Title
A DNA repair pathway can regulate transcriptional noise to promote cell fate transitions

Journal
Science, 373(6557)

ISSN
0036-8075

Authors
Desai, Ravi V
Chen, Xinyue
Martin, Benjamin
et al.

Publication Date
2021-08-20

DOI
10.1126/science.abc6506

Peer reviewed

HHS Public Access

Author manuscript

Science. Author manuscript; available in PMC 2022 August 20.

Published in final edited form as:

Science. 2021 August 20; 373(6557): . doi:10.1126/science.abc6506.

A DNA repair pathway can regulate 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.

Abstract

INTRODUCTION: Fluctuations have long been known to dynamically shape microstate distributions in physical systems. Throughout engineering, “dithering” approaches that modulate

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

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.

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/373/6557/eabc6506/suppl/DC1

Materials and Methods
Supplementary Text
Figs. S1 to S32
Tables S1 to S8
References (80–97)
MDAR Reproducibility Checklist

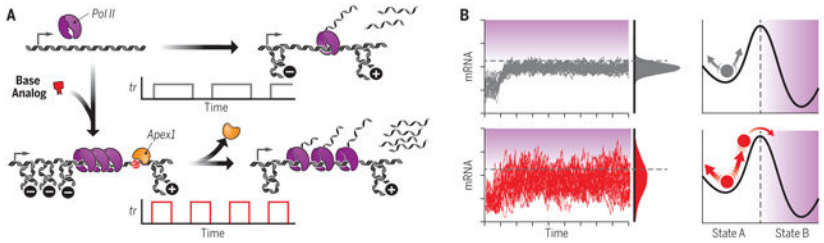
fluctuations are used to enhance inefficient processes and, in chemistry, thermal fluctuations are amplified (e.g., by Bunsen burners) to accelerate reactions. In biology, a long-standing question is whether stochastic expression fluctuations originating from episodic transcription “bursts” play any physiologic role.

RATIONALE: Stochastic fluctuations (noise), measured by the coefficient of variation, scale inversely with mean expression level. For example, transcriptional activators that increase the mean lead to decreased noise, whereas stressors that decrease the mean increase noise. However, this 1/mean “Poisson” scaling of transcriptional noise can be broken by certain processes (e.g., feedback) and, curiously, by small molecules such as pyrimidine nucleobases. We set out to determine the mechanism of action of nucleobases that amplify transcriptional noise independently of mean and explored their potential functional role. Specifically, we tested whether a noise-amplifying pyrimidine nucleotide and its naturally occurring base analogs decouple noise from the mean by disruption of a putative cellular noise control mechanism (i.e., a noise thermostat).

RESULTS: We found that DNA surveillance and repair machinery decouple transcriptional noise from mean expression levels, homeostatically changing noise independently of mean, and this potentiates cell fate transitions in stem cells. Specifically, during removal of modified nucleotide substrates (e.g., idoxuri-dine) and naturally occurring nucleotide analogs [e.g., 5-hydroxymethylcytosine (hmC) and 5-hydroxymethyluridine (hmU)], transcriptional noise is amplified globally across the transcriptome. The amplified transcriptional noise is intrinsic (i.e., not cell extrinsic), independent of changes in the mean (i.e., occurs with minimal change in mean), and distinct from a stress response. Forward genetic screening identified AP endonuclease 1 (Apex1), a member of the base excision repair (BER) DNA surveillance pathway, as the essential mediator of homeostatic noise amplification, and up-regulation of BER enzymes upstream of Apex1 (e.g., glycosylases) also amplified noise. Single-molecule and live-cell imaging showed that this homeostatic noise amplification originated from shorter-duration, but higher-intensity, transcriptional bursts. Mechanistically, Apex1 amplified noise by altering DNA topology, i.e., by increasing negative DNA supercoiling, which impedes transcription but upon repair accelerates transcription, thereby homeostatically maintaining mean expression levels. We call this mechanism “discordant transcription through repair (“DiThR,” pronounced “dither”). Computational modeling predicted that DiThR could increase responsiveness to fate-determining stimuli and, indeed, experimental activation of DiThR potentiated both differentiation of embryonic stem cells into neural ectodermal cells and reprogramming of differentiated fibroblasts into induced pluripotent stem cells.

CONCLUSION: Our data reveal that a DNA surveillance pathway uses the biomechanical link between supercoiling and transcription to homeostatically amplify transcriptional fluctuations. The resulting increase in expression excursions, or outliers, increases cellular responsiveness to diverse fate specification signals. Thus, DNA-processing activities that interrupt transcription could function in fate determination and may explain why naturally occurring base modifications, such as the oxidized nucleobase hmU, are enriched in embryonic stem cell DNA. The existence of a DiThR pathway that orthogonally regulates transcriptional fluctuations suggests that cells evolved mechanisms to exploit noise for the functional regulation of fate transitions and highlights the potential to harness these endogenous pathways for cellular reprogramming.

Graphical Abstract



Cellular “dither”: DiThR amplifies transcriptional fluctuations to facilitate fate transitions. (**A**) Transcription (tr) induces supercoiling (+ and −), but during BER, Apex1 increases supercoiling, impeding transcription. Upon completion of DNA repair and Apex1 removal, transcription is accelerated, generating shorter, more intense transcriptional “bursts.” (**B**) The altered transcriptional bursting amplifies expression noise and facilitates cell fate transitions during development and reprogramming.

Abstract

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 refer to as “discordant transcription through repair” (“DiThR,” which is 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 that 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 $\Delta\text{variance} \neq \Delta\text{mean}$.

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”), which are modulated by gene-regulatory circuits, enable probabilistic fate selection (Fig. 1A) in diverse biological systems (9–12). It is unclear 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 independently of mean and, given the detrimental effects of noise, whether such putative mechanisms might be regulated “on demand” to potentiate cell fate specification.

Nongenetic variability or noise in gene expression, often quantified by the 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) that are initiated by promoter toggling between ON and OFF states (Fig. 1B). The two-state random telegraph model describes this bursting using 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) such as 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 the coefficient of variation (CV, calculated as σ/μ, where σ is the transcriptional fluctuation; 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 σ/μ independently of the change in 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), that consistently increased the 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 medium (2i/LIF), after filtering and normalization using Seurat (27), showed that IdU amplified cell-to-cell variability in transcript abundance (i.e., transcript noise) for virtually all genes across the genome (4578 genes were 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 μ, 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 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 to quantify the 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 4971 genes analyzed, 945 (~20%) were classified as highly variable, whereas 113 (~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, showed that IdU induced a significant increase in transcript variability (expression noise) but comparatively little change in mean expression.

To determine whether certain characteristics could explain a gene's potential for noise enhancement, we examined gene length, promoter AT content, gene body AT content, number of exons, TATA box inclusion, and 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 that gene topology influences susceptibility to noise enhancement (fig. S4). Ontology analysis of highly variable genes showed enrichment of housekeeping pathways, along with pluripotency maintenance factors, particularly Sox2, Oct4, Nanog, and Klf4 (Fig. 1H and fig. S5). Because these pluripotency maintenance factors are key influencers of cell fate specification, we focused on the molecular mechanisms driving their amplified transcript noise.

