Cite as: R. V. Desai *et al.*, *Science* 10.1126/science.abc6506 (2021).

A DNA-repair pathway can affect transcriptional noise to promote cell fate transitions

Ravi V. Desai[1,2], Xinyue Chen[1], Benjamin Martin[1,3], Sonali Chaturvedi[1], Dong Woo Hwang[4], Weihan Li[4], Chen Yu[5], Sheng Ding[5,6], Matt Thomson[7], Robert H. Singer[4], Robert A. Coleman[4], Maike M. K. Hansen[3], Leor S. Weinberger[1,8,9]*

[1]Gladstone/UCSF Center for Cell Circuitry, Gladstone Institutes, San Francisco, CA 94158, USA. [2]Medical Scientist Training Program and Tetrad Graduate Program, University of California, San Francisco, CA 94158, USA. [3]Institute for Molecules and Materials, Radboud University, 6525 AJ Nijmegen, the Netherlands. [4]Department of Anatomy and Structural Biology, Albert Einstein College of Medicine, Bronx, NY 10461, USA. [5]Gladstone Institute of Cardiovascular Disease, Gladstone Institutes, San Francisco, CA 94158, USA. [6]School of Pharmaceutical Sciences, Tsinghua University, Beijing 100084, China. [7]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, 91125, USA. [8]Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94158, USA. [9]Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94158, USA.

*Corresponding author. Email: leor.weinberger@gladstone.ucsf.edu

Stochastic fluctuations in gene expression ('noise') are often considered detrimental, but fluctuations can also be exploited for benefit (e.g., dither). We show here that DNA base-excision repair amplifies transcriptional noise to facilitate cellular reprogramming. Specifically, the DNA-repair protein Apex1, which recognizes both naturally occurring and unnatural base modifications, amplifies expression noise while homeostatically maintaining mean-expression levels. This amplified expression noise originates from shorter duration, higher intensity, transcriptional bursts generated by Apex1-mediated DNA supercoiling. The remodeling of DNA topology first impedes and then accelerates transcription to maintain mean levels. This mechanism, which we term Discordant Transcription through Repair (DiThR; pronounced /'dither'/), potentiates cellular reprogramming and differentiation. Our study reveals a potential functional role for transcriptional fluctuations mediated by DNA base modifications in embryonic development and disease.

Fluctuations have been recognized to dynamically shape the distribution of microstates a system adopts, and modulation of fluctuations has been harnessed throughout engineering and the sciences (*1*). In chemistry, thermal fluctuations accelerate reactions (*2*); in engineering, amplification of electrical, acoustic, or mechanical fluctuations (i.e., 'dither', from the Middle English "didderen" meaning to "tremble") is used for signal recovery (*3*), and in neuroscience, electrophysiological fluctuations (*4*) are clinically amplified to improve sensorimotor function (*5*). Such dither approaches break Poisson dependency so that Δvariance ≠ Δmean.

Biological organisms may maximize fitness by harnessing putative fluctuations to enable probabilistic 'bet-hedging' decisions (*6–8*). Intrinsic molecular fluctuations in gene expression (i.e., stochastic 'noise'), modulated by gene-regulatory circuits, enables probabilistic fate selection (Fig. 1A) in diverse biological systems (*9–12*). Open questions remain as to whether cellular noise control is limited to locus-specific gene-regulatory circuits or if generalized noise-modulation mechanisms exist. Specifically, it is unclear how such generalized noise-modulation mechanisms might orthogonally tune noise independent of mean, and, given the detrimental effects of noise, if such putative mechanisms might be regulated 'on-demand' to potentiate cell-fate specification.

Non-genetic variability or noise in gene expression, often quantified by measurement of cell-to-cell variability in reporter expression, can arise from intrinsic and extrinsic sources. In mammalian cells, intrinsic noise originates from episodic transcriptional 'bursts' (*13–16*) initiated by promoter toggling between ON and OFF states (Fig. 1B). The two-state random-telegraph model describes this bursting through two parameters: (i) the fraction of time a promoter is active ($K_{ON}/[K_{ON}+K_{OFF}]$), and (ii) the number of transcripts produced during the ON state (burst size, K_{TX}/K_{OFF}) (*17*). These bursting parameters are tuned by regulatory machinery (*18*) like histone acetyltransferases, which can increase burst frequency by facilitating nucleosome clearance from promoters thereby increasing mean transcript abundance (*19*). Increases in mean expression (μ) are typically accompanied by a stereotypical reduction in noise measured by coefficient of variation, CV, (σ/μ) (Fig. 1B), whereas stressors that decrease mean are typically accompanied by an increase in noise (*20–22*). This 1/μ scaling of noise can be broken by gene-regulatory circuits such as feedback and feedforward loops (*23*), and some small-molecule pharmaceuticals can modulate transcriptional fluctuations/noise (σ/μ) independent of change in mean (μ) (*24*, *25*). Because some molecules can amplify expression noise of diverse unrelated promoters (*24*, *26*), we tested whether these molecules might function by disrupting or enhancing a putative cellular noise-control mechanism.

A series of screens (fig. S1) identified one compound, 5′-iodo-2’-deoxyuridine (IdU), which consistently increased expression noise of multiple transcriptional reporter constructs in diverse cell types. To test the generality of this noise amplification effect we used mouse embryonic stem cells (mESCs). Single-cell RNA sequencing (scRNA-seq) of mESCs cultured in pluripotency-maintenance media (2i/LIF) showed—after filtering and normalization using Seurat (*27*)—that IdU amplified cell-to-cell variability in transcript abundance (i.e., transcript noise) for virtually all genes across the genome—4,578 genes analyzed (Fig. 1C)—with little alteration in mean transcript abundance for most genes, as analyzed by either CV^2 (σ^2/μ^2) or variance (σ^2) versus mean (Fig. 1, C and D). To account for the Poisson scaling of variance on mean, transcript noise was also quantified using the Fano factor (σ^2/μ), which measures how noise deviates from Poisson scaling ($\sigma^2/\mu = 1$) (*28*, *29*). Despite mean-expression levels exhibiting minimal changes (Fig. 1E), the Fano factor increased for > 90% of genes (Fig. 1F). Long noncoding RNAs (lncRNAs) also exhibited noise enhancement and weakly expressed genes showed a slightly greater change in Fano (fig. S2). These results of a global increase in transcript noise with little change in mean abundance are in stark contrast to the effects of transcriptional activators or cellular stressors that alter noise in a stereotypic manner together with changes in mean number of transcripts (*20*, *30*).

To account for technical noise and quantify statistical significance of changes in noise and mean, we used an established Bayesian hierarchical model (BASiCS) (*31*) to create probabilistic, gene-specific estimates of both mean expression and cell-to-cell transcript variability. Of the 4,971 genes analyzed, 945 genes were classified (~20%) as highly variable, whereas 113 genes (~2%) showed a significant change in mean expression (Fig. 1, G and H). Bulk RNA-seq measurements of mean abundances confirmed the scRNA-seq findings (fig. S3). Thus, analyses from two methods (Seurat and BASiCS) show that IdU induced a significant increase in transcript variability (expression noise) but comparatively little change in mean expression.

To examine if certain characteristics could explain a gene’s potential for noise enhancement we examined (i) gene length, (ii) promoter and (iii) gene-body AT content, (iv) number of exons, (v) TATA-box inclusion and (vi) strand orientation. None of these characteristics exhibited predictive power (fig. S4). However, genes susceptible to high noise enhancement were preferentially located within the interior of topologically associated domains (TADs), suggesting gene topology influences susceptibility to noise enhancement (fig. S4). Ontology analysis of highly variable genes showed enrichment of house-keeping pathways along with pluripotency maintenance factors, particularly Sox2, Oct4, Nanog and Klf4 (Fig. 1H and fig. S5). As these pluripotency maintenance factors are key influencers of cell-fate specification, we focused on the molecular mechanisms driving their amplified transcript noise.

We tested whether the enhanced variability arose from extrinsic factors, which include cell-cycle phase and cell-type identity (*32*). Cells within the scRNA-seq dataset were computationally assigned a cycle stage (G1, S, G2/M) (*33*) which showed that Nanog, Oct4, Sox2 and Klf4 were highly variable in each cell-cycle phase, indicating that their variability is not cell-cycle dependent (fig. S6). Moreover, pseudo-time analysis showed no bifurcations, indicating transcriptional variability was not due to a differentiation-induced mixture of cell-types (fig. S7).

Extrinsic variability may also arise from the coordinated propagation of noise through gene-regulatory networks (*34*) and can be measured by gene-to-gene correlation matrices (*35*, *36*). If the increase in global transcript noise is extrinsic, expression correlation between network partners would increase or remain unchanged. Analysis of gene-to-gene correlation matrices showed that ~80% of gene-gene pairs lost correlation strength after IdU treatment (Fig. 2A and fig. S8), indicating that enhanced expression noise is uncorrelated and not consistent with an extrinsic noise source. Exclusion of these extrinsic noise sources indicates that IdU must amplify intrinsic noise arising from stochastic fluctuations in either transcript production (promoter toggling) or degradation.

To test whether a change in promoter toggling could account for IdU-enhanced noise, we used single-molecule RNA FISH (smRNA-FISH) to count both nascent and mature transcripts of Nanog, a master regulator of pluripotency. Transcript counting was performed on a mESC line in which both endogenous alleles of Nanog are fused to eGFP at the C terminus. This fusion does not alter mRNA or protein half-life or impair differentiation potential (*37*). smRNA-FISH probes to eGFP (3` end of transcript) were used to count mature transcripts, while probes to the first intron of Nanog (5` end of transcript) were used to identify active transcriptional centers (TCs) and explicitly measure the number of mRNAs actively transcribed at the start of the gene. To minimize extrinsic noise, downstream analyses were limited to cells of similar size (fig. S9A). Consistent with scRNA-seq, smRNA-FISH showed a large increase in cell-to-cell variability of mature Nanog transcript abundance (~2-fold increase in Fano) with little change in mean abundance (Fig. 2B). Quantification of nascent Nanog transcript abundances using either intron or exon probes showed that fewer cells possessed active transcriptional centers (TCs) in the presence of IdU (Fig. 2C), whereas the number of nascent (i.e., unspliced) and mature (i.e., spliced) mRNAs at each TC was increased (Fig. 2D and fig. S9, B and C). Fitting of the two-state random-telegraph model to smRNA-FISH data revealed that increased

variability resulted from shortened burst duration (increased k_{OFF}) and amplified transcription rate (higher k_{TX}) (fig. S9D and table S2). To directly visualize the effect of IdU on burst duration, we performed live-cell imaging of transcription using p21-MS2 reporter cells (*38*). IdU generated shorter transcriptional bursts (increased k_{OFF}) while total mRNA output remained unchanged (fig. S10), further validating the reciprocal changes in burst kinetics seen with smRNA-FISH data. These results validate previous predictions (*24*, *29*) that enhanced noise could arise from reciprocal changes in transcriptional burst duration ($1/k_{OFF}$) and intensity (k_{TX}).

