Analog epigenetic cell memory by graded DNA methylation

Sebastian Palacios[1,2,3†], Simone Bruno[1,4†], Ron Weiss[2,5], Elia Salibi[1], Andrew Kane[2], Katherine Ilia[2,3], and Domitilla Del Vecchio[1,2*]

[1] Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

[2] Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

[3] Institute for Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

[4] Department of Data Science, Dana-Farber Cancer Institute, Boston, MA, 02115, USA.

[5] Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

[†] These authors contributed equally to this work

[*] Corresponding author: ddv@mit.edu

Summary. Long-term maintenance of gene expression states is required to conserve phenotypic differences among cell types [1] and is often compromised in disease [2] and aging [3]. The state of a gene where transcription is completely inhibited is preserved by chromatin modifications such as DNA methylation and histone H3 lysine 9 trimethylation (H3K9me3) [4, 5]. But how any other gene expression state may be stably maintained remains uncharted. Here, we demonstrate that cells can maintain long-term memory of all gene expression levels, a principle that we call analog epigenetic memory. We found through time-course analysis of gene expression and DNA methylation with clonal resolution that the grade of DNA methylation in the gene's promoter remains stable over time and inversely correlates with gene expression level. Furthermore, we observed that H3K9me3 is unable to propagate gene expression states without DNA methylation and altering the latter hampers gene expression level maintenance. Analysis of the chromatin state further revealed that DNA methylation mediates the recruitment of H3K9me3, but the latter does not mediate the former. Lack of this mediation is necessary for analog memory according to our nucleosome modification model. These results show that DNA methylation grade is the causal determinant of analog epigenetic cell memory. Our study uncovers the molecular mechanisms that enable cells to keep fine-grained differences in gene expression from the same genetic sequence, providing foundations for both understanding and engineering epigenetic cellular memory.

Introduction

Epigenetic cell memory allows cells to maintain different gene expression states despite a common genetic code [1]. In addition to being required for conserving phenotypic differences among distinct cell types [6], epigenetic memory is critical for the immune system [7], for the brain's function [8], and it could bring new tools for mammalian cell engineering [9]. Covalent modifications to DNA and histones have been implicated in epigenetic cell memory [10–13] because they affect gene expression by altering transcription [14–18] and can propagate through subsequent cell divisions [14, 19, 20]. In particular, silenced gene expression states characterized by complete lack of transcription can be stably maintained by histone H3 lysin 9 tri-methylation (H3K9me3) [4], DNA methylation [5], or histone H3 lysine 27 tri-methylation (H3K27me3) [21]. H3K27me3 also mediates a bistable epigenetic switch where gene expression is binary and can be in either of two stable states: silenced or fully active [21]. This form of digital memory can be explained by a bistable model of mutual antagonism between activating and repressive histone modifications [22, 23].

Although stable propagation of the silenced gene state has been demonstrated and causally linked to histone modifications and DNA methylation [4, 5, 21], whether any other gene expression state can be stably maintained has not been characterized. Yet, cells should have mechanisms allowing them to maintain memory of any gene expression level since cell types often express cell fate specific genes in graded ways [24–26]. Are then cells able to maintain long-term memory of any gene expression level? What molecular mechanisms may enable cells to achieve this property?

Models of *digital* epigenetic cell memory posit that only two gene expression states, active and silenced, can be stably maintained. This two-state memory model rules out the possibility that any other gene expression level can be maintained (Fig. 1A). As a result, the stationary distribution of gene expression levels across a population of cells has two peaks, one in correspondence to the active state and the other in correspondence to the silenced state, with no cells between them (Fig. 1A, right-side). By contrast, models of *analog* memory allow for the stable maintenance of all levels of gene expression (Fig. 1B). As such, the stationary distribution of gene expression levels can include cells expressing at any level (Fig. 1B, right-side). Although distributional data can rule out digital memory, time-resolved analysis of cell trajectories is required to conclude analog memory. In fact, the same distribution as in analog memory can be observed when cells expressing intermediate levels of gene expression are only in a transient state (Fig. 1C).

To gain an understanding of whether cells have analog memory and to identify the molecular mechanisms at

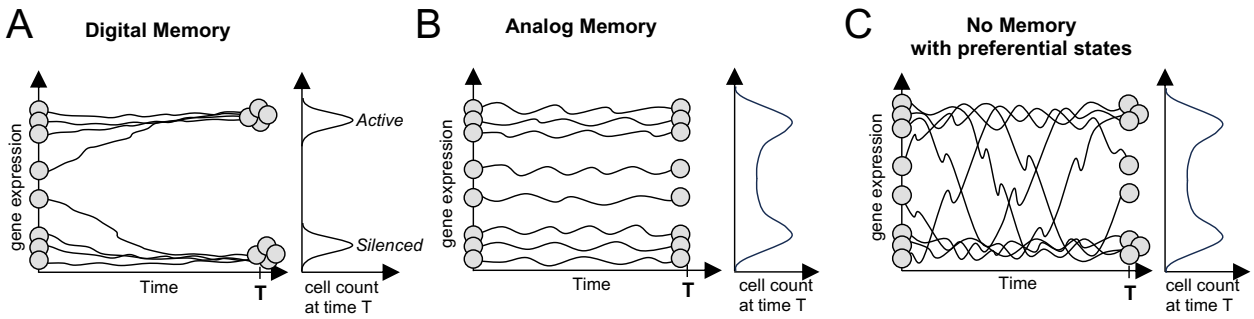

Fig. 1. Digital versus analog memory of gene expression. A Memory is digital when the cell can maintain in time either silenced or fully active gene expression states, but no other states. **B** Memory is analog when the cell can maintain any initial gene expression level. **C** Lack of memory with preferential states results in long-term gene expression trajectories that are independent of the initial state. However, preferential states are visited more often. The stationary distribution of gene expression across the cell population depicts the long-term behavior (here shown at time T) of the number of cells (x-axis) that express the gene at a given level (y-axis).

play, we have analyzed the dynamics of gene expression in mammalian cells bearing a chromosomally integrated reporter gene to which we selectively recruited epigenetic effectors of H3K9me3 and DNA methylation. We have first characterized the temporal dynamics and stationary distribution of gene expression via flowcytometry and the chromatin state of the reporter gene via chromatin and methylated DNA immunoprecipitation followed by qPCR. Once we used these data to restrict the possible behaviors of our nucleosome modification model, the model was able to predict analog memory. Following these predictions, we have performed time-course flow-cytometry analysis with clonal resolution starting from single cells (>90) for an extended period of time (>5 months), and measured the DNA methylation state of clones via bisulfite sequencing at multiple time points.

Results

A reporter gene to study gene expression dynamics following recruitment of epigenetic effectors

We first set out to develop an experimental system to systematically probe the role that histone and DNA methylation play in the establishment and maintenance of gene expression states. To this end, we constructed a reporter gene in an endogenous mammalian locus (Fig. 2A; Methods; Table SE1) [27]. The gene comprises the constitutive mammalian elongation factor 1a promoter (EF1a) driving the expression of fluorescent reporter (EBFP2). We flanked the gene with cHS4 chromatin insulators [28] for isolation of the reporter from other genes. The reporter comprises five binding sites upstream of the promoter for each of the DNA-binding proteins PhlF

and rTetR (Fig. SE.1A, B; Table SE2) [29], where the binding of PhlF and rTetR to DNA can be modulated using 2,4-diacetylphloroglucinol (DAPG) and doxycycline (dox), respectively. Moreover, the region upstream of the promoter comprises five guide RNA (gRNA) target sites, which enable the targeted recruitment of epigenetic effectors fused to dCas9 (Fig. 2A; Fig. SE.1C; Table SE2; Table SE3). The general experimental approach (Fig. 2B) consisted of expressing epigenetic effectors fused to PhlF, rTetR, or dCas9 (Fig. SE.1A-F and Table SE2) using transient transfection in cells bearing the reporter gene. Cells were co-transfected with a fluorescent transfection marker, which enabled the use of fluorescence-activated cell sorting to obtain cell populations with varying levels of epigenetic effector expression. The dynamics and stationary distributions of gene expression from the reporter gene were analyzed.

Cells express the gene at all levels post transient recruitment of DNMT3A

We first investigated the gene expression dynamics following the recruitment of DNA methylation writer DNA methyltransferase 3 alpha (DNMT3A) to the reporter gene. To this end, we expressed DNMT3A-dCas9 [30] (Fig. SE.1 C; Table SE2), a fusion of dCas9 and the catalytic domain of DNMT3A, via transient transfection targeted to five target sites upstream of the promoter in the reporter gene using a gRNA (Fig. 2C, D; Fig. SE.2A, B; Table SE3). Subsequently, we utilized fluorescence-activated cell sorting to obtain cell populations for various transfection levels and proceeded with time-course flow cytometry measurements. We observed that gene expression was repressed, reaching maximal repression approximately 11 days after transfection (Fig. 2C and Fig. SE.2A). After reaching maximal repression, the cell distributions remained stationary for all levels of DNMT3A-dCas9 (Fig. 2C, D and Fig. SE.2A-B). This, along with no detection of the transfection marker of DNMT3A-dCas9 by day 8 (Fig. SE.2B), indicates that cells had memory of the DNMT3A transient stimulation. No repression was observed for cells where a gRNA with a scrambled target sequence was co-transfected with DNMT3A-dCas9 or when dCas9 was targeted to the gene (Fig. SE.2C,D).

Although the epigenetic effector dosage increased the percent of cells in the silenced state, all dosages of DNMT3A-dCas9 led to stationary distributions with cells expressing the reporter at all levels of gene expression ranging from fully silenced to completely active (Fig. 2C, D and Fig. SE.2A). Experiments where we transiently transfected the same DNMT3A catalytic domain fused to rTetR instead of dCas9 (rTetR-XTEN80-DNMT3A) (Fig. SE.1D; Fig.SE.3; Table SE2) also led to stationary distributions of gene expression spanning the whole spectrum of levels, from silenced to fully active. These observations rule out digital memory and are consistent with analog memory (Fig. 1). Methylated DNA immunoprecipitation (MeDIP) and Chromatin immunoprecipitation (ChIP) followed by qPCR further showed an increase in DNA methylation and H3K9me3 levels, accompanied

by a decrease in H3K4me3 level in the reporter gene compared to the untransfected (active) cells (Fig. 2E; Methods; Table SE4). This is in accordance with reports showing the ability of methylated DNA to recruit writers for H3K9me3 [31] and indicate that these two chromatin modifications play a role in the temporal maintenance of gene expression states.