We investigated whether the enhanced variability arose from extrinsic factors, which included cell cycle phase and cell type identity (32). Cells within the scRNA-seq dataset were computationally assigned a cycle stage (G_1, S, or G_2/M) (33), which showed that Sox2, Oct4, Nanog, and Klf4 were highly variable in each cell cycle phase, indicating that their variability is not cell cycle dependent (fig. S6). Moreover, pseudotime analysis showed no bifurcations, indicating that transcriptional variability was not caused by 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, then the 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. Transcripts were counted in an mESC line in which both endogenous alleles of Nanog are fused to enhanced green fluorescence protein (eGFP) at the C terminus. This fusion does not alter mRNA or protein half-life nor does it impair differentiation potential (37). smRNA-FISH probes to eGFP (the 3′ end of the transcript) were used to count mature transcripts, and probes to the first intron of Nanog (the 5′ end of the 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 (about a twofold 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 had active 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}), whereas the 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}).

Although 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 and/or 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 the IdU-mediated increase in TC intensities results from amplified transcription rates (K_{TX}).

To determine whether 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 about threefold, 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 G_1-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 subpopulation of cells with nonphysiologic 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 posttranscriptional sources of gene expression variability (e.g., mRNA degradation, mRNA translation, and protein degradation).

When cultured in 2i/LIF medium, Nanog protein expression was unimodal and high, but when cultured in serum/LIF medium, mESCs exhibited bimodal expression with both a high-Nanog state and a low-Nanog state that predisposed 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-Nanog 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., to 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 twofold increase in the magnitude (intrinsic-CV^2) of fluctuations (Fig. 2F and fig. S13A), and autocorrelation 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 on the basis of 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 an 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 twofold 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 the 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 the expression of 25 genes involved in nucleoside metabolism and DNA repair using CRISPRi [three guide RNAs (gRNAs)/gene; table S4] and quantified how this affected the noise enhancement of IdU. We identified two genes, AP endonuclease 1 (Apex1) and thymidine kinase 1 (Tk1), the depletion of which abrogated noise enhancement (Fig. 3C). Gene depletion was confirmed by reverse transcription quantitative polymerase chain reaction (fig. S16).

Tk1 adds a requisite gamma-phosphate group to diphosphate nucleotides before 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 of this, a combination of 10 μM 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 the S/G_2 cell cycle phases, when levels of IdU incorporation are highest, displayed increased levels of transcriptional noise enhancement

(fig. S6). The reduction in Nanog noise with the 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 (also known as Ref-1 or Ape1) has a pivotal role in the BER pathway because 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 confirmed that IdU treatment increased Apex1 recruitment to the Nanog promoter (fig. S17B). To determine whether alternate activators of BER also enhanced noise, we subjected cells to oxidative stress [with hydrogen peroxide (H_2O_2)] and alkylation damage stress [with methyl methanesulfonate (MMS)]. Similar to IdU, hydrogen peroxide, and methyl methanesulfonate also enhanced gene expression noise without altering the mean level of expression (fig. S18, A and B). By contrast, cells subjected to ultraviolet (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 specific ability of BER to modulate gene expression noise.

Because BER is initiated by a family of DNA glycosylases that recognize and excise modified bases to create AP sites, we investigated whether perturbation of glycosylases affects gene expression noise. Individual depletion of either uracil-DNA glycosylase or thymine-DNA glycosylase failed to ablate IdU noise enhancement (fig. S19A), presumably because of the overlapping and compensatory action of glycosylase family members in base removal (51, 52). However, overexpression of either uracil-DNA glycosylase or methylpurine-DNA glycosylase 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 (knock out) 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 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 whether Apex1 binding rather than enzymatic activity was responsible for noise enhancement, we expressed a catalytically inactive mutant of Apex1 (56) in cells that had endogenous Apex1 depleted. We found

that catalytically inactive Apex1 partially rescued IdU-mediated noise enhancement (Fig. 3E). These data, together with evidence that supercoiling sets mechanical bounds on transcriptional bursting (57, 58), drove us to investigate whether Apex1 recruitment affects supercoiling.

To measure supercoiling levels, we used a psoralen–cross-linking 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 before bTMP incubation (60). IdU treatment significantly increased genomic supercoiling, as demonstrated by an approximately twofold 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 dimethyl sulfoxide (DMSO) control level, indicating that IdU alone in uncoiled DNA does not increase intercalation (fig. S20).

If DNA topology influences transcriptional bursting, then 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 Top1 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). Together with psoralen–cross-linking 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 nonproductive state (ON*), whereas unbinding of Apex1 allows mRNA production to resume with an amplified transcription rate that is proportional to time spent in the nonproductive state (fig. S25 and supplementary text 6); that is, the longer the residence time in the nonproductive 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). Consistent with this hypothesis, the model

accurately predicted noise enhancement mediated by topoisomerase inhibition (fig. S24C and supplementary text 6). The ability to render a gene transcriptionally nonproductive 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 call this model, and the associated phenomenon, “discordant transcription through repair” (“DiThR,” pronounced “dither”) because of the large discordance in pre-repair versus 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 there are parameter regimes where noise can be regulated independently 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 among these nucleoside analogs may reflect known differences in incorporation rates (69). Sensitivity analysis indicated that in genes for which $K_{OFF} >> K_{ON}$ (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, because all 98 of the up-regulated genes resided 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 the maintenance of gene expression homeostasis across a broad range of genes (fig. S29 and supplementary text 8).

We next investigated whether amplified transcriptional noise could improve responsiveness to external stimuli (e.g., fate specification signals), as was 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 verify these predictions experimentally, we tested whether IdU could potentiate both differentiation of mESCs into the neural ectodermal lineage and reprogramming of differentiated cells into pluripotent stem cells. To assess potentiation of differentiation, mESCs cultured in neural ectodermal–specifying medium were stained for CD24, an established marker of the neural ectoderm lineage (74, 75). Addition of 4 μM IdU for the

first 48 hours of differentiation yielded about a fourfold and twofold increase in CD24(+) cells at 48 hours and 96 hours into differentiation, respectively (fig. S30).

To look for potentiation of dedifferentiation, three cellular reprogramming systems were used. The first assay used 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 medium 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 accelerated activation of the pluripotency program. The second reprogramming assay used Oct4-GFP primary MEFs transduced with retroviral vectors expressing Oct4, Sox2, Klf4, and c-Myc (OKSM). IdU supplementation for the 48 hours immediately after transduction caused an ~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 by about fourfold (Fig. 4G and fig. S32C). Short hairpin RNA–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 use the biomechanical link between supercoiling and transcription to homeostatically enhance noise without altering the 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 interesting implications for the role of naturally occurring oxidized nucleobases (e.g., hmU) in cell fate determination, particularly because these base modifications are found at higher frequencies in ESC DNA (45). Mechanistic insights 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 because they will exhibit increased mean). Additionally, genes most susceptible to transcriptional noise enhancement tend to lie far from TAD boundaries. Because TAD boundaries largely overlap with supercoiling domain boundaries (76), and transcription-induced supercoiling may directly contribute to the formation of TADs (77), we reasoned 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 have either 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.