While longer polymerase dwell times or slowed polymerase elongation could be alternate hypotheses for the increase in nascent RNA detected by smFISH, these hypotheses are inconsistent with the simultaneous shortening of burst duration and maintenance of transcript output observed with MS2 imaging. The slowed/stalled polymerase hypothesis is also not consistent with the equivalent increase in both intron and exon probe intensities at TCs (fig. S9E). These data instead suggest that IdU-mediated increase in TC intensities result from amplified transcription rates (k_{TX}).

To test if enhanced transcript variability transmitted to protein abundances, we performed flow-cytometric analysis of Nanog-GFP reporter protein. In IdU-treated cells, the Nanog protein Fano factor increased by ~3-fold, with little change in mean, indicating that mRNA variability from altered promoter toggling indeed resulted in changes to protein noise (Fig. 2E). The increase in protein noise showed no dependency on cell-cycle (fig. S11, C and D) despite G1-to-S cell-cycle progression being slightly slowed by IdU treatment (fig. S11, A and B). Consistent with the extrinsic noise analysis above, there was no evidence of aneuploidy after IdU treatment (fig. S11A), precluding the possibility that increased noise results from a sub-population of cells with non-physiologic gene-copy numbers. Inhibition of transcription with Actinomycin D completely abrogated IdU enhancement of Nanog-GFP noise (fig. S12, A and B), indicating that IdU minimally perturbs post-transcriptional sources of gene-expression variability (e.g., mRNA degradation, mRNA translation, protein degradation).

When cultured in 2i/LIF, Nanog protein expression is unimodal and high, but when cultured in serum/LIF, mESCs exhibit bimodal expression with both a high Nanog state and low Nanog state that predisposes a cell toward differentiation (fig. S12C) (*39*). Given that IdU-induced amplification of Nanog variability arose from an intrinsic source of noise (i.e., changes in transcriptional bursting), we next tested a previous theoretical prediction that increased transcriptional noise would drive greater excursions from the high Nanog state into the low state (*40*). IdU treatment did indeed generate greater excursions into the low Nanog state for mESCs cultured in serum/LIF (fig. S12C), verifying theoretical predictions. This result demonstrated how modulation of transcriptional bursting can drive Nanog state-switching.

To verify that enhanced noise is not a population-level phenomenon brought on by differential responses to IdU in distinct cellular subpopulations (i.e., verify 'ergodicity' and that individual cells exhibit increased fluctuations), we used live-cell time-lapse imaging to quantify both the magnitude (intrinsic-CV^2) and frequency content (1/half-autocorrelation time) of Nanog-GFP fluctuations. Single-cell tracking of individual cells showed that IdU induced a 2-fold increase in the magnitude (intrinsic-CV^2) of fluctuations (Fig. 2F and fig. S13A), and auto-correlation analysis of detrended trajectories showed a broadening of the frequency distribution to higher spectra, indicating reduced stability (increased lability) of protein expression levels (fig. S13B). These higher frequency fluctuations are consistent with amplification of a non-genetic, intrinsic source of noise (*41*, *42*) because genetic sources of cellular heterogeneity, such as promoter mutations, would lead to longer retention of protein states (decreased lability) (*43*). *In silico* sorting of cells based on starting Nanog expression verified that noise enhancement was not dependent on the initial state of expression (fig. S14). Fluctuations in promoter toggling therefore drive individual cells to dynamically explore a larger state-space of Nanog expression. To further validate that IdU perturbs an intrinsic source of noise, we used a mESC line in which the two endogenous alleles of Sox2 are tagged with distinct fluorophores, which enables quantification of the intrinsic and extrinsic components of noise. Treatment with IdU increased Sox2 intrinsic noise greater than 2-fold across all expression levels (fig. S15) further validating that IdU enhances intrinsic noise.

To pinpoint the molecular mechanism, 14 nucleoside analogs (table S3) were screened for noise enhancement effects. 5′-bromo-2'-deoxyuridine (BrdU), 5-hydroxymethylcytosine (hmC), and 5-hydroxymethyluridine (hmU) also increased Nanog Fano factor to varying degrees (Fig. 3A). hmU and hmC are naturally produced by the Ten Eleven Translocation (Tet) family of enzymes during oxidation of thymine and methylated cytosine respectively (*44*, *45*). Given that these base modifications are removed through base-excision repair (BER), we surmised that their incorporation and removal from genomic DNA may cause noise enhancement (Fig. 3B) (*46*, *47*). To test this, we suppressed expression of 25 genes involved in nucleoside metabolism and DNA repair using CRISPRi (3 gRNAs/gene; table S4), and quantified how this affected IdU's noise enhancement. We identified 2 genes: AP Endonuclease 1 (Apex1) and thymidine kinase 1 (Tk1) whose depletion abrogated noise enhancement (Fig. 3C). Gene depletion was confirmed by RT-qPCR (fig. S16).

Tk1 adds a requisite gamma-phosphate group to diphosphate nucleotides prior to genomic incorporation (Fig. 3B) (*48*). Our results indicated that phosphorylation of IdU by

Tk1 and subsequent incorporation of phosphorylated IdU into the genome may be necessary for noise enhancement. As validation, combination of 10μM IdU with increasing amounts of thymidine, a competitive substrate of Tk1, returned Nanog noise to baseline levels (fig. S17A), indicating that noise enhancement is dose-dependent on IdU incorporation. scRNA-seq analysis also showed that cells in S/G2 cell-cycle phases—when levels of IdU incorporation are highest—display increased levels of transcriptional noise-enhancement (fig. S6). The reduction in Nanog noise with addition of exogenous thymidine indicates that IdU-induced noise amplification is not a generic effect of nucleotide imbalances (i.e., excess pyrimidine bases) within the cell.

Apex1 (a.k.a., Ref-1, Ape1) has a pivotal role in the BER pathway as it incises DNA at apurinic and apyrimidinic (AP) sites through its endonuclease domain, allowing for subsequent removal of the sugar backbone and patching of the gap (*49*, *50*). Chromatin immunoprecipitation (ChIP) confirmed that IdU treatment increased Apex1 recruitment to the Nanog promoter (fig. S17B). To test if alternate activators of BER also enhanced noise, we subjected cells to oxidative stress (H_2O_2) and alkylation damage stress (MMS). Similar to IdU, H_2O_2 and MMS also enhanced gene-expression noise without altering mean level of expression (fig. S18, A and B). In contrast, cells subjected to UV radiation (an activator of nucleotide excision repair, which shuts off global transcription), exhibited decreases in both mean and noise (Fano factor) (fig. S18, C and D), markedly differing from BER-mediated noise enhancement. These results further demonstrate the unique ability of BER to modulate gene-expression noise.

Since BER is initiated by a family of DNA glycosylases that recognize and excise modified bases to create AP sites, we tested if perturbation of glycosylases affected gene-expression noise. Individual depletion of either Uracil-DNA glycosylase (Ung) or Thymine-DNA glycosylase (Tdg) failed to ablate IdU noise enhancement (fig. S19A), presumably due to the overlapping and compensatory action of glycosylase family members in base removal (*51*, *52*). However, overexpression of either Uracil-DNA glycosylase (Ung) or Methylpurine-DNA glycosylase (Mpg) alone increased Nanog expression noise in the absence of IdU (fig. S19, B and C). These data suggest that noise-without-mean amplification is an inherent property of BER that occurs for endogenous modifications of both purine and pyrimidine bases.

To further confirm that Apex1 is necessary for noise-enhancement, we attempted to inactivate (knockout) Apex1 in mESCs, but this was lethal. As an alternative, we used a small-molecule, catalytic inhibitor (CRT0044876) specific for the Apex1 endonuclease domain (*53*). Contrary to the effect of Apex1 depletion, the combination of CRT0044876 with IdU synergistically increased Nanog expression noise, without significantly changing the mean (Fig. 3D).

The contrasting effects of Apex1 depletion and catalytic inhibition implied that physical binding rather than enzymatic activity of the protein modulates transcriptional bursting. Apex1 is known to induce helical distortions and local supercoiling to identify mismatched bases (*54*, *55*) and catalytically inactive (CI) Apex1 mutants bind DNA with higher affinity (*56*). This suggests that CRT0044876 may lengthen Apex1's residence time on DNA thus amplifying topological reformations. We verified that inhibition of Apex1 endonuclease activity with CRT0044876 did not inhibit IdU-mediated enhancement of Apex1 recruitment to the Nanog promoter (fig. S17B). To further test if Apex1 binding rather than enzymatic activity was responsible for noise enhancement, we expressed a catalytically inactive (CI) mutant of Apex1 (*56*) in cells that had endogenous Apex1 depleted. We found that CI-Apex1 partially rescued IdU-mediated noise enhancement (Fig. 3E). These data together with evidence that supercoiling sets mechanical bounds on transcriptional bursting (*57*, *58*), pushed us to test if Apex1 recruitment impacts supercoiling.

To measure supercoiling levels, we used a psoralen-crosslinking assay in which mESCs are incubated with biotinylated-trimethylpsoralen (bTMP), which preferentially intercalates into negatively supercoiled DNA (*59*). To eliminate DNA replication as a contributor of supercoiling, aphidicolin is added to inhibit DNA polymerases prior to bTMP incubation (*60*). IdU treatment significantly increased genomic supercoiling as demonstrated by a ~2-fold increase in bTMP intercalation (Fig. 3F). The combination of IdU and CRT0044876 further increased intercalation, suggesting that supercoiling is correlated with noise enhancement through increased Apex1-DNA interactions (Fig. 3F). IdU treatment followed by a short incubation with bleomycin (which decreases supercoiling through double-stranded breaks) reduced bTMP intercalation below the DMSO control level, indicating IdU alone in uncoiled DNA does not increase intercalation (fig. S20).