In order to discriminate the roles of DNA methylation and H3K9me3 in the maintenance of gene expression states, we investigated the gene expression dynamics following recruitment of writers for H3K9me3 to the gene. To this end, we fused DNA-binding protein PhlF to KRAB (Fig. SE.1E and Table SE2) [32], an epigenetic effector that promotes H3K9 trimethylation [33, 34]. We thus expressed PhlF-KRAB via transient transfection in the cells bearing the reporter gene, performed fluorescence-activated cell sorting to obtain cell populations across different transfection levels, and performed time-course flow cytometry measurements (Fig. 2F,G and Fig. SE.4A, B). We observed rapid repression of gene expression for cells transfected with PhlF-KRAB, while no repression for cells transfected with PhlF without fusion to KRAB. Irrespective of the dosages of PhlF-KRAB, gene expression reverted to the active gene state within 4 days after reaching maximal repression (Fig. 2F,G and Fig. SE.4A). This aligns with previous reports where KRAB-based epigenetic editing resulted in temporary repression while failing to enable long-term silencing [30, 35]. Further experiments with transient expression of PhlF-KRAB also resulted in temporary repression of the reporter gene while no long-term silencing was observed (Fig. SE.4C-E). To determine what chromatin modifications emerged as a result of KRAB treatment, we performed MeDIP and ChIP followed by qPCR on cells repressed with PhlF-KRAB (Fig. 2H and Methods). We observed that H3K9me3 level, but not DNA methylation level, increased in the reporter gene relative to the untransfected cells (active), while the level of H3K4me3 decreased. Therefore, although it is possible for H3K9me3 to mediate the catalysis of DNA methylation [36, 37], the data in Fig. 2H indicate that this is not the case in the reporter gene, consistent with what was found elsewhere [5]. Taken together, these data indicate that an increase in H3K9me3 in the absence of DNA methylation is unable to permanently downregulate gene expression in the cell population. Therefore, DNA methylation but not H3K9me3 emerges as key mediator of the temporal maintenance of gene expression states in cells silenced with DNMT3A (Fig. 2D,E).

The above results are recapitulated by a nucleosome modification model that includes H3K9me3 and H3K4me3, their autocatalysis [11, 19, 20, 38] and mutual inhibition [39], DNA methylation, its mutual inhibition with H3K4me3 [12, 40, 41], and the recruitment of H3K9me3 writers by methylated DNA [31] but no recruitment of DNMTs by H3K9me3 (SI Sections S.1, S.2, S.3, and Fig. SM.4). Consistent with the data, this model produces a stationary distribution of fluorescence in the DNMT3A-treated cell population that includes cells expressing at all levels (Fig. SM.4D), which is incompatible with digital memory (Fig. 1A). This model differs from the one originally published in [42] due to the lack of recruitment of DNA methylation writer DNMT3A by H3K9me3. We therefore investigated

with this new model, lacking H3K9me3-mediated recruitment of DNMTs, whether any gene expression state could be temporally maintained, a hallmark of analog memory (Fig. 1B).

The nucleosome modification model lacking H3K9me3-mediated recruitment of DNA methylation writers predicts analog memory

First, we performed a parametric analysis of the original nucleosome modification model [42] to determine the effect of the H3K9me3-mediated recruitment of DNA methylation writers on the stationary distribution of gene expression. The model predicts that when this recruitment is present and sufficiently strong, cell trajectories always converge to either of two states: silenced or fully active, giving a stationary distribution of gene expression consistent with digital memory. On the contrary, when this recruitment is absent, and DNA methylation decay rate is negligible, the model predicts that any gene expression state is stably maintained (SI Figs. SM.1A,B). Therefore, lack of H3K9me3-mediated recruitment of DNA methylation is necessary for analog memory.

Overall, to recapitulate the data of Fig. 2C, the model had to satisfy the following requirements. First, DNA demethylation rate is negligible on the time scales of our experiments, which is consistent with the reported high efficiency of DNMT1-mediated maintenance of methylated CpGs [43]. Second, H3K9me3 does not mediate DNA methylation establishment, which is supported by the ChIP and MeDIP data of Fig. 2E,H. Third, mutual inhibition between repressive (H3K9me3 and DNA methylation) and activating (H3K4me3) modifications is sufficiently strong (ε' sufficiently large, SI Section S.4). With the first two conditions, the nucleosome modification model simplifies to that illustrated in Fig. 2I, and with the additional third condition it recapitulates the data of both Fig. 2C and Fig. 2F (SI Fig. SM.4).

In this model, in the absence of epigenetic effectors, DNA methylation grade (CpGme), defined by the mean fraction of the number of CpGs methylated in the gene, remains approximately constant in time since DNA demethylation rate is negligible and DNA methylation establishment is absent (no epigenetic effector and no recruitment of it by H3K9me3). Because DNA methylation grade inversely modulates H3K4me3 level (Fig. 2J, top), and H3K4me3 is proportional to gene expression level (SI Section S.2), then also gene expression level should remain approximately constant in time. Stochastic simulations of the system confirmed that cells initially expressing the gene at any level corresponding to a given DNA methylation grade give rise to a population of cells that keep the level of gene expression and do not approach the active or silenced states (Fig. 2K), confirming analog memory (Fig. 1B).

Interestingly, despite this graded dosage-dependence between DNA methylation grade and gene expression,

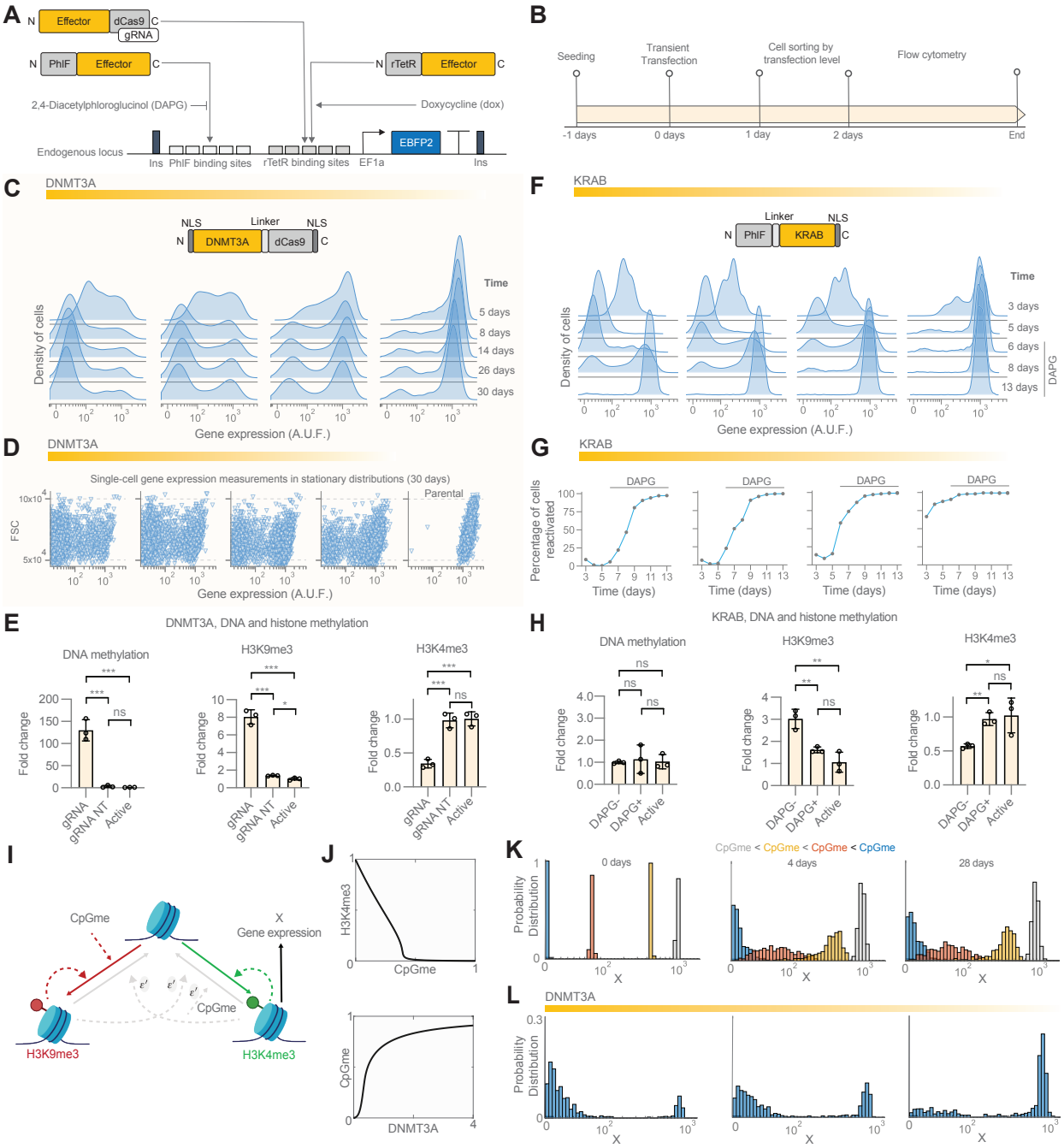

Fig. 2. Continued on the following page.

the model allows for a bimodal gene expression distribution in the cell population after transient DNMT3A-dCas9 transfection (Fig. 2L), in agreement with the experimental data (Fig. 2C). This is achieved by virtue of the third condition listed above, which leads to an ultrasensitive response of DNA methylation grade to the DNMT3A-dCas9 transient dosage (Fig. 2J, bottom and SI Fig. SM.1D,E).

Fig. 2. Gene expression dynamics after DNMT3A recruitment is inconsistent with digital memory. A Schematic of the reporter gene integrated into an endogenous mammalian locus using site-specific chromosomal integration. A mammalian constitutive promoter (EF1a) drives the expression of fluorescent protein EBFP2. Upstream binding sites enable the targeted recruitment of epigenetic effectors, which are fused to DNA-binding proteins rTetR, PhlF or dCas9. The reporter gene is flanked by chromatin insulators for isolation from other genes. **B** Experimental overview describing transient transfection into cells bearing the reporter gene, fluorescence-activated cell sorting based on transfection level, and time-course flow cytometry measurements. **C** Gene expression dynamics of DNMT3A silencing (DNMT3A-dCas9) of the reporter gene according to the experimental timeline shown in panel B. Shown are distributions from flow cytometry measurements of the reporter gene (EBFP2) for four different levels of transfection. DNMT3A-dCas9 was targeted to 5 target sites upstream of the promoter and a scrambled gRNA target sequence was used as control (Fig. SE.2A, B, Table S3). Data shown are from a representative replicate from 3 independent replicates. **D** Single-cell gene expression measurements using flow cytometry of DNMT3A-silenced cells corresponding to the cells shown in Panel C (30 days). Parental refers to the untransfected cells bearing the reporter gene. Data is from a representative replicate from 3 independent replicates. **E** MeDIP-qPCR and ChIP-qPCR analysis 14 days after transfection of DNMT3A-dCas9 and cell sorting for high levels of transfection. The promoter region was analyzed. Data shown is from from three independent replicates. Reported are fold changes and their mean using the standard $\Delta\Delta C_t$ method with respect to the active state. Error bars are s.d. of the mean. DNMT3A-dCas9 was targeted to 5 target sites upstream of the promoter (gRNA). A scrambled gRNA target sequence (gRNA NT) was used as control. *P$\leq$0.05, **P$\leq$0.01, ***P$\leq$0.001, unpaired two-tailed t-test.**F** Gene expression dynamics of KRAB repression (PhlF-KRAB) according to the experimental timeline shown in panel B. Shown are distributions from flow cytometry measurements of the reporter gene (EBFP2) for four different levels of transfection. One independent replicate was measured per day. Data shown are from 3 independent replicates. **G** Percentage of reactivated cells ($400 - 10^5$ gene expression A.U.F.) corresponding to cell populations shown in panel F. **H** MeDIP-qPCR and ChIP-qPCR analysis 6 days after transfection of PhlF-KRAB and cell sorting for high levels of transfection. The analysis is of the promoter region. Data is from three independent replicates. Shown are fold changes and their mean as determined by the standard $\Delta\Delta C_t$ method with respect to the active state. Error bars are s.d. of the mean. P$\leq$0.05, **P$\leq$0.01, ***P$\leq$0.001, unpaired two-tailed t-test. **I** Simplified chromatin modification circuit diagram obtained when KRAB = 0, DNMT3A = 0, TET1 = 0, and H3K9me3 does not mediate the catalysis of *de novo* DNA methylation. See SI Fig.SM.1C. **J** Top plot: dose-response curve for the (CpGme,H3K4me3) pair. Bottom plot: dose-response curve for the (DNMT3A,CpGme) pair. See SI Figs. SM.1D and SM.3. **K** Probability distributions of gene expression for the system represented by reactions listed in SI Tables SM.1 and SM.3. See SI Figs. SM.1B and SM.2. **L** Probability distributions of gene expression for the system after t = 28 days, obtained as described in panel I. In panel J and L, the DNMT3A dynamics are modeled as a pulse that exponentially decreases over time. In our model, α' is the normalized rate of DNA methylation establishment through repressive histone modifications, μ' is ratio between the DNA methylation erasure rate and the activating histone modification erasure rate, and ε (ε') is the parameter that scales the ratio between the basal (recruited) erasure rate and the auto-catalysis rate of each modification. See SI Fig. SM.1E and SM.3.

