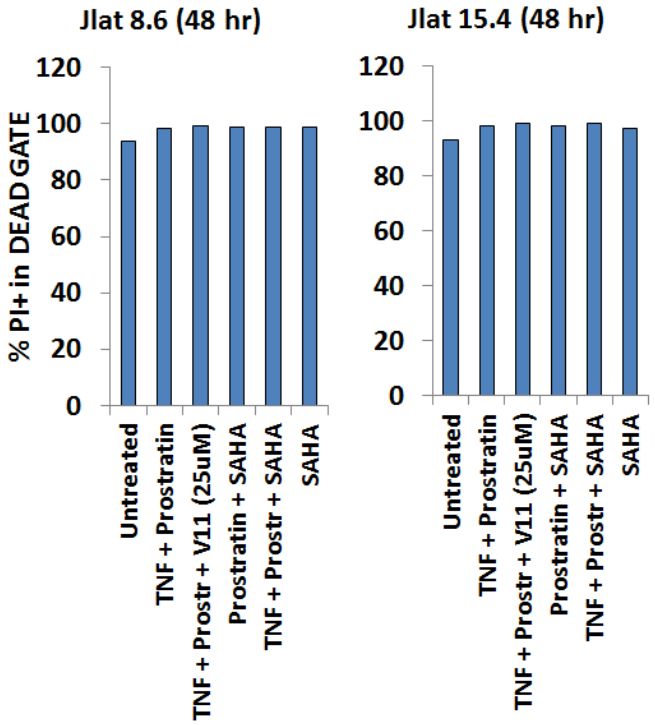


Figure S11: Propidium Iodide (PI) control stain confirms that flow cytometry SSC faithfully measures cell death. The PI stain was applied post 48 hr drug treatment and shows that cells outside of the LIVE gate determined by the forward versus side scatter (SSC) cytometry data are dead by PI stain.

Online Files:

Table S1: (Excel file) Includes the 85 Noise Enhancers with CompoundID, name, ΔFL, and ΔCV^2, for both d2GFP and mCherry reporters. In addition, the %Reactivation with TNF and Prostratin are included (values from Figure 3A).

Movie S1: (AVI file) Time-lapse fluorescence microscopy movie of LTR-d2GFP polyclonal cells for 20 hours, 24 hours post-treatment with noise enhancer V1. The movie plays on a loop and displays 20 unique x-y locations sampled over time in a 4x5 array. Tracked single-cells are marked with a red dot for quality control and post-processing (*2*, *31*).

Table S2: Table of compounds commonly referenced in Figs. 3-4 and Supporting Information.

NE ID #	Compound	Description
V1	docetaxel	Anti-mitotic chemotherapy drug, **microtubule inhibitor**
V2	ethinyl estradiol	**Evidence for Inhibition, Activation, and Binding** of Estrogen Receptor 1 (ESR1)
V3	estramustine	**Microtubule Depolarization** via binding of tubulin and associated proteins
V4	felbinac	Antiinflammatory
V6	bezafibrate	Antilipemic agent that lowers cholesterol and triglycerides.
V7	mebendazole	**Microtuble Inhibitor**
V8	thiamylal sodium	
V9	mercaptopurine	Purine nucleotide synthesis inhibitor and **alters DNA/RNA synthesis**
V10	dutasteride	**Inhibits** the conversion of testosterone into dihydrotestosterone (DHT), **an estrogen antagonist**
V11	cetirizine hydrochloride	**Anti-histamine. Inhibitor of CCL11.**
V12	acetophenazine maleate	Inhibitor of Dopamine Receptor D1 and D2
V13	oxytetracycline	**Antibiotic, binds ribosomes and modulator of translational. Inhibitor of CCL5.**
V14	artemisinin	Antimalarial. Inhibitor of CDK2 (control of cell cycle) and SP1
Literature		
Reported NEs	JQ1	Bromodomain Protein Inhibitor
	AZA	DNA Methyltransferase Inhibitor
	TSA	Histone Deacytalase Inhibitor (HDACi)
	SAHA	Histone Deacytalase Inhibitor (HDACi)
	VPA	Histone Deacytalase Inhibitor (HDACi)
	MS-275	Histone Deacytalase Inhibitor (HDACi)
Activators	TNF	Activator of NF-Kb
	Prostratin	PKC Agonist, Activator of NF-Kb
	PMA	PKC Agonist, Activator of NF-Kb