Methods summary

Quantification of cell-to-cell variability in gene expression was performed using the following techniques: (i) single-cell RNA-seq, (ii) single-molecule RNA FISH, (iii) live-cell imaging of RNA transcription with p21-MS2 reporter cell line, (iv) flow cytometry, and (v) live-cell imaging of Nanog-GFP protein expression. For scRNA-seq, mESCs treated with DMSO or 10 μM IdU for 24 hours were prepared for sequencing using 10× genomics specifications. Quality control, normalization, and variability analysis of scRNA-seq data were performed using two packages: Seurat and BASiCS. For smRNA-FISH, probes for the first exon, first intron, and 3′ GFP fusion of the Nanog transcript were developed using the designer tool from Stellaris. Nanog-GFP mESCs treated with either DMSO or 10 μM IdU were stained with Nanog mRNA probes and imaged on a Zeiss spinning-disk microscope. RNA spot counting and transcriptional center analysis were performed with FISH-quant. For live-cell imaging of transcription, p21-MS2 reporter U2OS cells were pretreated with Nutlin-3 and either DMSO or 10 μM IdU for 48 hours before imaging. Imaging was performed on a wide-field Olympus microscope for 118 min with 1-min intervals between frames. The cumulative transcription occurring in each cell was calculated based on the normalized transcriptional activity over 118 minutes of imaging. Flow cytometry was used to quantify single-cell variability in protein expression. To filter out gene expression variability arising from cell size heterogeneity, the smallest possible forward and side scatter region containing at least 3000 cells was used to isolate cells of similar size and shape for all analyses. To quantify single-cell fluctuations in Nanog protein expression over time, live-cell time lapse microscopy of Nanog-GFP mESCs was performed using a Zeiss spinning-disk microscope with imaging commencing immediately after addition of either DMSO or 10 μM IdU. Cell segmentation, tracking, and GFP quantification were performed using CellProfiler. Detrended fluorescence trajectories were used for noise autocorrelation and noise magnitude calculations.

CRISPRi screening for genetic controllers of transcriptional noise was performed in an arrayed fashion with Nanog-GFP mESCs stably expressing dCas9-Krab::mCherry. A total of 25 genes (three gRNAs/gene) were depleted through individual transduction of cells with gRNA lentiviral constructs harboring a blue fluorescent protein reporter. Forty-eight hours after infection, each population of cells expressing a unique gRNA was treated with either DMSO or 10 μM IdU for 24 hours followed by flow cytometric analysis of Nanog-GFP expression.

Activation of the BER pathway was performed by chemical treatment of mESCs (H_2O_2 and MMS) and overexpression of DNA glycosylases (Mpg and Ung). Nanog-GFP mESCs were treated with H_2O_2 and MMS for 1 hour and 24 hours, respectively, before flow cytometric analysis. For overexpression of DNA glycosylases, Nanog-GFP mESCs were transduced with lentiviral constructs harboring doxycycline-inducible cassettes for either Mpg or Ung with an mCherry reporter. Flow-cytometric analysis of Nanog-GFP expression was performed on transduced cells after 24 hours of doxycycline induction. Measurement of

negative supercoiling in mESCs was performed with bTMP staining. Before bTMP staining, mESCs were treated with 1 μM aphidicolin for 2 hours to remove the confounding effect of DNA replication on genomic supercoiling. After bTMP staining, UV cross-linking was performed for 15 minutes using 365 nm light. Imaging was performed on a Zeiss spinning-disk microscope. Nuclear segmentation using 4′,6-diamidino-2-phenylindole (DAPI) signal and single-cell quantification of bTMP staining intensity were performed using CellProfiler.

To elucidate the effect of Apex1 recruitment on transcriptional bursting, five computational models of increasing complexity were constructed based on the random-telegraph model of transcription. For each model, an associated stochastic reaction scheme was numerically solved using Gillespie's stochastic simulation algorithm to identify the model that best recapitulated experimental data. The effective kinetic rates of Nanog transcription in the control (DMSO) condition from smRNA-FISH data were used as the starting point for all simulations and were considered constant for all models. For each model, k_{repair} was the single degree of freedom. Identification of the model that best fits experimental data was based on maximum likelihood estimation of k_{repair} for each model, followed by minimization of Akaike information criterion.

To assess whether increased fluctuations in gene expression can promote cell fate transitions, three cellular reprogramming assays were tested: (i) Nanog-GFP secondary MEFs harboring stably integrated doxycycline-inducible cassettes for Oct4, Sox2, and Klf4; (ii) Oct4-GFP primary MEFs transduced with lentiviral vectors encoding Oct4, Sox2, Klf4, and c-Myc; and (iii) Oct4-GFP MEFs harboring stably integrated doxycycline-inducible cassettes for Oct4, Sox2, Klf4, and c-Myc. For each system, IdU-mediated noise enhancement was implemented for the first 48 hours of reprogramming. Induced pluripotent stem cell formation was assessed using alkaline phosphatase staining and flow-cytometric measurement of fluorescent reporter levels. A detailed account of all methods used in this study is provided in the supplementary materials.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank M. Simpson, B. Bruneau, J. Weissman, G. Balazsi, and members of the Weinberger laboratory for thoughtful discussions and suggestions; K. Claiborn for editing; G. Maki for graphics support; N. Raman in the Gladstone Institute Flow Cytometry Facility (NIH S10 RR028962, P30 AI027763, DARPA, and the James B. Pendleton Charitable Trust) for technical assistance; the Gladstone Assay Development and Drug Discovery Core for technical assistance with drug screening; K. Thorn and D. Larson in the UCSF Nikon Imaging Center (NIH S10 1S10OD017993-01A1) for technical assistance with imaging; M. Jost and J. Weissman for CRISPRi reagents; and the Gladstone Institute Genomics Core for technical assistance with single-cell RNA-sequencing. The dual-tagged Sox2 mESCs were a kind donation from B. Bruneau and E. Nora. The Oct4-GFP reprogrammable MEFs (harbor stably integrated OKSM factors) were a kind donation from S. 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 a Dutch Research Council (NWO) ENW-XS award (OCENW.XS3.055). L.S.W. acknowledges support from a Bowes Distinguished Professorship, Alfred P. Sloan Research Fellowship,

Pew Scholars in the Biomedical Sciences Program, NIH award R01AI109593, and the NIH Director's New Innovator Award (OD006677) and Pioneer Award (OD17181) programs.

Data and materials availability:

The raw and processed sequencing data reported herein have been deposited onto the Gene Expression Omnibus under 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 are archived on Zenodo (79). Reagents, including plasmids and cell lines, are available from the corresponding author upon request.