If DNA topology influences transcriptional bursting, additional modifiers of supercoiling should also affect Nanog noise. Topoisomerase 1 and 2a (Top1 and Top2a, respectively) relax coiled DNA through the introduction of single- and double-stranded breaks, respectively. Depletion of Top1 and Top2a by CRISPRi increased Nanog protein variability (fig. S21A). Inhibition of topoisomerase activity with the small-molecule inhibitors topotecan and etoposide recapitulated these effects (fig. S21B). Furthermore, overexpression of Topoisomerase 1 partially ablated IdU-mediated noise enhancement (fig. S21C). However, depletion of chromatin remodeling proteins known to interact with BER machinery failed to modulate IdU-mediated noise enhancement, suggesting that histone repositioning, a reported modulator of transcriptional noise, is not a major contributor to BER-mediated noise enhancement (fig. S21D). Taken together with

psoralen-crosslinking data, these results indicate that Apex1-induced supercoiling is a significant driver of noise-without-mean amplification.

To understand the mechanism by which Apex1 might increase transcriptional noise without altering mean expression, we developed a series of minimalist computational models to account for the experimental data (supplementary text 2). Monte Carlo simulations of each model using smRNA-FISH data for parameterization (table S2) indicated that a model incorporating transcription-coupled base excision best accounts for noise-without-mean amplification (Fig. 4, A and B; figs. S22 to S24; and supplementary text 5). In this model, Apex1 binding triggers entry to a negatively supercoiled transcriptionally non-productive state (ON*), whereas unbinding of Apex1 allows mRNA production to resume with an amplified transcription rate that is proportional to time spent in the non-productive state (fig. S25 and supplementary text 6); i.e., the longer the residence time in the non-productive state, the stronger the enhancement of transcription rate once repair is complete (a feedforward loop). This feedforward effect may originate from the increased negative supercoiling during repair which can facilitate a proportionate increase in upstream binding of transcriptional machinery (*61–68*). In line with this hypothesis, the model accurately predicts noise enhancement mediated by topoisomerase inhibition (fig. S24C and supplementary text 6). The ability to render a gene transcriptionally non-productive while also stimulating recruitment of transcriptional resources points to a homeostatic mechanism: the BER pathway maintains gene-expression homeostasis (i.e., mean) by amplifying transcriptional fluctuations through reciprocal modulation of burst intensity and duration (Fig. 4B). We term this model, and the associated phenomenon, Discordant Transcription through Repair (DiThR) due to the large discordance in pre- vs. post-repair transcriptional activity.

Sensitivity analysis of the DiThR model revealed that orthogonal modulation of Nanog mean and noise is possible within a large portion of the parameter space (fig. S26, A and B, and supplementary text 7). As validation, we tested the effect of 96 concentration combinations (table S6) of IdU and CRT0044876 to perturb the rates of Apex1 binding and unbinding respectively. The experimental results confirmed model predictions, showing that Nanog noise could be tuned independently of the mean (Fig. 4C). Testing of BrdU and hmU further validated that parameter regimes exist where noise can be regulated independent of mean (fig. S27). The hmU data in particular showed that the BER pathway can amplify noise while maintaining mean expression when removing a naturally occurring base modification. The different concentration thresholds for noise enhancement amongst these nucleoside analogs may reflect known differences in incorporation rates (*69*). Sensitivity analysis indicated that for genes whose $K_{OFF} >> K_{ON}$ (i.e., lowly expressed genes), IdU treatment would increase mean abundance (fig. S26E). This prediction was verified experimentally with bulk RNA-seq measurements of transcript abundance in mESCs treated with IdU, as all 98 of the up-regulated genes reside within the lowest expression regime (fig. S3).

To test whether the DiThR model applies to additional genes, mRNA distributions from the scRNA-seq dataset were fit to a Poisson-beta model (two-state model) allowing for estimation of K_{ON}, K_{OFF}, and K_{TX} (*70*). A consistent pattern emerged for genes classified as highly variable: 80% exhibited increased rates of promoter inactivation (K_{OFF}) and 84% had increased transcription rates (K_{TX}) (fig. S28). Alignment of these rate estimates with predictions from the model revealed that amplification of transcription after BER appears to be a unified mechanism for maintenance of gene-expression homeostasis across a broad range of genes (fig. S29 and supplementary text 8).

We next asked if amplified transcriptional noise could improve responsiveness to external stimuli (e.g., fate-specification signals) as previously suggested (*71–73*). Numerical simulations of the DiThR model predicted that IdU-mediated amplification of transcriptional noise could increase responsiveness to activation stimuli (Fig. 4D).

To experimentally verify these predictions, we tested if IdU could potentiate both differentiation of mESCs into the neural ectodermal (NE) lineage and reprogramming of differentiated cells into pluripotent stem cells. To assess potentiation of differentiation, mESCs cultured in NE-specifying media were stained for CD24, an established marker of the neural ectoderm lineage (*74*, *75*). Addition of 4uM IdU for the first 48 hours of differentiation yielded an ~4-fold and ~2-fold increase in CD24(+) cells at 48 hours and 96 hours into differentiation, respectively (fig. S30).

To assess for potentiation of dedifferentiation, three cellular reprogramming systems were used. The first assay utilized Nanog-GFP mouse embryonic fibroblasts (2` MEFs) that harbor stably integrated, doxycycline-inducible cassettes for three of the Yamanaka factors: Oct4, Sox2, and Klf4 (OSK). As confirmation that IdU acts as a noise-enhancer in this system, treatment of secondary MEFs with IdU for 48 hours in standard MEF media caused increased variability in Nanog protein expression (fig. S31A) with no changes in cell-cycle progression (fig. S31B). IdU supplementation for the first 48 hours of a 10-day reprogramming course enhanced the formation of pluripotent colonies as measured by alkaline phosphatase staining (Fig. 4E). Bulk RNA-seq at days 2 and 5 of reprogramming (fig. S31C) and flow-cytometric analysis at day 10 (Fig. 4F) demonstrated that early-stage noise-enhancement accelerated activation of the pluripotency program. The second reprogramming assay utilized Oct4-GFP primary MEFs that were transduced with retroviral vectors expressing

Oct4, Sox2, Klf4, and c-Myc (OKSM). IdU supplementation for the 48 hours immediately after transduction caused a ~2.4-fold increase in the number of Oct4-GFP(+) colonies (fig. S32A). As a third test of reprogramming, Oct4-GFP MEFs that harbor stably integrated, doxycycline-inducible cassettes for the OKSM factors were used. Similar to observations in the previous two cell lines, IdU treatment enhanced reprogramming efficiency ~4-fold (Fig. 4G and fig. S32C). Importantly, shRNA-mediated depletion of Apex1 (fig. S32B) in these MEFs ablated the enhanced reprogramming efficiency observed with IdU treatment (Fig. 4G and fig. S32C), thus demonstrating how BER-mediated amplification of intrinsic gene expression fluctuations is necessary to potentiate responsiveness to fate-specification signals.

Our data reveal that a DNA-surveillance pathway can utilize the biomechanical link between supercoiling and transcription to homeostatically enhance noise without altering mean-expression levels. This homeostatic noise-without-mean amplification (DiThR mechanism) appears to increase cellular responsiveness to multiple types of fate-specification signals. This raises intriguing implications for the role of naturally occurring oxidized nucleobases (e.g., hmU) in cell-fate determination, particularly since these base modifications are found at higher frequencies in embryonic stem-cell DNA (*45*). Mechanistic insight from modeling and experimental perturbation of Apex1 suggest that homeostatic (i.e., orthogonal) noise amplification may also apply to other DNA-processing activities that interrupt transcription. Homeostatic noise amplification cannot occur for all promoters (i.e., promoters with $K_{OFF} >> K_{ON}$ are precluded as they will exhibit increased mean). Additionally, genes most susceptible to transcriptional noise-enhancement tend to lie far from TAD boundaries. Since TAD boundaries are largely overlapping with supercoiling domain boundaries (*76*) and transcription-induced supercoiling may directly contribute to the formation of TADs (*77*), we reason that TAD boundaries may maintain a constantly high level of supercoiling, thus offering a narrow dynamic range for noise enhancement. Propagation of transcriptional variability to the protein level likely depends on protein half-lives and thus may not occur for a large swath of proteins. The proteins monitored in this study either have naturally short half-lives (Nanog) or PEST tags (e.g., d_2GFP) which minimizes the buffering of transcriptional bursts conferred by longer protein half-lives (*78*). The ability to independently control the mean and variance of gene expression may indicate that cells can amplify transcriptional noise for fate exploration and specification.