References

1. A. Singh, B. Razooky, C. D. Cox, M. L. Simpson, L. S. Weinberger, Transcriptional bursting from the HIV-1 promoter is a significant source of stochastic noise in HIV-1 gene expression. *Biophys. J.* **98**, L32–L34 (2010). Medline doi:10.1016/j.bpj.2010.03.001

2. R. D. Dar, B. S. Razooky, A. Singh, T. V. Trimeloni, J. M. McCollum, C. D. Cox, M. L. Simpson, L. S. Weinberger, Transcriptional burst frequency and burst size are equally modulated across the human genome. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17454–17459 (2012). Medline doi:10.1073/pnas.1213530109

3. A. Jordan, D. Bisgrove, E. Verdin, HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *EMBO J.* **22**, 1868–1877 (2003). Medline doi:10.1093/emboj/cdg188

4. H. C. Yang, S. Xing, L. Shan, K. O'Connell, J. Dinoso, A. Shen, Y. Zhou, C. K. Shrum, Y. Han, J. O. Liu, H. Zhang, J. B. Margolick, R. F. Siliciano, Small-molecule screening using a human primary cell model of HIV latency identifies compounds that reverse latency without cellular activation. *J. Clin. Invest.* **119**, 3473–3486 (2009). Medline

5. J. R. S. Newman, S. Ghaemmaghami, J. Ihmels, D. K. Breslow, M. Noble, J. L. DeRisi, J. S. Weissman, Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* **441**, 840–846 (2006). Medline doi:10.1038/nature04785

6. C. I. Bliss, The toxicity of poisons applied jointly. *Ann. Appl. Biol.* **26**, 585–615 (1939). doi:10.1111/j.1744-7348.1939.tb06990.x

7. D. T. Gillespie, Exact stochastic simulation of coupled chemical-reactions. *J. Phys. Chem.* **81**, 2340–2361 (1977). doi:10.1021/j100540a008

8. M. L. Simpson, C. D. Cox, G. S. Sayler, Frequency domain chemical Langevin analysis of stochasticity in gene transcriptional regulation. *J. Theor. Biol.* **229**, 383–394 (2004). Medline doi:10.1016/j.jtbi.2004.04.017

9. C. D. Cox, J. M. McCollum, M. S. Allen, R. D. Dar, M. L. Simpson, Using noise to probe and characterize gene circuits. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 10809–10814 (2008). Medline doi:10.1073/pnas.0804829105

10. J. H. Zhang, T. D. Chung, K. R. Oldenburg, A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screen.* **4**, 67–73 (1999). Medline doi:10.1177/108705719900400206

11. R. Skupsky, J. C. Burnett, J. E. Foley, D. V. Schaffer, A. P. Arkin, HIV promoter integration site primarily modulates transcriptional burst size rather than frequency. *PLOS Comput. Biol.* **6**, e1000952 (2010). Medline doi:10.1371/journal.pcbi.1000952

12. R. D. Dar, D. K. Karig, J. F. Cooke, C. D. Cox, M. L. Simpson, Distribution and regulation of stochasticity and plasticity in *Saccharomyces cerevisiae*. *Chaos* **20**, 037106 (2010). Medline doi:10.1063/1.3486800

13. M. B. Elowitz, A. J. Levine, E. D. Siggia, P. S. Swain, Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002). Medline doi:10.1126/science.1070919

14. L. Duan, I. Ozaki, J. W. Oakes, J. P. Taylor, K. Khalili, R. J. Pomerantz, The tumor suppressor protein p53 strongly alters human immunodeficiency virus type 1 replication. *J. Virol.* **68**, 4302–4313 (1994). Medline

15. J. Bargonetti, A. Chicas, D. White, C. Prives, p53 represses Sp1 DNA binding and HIV-LTR directed transcription. *Cell Mol Biol (Noisy-le-grand)* **43**, 935–949 (1997). Medline

16. E. J. Duh, W. J. Maury, T. M. Folks, A. S. Fauci, A. B. Rabson, Tumor necrosis factor alpha activates human immunodeficiency virus type 1 through induction of nuclear factor binding to the NF-kappa B sites in the long terminal repeat. *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5974–5978 (1989). Medline doi:10.1073/pnas.86.15.5974

17. G. Poli, P. Bressler, A. Kinter, E. Duh, W. C. Timmer, A. Rabson, J. S. Justement, S. Stanley, A. S. Fauci, Interleukin 6 induces human immunodeficiency virus expression in infected monocytic cells alone and in synergy with tumor necrosis factor alpha by transcriptional and post-transcriptional mechanisms. *J. Exp. Med.* **172**, 151–158 (1990). Medline doi:10.1084/jem.172.1.151