High and intermediate gene expression levels are stably maintained and inversely correlate with DNA methylation grade

Following the model predictions, we hypothesized that epigenetic memory could be analog, thereby enabling cells to maintain any gene expression level. To test this hypothesis, we first sought to investigate whether cells with the gene in a state different from silenced within DNMT3A-treated populations exhibited long-term stability of gene expression. To this end, we performed fluorescence-activated cell sorting of the DNMT3A-silenced cells from low, intermediate, and high gene expression levels (Fig. 3A, Methods). We then performed time-course flow cytometry measurements for 29 days to analyze the long-term gene expression dynamics. The cells with low and high levels of gene expression maintained an approximately stable distribution of gene expression levels after cell

212 sorting, as expected by a digital memory model. Nonetheless, the cells with intermediate levels of expression,
213 after an initial widening of the expression distribution, also kept a stable gene expression distribution across the
214 population for the entire time-course (Fig. 3B-D). This is incompatible with digital memory (Fig. 1A). This also
215 rules out a case of no memory with preferential states (Fig. 1C) since it demonstrates that cells with intermediate
216 gene expression levels are not *en route* to the silenced or active states, but rather are in a stable epigenetic
217 configuration. To characterize the epigenetic state of these cells, we performed ChIP followed by qPCR of the
218 low, intermediate, and active cell populations (Fig. 3E, Methods). We found an increase of H3K9me3 in the low
219 population when compared to the intermediate and active cell populations. Accordingly, we observed a decrease
220 of H3K4me3 in the low and intermediate expressing cells when compared to the active cells. These results are
221 consistent with the known role of H3K9me3 as a chromatin modification associated with repression and H3K4me3
222 as a chromatin modification associated with active transcription at promoters [44].

223 Because we pinpointed DNA methylation as the critical mediator of the maintenance of gene expression states,
224 we performed targeted bisulfite sequencing of the reporter gene for each of the three cell populations: low,
225 intermediate, and active (Fig. 3F, Methods). While the active state is essentially devoid of methylated CpGs, the
226 probability of finding any given CpG methylated in the promoter and rTetR binding sites in the intermediate state
227 is distinctly smaller than in the low state and higher than in the active state. From these data, we computed the
228 probability distributions of the fraction of methylated CpGs in the promoter and rTetR binding sites, for each of the
229 low, intermediate, and active states (Fig. 3G, Methods - Section "Quantification and statistical analysis"). These
230 distributions reinforce that the intermediately expressing cells have intermediate levels of DNA methylation. In
231 particular, the mean value of the fraction of methylated CpGs, which we define here as the DNA methylation
232 grade, is intermediate in cells with intermediate expression levels compared to that of the low and active cells
233 and has a loglinear negative correlation (p-value <0.05) with the mean gene expression level (Fig. 3H). All these
234 findings are recapitulated by our model when H3K9me3 does not mediate the recruitment of DNA methylation
235 writers (SI Fig. SM.1B and Fig. SM.2).

236 We tested further the hypothesis that epigenetic memory could be analog by subdividing according to gene
237 expression level a population of cells that had been previously sorted from intermediate gene expression in
238 DNMT3A-silenced cells (Fig. 3I, Methods). If epigenetic memory were analog, cells sorted from different gene
239 expression levels should approximately maintain these distinct levels over time. Each of the four resulting cell
240 populations maintained distinct distributions of gene expression levels according to the level on the day of sorting
241 (Fig. 3J-L). We also performed fluorescence-activated cell

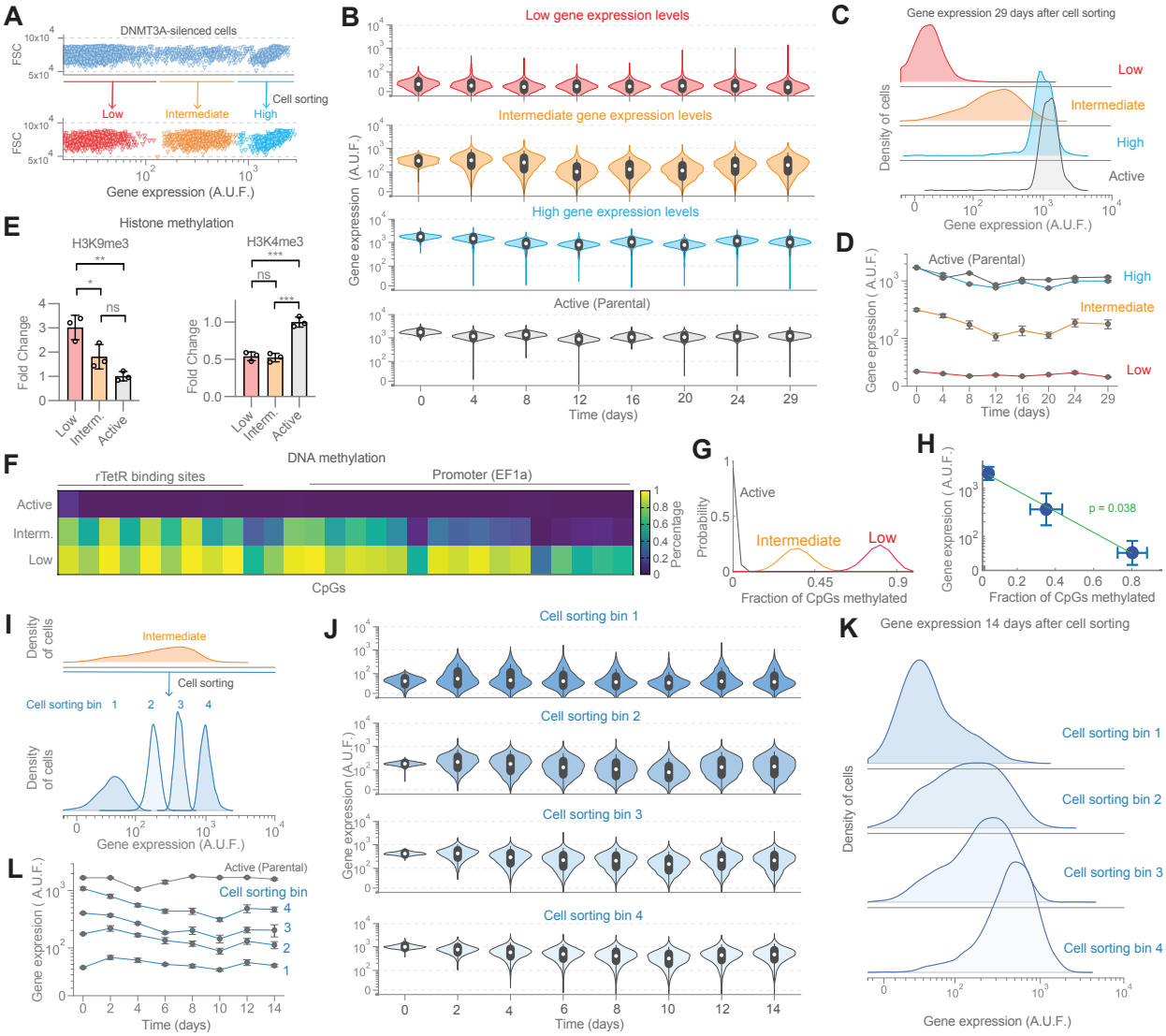

Fig. 3. Continued on the following page.

Fig. 3. Gene expression states different from silenced are stably maintained and inversely correlate with DNA methylation grade. A Fluorescence-activated cell sorting of DNMT3A-silenced cells (DNMT3A-dCas9) according to low, intermediate, and high levels of gene expression (EBFP2). Shown are single-cell gene expression measurements on the day of cell sorting using flow cytometry. Data are from a representative replicate from three independent replicates. **B** Time-course flow cytometry analysis of the dynamics of gene expression exhibited by cells obtained in panel A. Data are from one representative replicate from three independent replicates. **C** Density of cells from flow cytometry measurements obtained on the last time point in panel B (29 days after cell sorting). Data shown are from one representative replicate from three independent replicates. **D** Dynamics of gene expression exhibited by cells shown in panels A-C. Shown is the mean of geometric mean fluorescence from three independent replicates. Error bars are s.d. of the mean. **E** ChIP-qPCR for H3K9me3 and H3K4me3 of the promoter in the reporter gene (untransfected cells bearing the reporter gene). Shown are fold changes and corresponding mean from three independent replicates as determined by the $\Delta\Delta C_t$ method with respect to the active state. Error bars are s.d. *P$\leq$0.05, **P$\leq$0.01, ***P$\leq$0.001, unpaired two-tailed t-test. **F** Heatmap of targeted bisulfite sequencing of the reporter gene. Shown is the mean percentage of DNA methylation for each CpG from three independent replicates. **G** Probability distribution of the fraction of CpGs methylated (Methods). **H** Correlation plot between the mean fraction of CpGs methylated and mean gene expression level. The green line represents the best-fit line $y = -0.5059x + 0.5497$ obtained by linear regression. The p-value associated with the slope coefficient is p $= 0.038 < 0.05$ (Methods). **I** Fluorescence-activated cell subsorting of DNMT3A-silenced cells (DNMT3A-dCas9) exhibiting intermediate levels of gene expression (EBFP2). Data is from gene expression measurements using flow cytometry on the day of cell sorting. Data is from a representative replicate from three independent replicates. **J** Flow cytometry-based time-course analysis of the dynamics of gene expression exhibited by cells obtained in panel I. Data is from one representative replicate from three independent replicates. **K** Density of cells from flow cytometry measurements obtained on the last time point in panel J (14 days after cell sorting). Data shown are from one representative replicate from three independent replicates. **L** Dynamics of gene expression exhibited by cells shown in panels I-K. Shown is the mean of geometric mean fluorescence from three independent replicates. Error bars are s.d. of the mean.

sorting from cells that had been silenced with DNMT3A fused to rTetR (rTetR-XTEN80-DNMT3A) instead of dCas9 (Fig. SE.5A-D; Methods), and the same results were obtained. These observations support the hypothesis that epigenetic cell memory is analog (Fig. 1B) in agreement with the model (SI Fig. SM.5), and that DNA methylation grade is likely the critical mediator to the stable maintainance of gene expression levels.