REFERENCES AND NOTES

1. L. Boltzmann, Weitere Studien uber das Wärmegleichgewicht unter Gasmolekulen, Sitzungsber. Kais. Akad. Wiss. Wien Math. *Naturwiss* **66**, 275–370 (1872).
2. S. Arrhenius, Über die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren. *Z. Phys. Chem.* **4**, 226–248 (1889). doi:10.1515/zpch-1889-0416
3. L. Roberts, Picture coding using pseudo-random noise. *IEEE Trans. Inf. Theory* **8**, 145–154 (1962). doi:10.1109/TIT.1962.1057702
4. P. Fatt, B. Katz, Some observations on biological noise. *Nature* **166**, 597–598 (1950). doi:10.1038/166597a0 Medline
5. A. A. Priplata, J. B. Niemi, J. D. Harry, L. A. Lipsitz, J. J. Collins, Vibrating insoles and balance control in elderly people. *Lancet* **362**, 1123–1124 (2003). doi:10.1016/S0140-6736(03)14470-4 Medline
6. D. Cohen, Optimizing reproduction in a randomly varying environment. *J. Theor. Biol.* **12**, 119–129 (1966). doi:10.1016/0022-5193(66)90188-3 Medline
7. A. Arkin, J. Ross, H. H. McAdams, Stochastic kinetic analysis of developmental pathway bifurcation in phage lambda-infected Escherichia coli cells. *Genetics* **149**, 1633–1648 (1998). doi:10.1093/genetics/149.4.1633 Medline
8. J. L. Spudich, D. E. Koshland Jr., Non-genetic individuality: Chance in the single cell. *Nature* **262**, 467–471 (1976). doi:10.1038/262467a0 Medline
9. P. B. Gupta, C. M. Fillmore, G. Jiang, S. D. Shapira, K. Tao, C. Kuperwasser, E. S. Lander, Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* **146**, 633–644 (2011). doi:10.1016/j.cell.2011.07.026 Medline
10. L. S. Weinberger, J. C. Burnett, J. E. Toettcher, A. P. Arkin, D. V. Schaffer, Stochastic gene expression in a lentiviral positive-feedback loop: HIV-1 Tat fluctuations drive phenotypic diversity. *Cell* **122**, 169–182 (2005). doi:10.1016/j.cell.2005.06.006 Medline
11. A. Raj, A. van Oudenaarden, Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell* **135**, 216–226 (2008). doi:10.1016/j.cell.2008.09.050 Medline
12. J. R. Chubb, T. Trcek, S. M. Shenoy, R. H. Singer, Transcriptional pulsing of a developmental gene. *Curr. Biol.* **16**, 1018–1025 (2006). doi:10.1016/j.cub.2006.03.092 Medline
13. I. Golding, J. Paulsson, S. M. Zawilski, E. C. Cox, Real-time kinetics of gene activity in individual bacteria. *Cell* **123**, 1025–1036 (2005). doi:10.1016/j.cell.2005.09.031 Medline
14. T. B. Kepler, T. C. Elston, Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations. *Biophys. J.* **81**, 3116–3136 (2001). doi:10.1016/S0006-3495(01)75949-8 Medline
15. A. Raj, C. S. Peskin, D. Tranchina, D. Y. Vargas, S. Tyagi, Stochastic mRNA synthesis in mammalian cells. *PLOS Biol.* **4**, e309 (2006). doi:10.1371/journal.pbio.0040309 Medline
16. W. J. Blake, M. KAErn, C. R. Cantor, J. J. Collins, Noise in eukaryotic gene expression. *Nature* **422**, 633–637 (2003). doi:10.1038/nature01546 Medline
17. J. Peccoud, B. Ycart, Markovian modeling of gene-product synthesis. *Theor. Popul. Biol.* **48**, 222–234 (1995). doi:10.1006/tpbi.1995.1027
18. J. Rodriguez, D. R. Larson, Transcription in living cells: Molecular mechanisms of bursting. *Annu. Rev. Biochem.* **89**, 189–212 (2020). doi:10.1146/annurev-biochem-011520-105250 Medline
19. D. Nicolas, B. Zoller, D. M. Suter, F. Naef, Modulation of transcriptional burst frequency by histone acetylation. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 7153–7158 (2018). doi:10.1073/pnas.1722330115 Medline
20. A. Bar-Even, J. Paulsson, N. Maheshri, M. Carmi, E. O'Shea, Y. Pilpel, N. Barkai, Noise in protein expression scales with natural protein abundance. *Nat. Genet.* **38**, 636–643 (2006). doi:10.1038/ng1807 Medline
21. R. D. Dar, S. M. Shaffer, A. Singh, B. S. Razooky, M. L. Simpson, A. Raj, L. S. Weinberger, Transcriptional bursting explains the noise-versus-mean relationship in mRNA and protein levels. *PLOS ONE* **11**, e0158298 (2016). doi:10.1371/journal.pone.0158298 Medline
22. 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). doi:10.1038/nature04785 Medline
23. U. Alon, *An Introduction to Systems Biology: Design Principles of Biological Circuits.* (Chapman & Hall/CRC, 2007).
24. R. D. Dar, N. N. Hosmane, M. R. Arkin, R. F. Siliciano, L. S. Weinberger, Screening for noise in gene expression identifies drug synergies. *Science* **344**, 1392–1396 (2014). doi:10.1126/science.1250220 Medline
25. M. M. K. Hansen, W. Y. Wen, E. Ingerman, B. S. Razooky, C. E. Thompson, R. D. Dar, C. W. Chin, M. L. Simpson, L. S. Weinberger, A post-transcriptional feedback mechanism for noise suppression and fate stabilization. *Cell* **173**, 1609–1621.e15 (2018). doi:10.1016/j.cell.2018.04.005 Medline
26. Y. Li, Y. Shan, R. V. Desai, K. H. Cox, L. S. Weinberger, J. S. Takahashi, Noise-driven

cellular heterogeneity in circadian periodicity. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 10350–10356 (2020). doi:10.1073/pnas.1922388117 Medline
27. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018). doi:10.1038/nbt.4096 Medline
28. M. M. K. Hansen, R. V. Desai, M. L. Simpson, L. S. Weinberger, Cytoplasmic amplification of transcriptional noise generates substantial cell-to-cell variability. *Cell Syst.* **7**, 384–397.e6 (2018). doi:10.1016/j.cels.2018.08.002 Medline
29. B. Munsky, G. Neuert, A. van Oudenaarden, Using gene expression noise to understand gene regulation. *Science* **336**, 183–187 (2012). doi:10.1126/science.1216379 Medline
30. D. M. Suter, N. Molina, D. Gatfield, K. Schneider, U. Schibler, F. Naef, Mammalian genes are transcribed with widely different bursting kinetics. *Science* **332**, 472–474 (2011). doi:10.1126/science.1198817 Medline
31. N. Eling, A. C. Richard, S. Richardson, J. C. Marioni, C. A. Vallejos, Correcting the mean-variance dependency for differential variability testing using single-cell RNA sequencing data. *Cell Syst.* **7**, 284–294.e12 (2018). doi:10.1016/j.cels.2018.06.011 Medline
32. M. Kaern, T. C. Elston, W. J. Blake, J. J. Collins, Stochasticity in gene expression: From theories to phenotypes. *Nat. Rev. Genet.* **6**, 451–464 (2005). doi:10.1038/nrg1615 Medline
33. A. Scialdone, K. N. Natarajan, L. R. Saraiva, V. Proserpio, S. A. Teichmann, O. Stegle, J. C. Marioni, F. Buettner, Computational assignment of cell-cycle stage from single-cell transcriptome data. *Methods* **85**, 54–61 (2015). doi:10.1016/j.ymeth.2015.06.021 Medline
34. J. M. Pedraza, A. van Oudenaarden, Noise propagation in gene networks. *Science* **307**, 1965–1969 (2005). doi:10.1126/science.1109090 Medline
35. R. Bargaje, K. Trachana, M. N. Shelton, C. S. McGinnis, J. X. Zhou, C. Chadick, S. Cook, C. Cavanaugh, S. Huang, L. Hood, Cell population structure prior to bifurcation predicts efficiency of directed differentiation in human induced pluripotent cells. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 2271–2276 (2017). doi:10.1073/pnas.1621412114 Medline
36. M. Mojtahedi, A. Skupin, J. Zhou, I. G. Castaño, R. Y. Y. Leong-Quong, H. Chang, K. Trachana, A. Giuliani, S. Huang, Cell fate decision as high-dimensional critical state transition. *PLOS Biol.* **14**, e2000640 (2016). doi:10.1371/journal.pbio.2000640 Medline
37. C. Sokolik, Y. Liu, D. Bauer, J. McPherson, M. Broeker, G. Heimberg, L. S. Qi, D. A. Sivak, M. Thomson, Transcription factor competition allows embryonic stem cells to distinguish authentic signals from noise. *Cell Syst.* **1**, 117–129 (2015). doi:10.1016/j.cels.2015.08.001 Medline
38. L. A. Carvajal, D. B. Neriah, A. Senecal, L. Benard, V. Thiruthuvanathan, T. Yatsenko, S.-R. Narayanagari, J. C. Wheat, T. I. Todorova, K. Mitchell, C. Kenworthy, V. Guerlavais, D. A. Annis, B. Bartholdy, B. Will, J. D. Anampa, I. Mantzaris, M. Aivado, R. H. Singer, R. A. Coleman, A. Verma, U. Steidl, Dual inhibition of MDMX and MDM2 as a therapeutic strategy in leukemia. *Sci. Transl. Med.* **10**, eaao3003 (2018). doi:10.1126/scitranslmed.aao3003 Medline
39. E. Abranches, A. M. V. Guedes, M. Moravec, H. Maamar, P. Svoboda, A. Raj, D. Henrique, Stochastic NANOG fluctuations allow mouse embryonic stem cells to explore pluripotency. *Development* **141**, 2770–2779 (2014). doi:10.1242/dev.108910 Medline
40. T. Kalmar, C. Lim, P. Hayward, S. Muñoz-Descalzo, J. Nichols, J. Garcia-Ojalvo, A. Martinez Arias, Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. *PLOS Biol.* **7**, e1000149 (2009). doi:10.1371/journal.pbio.1000149 Medline
41. D. W. Austin, M. S. Allen, J. M. McCollum, R. D. Dar, J. R. Wilgus, G. S. Sayler, N. F. Samatova, C. D. Cox, M. L. Simpson, Gene network shaping of inherent noise spectra. *Nature* **439**, 608–611 (2006). doi:10.1038/nature04194 Medline
42. N. Rosenfeld, J. W. Young, U. Alon, P. S. Swain, M. B. Elowitz, Gene regulation at the single-cell level. *Science* **307**, 1962–1965 (2005). doi:10.1126/science.1106914 Medline
43. A. Sigal, R. Milo, A. Cohen, N. Geva-Zatorsky, Y. Klein, Y. Liron, N. Rosenfeld, T. Danon, N. Perzov, U. Alon, Variability and memory of protein levels in human cells. *Nature* **444**, 643–646 (2006). doi:10.1038/nature05316 Medline
44. S. Ito, A. C. D'Alessio, O. V. Taranova, K. Hong, L. C. Sowers, Y. Zhang, Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466**, 1129–1133 (2010). doi:10.1038/nature09303 Medline
45. T. Pfaffeneder, F. Spada, M. Wagner, C. Brandmayr, S. K. Laube, D. Eisen, M. Truss, J. Steinbacher, B. Hackner, O. Kotljarova, D. Schuermann, S. Michalakis, O. Kosmatchev, S. Schiesser, B. Steigenberger, N. Raddaoui, G. Kashiwazaki, U. Müller, C. G. Spruijt, M. Vermeulen, H. Leonhardt, P. Schär, M. Müller, T. Carell, Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. *Nat. Chem. Biol.* **10**, 574–581 (2014). doi:10.1038/nchembio.1532 Medline
46. Y. F. He, B.-Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C.-X. Song, K. Zhang, C. He, G.-L. Xu, Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**, 1303–1307 (2011). doi:10.1126/science.1210944 Medline
47. L. Shen, H. Wu, D. Diep, S. Yamaguchi, A. C. D'Alessio, H.-L. Fung, K. Zhang, Y. Zhang, Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell* **153**, 692–706 (2013). doi:10.1016/j.cell.2013.04.002 Medline
48. E. S. J. Arnér, S. Eriksson, Mammalian deoxyribonucleoside kinases. *Pharmacol. Ther.* **67**, 155–186 (1995). doi:10.1016/0163-7258(95)00015-9 Medline
49. B. Demple, T. Herman, D. S. Chen, Cloning and expression of APE, the cDNA encoding the major human apurinic endonuclease: Definition of a family of DNA repair enzymes. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 11450–11454 (1991). doi:10.1073/pnas.88.24.11450 Medline
50. T. Lindahl, R. D. Wood, Quality control by DNA repair. *Science* **286**, 1897–1905 (1999). doi:10.1126/science.286.5446.1897 Medline
51. U. Müller, C. Bauer, M. Siegl, A. Rottach, H. Leonhardt, TET-mediated oxidation of methylcytosine causes TDG or NEIL glycosylase dependent gene reactivation. *Nucleic Acids Res.* **42**, 8592–8604 (2014). doi:10.1093/nar/gku552 Medline
52. M. L. Hegde, T. K. Hazra, S. Mitra, Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells. *Cell Res.* **18**, 27–47 (2008). doi:10.1038/cr.2008.8 Medline
53. S. Madhusudan, F. Smart, P. Shrimpton, J. L. Parsons, L. Gardiner, S. Houlbrook, D. C. Talbot, T. Hammonds, P. A. Freemont, M. J. E. Sternberg, G. L. Dianov, I. D. Hickson, Isolation of a small molecule inhibitor of DNA base excision repair. *Nucleic Acids Res.* **33**, 4711–4724 (2005). doi:10.1093/nar/gki781 Medline
54. O. A. Kladova, M. Bazlekowa-Karaban, S. Baconnais, O. Piétrement, A. A. Ishchenko, B. T. Matkarimov, D. A. Iakovlev, A. Vasenko, O. S. Fedorova, E. Le Cam, B. Tudek, N. A. Kuznetsov, M. Saparbaev, The role of the N-terminal domain of human apurinic/apyrimidinic endonuclease 1, APE1, in DNA glycosylase stimulation. *DNA Repair* **64**, 10–25 (2018). doi:10.1016/j.dnarep.2018.02.001 Medline
55. C. D. Mol, T. Izumi, S. Mitra, J. A. Tainer, DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination [corrected]. *Nature* **403**, 451–456 (2000). doi:10.1038/35000249 Medline
56. D. R. McNeill, D. M. Wilson 3rd, A dominant-negative form of the major human abasic endonuclease enhances cellular sensitivity to laboratory and clinical DNA-damaging agents. *Mol. Cancer Res.* **5**, 61–70 (2007). doi:10.1158/1541-7786.MCR-06-0329 Medline
57. S. Chong, C. Chen, H. Ge, X. S. Xie, Mechanism of transcriptional bursting in bacteria. *Cell* **158**, 314–326 (2014). doi:10.1016/j.cell.2014.05.038 Medline
58. S. A. Sevier, D. A. Kessler, H. Levine, Mechanical bounds to transcriptional noise. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 13983–13988 (2016). doi:10.1073/pnas.1612651113 Medline
59. S. Corless, N. Gilbert, Investigating DNA supercoiling in eukaryotic genomes. *Brief. Funct. Genomics* **16**, 379–389 (2017). doi:10.1093/bfgp/elx007 Medline
60. K. N. Babos, K. E. Galloway, K. Kisler, M. Zitting, Y. Li, Y. Shi, B. Quintino, R. H. Chow, B. V. Zlokovic, J. K. Ichida, Mitigating antagonism between transcription and proliferation allows near-deterministic cellular reprogramming. *Cell Stem Cell* **25**, 486–500.e9 (2019). doi:10.1016/j.stem.2019.08.005 Medline
61. P. Guptasarma, Cooperative relaxation of supercoils and periodic transcriptional initiation within polymerase batteries. *BioEssays* **18**, 325–332 (1996). doi:10.1002/bies.950180411 Medline
62. S. Kim, B. Beltran, I. Irnov, C. Jacobs-Wagner, Long-distance cooperative and antagonistic RNA polymerase dynamics via DNA supercoiling. *Cell* **179**, 106–119.e16 (2019). doi:10.1016/j.cell.2019.08.033 Medline
63. F. Kouzine, S. Sanford, Z. Elisha-Feil, D. Levens, The functional response of upstream DNA to dynamic supercoiling in vivo. *Nat. Struct. Mol. Biol.* **15**, 146–154 (2008). doi:10.1038/nsmb.1372 Medline