REFERENCES AND NOTES

1. Boltzmann L, Weitere Studien uber das Wärmegleichgewicht unter Gasmolekulen, Sitzungsber. Kais. Akad. Wiss. Wien Math. Naturwiss 66, 275–370 (1872).
2. Arrhenius S, Über die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Säuren. Z. Phys. Chem 4, 226–248 (1889). doi: 10.1515/zpch-1889-0416
3. Roberts L, Picture coding using pseudo-random noise. IEEE Trans. Inf. Theory 8, 145–154 (1962). doi: 10.1109/TIT.1962.1057702
4. Fatt P, Katz B, Some observations on biological noise. Nature 166, 597–598 (1950). doi: 10.1038/166597a0; [PubMed: 14780165]
5. Priplata AA, Niemi JB, Harry JD, Lipsitz LA, Collins JJ, Vibrating insoles and balance control in elderly people. Lancet 362, 1123–1124 (2003). doi: 10.1016/S0140-6736(03)14470-4; [PubMed: 14550702]
6. Cohen D, Optimizing reproduction in a randomly varying environment. J. Theor. Biol 12, 119–129 (1966). doi: 10.1016/0022-5193(66)90188-3; [PubMed: 6015423]
7. Arkin A, Ross J, McAdams HH, 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; [PubMed: 9691025]
8. Spudich JL, Koshland DE Jr., Non-genetic individuality: Chance in the single cell. Nature 262, 467–471 (1976). doi: 10.1038/262467a0; [PubMed: 958399]
9. Gupta PB et al. , 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; [PubMed: 21854987]
10. Weinberger LS, Burnett JC, Toettcher JE, Arkin AP, Schaffer DV, 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; [PubMed: 16051143]
11. Raj A, van Oudenaarden A, Nature, nurture, or chance: Stochastic gene expression and its consequences. Cell 135, 216–226 (2008). doi: 10.1016/j.cell.2008.09.050; [PubMed: 18957198]
12. Chubb JR, Trcek T, Shenoy SM, Singer RH, Transcriptional pulsing of a developmental gene. Curr. Biol 16, 1018–1025 (2006). doi: 10.1016/j.cub.2006.03.092; [PubMed: 16713960]
13. Golding I, Paulsson J, Zawilski SM, Cox EC, Real-time kinetics of gene activity in individual bacteria. Cell 123, 1025–1036 (2005). doi: 10.1016/j.cell.2005.09.031; [PubMed: 16360033]
14. Kepler TB, Elston TC, Stochasticity in transcriptional regulation: Origins, consequences, and mathematical representations. Biophys. J 81, 3116–3136 (2001). doi: 10.1016/S0006-3495(01)75949-8; [PubMed: 11720979]
15. Raj A, Peskin CS, Tranchina D, Vargas DY, Tyagi S, Stochastic mRNA synthesis in mammalian cells. PLOS Biol 4, e309 (2006). doi: 10.1371/journal.pbio.0040309; [PubMed: 17048983]
16. Blake WJ, KAErn M, Cantor CR, Collins JJ, Noise in eukaryotic gene expression. Nature 422, 633–637 (2003). doi: 10.1038/nature01546; [PubMed: 12687005]
17. Peccoud J, Ycart B, Markovian modeling of gene-product synthesis. Theor. Popul. Biol 48, 222–234 (1995). doi: 10.1006/tpbi.1995.1027

18. Rodriguez J, Larson DR, Transcription in living cells: Molecular mechanisms of bursting. Annu. Rev. Biochem 89, 189–212 (2020). doi: 10.1146/annurev-biochem-011520-105250; [PubMed: 32208766]
19. Nicolas D, Zoller B, Suter DM, Naef F, Modulation of transcriptional burst frequency by histone acetylation. Proc. Natl. Acad. Sci. U.S.A 115, 7153–7158 (2018). doi: 10.1073/pnas.1722330115; [PubMed: 29915087]
20. Bar-Even A et al. , Noise in protein expression scales with natural protein abundance. Nat. Genet 38, 636–643 (2006). doi: 10.1038/ng1807; [PubMed: 16715097]
21. Dar RD et al. , Transcriptional bursting explains the noise-versus-mean relationship in mRNA and protein levels. PLOS ONE 11, e0158298 (2016). doi: 10.1371/journal.pone.0158298; [PubMed: 27467384]
22. Newman JRS et al. , Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. Nature 441, 840–846 (2006). doi: 10.1038/nature04785; [PubMed: 16699522]
23. Alon U, An Introduction to Systems Biology: Design Principles of Biological Circuits (Chapman & Hall/CRC, 2007).
24. Dar RD, Hosmane NN, Arkin MR, Siliciano RF, Weinberger LS, Screening for noise in gene expression identifies drug synergies. Science 344, 1392–1396 (2014). doi: 10.1126/science.1250220; [PubMed: 24903562]
25. Hansen MMK et al. , 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; [PubMed: 29754821]
26. Li Y et al. , Noise-driven cellular heterogeneity in circadian periodicity. Proc. Natl. Acad. Sci. U.S.A 117, 10350–10356 (2020). doi: 10.1073/pnas.1922388117; [PubMed: 32358201]
27. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R, Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat. Biotechnol 36, 411–420 (2018). doi: 10.1038/nbt.4096; [PubMed: 29608179]
28. Hansen MMK, Desai RV, Simpson ML, Weinberger LS, 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; [PubMed: 30243562]
29. Munsky B, Neuert G, van Oudenaarden A, Using gene expression noise to understand gene regulation. Science 336, 183–187 (2012). doi: 10.1126/science.1216379; [PubMed: 22499939]
30. Suter DM et al. , Mammalian genes are transcribed with widely different bursting kinetics. Science 332, 472–474 (2011). doi: 10.1126/science.1198817; [PubMed: 21415320]
31. Eling N, Richard AC, Richardson S, Marioni JC, Vallejos CA, 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; [PubMed: 30172840]
32. Kaern M, Elston TC, Blake WJ, Collins JJ, Stochasticity in gene expression: From theories to phenotypes. Nat. Rev. Genet 6, 451–464 (2005). doi: 10.1038/nrg1615; [PubMed: 15883588]
33. Scialdone A et al. , Computational assignment of cell-cycle stage from single-cell transcriptome data. Methods 85, 54–61 (2015). doi: 10.1016/j.ymeth.2015.06.021; [PubMed: 26142758]
34. Pedraza JM, van Oudenaarden A, Noise propagation in gene networks. Science 307, 1965–1969 (2005). doi: 10.1126/science.1109090; [PubMed: 15790857]
35. Bargaje R et al. , 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; [PubMed: 28167799]
36. Mojtahedi M et al. , Cell fate decision as high-dimensional critical state transition. PLOS Biol 14, e2000640 (2016). doi: 10.1371/journal.pbio.2000640; [PubMed: 28027308]
37. Sokolik C et al. , 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; [PubMed: 26405695]
38. Carvajal LA et al. , Dual inhibition of MDMX and MDM2 as a therapeutic strategy in leukemia. Sci. Transl. Med 10, eaao3003 (2018). doi: 10.1126/scitranslmed.aao3003; [PubMed: 29643228]