18. C. Van Lint, C. A. Amella, S. Emiliani, M. John, T. Jie, E. Verdin, Transcription factor binding sites downstream of the human immunodeficiency virus type 1 transcription start site are important for virus infectivity. *J. Virol.* **71**, 6113–6127 (1997). Medline

19. A. el Kharroubi, E. Verdin, Protein-DNA interactions within DNase I-hypersensitive sites located downstream of the HIV-1 promoter. *J. Biol. Chem.* **269**, 19916–19924 (1994). Medline

20. D. Katagiri, H. Hayashi, A. F. Victoriano, T. Okamoto, K. Onozaki, Estrogen stimulates transcription of human immunodeficiency virus type 1 (HIV-1). *Int. Immunopharmacol.* **6**, 170–181 (2006). Medline doi:10.1016/j.intimp.2005.07.017

21. S. N. Asin, A. M. Heimberg, S. K. Eszterhas, C. Rollenhagen, A. L. Howell, Estradiol and progesterone regulate HIV type 1 replication in peripheral blood cells. *AIDS Res. Hum. Retroviruses* **24**, 701–716 (2008). Medline doi:10.1089/aid.2007.0108

22. M. R. Fernandez-Fernandez, D. B. Veprintsev, A. R. Fersht, Proteins of the S100 family regulate the oligomerization of p53 tumor suppressor. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 4735–4740 (2005). Medline doi:10.1073/pnas.0501459102

23. M. R. Fernandez-Fernandez, T. J. Rutherford, A. R. Fersht, Members of the S100 family bind p53 in two distinct ways. *Protein Sci.* **17**, 1663–1670 (2008). Medline doi:10.1110/ps.035527.108

24. Y. Liu, D. C. Buck, K. A. Neve, Novel interaction of the dopamine D2 receptor and the Ca^{2+} binding protein S100B: Role in D2 receptor function. *Mol. Pharmacol.* **74**, 371–378 (2008). Medline doi:10.1124/mol.108.044925

25. D. P. Bednarik, J. A. Cook, P. M. Pitha, Inactivation of the HIV LTR by DNA CpG methylation: Evidence for a role in latency. *EMBO J.* **9**, 1157–1164 (1990). Medline

26. J. Blazkova, K. Trejbalova, F. Gondois-Rey, P. Halfon, P. Philibert, A. Guiguen, E. Verdin, D. Olive, C. Van Lint, J. Hejnar, I. Hirsch, CpG methylation controls reactivation of HIV from latency. *PLoS Pathog.* **5**, e1000554 (2009). Medline doi:10.1371/journal.ppat.1000554

27. J. K. Chan, D. Bhattacharyya, K. G. Lassen, D. Ruelas, W. C. Greene, Calcium/calcineurin synergizes with prostratin to promote NF-κB dependent activation of latent HIV. *PLoS ONE* **8**, e77749 (2013). Medline doi:10.1371/journal.pone.0077749

28. G. Marzio, M. Tyagi, M. I. Gutierrez, M. Giacca, HIV-1 tat transactivator recruits p300 and CREB-binding protein histone acetyltransferases to the viral promoter. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 13519–13524 (1998). Medline doi:10.1073/pnas.95.23.13519

29. J. S. Handen, H. F. Rosenberg, Suppression of HIV-1 transcription by beta-chemokines RANTES, MIP1-alpha, and MIP-1beta is not mediated by the NFAT-1 enhancer element. *FEBS Lett.* **410**, 301–302 (1997). Medline doi:10.1016/S0014-5793(97)00654-6

30. D. Boehm, V. Calvanese, R. D. Dar, S. Xing, S. Schroeder, L. Martins, K. Aull, P. C. Li, V. Planelles, J. E. Bradner, M. M. Zhou, R. F. Siliciano, L. Weinberger, E. Verdin, M. Ott, BET bromodomain-targeting compounds reactivate HIV from latency via a Tat-independent mechanism. *Cell Cycle* **12**, 452–462 (2013). Medline doi:10.4161/cc.23309

31. L. S. Weinberger, R. D. Dar, M. L. Simpson, Transient-mediated fate determination in a transcriptional circuit of HIV. *Nat. Genet.* **40**, 466–470 (2008). Medline doi:10.1038/ng.116