64. J. Liu, F. Kouzine, Z. Nie, H.-J. Chung, Z. Elisha-Feil, A. Weber, K. Zhao, D. Levens, The FUSE/FBP/FIR/TFIIH system is a molecular machine programming a pulse of c-myc expression. *EMBO J.* **25**, 2119–2130 (2006). doi:10.1038/sj.emboj.7601101 Medline
65. M. Mizutani, T. Ohta, H. Watanabe, H. Handa, S. Hirose, Negative supercoiling of DNA facilitates an interaction between transcription factor IID and the fibroin gene promoter. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 718–722 (1991). doi:10.1073/pnas.88.3.718 Medline
66. S. S. Teves, S. Henikoff, Transcription-generated torsional stress destabilizes nucleosomes. *Nat. Struct. Mol. Biol.* **21**, 88–94 (2014). doi:10.1038/nsmb.2723 Medline
67. M. Bazlekowa-Karaban, P. Prorok, S. Baconnais, S. Taipakova, Z. Akishev, D. Zembrzuska, A. V. Popov, A. V. Endutkin, R. Groisman, A. A. Ishchenko, B. T. Matkarimov, A. Bissenbaev, E. Le Cam, D. O. Zharkov, B. Tudek, M. Saparbaev, Mechanism of stimulation of DNA binding of the transcription factors by human apurinic/apyrimidinic endonuclease 1, APE1. *DNA Repair* **82**, 102698 (2019). doi:10.1016/j.dnarep.2019.102698 Medline
68. J. F. Breit, K. Ault-Ziel, A.-B. Al-Mehdi, M. N. Gillespie, Nuclear protein-induced bending and flexing of the hypoxic response element of the rat vascular endothelial growth factor promoter. *FASEB J.* **22**, 19–29 (2008). doi:10.1096/fj.07-8102com Medline
69. B. Leuner, E. R. Glasper, E. Gould, Thymidine analog methods for studies of adult neurogenesis are not equally sensitive. *J. Comp. Neurol.* **517**, 123–133 (2009). doi:10.1002/cne.22107 Medline
70. M. Delmans, M. Hemberg, Discrete distributional differential expression (D3E)—A tool for gene expression analysis of single-cell RNA-seq data. *BMC Bioinformatics* **17**, 110 (2016). doi:10.1186/s12859-016-0944-6 Medline
71. W. J. Blake, G. Balázsi, M. A. Kohanski, F. J. Isaacs, K. F. Murphy, Y. Kuang, C. R. Cantor, D. R. Walt, J. J. Collins, Phenotypic consequences of promoter-mediated transcriptional noise. *Mol. Cell* **24**, 853–865 (2006). doi:10.1016/j.molcel.2006.11.003 Medline
72. H. H. Chang, M. Hemberg, M. Barahona, D. E. Ingber, S. Huang, Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. *Nature* **453**, 544–547 (2008). doi:10.1038/nature06965 Medline
73. G. M. Süel, R. P. Kulkarni, J. Dworkin, J. Garcia-Ojalvo, M. B. Elowitz, Tunability and noise dependence in differentiation dynamics. *Science* **315**, 1716–1719 (2007). doi:10.1126/science.1137455 Medline
74. J. Pruszak, W. Ludwig, A. Blak, K. Alavian, O. Isacson, CD15, CD24, and CD29 define a surface biomarker code for neural lineage differentiation of stem cells. *Stem Cells* **27**, 2928–2940 (2009). doi:10.1002/stem.211 Medline
75. S. Semrau, J. E. Goldmann, M. Soumillon, T. S. Mikkelsen, R. Jaenisch, A. van Oudenaarden, Dynamics of lineage commitment revealed by single-cell transcriptomics of differentiating embryonic stem cells. *Nat. Commun.* **8**, 1096 (2017). doi:10.1038/s41467-017-01076-4 Medline
76. C. Naughton, N. Avlonitis, S. Corless, J. G. Prendergast, I. K. Mati, P. P. Eijk, S. L. Cockroft, M. Bradley, B. Ylstra, N. Gilbert, Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. *Nat. Struct. Mol. Biol.* **20**, 387–395 (2013). doi:10.1038/nsmb.2509 Medline
77. D. Racko, F. Benedetti, J. Dorier, A. Stasiak, Transcription-induced supercoiling as the driving force of chromatin loop extrusion during formation of TADs in interphase chromosomes. *Nucleic Acids Res.* **46**, 1648–1660 (2018). doi:10.1093/nar/gkx1123 Medline
78. A. Singh, Transient changes in intercellular protein variability identify sources of noise in gene expression. *Biophys. J.* **107**, 2214–2220 (2014). doi:10.1016/j.bpj.2014.09.017 Medline
79. R. Desai, B. Martin, scRNA-seq and modeling code for: A DNA-repair pathway can affect transcriptional noise to promote cell fate transitions, Zenodo (2021); https://doi.org/10.5281/zenodo.4891977.
80. E. P. Nora, A. Goloborodko, A.-L. Valton, J. H. Gibcus, A. Uebersohn, N. Abdennur, J. Dekker, L. A. Mirny, B. G. Bruneau, Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization. *Cell* **169**, 930–944.e22 (2017). doi:10.1016/j.cell.2017.05.004 Medline
81. P. Angerer, L. Haghverdi, M. Büttner, F. J. Theis, C. Marr, F. Buettner, destiny: Diffusion maps for large-scale single-cell data in R. *Bioinformatics* **32**, 1241–1243 (2016). doi:10.1093/bioinformatics/btv715 Medline
82. F. Mueller, A. Senecal, K. Tantale, H. Marie-Nelly, N. Ly, O. Collin, E. Basyuk, E. Bertrand, X. Darzacq, C. Zimmer, FISH-quant: Automatic counting of transcripts in 3D FISH images. *Nat. Methods* **10**, 277–278 (2013). doi:10.1038/nmeth.2406 Medline
83. H. Ochiai, T. Sugawara, T. Sakuma, T. Yamamoto, Stochastic promoter activation affects Nanog expression variability in mouse embryonic stem cells. *Sci. Rep.* **4**, 7125 (2014). doi:10.1038/srep07125 Medline
84. C. McQuin, A. Goodman, V. Chernyshev, L. Kamentsky, B. A. Cimini, K. W. Karhohs, M. Doan, L. Ding, S. M. Rafelski, D. Thirstrup, W. Wiegraebe, S. Singh, T. Becker, J. C. Caicedo, A. E. Carpenter, CellProfiler 3.0: Next-generation image processing for biology. *PLOS Biol.* **16**, e2005970 (2018). doi:10.1371/journal.pbio.2005970 Medline
85. M. A. Horlbeck, L. A. Gilbert, J. E. Villalta, B. Adamson, R. A. Pak, Y. Chen, A. P. Fields, C. Y. Park, J. E. Corn, M. Kampmann, J. S. Weissman, Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *eLife* **5**, e19760 (2016). doi:10.7554/eLife.19760 Medline
86. B. S. Razooky, A. Pai, K. Aull, I. M. Rouzine, L. S. Weinberger, A hardwired HIV latency program. *Cell* **160**, 990–1001 (2015). doi:10.1016/j.cell.2015.02.009 Medline
87. L. V. Sharova, A. A. Sharov, T. Nedorezov, Y. Piao, N. Shaik, M. S. H. Ko, Database for mRNA half-life of 19 977 genes obtained by DNA microarray analysis of pluripotent and differentiating mouse embryonic stem cells. *DNA Res.* **16**, 45–58 (2009). doi:10.1093/dnares/dsn030 Medline
88. W. Wongpaiboonwattana, M. P. Stavridis, Neural differentiation of mouse embryonic stem cells in serum-free monolayer culture. *J. Vis. Exp.* (99): e52823 (2015). doi:10.3791/52823 Medline
89. M. T. Walsh, X. Huang, Measurement of incorporation kinetics of non-fluorescent native nucleotides by DNA polymerases using fluorescence microscopy. *Nucleic Acids Res.* **45**, e175–e175 (2017). doi:10.1093/nar/gkx833 Medline
90. B. Zoller, D. Nicolas, N. Molina, F. Naef, Structure of silent transcription intervals and noise characteristics of mammalian genes. *Mol. Syst. Biol.* **11**, 823 (2015). doi:10.15252/msb.20156257 Medline
91. K. Bahar Halpern, I. Caspi, D. Lemze, M. Levy, S. Landen, E. Elinav, I. Ulitsky, S. Itzkovitz, Nuclear retention of mRNA in mammalian tissues. *Cell Rep.* **13**, 2653–2662 (2015). doi:10.1016/j.celrep.2015.11.036 Medline
92. D. A. Gilchrist, G. Dos Santos, D. C. Fargo, B. Xie, Y. Gao, L. Li, K. Adelman, Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell* **143**, 540–551 (2010). doi:10.1016/j.cell.2010.10.004 Medline
93. K. Tantale, F. Mueller, A. Kozulic-Pirher, A. Lesne, J.-M. Victor, M.-C. Robert, S. Capozi, R. Chouaib, V. Bäcker, J. Mateos-Langerak, X. Darzacq, C. Zimmer, E. Basyuk, E. Bertrand, A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nat. Commun.* **7**, 12248 (2016). doi:10.1038/ncomms12248 Medline
94. J. Zhang, T. Zhou, Promoter-mediated transcriptional dynamics. *Biophys. J.* **106**, 479–488 (2014). doi:10.1016/j.bpj.2013.12.011 Medline
95. D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361 (1977). doi:10.1021/j100540a008
96. 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). doi:10.1073/pnas.1213530109 Medline
97. K. P. Burnham, D. R. Anderson, K. P. Burnham, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (Springer, ed. 2, 2002).