39. Abranches E et al. , Stochastic NANOG fluctuations allow mouse embryonic stem cells to explore pluripotency. Development 141, 2770–2779 (2014). doi: 10.1242/dev.108910; [PubMed: 25005472]
40. Kalmar T et al. , Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells. PLOS Biol 7, e1000149 (2009). doi: 10.1371/journal.pbio.1000149; [PubMed: 19582141]
41. Austin DW et al. , Gene network shaping of inherent noise spectra. Nature 439, 608–611 (2006). doi: 10.1038/nature04194; [PubMed: 16452980]
42. Rosenfeld N, Young JW, Alon U, Swain PS, Elowitz MB, Gene regulation at the single-cell level. Science 307, 1962–1965 (2005). doi: 10.1126/science.1106914; [PubMed: 15790856]
43. Sigal A et al. , Variability and memory of protein levels in human cells. Nature 444, 643–646 (2006). doi: 10.1038/nature05316; [PubMed: 17122776]
44. Ito S et al. , 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; [PubMed: 20639862]
45. Pfaffeneder T et al. , Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. Nat. Chem. Biol 10, 574–581 (2014). doi: 10.1038/nchembio.1532; [PubMed: 24838012]
46. He YF et al. , Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science 333, 1303–1307 (2011). doi: 10.1126/science.1210944; [PubMed: 21817016]
47. Shen L et al. , 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; [PubMed: 23602152]
48. Arnér ESJ, Eriksson S, Mammalian deoxyribonucleoside kinases. Pharmacol. Ther 67, 155–186 (1995). doi: 10.1016/0163-7258(95)00015-9; [PubMed: 7494863]
49. Demple B, Herman T, Chen DS, 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; [PubMed: 1722334]
50. Lindahl T, Wood RD, Quality control by DNA repair. Science 286, 1897–1905 (1999). doi: 10.1126/science.286.5446.1897; [PubMed: 10583946]
51. Müller U, Bauer C, Siegl M, Rottach A, Leonhardt H, 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; [PubMed: 24948610]
52. Hegde ML, Hazra TK, Mitra S, 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; [PubMed: 18166975]
53. Madhusudan S et al. , Isolation of a small molecule inhibitor of DNA base excision repair. Nucleic Acids Res 33, 4711–4724 (2005). doi: 10.1093/nar/gki781; [PubMed: 16113242]
54. Kladova OA et al. , The role of the N-terminal domain of human apurinic/apyrimidinic endonuclease 1, APE1, in DNA glycosylase stimulation. DNA Repair (Amst.) 64, 10–25 (2018). doi: 10.1016/j.dnarep.2018.02.001; [PubMed: 29475157]
55. Mol CD, Izumi T, Mitra S, Tainer JA, 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; [PubMed: 10667800]
56. McNeill DR, Wilson DM 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; [PubMed: 17259346]
57. Chong S, Chen C, Ge H, Xie XS, Mechanism of transcriptional bursting in bacteria. Cell 158, 314–326 (2014). doi: 10.1016/j.cell.2014.05.038; [PubMed: 25036631]
58. Sevier SA, Kessler DA, Levine H, Mechanical bounds to transcriptional noise. Proc. Natl. Acad. Sci. U.S.A 113, 13983–13988 (2016). doi: 10.1073/pnas.1612651113; [PubMed: 27911801]
59. Corless S, Gilbert N, Investigating DNA supercoiling in eukaryotic genomes. Brief. Funct. Genomics 16, 379–389 (2017). doi: 10.1093/bfgp/elx007; [PubMed: 28444308]

60. Babos KN et al. , 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; [PubMed: 31523028]
61. Guptasarma P, Cooperative relaxation of supercoils and periodic transcriptional initiation within polymerase batteries. BioEssays 18, 325–332 (1996). doi: 10.1002/bies.950180411; [PubMed: 8967901]
62. Kim S, Beltran B, Irnov I, Jacobs-Wagner C, 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; [PubMed: 31539491]
63. Kouzine F, Sanford S, Elisha-Feil Z, Levens D, The functional response of upstream DNA to dynamic supercoiling in vivo. Nat. Struct. Mol. Biol 15, 146–154 (2008). doi: 10.1038/nsmb.1372; [PubMed: 18193062]
64. Liu J et al. , 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; [PubMed: 16628215]
65. Mizutani M, Ohta T, Watanabe H, Handa H, Hirose S, 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; [PubMed: 1992462]
66. Teves SS, Henikoff S, Transcription-generated torsional stress destabilizes nucleosomes. Nat. Struct. Mol. Biol 21, 88–94 (2014). doi: 10.1038/nsmb.2723; [PubMed: 24317489]
67. Bazlekowa-Karaban M et al. , 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; [PubMed: 31518879]
68. Breit JF, Ault-Ziel K, Al-Mehdi A-B, Gillespie MN, 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; [PubMed: 17766324]
69. Leuner B, Glasper ER, Gould E, Thymidine analog methods for studies of adult neurogenesis are not equally sensitive. J. Comp. Neurol 517, 123–133 (2009). doi: 10.1002/cne.22107; [PubMed: 19731267]
70. Delmans M, Hemberg M, 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; [PubMed: 26927822]
71. Blake WJ et al. , Phenotypic consequences of promoter-mediated transcriptional noise. Mol. Cell 24, 853–865 (2006). doi: 10.1016/j.molcel.2006.11.003; [PubMed: 17189188]
72. Chang HH, Hemberg M, Barahona M, Ingber DE, Huang S, Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. Nature 453, 544–547 (2008). doi: 10.1038/nature06965; [PubMed: 18497826]
73. Süel GM, Kulkarni RP, Dworkin J, Garcia-Ojalvo J, Elowitz MB, Tunability and noise dependence in differentiation dynamics. Science 315, 1716–1719 (2007). doi: 10.1126/science.1137455; [PubMed: 17379809]
74. Pruszak J, Ludwig W, Blak A, Alavian K, Isacson O, 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; [PubMed: 19725119]
75. Semrau S et al. , 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; [PubMed: 29061959]
76. Naughton C et al. , Transcription forms and remodels supercoiling domains unfolding large-scale chromatin structures. Nat. Struct. Mol. Biol 20, 387–395 (2013). doi: 10.1038/nsmb.2509; [PubMed: 23416946]
77. Racko D, Benedetti F, Dorier J, Stasiak A, 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; [PubMed: 29140466]

78. Singh A, 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; [PubMed: 25418106]

79. Desai R, Martin B, scRNA-seq and modeling code for: A DNA-repair pathway can regulate transcriptional noise to promote cell fate transitions, Zenodo (2021); doi: 10.5281/zenodo.4891977