All gene expression levels are maintained and inversely correlate with a stable DNA methylation grade

To further test the hypothesis that epigenetic memory operates on an analog basis and that DNA methylation grade mediates the maintenance of gene expression level, we set out to characterize the long-term dynamics of gene expression with extensive sampling of single cells from DNMT3A-silenced cell populations followed by monoclonal expansion and analysis. This approach provides a high-resolution evaluation of gene expression trajectories generated exclusively from one initial DNA methylation profile. We therefore performed fluorescence-activated single-cell sorting from three distinct cell populations that had been previously sorted from a DNMT3A-silenced population (Fig. 4A). Specifically, we selected eight different consecutive regions of gene expression levels spanning across the low (2 regions), intermediate (4 regions), and high (2 regions) gene expression levels (Methods). We sampled single cells from each of those regions and performed clonal expansion, which resulted in $\geq$10 monoclonal populations from each of the eight gene expression regions and a total of 93 monoclonal

populations. Flow cytometry was performed for each of the monoclonal populations 14 days after the initial single-cell sort (Fig. 4B). The monoclonal populations showed a continuum of gene expression levels ranging from background expression level to active gene expression level. (SI Section S.4, Figs. SM.2 and SM.3A). To test whether the gene expression levels from the monoclonal populations could be maintained in time, we selected 8 monoclonal populations (highlighted in Fig. 4B), spanning the gene expression range from the 93 clones, and performed time-course flow cytometry measurements up to 161 days (> 5months) (Fig. 4C) after the initial single-cell sort. In these measurements, we observed that the gene expression distributions remained approximately stable over time, validating the initial hypothesis that cells can maintain memory of any gene expression level, which defines analog memory (Fig. 1B).

We next investigated DNA methylation grade as a determinant of the maintenance of gene expression level. To this end, we performed targeted bisulfite sequencing of the reporter gene in each of the eight populations at two different time points (Fig. 4D,E; Methods). These data indicate that clones (7 and 8) selected from cells in the high gene expression distribution (Fig. 3A) are largely devoid of methylated CpGs. By contrast, clones (1 and 2) selected from cells within the low distribution have the majority of CpGs methylayed. The remaining clones (3,4,5, and 6), selected from the intermediate distribution, display a decreasing number of CpGs that are methylated with high probability as we move from clone 3 to clone 6 (Fig. 4D). These trends inversely correlate with the gene expression levels in the clones (Fig. 4E). To acquire a quantitative understanding of the graded nature of this variation, we computed from the bisulfite sequencing data the probability distribution of the fraction of methylated CpGs in the gene's promoter and rTetR binding sites (See Methods - Section "Quantification and statistical analysis"). This confirmed a progressive and gradual shift of the distributions through the eight clonal populations (Fig. 4F), in accordance with model predictions (Fig. 4G). The correlation between the mean gene expression level of each clonal population and the mean number of methylated CpGs is loglinear (p $= 5.34 \cdot 10^{-5} < 0.05$, Fig. 4H,I), confirming a gradual increase of the DNA methylation grade through the clonal populations from high to low gene expression levels. Furthermore, the distribution of the fraction of methylated CpGs remained approximately unchanged between the two time points where bisulfite sequencing was performed (Fig. 4F, G. Compare solid and dashed lines), indicating that the DNA methylation grade is stably maintained.

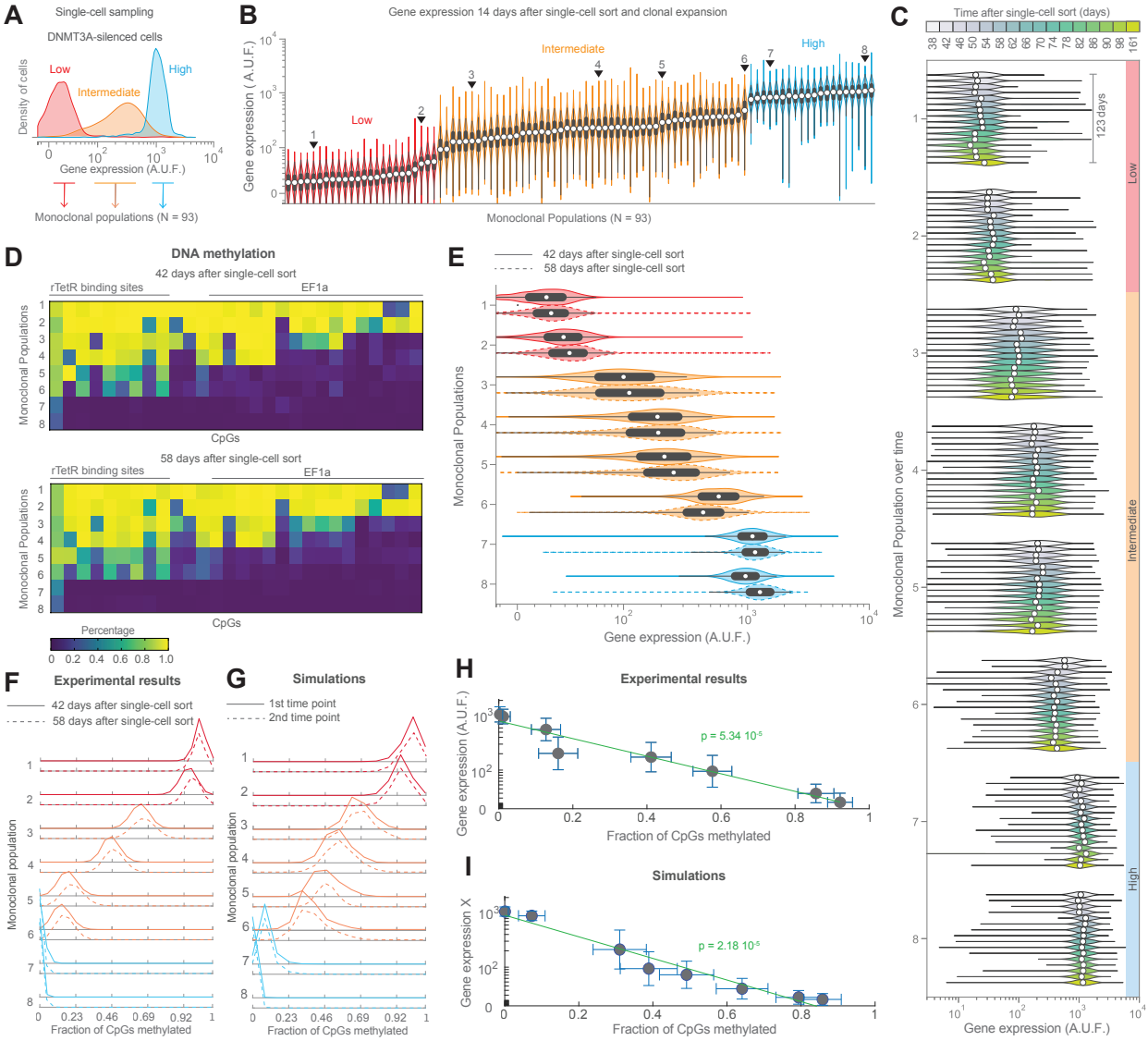

Fig. 4. Continued on the following page.

Fig. 4. Cell memory is analog and stable DNA methylation grades inversely correlate with gene expression level. A Density of cell populations from where monoclonal populations were obtained. The cell populations had stably maintained low, intermediate, and high levels of gene expression (EBFP2). Single cells were obtained from across the three different cell populations. After monoclonal expansion, the total number of monoclonal populations obtained was N = 93. **B** Violin plots of the monoclonal populations (N = 93) measured 14 days after single-cell sorting from the cell populations shown on panel A. Monoclonal populations were positioned on the x-axis according to geometric mean of gene expression. Clones indicated with numbers 1-8 were selected for further analysis. Colors represent the cell population of origin as shown in panel A. **C** Flow cytometry-based time course analysis of gene expression dynamics of the eight monoclonal populations selected in panel B. **D** Bisulfite sequencing of the reporter gene in the eight monoclonal population at two different time points. **E** Flow cytometry measurements of the reporter gene in the eight monoclonal population at two different time points corresponding to those shown on panel D. **F** Probability distribution of the fraction of CpGs methylated (See "Methods" section for details on deriving these plots from data in panel E). The standard deviation for each clone is the following: 0.1752, 0.4707, 0.9492, 1.2714, 1.3101, 1.262, 1.1683, and 0.8095, ordered from clone 1 to clone 8. **G** Simulations of the probability distributions of the system represented by the reactions listed in SI Tables SM.5 and SM.3. The distributions are obtained computationally using SSA [45] and, on the x-axis we have the fraction of CpGs methylated. The standard deviation for each clone is: 0.1346, 0.5472, 1.2585, 1.1189, 1.0431, 1.1117, 0.8512, and 0.8259, ordered from clone 1 to clone 8. The parameter values used for these simulations are listed in SI Section S.9. **H** Correlation plot between the mean fraction of CpGs methylated and mean gene expression level at 42 days after single-cell sort. The green line represents the best linear fit $y = -0.3756x + 0.4358$. The p-value associated with the slope coefficient is p $= 5.34 \cdot 10^{-5} < 0.05$ (See "Methods" section for details on the linear regression). **I** Correlation plot between the mean fraction of CpGs methylated and mean gene expression level, X, obtained via simulations. The green line represents the best linear fit $y = -0.4673x + 0.4580$. The p-value associated with the slope coefficient is p $= 2.18 \cdot 10^{-5} < 0.05$ (See "Methods" section for details on the linear regression).

Further analysis of the clonal distributions of the fraction of methylated CpGs revealed that the variability of the number of methylated CpGs about their mean value, measured by the standard deviation, is larger for intermediate DNA methylation grades when compared to low and high methylation grades (Fig. 4F). The mathematical model recapitulates this property (Fig. 4G) when it includes the two following additional known molecular interactions: (a) H3K9me3 helps recruit DNA methylation maintenance enzyme (DNMT1) to the gene [46] and (b) DNMT1 can cause *de novo* DNA methylation, although at very low rate [47, 48] (see SI Section S.5 for model derivation). With these interactions, the errors made by DNMT1 in maintaining methylation are less probable for a fully methylated gene due to higher recruitment rate of DNMT1 via H3K9me3. Similarly, errors made by DNMT1 in *de novo* methylating additional CpGs are less probable for a gene with no methylation since DNMT1 recruitment is practically absent (see SI Section S.5 and Fig. SM.6 for model analysis.) Overall, these interactions explain the tighter distributions observed for highly and lowly methylated gene states (Fig. 4F).

The genome engineering system that we utilized in this study has been demonstrated to be highly precise [27]. Nevertheless, we still sought to verify that the different clones all had one genome copy of the reporter gene to rule out the possibility that the variation of gene expression observed across the clones could be due to variability in the genomic integration copy number. To this end, we performed copy number variation analysis using digital PCR (dPCR), which ruled out this possibility (SE.5E and Methods). Taken together, these data establish that the DNA methylation grade is maintained by the clonal populations and that this grade fine-tunes the gene expression level.