ACKNOWLEDGMENTS

We thank Michael Simpson, Benoit Bruneau, Jonathan Weissman and the Weinberger lab for thoughtful discussions and suggestions. We thank Kathryn Claiborn for editing and Giovanni Maki for graphics support. We thank Gabor Balazsi and Michael Simpson for helpful discussions. We acknowledge the technical assistance of Nandhini Raman in the Gladstone Institute Flow Cytometry Facility (NIH S10 RR028962, P30 AI027763, DARPA, and the James B. Pendleton Charitable Trust) and the Gladstone Assay Development and Drug Discovery Core. We also acknowledge Kurt Thorn and DeLaine Larson in the UCSF Nikon Imaging Center (NIH S10 1S10OD017993-01A1). We are grateful to the Gladstone Institute Genomics Core for technical assistance with single-cell

RNA-sequencing. The dual-tagged Sox2 mESCs were a kind donation from Benoit Bruneau and Elphege Nora. We thank Marco Jost and Jonathan Weissman for CRISPRi reagents. The Oct4-GFP reprogrammable MEFs (harbor stably integrated OKSM factors) were a kind donation from Shangqin Guo. **Funding:** R.V.D. is supported by an NIH/NICHD F30 fellowship (HD095614-03). R.A.C. acknowledges support from NIH award 1R01GM126045-05. R.H.S. acknowledges support from NIH awards NS083085 and 1R35GM136296. M.M.K.H. acknowledges support from the Dutch Research council (NWO) ENW-XS award (OCENW.XS3.055). L.S.W. acknowledges support from the Bowes Distinguished Professorship, Alfred P. Sloan Research Fellowship, Pew Scholars in the Biomedical Sciences Program, NIH awards R01AI109593, and the NIH Director's New Innovator Award (OD006677) and Pioneer Award (OD17181) programs. **Author contributions:** R.V.D., M.T., and L.S.W. conceived and designed the study. R.V.D., B.M., and M.T. analyzed the sequencing data. R.V.D., X.C., C.U., S.D., and L.S.W conceived and designed the cellular reprogramming experiments. X.C., D.W.H., W.L., R.H.S., R.A.C., and L.S.W conceived and designed the MS2 imaging experiments. R.V.D., X.C., S.C., D.W.H., W.L., and C.U. performed the experiments. R.V.D., X.C., B.M., M.T., R.A.C., M.M.K.H., and L.S.W. analyzed data. R.V.D., M.M.K.H., B.M., and L.S.W. constructed and analyzed the mathematical models. R.V.D., M.M.K.H., and L.S.W. wrote the manuscript. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Raw and processed sequencing data reported in this paper have been deposited onto GEO, accession number: GSE176044. Custom code for analysis of scRNA-seq data and mathematical modeling are available on GitHub at https://github.com/weinbergerlab-ucsf/Code_Desai_et_al and archived on Zenodo (79). Reagents, including plasmids and cell-lines, are available from the corresponding author upon request.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/cgi/content/full/science.abc6506/DC1
Materials and Methods
Supplementary Text
Figs. S1 to S32
Tables S1 to S8
References (*80*–*97*)
MDAR Reproducibility Checklist

31 May 2020; accepted 8 July 2021
Published online 22 July 2021
10.1126/science.abc6506

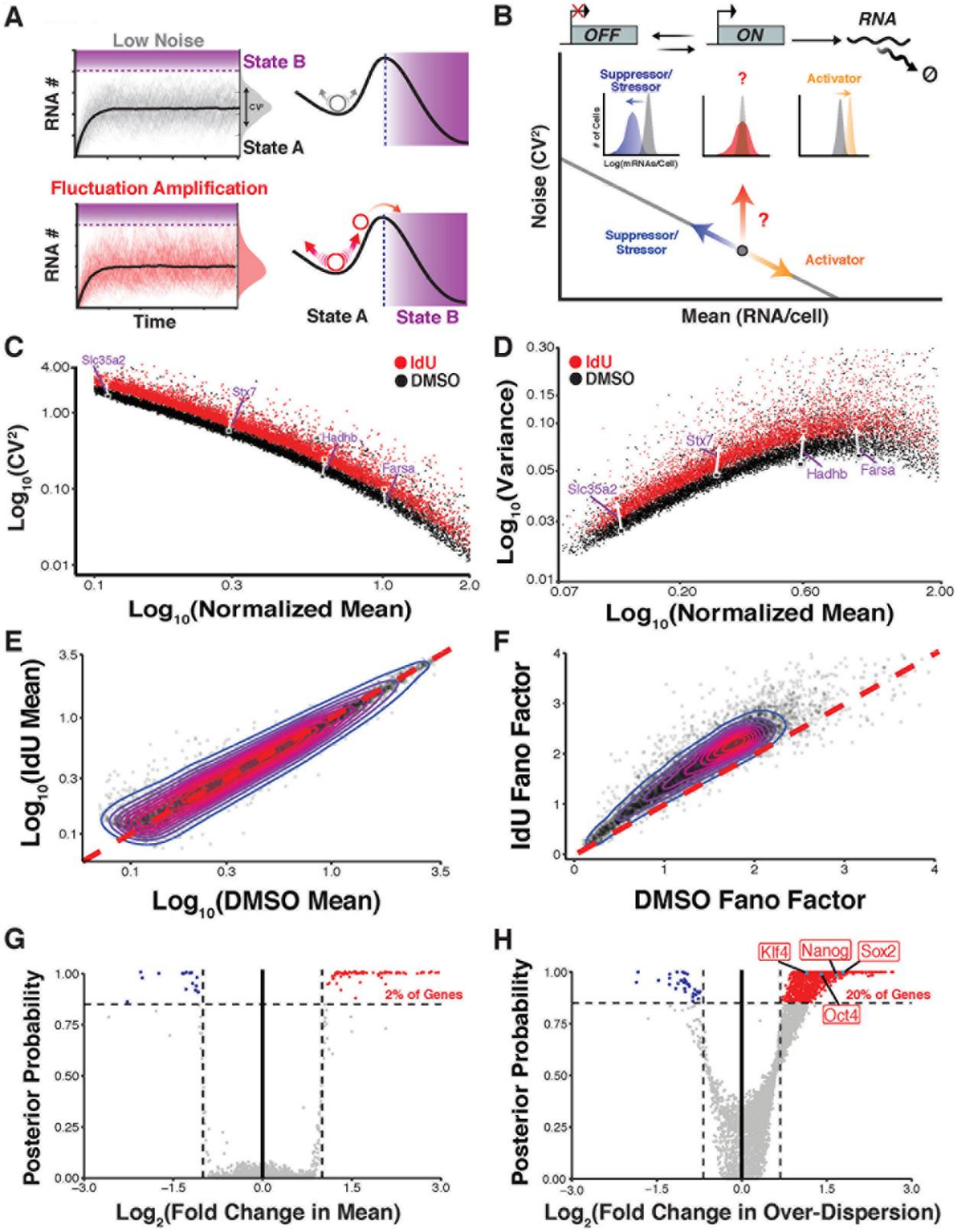

Fig. 1. Genome-wide amplification of cell-to-cell mRNA variability (i.e., 'noise') independent of mean. (**A**) (Left) Monte-Carlo simulations of the two-state random-telegraph model of transcription showing low noise and higher noise trajectories with matched mean expression levels. Coefficient of variation (σ^2/μ^2, CV^2) quantifies magnitude of fluctuations. (Right) The predicted facilitation of state transitions through 'dithering'. (**B**) (Top) Schematic of two-state random-telegraph model of transcription. (Bottom) Schematic of mean vs. CV^2 for mRNA abundance with solid gray line representing Poisson, inverse scaling of CV^2 as a function of mean. Question mark symbolizes unknown noise-control mechanisms that amplify fluctuations independently of mean. Histograms depict expected shift in mRNA copy number distributions. (**C** to **F**) scRNA-seq of mESCs treated with DMSO (black) or 10μM IdU (red) for 24 hours. 812 and 744 transcriptomes from DMSO and IdU treatments, respectively, were analyzed. (C) Mean expression vs. CV^2 and (D) mean vs. variance. Four examples of housekeeping genes (purple) demonstrate how IdU increases expression fluctuations with minimal change in mean (white arrows). (E) Mean expression and (F) Fano factor (σ^2/μ) of genes in DMSO vs. IdU treatments. (**G** and **H**) BASiCS analysis of scRNA-seq data. (G) Fold change in mean vs. certainty (posterior probability) that gene is up- or down-regulated. With IdU treatment, 113 genes (red) were classified as differentially expressed (>2-fold change in mean with >85% probability). (H) Fold change in over-dispersion vs. certainty (posterior probability) that gene is highly- or lowly-variable. 945 genes (red) were classified as highly variable (>1.5-fold change in over-dispersion with >85% probability).