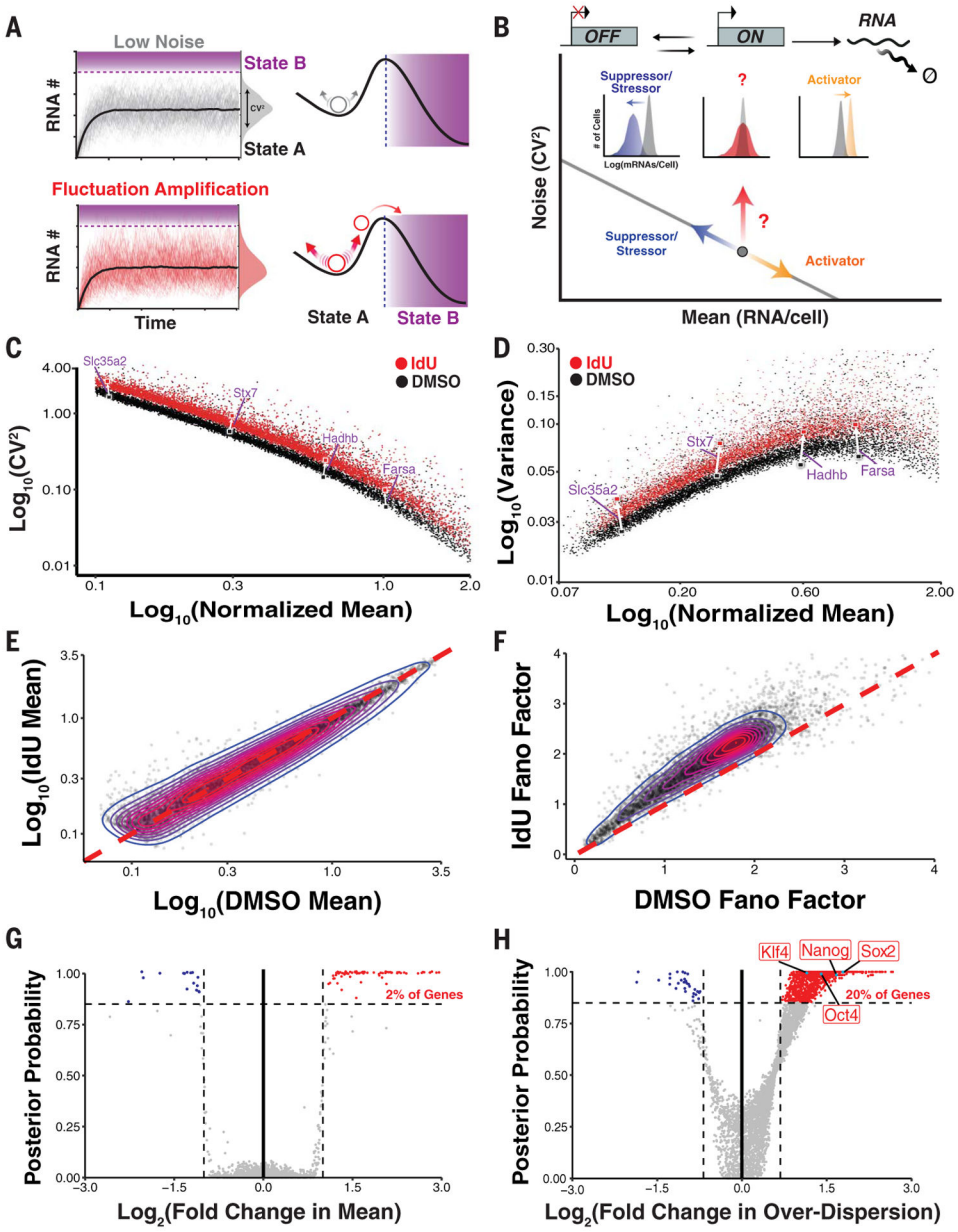


Fig. 1. Genome-wide amplification of cell-to-cell mRNA variability (noise) independently of mean.

(**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) Predicted facilitation of state transitions through dithering. (**B**) (Top) Schematic of two-state random-telegraph model of transcription. (Bottom) Schematic of mean versus CV^2 for mRNA abundance. Solid gray line indicates Poisson, inverse scaling of CV^2 as a function of mean. The 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 10 μM IdU (red) for 24 hours. A total of 812 and 744 transcriptomes from DMSO and

IdU treatments, respectively, were analyzed. (C) Mean expression versus CV^2 and (D) mean versus variance. Four examples of housekeeping genes (purple) demonstrate how IdU increases expression fluctuations with minimal change in mean (white arrows). (E and F) Mean expression (E) and Fano factor (F) (σ^2/μ) of genes in DMSO versus IdU treatments. (**G** and **H**) BASiCS analysis of scRNA-seq data. (G) Fold change in mean versus certainty (posterior probability) that a gene is up- or down-regulated. With IdU treatment, 113 genes (red) were classified as differentially expressed (more than a twofold change in mean with >85% probability). (H) Fold change in overdispersion versus certainty (posterior probability) that gene is highly or lowly variable. A total of 945 genes (red) were classified as highly variable (>1.5-fold change in overdispersion with >85% probability).