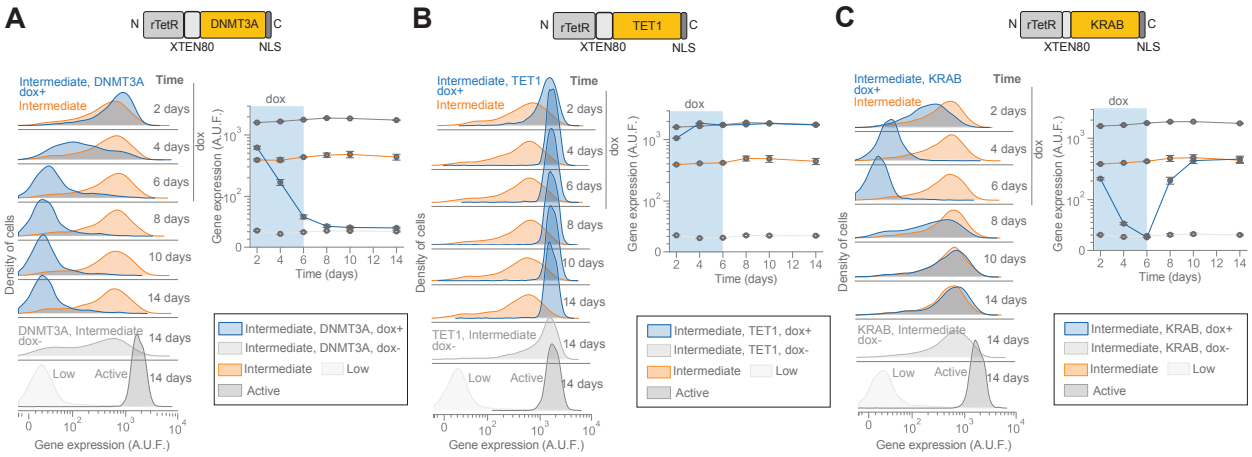

Fig. 5. DNA methylation writers and erasers impair maintenance of intermediate gene expression states. A Time-course of flow cytometry measurements after transient transfection of rTetR-XTEN80-DNMT3A in cells that have been stably maintaining intermediate levels of gene expression. **B** Time-course of flow cytometry measurements following transient transfection of rTetR-XTEN80-TET1 in cells that have been stably maintaining intermediate gene expression levels. **C** Time-course of flow cytometry measurements after transient transfection of rTetR-KRAB in cells with stable intermediate levels of gene expression. Densities of cells shown are from one representative replicate from three independent replicates. Gene expression is from the reporter gene (EBFP2). Shown is also the mean of geometric means from three independent replicates. Error bars are s.d. of the mean.

Writers and erasers of DNA methylation impair gene expression level maintenance

Next, we sought to establish a causal link between the DNA methylation grade and the level of gene expression. To this end, we investigated whether altering the DNA methylation grade in cells maintaining an intermediate level of gene expression would impair this maintenance. To enhance DNA methylation, we utilized rTetR-XTEN80-DNMT3A as we have shown before that DNMT3A adds DNA methylation to the gene (Fig. 2F) consistent with prior reports [30]. To remove DNA methylation, we engineered a fusion protein comprising rTetR fused to the catalytic domains of the ten eleven translocation protein 1 (TET1) [49, 50] to construct rTetR-XTEN80-TET1 (Fig. SE.1F and Table SE2). This catalytic domain has been shown to enable targeted DNA demethylation of mammalian genes [50]. Moreover, we fused rTeR to KRAB to construct rTetR-KRAB (SE.1G and Table SE2) in order to evaluate changes to gene expression when enhancing H3K9me3 without changing DNA methylation Fig. 2H). We transiently transfected these epigenetic effectors in cells that had been previously sorted for intermediate levels of gene expression from the DNMT3A-treated cells and exhibited intermediate DNA methylation grades. As before, we utilized fluorescence-activated cell sorting to obtain transfected cells and performed time-course flow cytometry measurements (Fig. 5 and Fig. SE.5F-H). By 10 days after transient transfection with rTetR-XTEN80-DNMT3A, the gene expression was repressed, remained silenced over time, and overlapped with the distribution of cells that had been previously sorted for low levels of gene expression post DNMT3A-treatment (Fig. 5A). Transient transfection with rTetR-XTEN80-TET1 resulted in nearly complete and permanent reactivation

of gene expression (Fig. 5B). By contrast, transient transfection with rTetR-KRAB did not result in any permanent change of gene expression. Specifically, 6 days after transfection, we observed levels of repression comparable to those of the low levels in DNMT3A-treated cells, but then expression reverted to the initial intermediate level and remained stable afterwards (Fig. 5C). These data indicate that transient recruitment of a DNA methylation writer or eraser to an intermediadely methylated, and partially active gene, can permanently silence or reactivate the gene. By contrast, transient recruitment of an H3K9me3 writer does not exert a permanent change in gene expression. Taken together, these data show that DNA methylation grade is the causal determinant of the maintenance of gene expression levels in accordance with the model predictions (Fig. SM.7).

Discussion

The extent to which epigenetic cell memory is an exclusive attribute of silenced gene states or whether it is possible for any other gene expression level is the question addressed by this paper. We found that cells can stably maintain any initially set gene expression level (Fig. 3B, J and Fig. 4B,C) and that this level inversely correlates with grades of DNA methylation in the gene's promoter and upstream operator sites (Fig. 3F-H and Fig. 4D, E, H). Each DNA methylation grade, in turn, also persists in time (Fig. 4F). Transient recruitment of DNA methylation writer or eraser enzymes to a gene that maintains an intermediate level of gene expression impairs maintenance (Fig. 5). Taken together, these results indicate that epigenetic cell memory is analog and that the persistence of DNA methylation grade enables it. Our nucloesome modification model, which includes DNA methylation and its interactions with histone modifications H3K9me3 and H3K4me3, recapitulates analog epigenetic cell memory under the critical conditions: (i) DNMT1-enabled DNA methylation maintenance error rate is very low, necessary for an approximately stable DNA methylation grade across cell divisions, and (ii) H3K9me3 does not mediate the catalysis of *de novo* DNA methylation (see SI Section S.3). Notably, failure of (ii), that is, a strong catalysis of DNA methylation mediated by H3K9me3, even when (i) holds, leads to digital memory (Figs. SM.1B and SM.2). Therefore, a model emerges that can explain the molecular interactions responsible for digital versus analog epigenetic cell memory (Fig. 6).

Although DNA methylation has been regarded as the embodiment of inheritance since the DNA methylation pattern of a gene can, in principle, propagate through subsequent cell divisions [14], DNA methylation is now emerging as a more dynamic process than originally thought [51, 52]. Indeed, DNA methylation does not operate in isolation but rather in orchestration with histone modifications [31, 37, 53–55], through mutual catalysis with H3K9me3 [31, 37, 46] and inhibition with H3K4me3 [54, 55]. The emergent positive feedback loop where

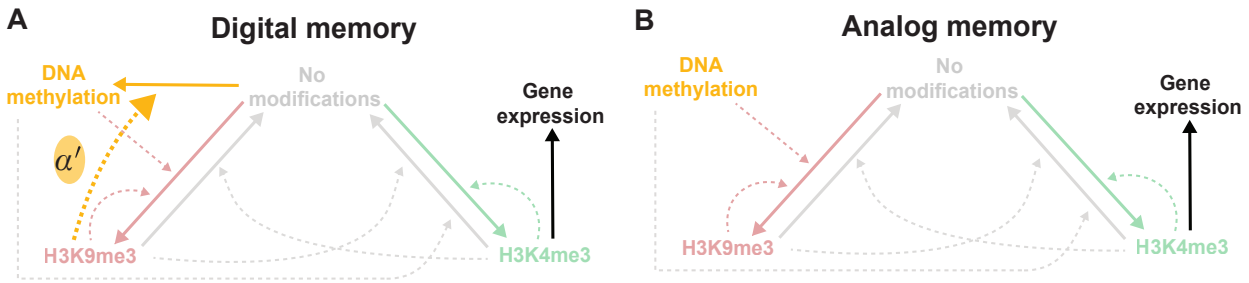

Fig. 6. The interactions between DNA methylation and histone modifications explain the difference between digital and analog epigenetic cell memory. A Digital memory. H3K9me3 recruits writer enzymes for DNA methylation (yellow dashed arrow marked by α'). Because of this interaction, a positive feedback loop emerges where H3K9me3 enhances DNA methylation and this, in turn, enhances H3K9me3. **B** Analog memory. H3K9me3 does not recruit writer enzymes for DNA methylation. The positive feedback loop between DNA methylation and H3K9me3 is broken. In the diagrams, we have assumed that no external epigenetic effectors (such as KRAB, DNMT3A, and TET1) are present. Refer to Fig. SM.1A, C for the complete diagram and model.

DNA methylation mediates the catalysis of H3K9me3 and this, in turn, mediates catalysis of *de novo* DNA methylation, can push an initially partially methylated gene to be fully methylated. The competitive interaction of DNA methylation with H3K4me3, in turn, can cause H3K4me3 to win over an initially low DNA methylation state and to revert this to the unmethylated state. When these interactions are sufficiently strong, then digital memory emerges (SI Section S.3, Fig. SM.2), and fully methylated and completely unmethylated states are the only ones that can be faithfully inherited.

Our MeDIP and ChIP data show that H3K9me3 does not catalyze DNA methylation in our reporter gene (Fig. 2H), consistent with previous reports [5]. With this, our nucleosome modification model predicts that all DNA methylation grades can be stably maintained (SI Fig.SM.1), which is confirmed by our bisulfite sequencing data (Fig. 4F). These data further show that the variability of the number of methylated GpGs about its mean level is larger for intermediate DNA methylation grades (Fig. 4F). The nucleosome modification model would not be able to recapitulate this if we had regarded the DNA methylation process as isolated from H3K9me3 and H3K4me3 (SI Section S.5 and Fig. SM.6). Specifically, the model recapitulates this feature when it includes the known facts that DNMT1 is recruited by H3K9me3 [46] and that DNMT1 can make mistakes, although at a very low rate [47, 48] (Fig. 4G and SI Section S.5, Fig. SM.6).

In summary, DNA methylation and histone modifications operate in orchestration and the strength of the resulting positive feedback loop discriminates between digital and analog memory (Fig. 6). The strength of this feedback loop (denoted by α' in Fig. 6A), in turn, likely depends on the biological context and may explain why in certain cell types the expression of fate specific genes appears digital while in others is clearly continuous [24, 56] and is associated with graded DNA methylation [57]. More research will be required to determine the extent to which these gradients of gene expression correlate with the DNA methylation grade at the gene's promoter and

are temporally maintained at the single cell or clonal level. Single-cell nucleosome, methylome, and transcriptome sequencing (scNMT-seq) [58] can be applied to determine the correlation between the DNA methylation grade and the expression level of a gene. However, acquiring temporal dynamic data of specific genes in single cells (or in clonal populations) will be required to investigate maintenance, and hence epigenetic memory, of DNA methylation and gene expression levels through subsequent cell divisions.

More broadly, analog epigenetic cell memory challenges the concept of cell type as a discrete entity emerging from a sequence of binary decisions that either silence or activate genes [56], as captured by the Waddington epigenetic landscape metaphore [59]. With analog memory, cell fate becomes a fluid entity that arises from a continuum spectrum of choices of gene expression levels. This view supports the possibility for the existence of uncharacterized graded cell types, with functions that have not been appreciated yet. On the other hand, the possibility of controlling the level at which a gene can be persistently expressed, by modulating DNA methylation grade, could allow the engineering of gradients of cell types, thereby increasing the sophistication and functionality of tissues that can be created. To this end, additional research will be needed in order to re-engineer the chromatin modification circuit such that we can precisely control DNA methylation grade by user-defined inputs, such as small signaling molecules and drugs.