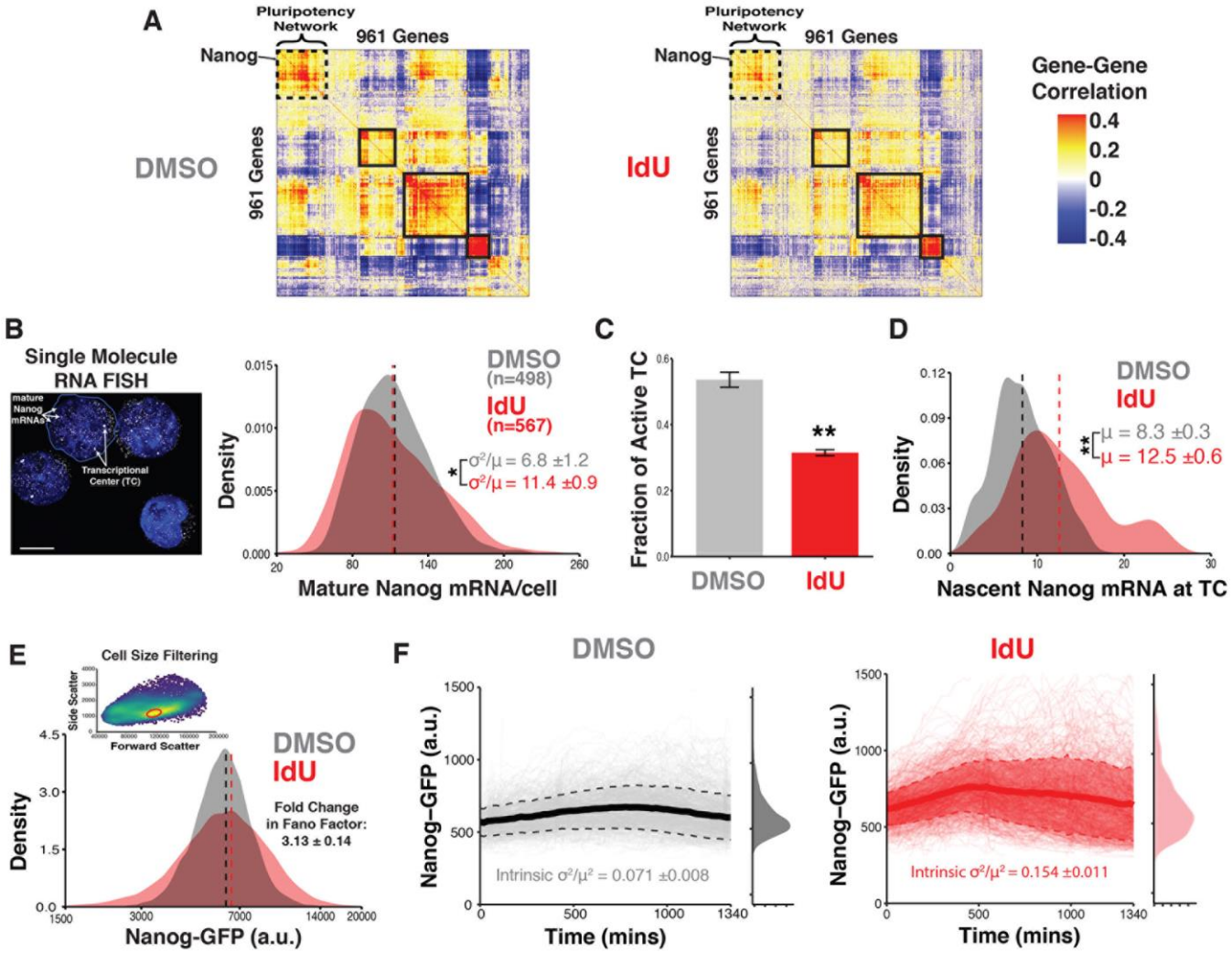

Fig. 2. Amplification of mRNA noise is not due to extrinsic sources, results from shorter but more intense transcriptional bursts, and propagates to protein levels. (**A**) Pearson correlations of expression for gene pairs in scRNA-seq dataset. Hierarchical clustering reveals networks of genes (highlighted in black rectangles) sharing similar correlation patterns. Dashed rectangle highlights network enriched with pluripotency factors like Nanog. (**B** to **D**) Results of smRNA-FISH used to count nascent and mature Nanog mRNA in Nanog-GFP mESCs treated with DMSO or 10μM IdU for 24 hours in 2i/LIF media. Data are from four biological replicates. (B) (Left) Representative micrograph (maximum intensity projection) of mESCs with DAPI staining in which Nanog transcripts are labeled with probe-set for eGFP. Bright foci correspond to transcriptional centers as verified by intron probe set. Scale bar is 5μm. (Right) Distributions of mature Nanog transcripts/cell. Dashed lines represent mean. Averaged Fano factors over all four replicates are reported (± SD), *p =0.0011 by a two-tailed, unpaired Student's *t* test. (C) Fraction of possible transcriptional centers that are active as detected by overlap of signal in exon and intron probe channels. Each cell is assumed to have 2 possible transcriptional centers (TCs). Data represent mean and SD, **p =6.9×10^{-5} by a two-tailed, unpaired Student's *t* test. (D) Distributions of nascent Nanog mRNA per TC. Average number of nascent mRNAs over all four replicates are reported, **p =1.0×10^{-4} by a two-tailed, unpaired Student's *t* test. (**E**) Representative flow cytometry distribution of Nanog-GFP expression in mESCs treated with DMSO or 10μM IdU for 24 hours in 2i/LIF. Dashed lines represent mean. Fold change in Fano factor (± SD) obtained from three biological replicates. Inset: Representative flow cytometry dot-plot showing conservative gating on forward and side scatter to filter extrinsic noise arising from cell size heterogeneity. (**F**) Time-lapse imaging of Nanog-GFP mESCs treated with either DMSO (n = 1513) or 10μM IdU (n = 1414) in 2i/LIF. Trajectories from two replicates of each condition are pooled, with solid and dashed lines representing mean and standard deviation of trajectories respectively. Distributions of Nanog-GFP represent expression at final time-point. Intrinsic-CV^2 of each detrended trajectory was calculated, with the average (± SD) of all trajectories reported.