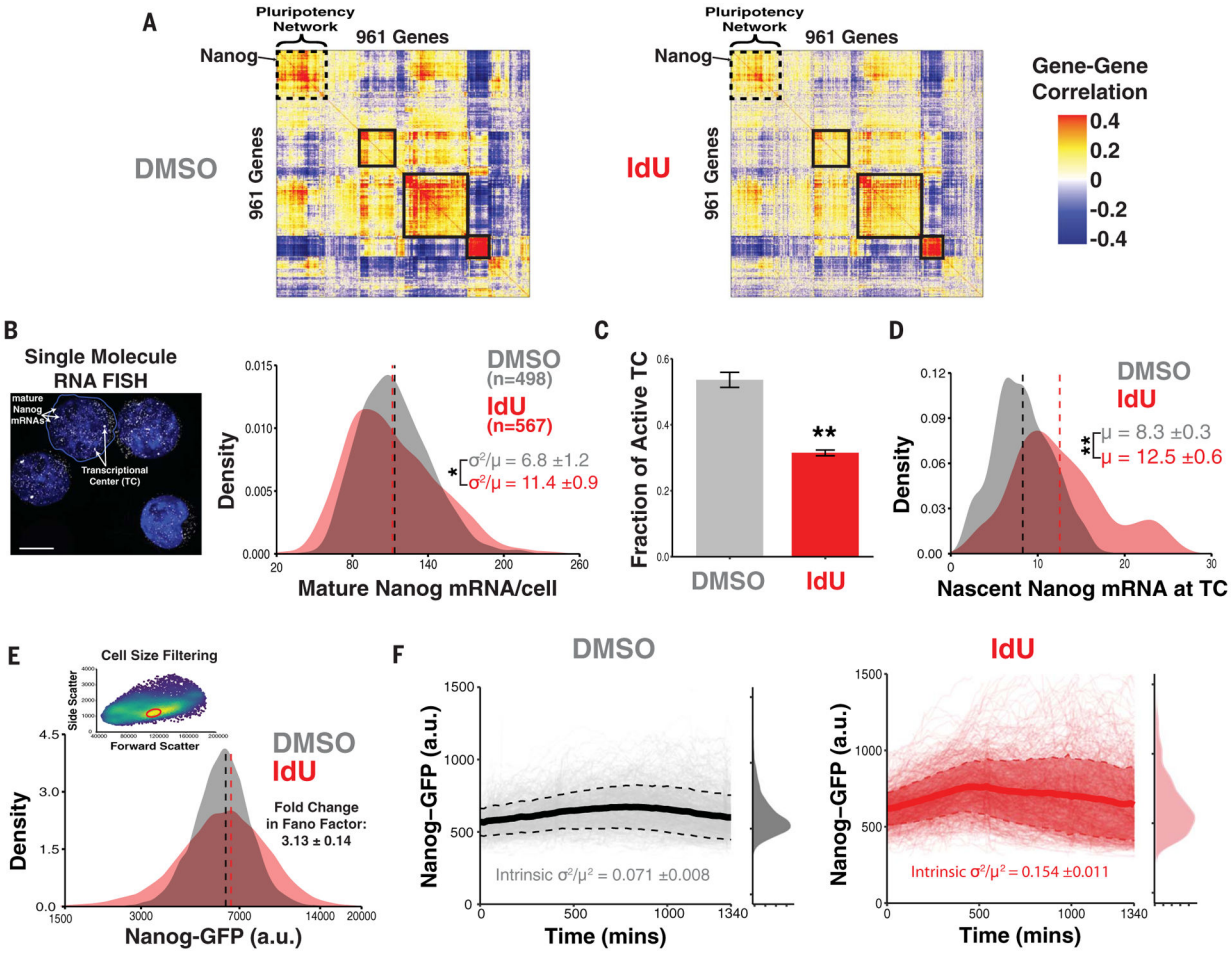


Fig. 2. Amplification of mRNA noise is not caused by extrinsic sources, results from shorter but more intense transcriptional bursts, and propagates to protein levels.

(**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 such as Nanog. (**B** to **D**) Results of smRNA-FISH used to count nascent and mature Nanog mRNA in Nanog-GFP mESCs treated with DMSO or 10 μM IdU for 24 hours in 2i/LIF medium. 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 TCs as verified by intron probe set. Scale bar, 5 μm. (Right) Distributions of mature Nanog transcripts per cell. Dashed lines represent mean. Averaged Fano factors over all four replicates are reported (±SD), $*P = 0.0011$, two-tailed, unpaired Student's *t* test. (C) Fraction of possible TCs that are active as detected by overlap of signal in exon and intron probe channels. Each cell is assumed to have two possible TCs. Data represent mean ± SD, $**P = 6.9 \times 10^{-5}$, 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, $**P = 1.0 \times 10^{-4}$, two-tailed, unpaired Student's *t* test. (**E**) Representative flow cytometry distribution of Nanog-GFP expression in mESCs treated with DMSO or 10 μM IdU for 24 hours in

2i/LIF medium. 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 10 μM IdU (n = 1414) in 2i/LIF medium. Trajectories from two replicates of each condition are pooled, with solid and dashed lines representing mean and SD of trajectories, respectively. Distributions of Nanog-GFP represent expression at the final time point. Intrinsic CV^2 of each detrended trajectory was calculated, with the average (±SD) of all trajectories reported.

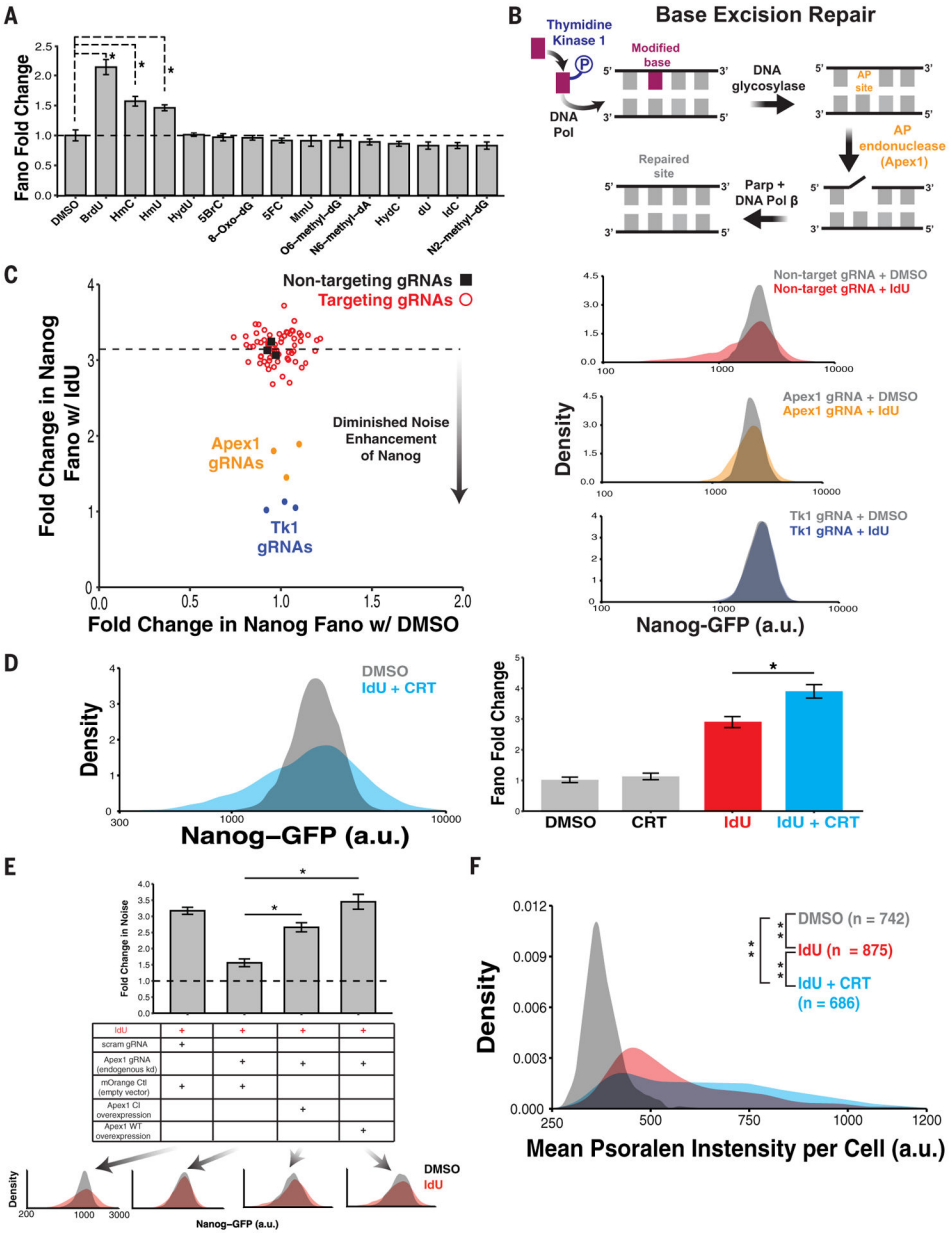


Fig. 3. Noise amplification independent of mean is caused by Apex1-mediated DNA repair.