In summary, our findings show that cells have the ability to maintain long-term memory of all gene expression levels, and that they can achieve this by stably maintaining DNA methylation grades. This indicates that graded cell types are biologically plausible, thereby spurring the need to revisit the notion of cell fate alltogether and paving the way for future tissue engineering tools.

Methods

Cell culture conditions

Cells were maintained in tissue culture plates (Corning) with growth media comprised of Dulbecco's Modified Eagle Medium (DMEM) (Corning) supplemented with 10% FBS (Corning), 1X non-essential amino acids (Gibco), 100 U/ml penicillin (Sigma), and 100 μg/ml streptomycin (Sigma) at 37℃ in a 5% CO2 incubator. The growth media was changed every 2-4 days. For routine passaging, the growth media was removed, followed by the addition of 0.25% trypsin (Corning) and incubation at 37℃ in a 5% CO2 incubator for 7 minutes. Trypsin was then neutralized with growth media at room temperature. The cells were then centrifuged for 3 minutes at 200g, the supernatant was discarded, and the cell pellet was resuspended in fresh growth media. The cells were then

443 diluted (ratio between 1 and 20) and reseeded into new tissue culture plates for subsequent cell culture.

444 **Transfections**

445 Cells were cultured in growth media at 37℃ in a 5% CO2 incubator. The cells were then treated with
446 0.25% trypsin (Corning) and incubated for 7 minutes at 37℃ in a 5% CO2 incubator. The cell suspension
447 was subsequently neutralized with growth media at room temperature. Cells were centrifuged for 3 min at
448 200g followed by removal of the supernatant and cell resuspension in fresh growth media. Cells were then
449 seeded in 12-well tissue culture plates (Corning) at 300,000 cells/well and cultured for 24 hours at 37℃ in a
450 5% CO2 incubator. Growth media was replaced with growth media without antibiotics. Transfection was then
451 performed using Lipofectamine LTX (Invitrogen) according to the manufacturer's instructions. After 24 hours of
452 incubation at 37℃ in a 5% CO2 incubator, replacement for routine growth media supplemented with penicillin and
453 streptomycin was performed. All transfections with rTetR-XTEN80-DNMT3A, rTetR-XTEN80-DNMT3A, rTetR-
454 XTEN80-TET1 and rTetR-KRAB were performed using 500ng of the fusion protein and 300ng of trasfection
455 marker (EYFP). Transfections with PhlF-KRAB that involved cell sorting were performed using 450ng of either
456 PhlF-KRAB or PhlF, and 450ng of the transfection marker (450ng). Experiments of PhlF-KRAB without cell
457 sorting were perfomed using 450ng of PhlF-KRAB and 450ng of the transfection marker (1:1 stochiometry), 45ng
458 of PhlF-KRAB and 450ng of the transfection marker (0.1:1 stochiometry), or 4.5ng of PhlF-KRAB and 450ng of
459 the transfection marker (0.01:1 stochiometry). An empty vector was used to maintain the DNA amounts constant
460 for the various stoichiometries. Transfections of DNMT3A-dCas9 that involved cell sorting were performed using
461 500ng of DNMT3A-dCas9, 600ng of the gRNA expression vector, and 300ng of the trasfection marker (EYFP). In
462 experiments using dCas9, 450ng of the gRNA expression vector, 300ng of either DNMT3A-dCas9 or dCas9, and
463 300ng of the transfection marker (EYFP) were used. All the experiments presented use 2μM of dox and 30μM of
464 DAPG.

465 **Chromosomal integration**

466 The reporter gene was integrated using site-specific chromosomal integration into an endogenous mammalian
467 locus (Rosa26) in CHO-K1 cells bearing a landing pad (monoallelic) for efficient and precise integration [27].
468 This locus had been previously identified by a homology-based computational search (BLAST) using the murine
469 Rosa26 locus as reference. EYFP is expressed from the landing pad exclusively before site-specific chromosomal
470 integration into the Rosa26 locus. The landing pad confers puromycin resistance exclusively after site-specific
471 chromosomal integration into the Rosa26 locus. For integration, 500ng of the chromosomal integration vector
472 and 500ng of the Bxb1 recombinase expression vector were transfected into the cells. After 24 hours, routine
473 cell culture was commenced with the supplementation of growth media with puromycin (8μg/ml) for 2 weeks.

474 Subsequently, a fluorescence-activated cell sorter (BD Aria) was utilized to obtain cells that were positive for
475 EBFP2 expression and negative for EYFP expression. Digital PCR was used to confirm 1 reporter gene per
476 genome (Fig. SE.5E; Methods). The cells were also assayed and confirmed to test negative for mycoplasma
477 contamination (MIT's Koch Institute Preclinical Modeling Core).

478 **DNA construction**

479 DNA cloning was performed using a combination of Golden Gate Assembly, Gibson Assembly, restriction
480 cloning, and PCR cloning. Expression vectors for the epigenetic effectors were cloned using Golden Gate
481 Assembly. The cloning was performed in a single reaction with the plasmids comprising a destination vector,
482 cHS4 chromatin insulator, EF1a promoter, 5' UTR, gene sequence, 3' UTR, and polyadenylation signal. Gene
483 sequences for rTetR-TET1, rTetR-XTEN80-TET1, rTeTR-DNMT3A, rTeTR-XTEN80-DNMT3A, PhlF-KRAB, rTetR-
484 KRAB, and PhlF were synthesized and cloned into a plasmid with compatible overhangs for Golden Gate Assembly
485 (Azenta Life Sciences). The expression vectors for TET1 and rTetR were cloned by performing PCR cloning using
486 the rTetR-XTEN80-TET1 expression vector as a template. For the reporter gene, the region comprising the PhlF
487 and rTetR binding sites was synthesized (Azenta Life Sciences) and subsequently cloned upstream of a vector
488 containing the EF1a promoter using restriction cloning. Golden Gate Assembly was then performed in a reaction
489 comprising the EF1a promoter with upstream DNA-binding sites, destination vector, upstream cHS4 chromatin
490 insulator, 5' UTR, EBFP2, 3' UTR, and polyadenylation signal. The resulting vector was cloned along with
491 the downstream cHS4 chromatin insulator into the chromosomal integration vector using Gibson Assembly. All
492 sequences were sequenced verified via Sanger sequencing (Azenta Life Sciences) or whole plasmid sequencing
493 (Primordium).

494 **Flow cytometry**

495 Cells were trypsinized using 0.25% trypsin (Corning) and incubation at 37℃ in a 5% CO2 incubator for 7
496 minutes. Following this, trypsin was neutralized by adding growth media at room temperature. Afterwards,
497 the cells were centrifuged at 200g for 3 minutes. The supernatant was removed, and the resulting cell pellet
498 was then resuspended in fresh growth media. The cell suspension was filtered through a filter cap into flow
499 cytometry tubes (Falcon). Flow cytometry data was then collected using a flow cytometer (BD LSRFortessa).
500 EBFP2 was measured using a 405nm laser and a 450/50 filter with 259 V, EYFP was measured using a 488nm
501 laser and 530/30 filter with 190 V, mKO2 was measured using a 561nm laser and 610/20 filter with 250 V,
502 iRFP was measured using a 637nm laser and 670/30 filter with 326 V. Cytoflow (version 1.2) was used for
503 analyzing all flow cytometry data and obtaining cell densities, means, standard deviations, and percentages of
504 flow cytometry measurements. For analysis, measurements were gated for single cells based on forward and

side scatter (Fig. SE.5I), and only gated measurements were used for subsequent analysis. After processing the flow cytometry data using Cytoflow, the data was plotted using either Cytoflow, GraphPad Prism or Matplotlib. Graphpad Prism was used for generating bar plots. For the correlation plot between mean EBFP2 fluorescence and DNA methylation, Cytoflow's gated measurements were imported into MATLAB, where the correlation was performed using the polyfit function.

Time-course flow cytometry

A fluorescence-activated cell sorter (BD Aria) was used to obtain cells with specific levels of transfection (EYFP) or gene expression (EBFP2). For transfection experiments that involved cell sorting, cell sorting was performed between 42-48 hours after transfection. When flow cytometry was performed on the day of cell sorting, a fraction of the cells that were sorted was utilized for flow cytometry measurements (BD LSRFortessa), and the rest of the cells were reseeded on tissue culture plates. For flow cytometry measurements, cells were trypsinized using 0.25% trypsin (Corning) followed by incubation at 37℃ in a 5% CO2 incubator. Trypsin neutralization was performed with growth media at room temperature. The cells were then centrifuged for 3 min at 200g. The supernatant was removed followed by resuspension of the cells in fresh in growth media. A fraction of the cells was utilized for the acquisition of flow cytometry measurements (BD LSRFortessa). The rest of the cells were diluted (ratio between 1 and 20) and reseeded into new tissue culture plates for subsequent measurements.

Cell sorting

Cells were trypsinized using 0.25% trypsin (Corning) followed by incubation at 37℃ in a 5% CO2 incubator for 7 minutes. Subsequently, trypsin was neutralized with growth media at room temperature. The cells were then centrifuged at 200g for 3 minutes. The supernatant was discarded and the cell pellet was resuspended in growth media supplemented with 2.5g/ml BSA and corresponding small molecules according to the experiment. The cell suspension was then filtered through a filter cap into flow cytometry tubes. Cell sorting was then performed using a fluorescence-activated cell sorter (BD Aria) approximately 42-48 hours after transfection, and cells were collected in growth media supplemented with penicillin/streptomycin and corresponding small molecules according to the experiment. EBFP2 was measured using a 405nm laser and a 450/50 filter with 332 V. EYFP was measured using a 488nm laser and 515/20 filter with 200 V. Cells were sorted either for transfection level (EYFP) or gene expression level (EBFP2). During cell sorting for EBFP2, a gate was also applied to sort exclusively for EYFP negative cells. Cell sorting to study the gene expression dynamics after epigenetic modulation using DNMT3A-dCas9 (Fig. 2C, D and Fig. SE.2A, B) used ranges 100-200, 222-445, 500-1000 and 1150-2300 EYFP A.U.F. Cell sorting to study the gene expression dynamics after epigenetic modulation using PhlF-KRAB (Fig. 2F, G and Fig. SE.4A, B) used ranges 25-100, 100-1000, 1000-10000, and 10000 -