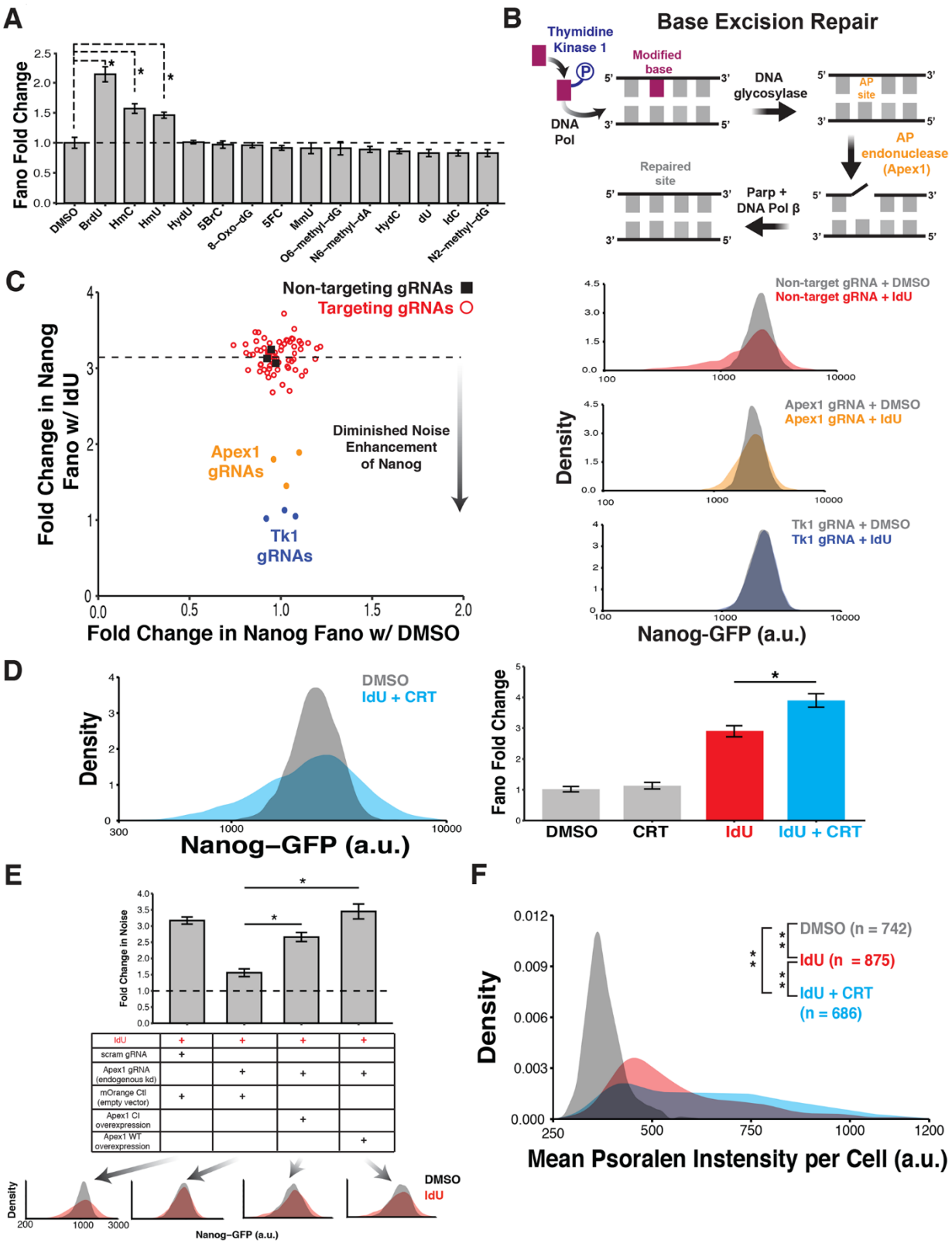

Fig. 3. Noise amplification independent of mean is due to Apex1-mediated DNA repair. (**A**) Screening of 14 additional nucleoside analogs. Nanog-GFP mESCs grown in 2i/LIF were supplemented with 10μM of nucleoside analog for 24 hours. Fano factor for Nanog protein expression was normalized to DMSO. Data represent mean (± SD) of biological replicates, *p < 0.01 by a Kruskal-Wallis test followed by Tukey's multiple comparison test. (**B**) Schematic of nucleoside analog incorporation into genomic DNA and removal through base excision repair pathway. (**C**) (Left) CRISPRi screening for genetic dependencies of IdU noise enhancement. Nanog-GFP mESCs stably expressing dCas9-KRAB-p2A-mCherry were transduced with a single gRNA expression vector with BFP reporter. 75 gRNAs (25 genes, 3 gRNAs/gene) were tested in addition to 3 non-targeting control gRNAs. Two days following transduction, each gRNA-expressing population of mESCs was treated with DMSO or 10μM IdU for 24 hours in 2i/LIF media. Nanog Fano factor for DMSO and IdU treatment of each gRNA population was normalized to Nanog Fano factor of non-targeting gRNA^{+}DMSO population. Each point represents a gRNA. Dashed horizontal line represents average noise enhancement of Nanog from IdU in the background of non-targeting gRNA expression (black squares). Depletion of Apex1 and Tk1 diminishes noise enhancement of Nanog from IdU. (Right) Representative flow cytometry distributions of Nanog expression for mESCs expressing non-targeting (top-right), Apex1 (middle-right), or Tk1 (bottom-right) gRNAs and treated with DMSO or 10μM IdU. (**D**) Combination of IdU and small-molecule inhibitor of the Apex1 endonuclease domain (CRT0044876). (Left) Representative flow cytometry distributions of Nanog expression for mESCs treated with DMSO or 10μM IdU + 100μM CRT0044876. (Right) mESCs were treated with DMSO, 100μM CRT0044876, 10μM IdU or 10μM IdU + 100μM CRT0044876 for 24 hours in 2i/LIF. Nanog Fano factor for each treatment was normalized to DMSO control. Data represent mean (± SD) of three biological replicates, *p = 0.0028 by a two-tailed, unpaired Student's *t* test. (**E**) Overexpression of wild-type (WT) or catalytically inactive (CI) Apex1 with simultaneous CRISPRi depletion of endogenous Apex1. (Top) Fold change in Nanog Fano factor for respective treatment condition described in rectangular grid. An mOrange empty vector was used as a transduction control. Nanog Fano factor for each treatment was normalized to mOrange control cells treated with DMSO. Data represent mean (± SD) of three biological replicates, *p < 0.005 by a two-tailed, unpaired Student's *t* test. (Bottom) Representative flow cytometry distributions of Nanog expression for each treatment condition. (**F**) Single-cell quantification of negative supercoiling levels using psoralen-crosslinking assay. mESCs were treated with DMSO, 10μM IdU or 10μM IdU + 100μM CRT0044876 for 24 hours in 2i/LIF. Distributions for nuclear intensities of bTMP staining are shown. Data are pooled from two biological replicates of each treatment, **p < 0.0001 by a Kruskal-Wallis test followed by Tukey's multiple comparison test.

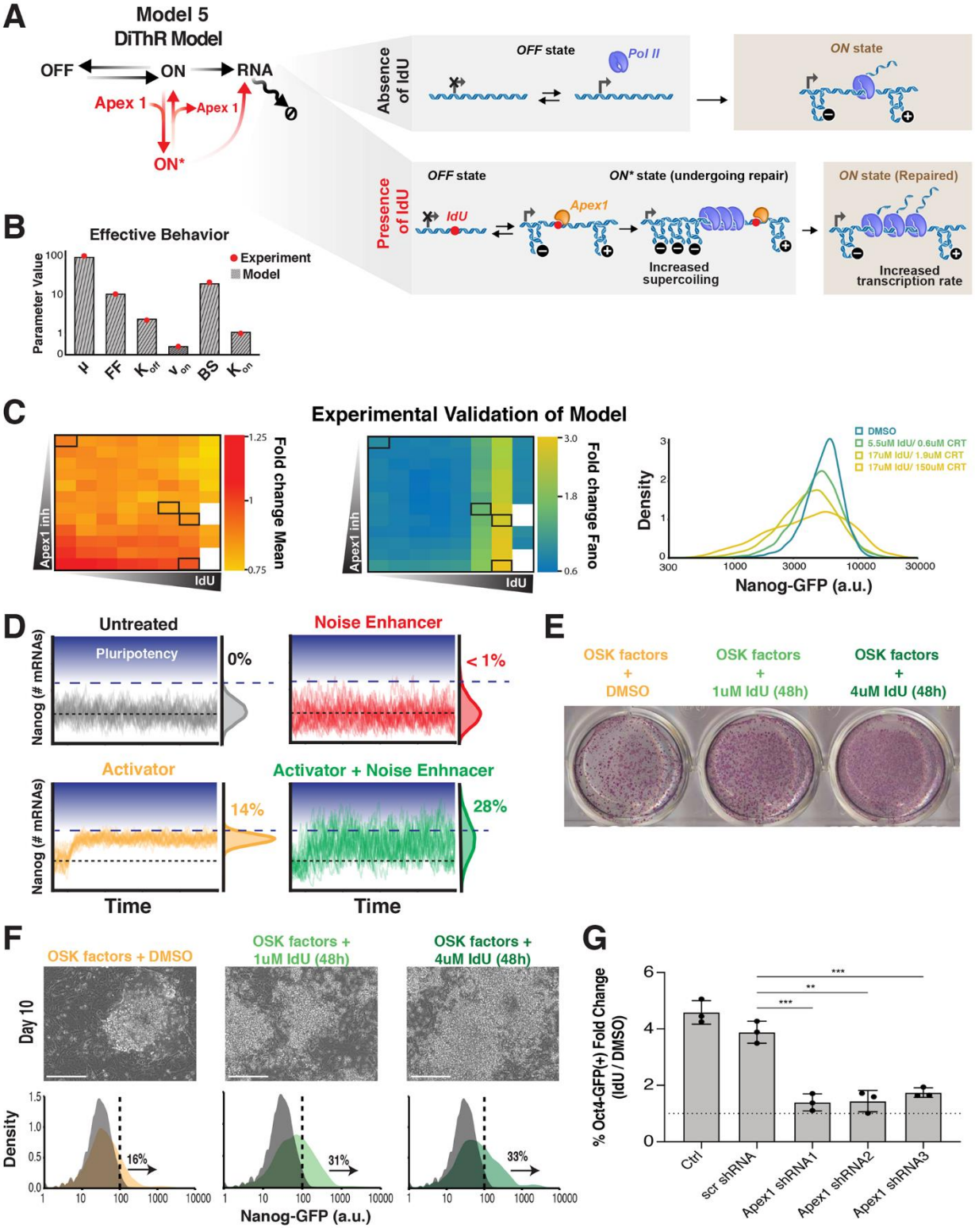

Fig. 4. A mechanistic model for Discordant Transcription through Repair (DiThR) and the phenotypic consequence of DiThR on potentiation of cellular reprogramming. (**A**) Detailed schematic of Model 5 (DiThR model) (see fig. S22 for schematic of models 1-4), which utilizes transcription-coupled base-excision. In the presence of IdU (bottom panel), Apex1 binding occurs when gene is transcriptionally permissive (ON state). Binding induces negative supercoiling which lengthens the time that a gene is transcriptionally non-productive (ON* state) while also facilitating recruitment of transcriptional resources. Upon repair completion, a higher transcriptional rate (that is proportional to time spent in ON* state) is reached. Mean expression is maintained with larger transcriptional fluctuations. (**B**) The effective behavior (mean Nanog mRNA [μ], Fano factor [FF], K_{off}, fraction of time active [v_{on}], burst size [BS], K_{on}) of the DiThR model is compared to experimentally derived values of each parameter (red dots) obtained from smRNA-FISH data. Absolute percentage error (APE) is calculated as described in supplementary text 5.2.2. Model 5 (DiThR model) best matches experimental data. (**C**) Testing of 96 concentration combinations of IdU and CRT0044876 to validate tunability of Nanog variability. IdU and CRT0044876 were used to increase binding and decrease unbinding of Apex1 respectively. Data represent average of two biological replicates. (Leftmost and Center Panels) 96-well heatmaps displaying fold change in Nanog mean and Fano factor for each drug combination as compared to DMSO (top-leftmost well). Insufficient number of cells for extrinsic noise filtering (<50,000) were recorded from white wells. (Rightmost Panel) Representative flow cytometry distributions from highlighted wells (black rectangles). (**D**) Simulations of the DiThR model for Nanog gene expression in the presence of DMSO (top left), IdU (top right), an activator (increased K_{ON}, decreased K_{OFF}) of promoter activity (bottom left) and an activator combined with IdU (bottom right). (**E**) Alkaline phosphatase staining for pluripotent stem cell colonies. Nanog-GFP secondary MEFs harboring stably-integrated, doxycycline-inducible cassettes for Oct4, Sox2, and Klf4 (OSK) were subjected to 10 days of doxycycline treatment in combination with DMSO (first well), 1μM IdU (second well), or 4μM IdU (third well) for the first 48 hours of reprogramming. (**F**) (Top) Micrographs of Nanog-GFP secondary MEFs at day 10 of doxycycline-induced reprogramming (scale bar = 100 μm). (Bottom) Flow cytometric analysis of Nanog-GFP activation at day 10 of reprogramming. Data are pooled from two replicates. (**G**) Fold change in percentage of Oct4-GFP(+) cells at day 10 of reprogramming between IdU and DMSO treatment conditions with and without Apex1 depletion. 4μM IdU or equivalent volume DMSO were present for first 48 hours of reprogramming. Data represent mean (± SD) of three biological replicates, **p = 0.0014, ***p < 0.001 by a two-tailed, unpaired Student's *t* test.

A DNA-repair pathway can affect transcriptional noise to promote cell fate transitions

Ravi V. Desai, Xinyue Chen, Benjamin Martin, Sonali Chaturvedi, Dong Woo Hwang, Weihan Li, Chen Yu, Sheng Ding, Matt Thomson, Robert H. Singer, Robert A. Coleman, Maike M.K. Hansen and Leor S. Weinberger

published online July 22, 2021