(**A**) Screening of 14 additional nucleoside analogs. Nanog-GFP mESCs grown in 2i/LIF medium were supplemented with a 10 μM concentration of nucleoside analog for 24 hours. Fano factor for Nanog protein expression was normalized to DMSO. Data represent mean (±SD) of biological replicates, $*P < 0.01$, Kruskal-Wallis test followed by Tukey's multiple comparisons test. (**B**) Schematic of nucleoside analog incorporation into genomic DNA and removal through the BER 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 blue fluorescent protein reporter. A total of 75 gRNAs (25 genes, with three gRNAs/gene) were tested, in addition to three nontargeting control gRNAs. Two days after transduction, each

gRNA-expressing population of mESCs was treated with DMSO or 10 μM IdU for 24 hours in 2i/LIF medium. The Nanog Fano factor for DMSO and IdU treatment of each gRNA population was normalized to the Nanog Fano factor for the nontargeting gRNA + DMSO population. Each point represents a gRNA. Dashed horizontal line represents average noise enhancement of Nanog from IdU in the background of nontargeting 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 nontargeting (top right), Apex1 (middle right), or Tk1 (bottom right) gRNAs and treated with DMSO or 10 μM 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 10 μM IdU + 100 μM CRT0044876. (Right) mESCs were treated with DMSO, 100 μM CRT0044876, 10 μM IdU, or 10 μM IdU + 100 μM CRT0044876 for 24 hours in 2i/LIF medium. The Nanog Fano factor for each treatment was normalized to the DMSO control. Data represent mean ± SD of three biological replicates, $*P = 0.0028$, 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 the rectangular grid. An mOrange empty vector was used as a transduction control. The Nanog Fano factor for each treatment was normalized to mOrange control cells treated with DMSO. Data represent mean ± SD of three biological replicates, $*P < 0.005$, 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 the psoralen–cross-linking assay. mESCs were treated with DMSO, 10 μM IdU, or 10 μM IdU + 100 μM CRT0044876 for 24 hours in 2i/LIF medium. Distributions for nuclear intensities of bTMP staining are shown. Data are pooled from two biological replicates of each treatment, $**P < 0.0001$, Kruskal-Wallis test followed by Tukey's multiple comparisons test.

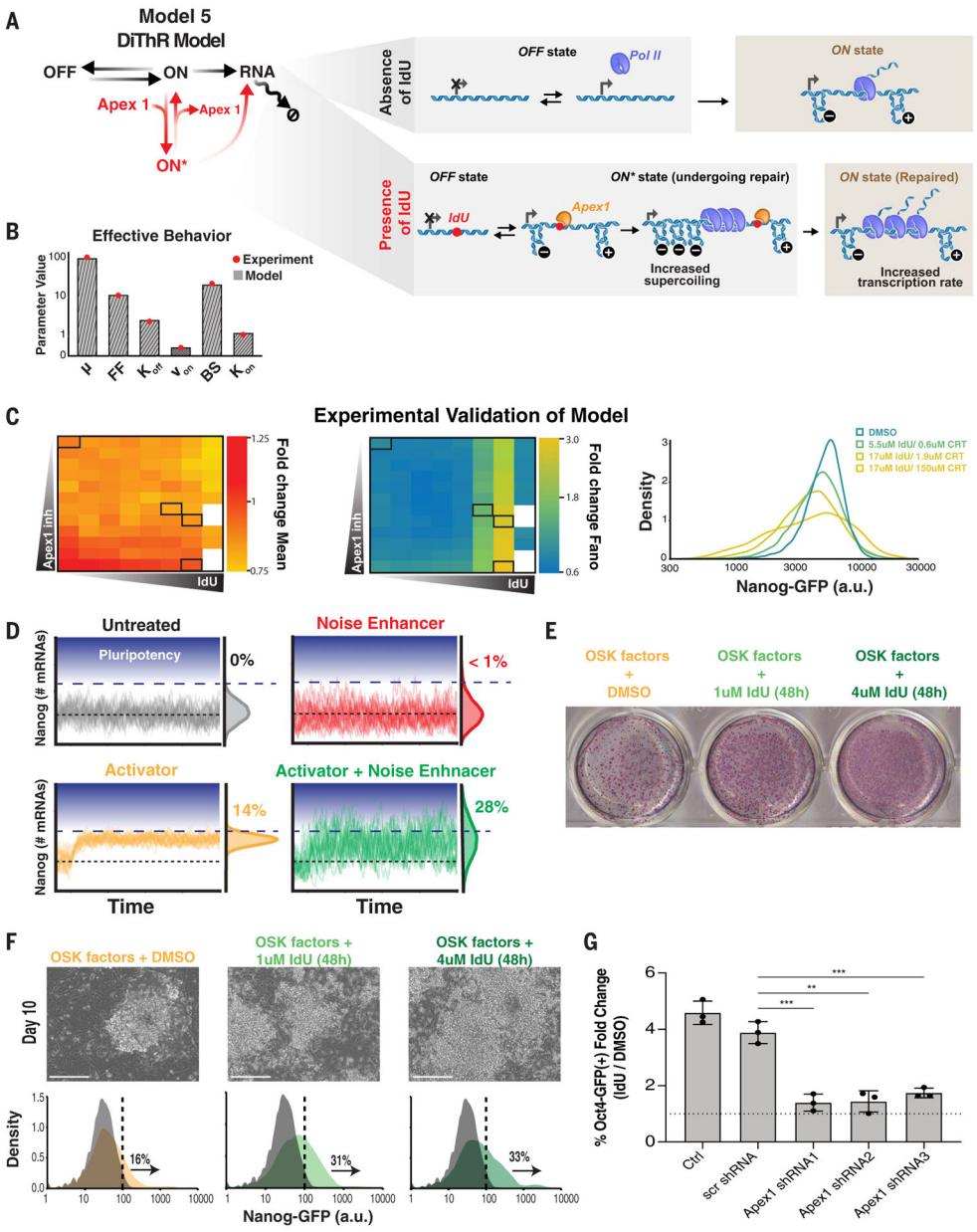


Fig. 4. Mechanistic model for DiThR and the phenotypic consequence of DiThR on potentiation of cellular reprogramming.

(**A**) Detailed schematic of model 5 (the DiThR model) (see fig. S22 for a schematic of models 1 to 4), which uses transcription-coupled base excision. In the presence of IdU (bottom panel), Apex1 binding occurs when the gene is transcriptionally permissive (i.e., in the ON state). Binding induces negative supercoiling, which lengthens the time that a gene is transcriptionally nonproductive (i.e., in the 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**) Effective behavior [μ, Fano factor (FF), K_{OFF}, fraction of time active (ν_{on}), burst size (BS), K_{ON}] of the DiThR model is compared with experimentally derived values of each parameter (red dots) obtained from

smRNA-FISH data. Absolute percentage error was calculated as described in supplementary text 5.2.2. Model 5 (the DiThR model) best matches experimental data. (**C**) Testing of 96 concentration combinations of IdU and CRT0044876 (Apex1 inh) to validate tunability of Nanog variability. IdU and CRT0044876 were used to increase binding and decrease unbinding of Apex1, respectively. Data represent the average of two biological replicates. Left and center panels are 96-well heatmaps displaying the fold change in Nanog mean and Fano factor for each drug combination compared with DMSO (top left well). An insufficient number of cells for extrinsic noise filtering (<50,000) was recorded from white wells. (Right 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), 1 μM IdU (second well), or 4 μM 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; 4 μM IdU or an equivalent volume DMSO was present for first 48 hours of reprogramming. Data represent mean ± SD of three biological replicates, $^{**}P = 0.0014$, $^{***}P < 0.001$, two-tailed, unpaired Student's *t* test.