536 100000 corresponding to EYFP A.U.F. Cell sorting to study histone and DNA methylation changes using MeDIP-
537 qPCR and ChIP-qPCR (Fig. 2E, H) used ranges 3000-100000 EYFP A.U.F. Cell sorting to demonstrate the
538 stability of DNMT3A-silenced (DNMT3A-dCas9) cells expressing intermediate levels of gene expression (Fig.
539 3A) was performed according to <100 A.U.F for low expression, 200-500 A.U.F for intermediate expression, and
540 >1000 A.U.F for high expression. Cells had been silenced with transfection of DNMT3A-dCas9 targeted to the
541 upstream binding sites of the promoter (Tables SE1 and SE3), cell sorting for 7000-100000 EYFP A.U.F., and
542 were maintained in culture for >2 weeks. Cell sorting to subdivide cell populations from stable intermediate
543 levels of gene expression (Fig. 3I) was performed according to <100 A.U.F for bin 1, 179-224 for bin 2, 400-
544 500 for bin 3, >1000 A.U.F for bin 4. Cells expressing intermediate levels of gene expression had been sorted
545 from DNMT3A-silenced cells (DNMT3A-dCas9) from the range 300-400 EBFP A.U.F. and maintained in culture
546 for >2 weeks. Cells had been silenced with transfection of DNMT3A-dCas9 targeted to the upstream binding
547 sites (Tables SE1 and SE3), cell sorting for 3000-100000 EYFP A.U.F. (transfection marker), and maintained in
548 culture for >2 weeks. Cell sorting for rTetR-XTEN80-DNMT3A (Fig. 5A) and rTetR-KRAB (Fig. 5C) transfections
549 in cells with stable intermediate levels of expression used 3000-100000 EYFP A.U.F. Cell sorting for rTetR-
550 XTEN80-TET1 (Fig. 5B) transfections in cells with stable intermediate levels of expression used 410-2800 EYFP
551 A.U.F. Cells expressing low and intermediate levels of gene expression had been sorted from DNMT3A-silenced
552 cells (DNMT3A-dCas9) using <100 A.U.F for low expression and 200-500 A.U.F for intermediate expression and
553 maintained in culture for >2 weeks. Cells had been silenced with transfection of DNMT3A-dCas9 targeted to
554 the upstream binding sites (Tables SE1 and SE3), cell sorting for 3000-100000 EYFP A.U.F., and maintained in
555 culture for >2 weeks. Cell sorting to study the gene expression dynamics after epigenetic modulation using rTetR-
556 XTEN80-DNMT3A (Fig. SE.3A) used ranges 60-445, 500-7000, and 7000-100000 EYFP A.U.F. Cell sorting to
557 demonstrate the stability of DNMT3A-silenced (rTetR-XTEN80-DNMT3A) cells expressing intermediate levels of
558 gene expression (Fig. SE.5) was performed according to <80 EBFP2 A.U.F for low expression, and ranges
559 304-380, 400-500, and 530-662 EBFP2 A.U.F for intermediate expression. The cells had been silenced with
560 transfection of rTetR-XTEN80-DNMT3A, supplementation of dox for 8 days, cell sorting for 60-445 EYFP A.U.F.,
561 and maintained in culture for >2 weeks.

562 **Single-cell sampling and expansion**

563 Single cell sorting was performed using a fluorescence-activated cell sorter (BD Aria) for various levels of
564 EBFP2 fluorescence. For the cells exhibiting low levels of fluorescence, single cells were obtained from range
565 21-30 A.U.F. and range 50-60 A.U.F. For the cells exhibiting intermediate levels of fluorescence, single cells
566 were obtained from range 200-240 A.U.F, range 254-305 A.U.F, range 320-384 A.U.F, and range 417-500 A.U.F.
567 For the cells exhibiting high levels of fluorescence, single cells were obtained from range 1000-1200 A.U.F and

range 1750-2100 A.U.F. From each range, 90 single cells well obtained for a total of 720 single cells, each seeded on a separate well of a 96-well tissue culture plate (Corning) with growth media supplemented with penicillin/streptomycin. Single cells were then cultured under standard cell culture conditions at 37℃ in a 5% CO2 incubator. Growth media was replaced independently for each well every four days. After 14 days of cell culture, 93 monoclonal populations were observed, representing $\geq$10 monoclonal populations for each of the 8 ranges sampled. The cells in Fig. 3 A-E were used for cell sorting at 29 days.

MeDIP-qPCR and ChIP-qPCR

MeDIP-qPCR was performed using the MagMeDIP-qPCR kit (Diagenode) according to the manufacturer's instructions. ChIP-qPCR was performed using the True MicroChIP-seq kit (C01010132, Diagenode) according to the manufacturer's instructions. All sonication was performed using a Bioruptor Pico (Diagenode). We utilized the anti-5mC antibody in the MagMeDIP-qPCR kit (Diagenode) for DNA methylation. We utilized antibodies for H3K4me3 (C15410003, Diagenode) and H3K9me3 (C15410056, Diagenode) for histone methylation. The qPCRs were performed using Power SYBR Green PCR Master Mix (Applied Biosystems) and primers (Table SE4) in a LightCycler 96 (Roche). Melt curves were used to confirm a single amplification product and the LightCycler 96 Application Software (Roche) was used to obtain the C_t values. For the reporter gene, the promoter region was amplified. β-Actin (H3K4me3) and IgF2 (H3K9me3 and DNA methylation) were used as reference endogenous genes [5]. Folds reported refer to the standard $\Delta\Delta C_t$ method.

Targeted bisulfite sequencing

Genomic DNA was extracted using the DNeasy Blood & Tissue Kit (Qiagen) according to the manufacture's instructions. Samples were then submitted to Zymo Research for conducting the rest of the assay. Assays targeting CpG sites along the chromosomal integration region comprising the continuous sequence covering all DNA binding sites, promoter and EBFP2 were developed using the Rosefinch tool from Zymo Research. Primers were used to amplify bisulfite-converted DNA from the samples, using the EZ DNA Methylation-Lightning Kit for conversion. Amplified products were pooled, barcoded, and prepared for sequencing on an Illumina MiSeq system with a V2 300bp kit, following paired-end sequencing. Sequencing reads were aligned to the chromosomal integration sequence using Bismark. The estimation of methylation levels for each sampled cytosine was calculated by dividing the count of reads identifying a cytosine (C) by the sum of reads identifying either a cytosine (C) or a thymine (T).

Digital PCR

Genomic DNA extraction was performed with the DNeasy Blood & Tissue Kit (Qiagen), according to the

manufacturer's instructions. Samples were then submitted to Azenta Life Sciences for conducting the rest of the assay. A NanoDrop spectrophotometer was first used to assess sample concentration and confirm purity of the samples. The samples were then loaded in triplicate for analysis on a QIAcuity 8.5k 96-Well Nanoplate using the Qiagen QIAcuity digital PCR system, with probes targeting EBFP2 and COG1. Copy number calculation was conducted using the QIAcuity Software Suite (version 2.2.0.26). Each sample was compared to a no-template control to verify the absence of external contamination.

Desing of gRNA sequences

Sequences for gRNAs were designed in Benchiling using the chromosomal integration sequence and the CHO-K1 genome sequence (National Library of Medicine (NCBI) RefSeq assembly GCF_000223135.1).

Protein structure predictions

Protein structure predictions were performed with AlphaFold2 [60] using MMseqs2 on the Google Colab Platform (ColabFold v1.5.5)[61] on GPUs. AlphaFold2 was used to perform structural predictions using 5 models for each protein. The top ranking predicted structure (according to plDDT score for single chains and pTM score for protein complexes) was then relaxed using Assisted Model Building with Energy Refinement (AMBER) relaxation. The protein structures of PhlF-KRAB, rTetR-KRAB, PhlF, and rTetR were predicted as multimers (homodimers). The protein structures of DNMT3A-dCas9, rTetR-XTEN80-DNMT3A, rTetR-XTEN80-TET1 were predicted as single chains. The structures of rTetR-XTEN80-TET1 and rTetR-XTEN80-DNMT3A were then aligned using the predicted homodimer of rTetR as a template in PyMOL (2.5.2). Minimal rotations ($\leq$1 per linker) were performed at the flexible linkers in PyMOL.

Quantification and statistical analysis

To determine the probability of a specific fraction of CpGs being methylated based on bisulfite sequencing data, which offers the probability of a specific CpG being methylated for the cell population, we employ the following method. This approach is applied to both bisulfite sequencing datasets presented in Figs. 3E and 4D, resulting in the corresponding plots in Figs. 3F and 4F, respectively.

Let us assume that the gene of interest has a total of N CpGs. Then, for any $0 \leq j \leq N$, having j CpGs methylated corresponds to the event where precisely j CpGs are methylated, and all other CpGs (i.e., $N - j$) are unmethylated. Noe, let us define the probability of the described event as P_j, where $0 \leq j \leq N$. Additionally, let the probability of CpG i being methylated be p_i, and the probability of CpG i being unmethylated be q_i, where $1 \leq i \leq N$. The bisulfite sequencing data provide us with p_i, and since the event "CpG unmethylated" is the

complement of "CpG methylated", q_i is known and can be expressed as $q_i = 1 - p_i$. Then, the probability that exactly j CpGs are methylated out of a total of N CpGs, where each CpG has its own probability of being methylated (p_i) or not being methylated (q_i), can be calculated by using the following formula:

$$P(\text{Exactly } j \text{ CpGs are methylated}) = \sum_{\substack{S \subseteq \{1,2,\ldots,N\} \\ |S|=j}} \left(\prod_{i \in S} p_i \cdot \prod_{k \in \{1,2,\ldots,N\} \setminus S} q_k \right). \tag{1}$$

More precisely, the sum ($\sum$) is taken over all possible subsets S of events with exactly j elements. For each subset S, we calculate the product of the probabilities of the events in S occurring ($\prod_{i \in S} p_i$). We then calculate the product of the probabilities of the events not in S occurring ($\prod_{k \in \{1,2,\ldots,N\} \setminus S} q_k$). The overall probability is the sum of these products for all subsets S. In simpler terms, the formula considers all possible combinations of j events occurring out of N events, multiplying the probabilities of those events occurring and the probabilities of the remaining events not occurring in each combination. The sum aggregates these probabilities for all valid combinations, providing the probability that exactly j events happen.

As illustrative example, let us consider the case in which the total number of CpG in the gene of interest are $N = 3$. Then, defining the probability of CpG i to be and not to be methylated as p_i and $q_i = 1 - p_i$, with $i = 1, 2, 3$, we then can calculate the probability of having j CpGs methylated, i.e., P_j, with $j = 0, 1, 2, 3$, using (1), obtaining $P_0 = q_1 q_2 q_3$, $P_1 = p_1 q_2 q_3 + q_1 p_2 q_3 + q_1 q_2 p_3$, $P_2 = p_1 p_2 q_3 + p_1 q_2 p_3 + q_1 p_2 p_3$, and $P_3 = p_1 p_2 p_3$. The probability distribution of the fraction of CpGs methyalated is then obtained by plotting (j, P_j) for $j = 0, 2, ..., N$.

Concerning the best-fit linear regression lines shown in Figs. 3G, 4H, and 4I, they were obtained using MATLAB polyfit function. Furthermore, the p-value associated with the slope coefficient was determined through the MATLAB fitlm function.

For statistical tests comparing percentages calculated from flow cytometry measurements or fold changes from MeDIP-qPCR and ChIP-qPCR, unpaired two sample t-tests were performed.

Computational analysis

All simulations conducted in this paper were done using Gillespie's Stochastic Simulation Algorithm (SSA) [45].

Author contributions

S.P., S.B., and D.D.V. conceptualized the study. D.D.V. directed and supervised the study. S.P., S.B., R.W., E.S., A.K. and D.D.V. contributed to experimental design and/or data analysis. S.P., E.S., A.K. and K.I. conducted or assisted experimental investigation. S.B. performed mathematical modeling and data analysis. S.P., S.B. and D.D.V. wrote the manuscript.

Data and materials availability

Data to evaluate the conclusions of the study will be in the manuscript and supplementary materials before publication. Original data is available from the corresponding author upon request. All sequences used in the study are to be included in the manuscript, and plasmids will be deposited in Addgene before publication. Targeted bisulfite sequencing data will be available from a public repository before publication.

Code availability

Code for mathematical modeling is available from the corresponding author upon request.

Acknowledgements

We thank Dr. James J. Collins for providing scientific comments and suggestions on the link between the model predictions and the experimental data, Kalon Overholt for providing useful references on DNA methylation and providing feedback on some of the figures, and the Biological Engineering Comm Lab for helping us refine some of the diagrams. We would like to thank the Koch Institute's Flow Cytometry Core at MIT where the cell sorting was performed. Furthermore, we would like to thank our funding: NSF-MODULUS, Award number 2027949.

Declaration of interests

The authors declare no competing interests.

References

1. Felsenfeld, G. A brief history of epigenetics. *Cold Spring Harb Perspect Biol.* **6,** a018200. https://doi.org/10.1101/cshperspect.a018200 (2012).
2. Nishiyama, A. & Nakanishi, M. Navigating the DNA Methylation Landscape of Cancer. *Trends in Genetics* **37,** 1012–1027. ISSN: 0168-9525. https://www.sciencedirect.com/science/article/pii/S016895252100130X (2021).
3. Jones, M., Goodman, S. & Kobor, M. DNA Methylation and Healthy Human Aging. *Aging Cell* **14,** 924–932 (Dec. 2015).
4. Saxton, D. S. & Rine, J. Distinct silencer states generate epigenetic states of heterochromatin. *Mol Cell* **82,** 3566–3579. https://doi.org/10.1016/j.molcel.2022.08.002 (2022).
5. Bintu, L. *et al.* Dynamics of epigenetic regulation at the single-cell level. *Science*. https://doi.org/10.1126/science.aab2956 (2016).
6. Holmberg, J. & Perlmann, T. Maintaining differentiated cellular identity. *Nat Rev Genet* **13,** 429–39. https://doi.org/10.1038/nrg3209 (2012).
7. Zhang, Q. & Cao, X. Epigenetic regulation of the innate immune response to infection. *Nat Rev Immunol* **19,** 417–432. https://doi.org/10.1038/s41577-019-0151-6 (2019).
8. Day, J. J. & Sweatt, J. D. DNA methylation and memory formation. *Nat Neurosci* **13,** 1319–23. https://doi.org/10.1038/nn.2666 (2010).
9. Thakore, P., Black, J., Hilton, I. & Gersbach, C. Editing the epigenome: technologies for programmable transcription and epigenetic modulation. *Nat Methods* **13,** 127–137 (Feb. 2016).
10. Kouzarides, T. Chromatin modifications and their function. *Cell* **128,** 693–705. https://doi.org/10.1016/j.cell.2007.02.005 (2007).
11. Allis, C. D., Caparros, M.-L., Jenuwein, T. & Reinberg, D. *Epigenetics* (Cold Spring Harbor Laboratory Press, Second Edition, 2015).
12. Huang, S., Litt, M. & Blakey, C. A. *Epigenetic Gene Expression and Regulation* (Academic Press, 2015).
13. Allis, C. D. & Jenuwein, T. The molecular hallmarks of epigenetic control. *Nat Rev Genet* **17,** 487–500. https://doi.org/10.1038/nrg.2016.59 (2016).
14. Li, E. & Zhang, Y. DNA methylation in mammals. *Cold Spring Harb Perspect Biol* **6.** https://doi.org/10.1101/cshperspect.a019133 (2014).

15. Cheng, X. Structural and functional coordination of DNA and histone methylation. *Cold Spring Harb Perspect Biol.* **6,** 693–705. https://doi.org/10.1101/cshperspect.a018747 (2014).

16. Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410,** 120–4. https://doi.org/10.1038/35065138 (2001).

17. Jones, P. L. *et al.* Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet* **19,** 187–191. https://doi.org/10.1038/561 (1998).

18. Schneider, R. *et al.* Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. *Nat Cell Biol* **6,** 73–7. https://doi.org/10.1038/ncb1076 (2003).

19. Hathaway, N. A. *et al.* Dynamics and memory of heterochromatin in living cells. *Cell* **149,** 1447–1460. https://doi.org/10.1016/j.cell.2012.03.052 (2012).

20. Zhang, T., Cooper, S. & Brockdorff, N. The interplay of histone modifications - writers that read. *EMBO Rep* **16,** 1467–1481. https://doi.org/10.15252/embr.201540945 (2015).

21. Angel, A., Song, J., Dean, C. & Howard, M. A Polycomb-based switch underlying quantitative epigenetic memory. *Nature* **476,** 105–108. https://doi.org/10.1038/nature10241 (2011).

22. Dodd, I. B., Micheelsen, M. A., Sneppen, K. & Thon, G. Theoretical Analysis of Epigenetic Cell Memory by Nucleosome Modification. *Cell* **129,** 813–822. https://doi.org/10.1016/j.cell.2007.02.053 (2007).

23. Ulm, E. A., Sleiman, S. F. & Chamberlin, H. M. Developmental functions for the Caenorhabditis elegans Sp protein SPTF-3. *Mech Dev* **128,** 428–41. https://doi.org/10.1016/j.mod.2011.08.003 (2011).

24. Yao, Z. *et al.* A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell* **184,** 3222–3241. https://doi.org/10.1016/j.cell.2021.04.021 (2021).

25. Stanley, G., Gokce, O., Malenka, R. C., Südhof, T. C. & Quake, S. R. Continuous and Discrete Neuron Types of the Adult Murine Striatum. *Neuron* **105,** 688–699. https://doi.org/10.1016/j.neuron.2019.11.004 (2020).

26. Phillips, J. W. *et al.* A repeated molecular architecture across thalamic pathways. *Nat Neurosci* **22,** 1925–1935. https://doi.org/10.1038/s41593-019-0483-3 (2019).

27. Gaidukov, L. *et al.* A multi-landing pad DNA integration platform for mammalian cell engineering. *Nucleic Acids Res.* **46,** 4072–4086 (May 2018).

28. Recillas-Targa, F. *et al.* Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 6883–6888 (May 2002).

29. Stanton, B. C. *et al.* Systematic transfer of prokaryotic sensors and circuits to mammalian cells. *ACS Synth. Biol.* **3,** 880–891 (Dec. 2014).

30. O'Geen, H. *et al.* dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression. *Nucleic Acids Res.* **45,** 9901–9916 (Sept. 2017).

31. Fuks, F. *et al.* The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J Biol Chem* **278,** 4035–4040. https://doi.org/10.1074/jbc.M210256200 (2003).

32. Margolin, J. F. *et al.* Krüppel-associated boxes are potent transcriptional repression domains. *Proc. Natl. Acad. Sci.* **91,** 4509–4513. 10.1073/pnas.91.10.4509 (1994).

33. Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G. & Rauscher, F. J. SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Annual Review of Physical Chemistry* **16,** 919–32. https://doi.org/10.1101/gad.973302 (2002).

34. Ecco, G., Imbeault, M. & Trono, D. KRAB zinc finger proteins. *Development* **144,** 2719–2729. https://doi.org/10.1242/dev.132605 (2017).

35. Nuñez, J. K. *et al.* Genome-wide programmable transcriptional memory by CRISPR-based epigenome editing. *Cell* **184,** 2503–2519 (2021).

36. Hervouet, E., Peixoto, P., Delage-Mourroux, R., Boyer-Guittaut, M. & Cartron, P.-F. Specific or not specific recruitment of DNMTs for DNA methylation, an epigenetic dilemma. *Clin Epigenetics,* 17. https://doi.org/10.1186/s13148-018-0450-y (2018).

37. Lopez-Serra, L. & Esteller, M. Proteins that bind methylated DNA and human cancer: reading the wrong words. *Br J Cancer* **98,** 1881–1885. https://doi.org/10.1038/sj.bjc.6604374 (2008).

38. Trerotola, M., Relli, V., Simeone, P. & Alberti, S. Epigenetic inheritance and the missing heritability. *Hum Genomics* **9,** 17. https://doi.org/10.1186/s40246-015-0041-3 (2015).

39. Upadhyay, A. K., Horton, J. R., Zhang, X. & Cheng, X. Coordinated methyl-lysine erasure: structural and functional linkage of a Jumonji demethylase domain and a reader domain. *Curr Opin Struct Biol* **21,** 750–60. https://doi.org/10.1016/j.sbi.2011.08.003 (2011).

40. Rasmussen, K. D. & Helin, K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes Dev.* **30,** 733–50. https://doi.org/10.1101/gad.276568.115 (2016).

41. Brown, D. A. *et al.* The SET1 Complex Selects Actively Transcribed Target Genes via Multivalent Interaction with CpG Island Chromatin. *Cell Rep* **20,** 2313–2327. https://doi.org/10.1016/j.celrep.2017.08.030 (2017).

42. Bruno, S., Williams, R. J. & Del Vecchio, D. Epigenetic cell memory: The gene's inner chromatin modification circuit. *PLOS Computational Biology* **18,** 1–27. https://doi.org/10.1371/journal.pcbi.1009961 (2022).

759 43. Jackson, M. *et al.* Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation
760 in embryonic stem cells. *Mol Cell Biol* **24.** https://doi.org/10.1128/MCB.24.20.8862-8871 (2004).

761 44. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **21,** 381–395.

762 45. Gillespie, D. T. Stochastic Simulation of Chemical Kinetics. *Annual Review of Physical Chemistry* **58,** 35–55.
763 https://doi.org/10.1146/annurev.physchem.58.032806.104637 (2007).

764 46. Ren, W. *et al.* Direct readout of heterochromatic H3K9me3 regulates DNMT1-mediated maintenance DNA
765 methylation. *Proceedings of the National Academy of Sciences of the United States of America* **117,**
766 18439–18447. https://doi.org/10.1073/pnas.2009316117 (2020).

767 47. Laird, C. D. *et al.* Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary
768 strands of individual DNA molecules. *Proceedings of the National Academy of Sciences of the United States*
769 *of America* **101,** 204–209. https://doi.org/10.1073/pnas.2536758100 (2004).

770 48. Riggs, A. & Xiong, Z. Methylation and epigenetic fidelity. *Proceedings of the National Academy of Sciences*
771 *of the United States of America* **101,** 4–5. https://doi.org/10.1073/pnas.2536758100 (2004).

772 49. Ito, S. *et al.* Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*
773 **333,** 1300–3. https://doi.org/10.1126/science.1210597 (2011).

774 50. Liu, X. S. *et al.* Editing DNA methylation in the mammalian genome. *Cell* **167,** 233–247 (Sept. 2016).

775 51. Mattei, A. L., Bailly, N. & Meissner, A. DNA methylation: a historical perspective. *Trends Genet* **38,** 676–707.
776 https://doi.org/10.1016/j.tig.2022.03.010 (2022).

777 52. Hong, T. *et al.* TET2 modulates spatial relocalization of heterochromatin in aged hematopoietic stem and
778 progenitor cells. *Nat Aging* **3,** 1387–1400. https://doi.org/10.1038/s43587-023-00505-y (2023).

779 53. Du, J., Johnson, L. M., Jacobsen, S. E. & Patel, D. J. DNA methylation pathways and their crosstalk with
780 histone methylation. *Nat Rev Mol Cell Biol* **16,** 519–32. https://doi.org/10.1038/nrm4043 (2015).

781 54. Ooi, S. K. *et al.* DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA.
782 *Nature* **448,** 714–717. https://doi.org/10.1038/nature05987 (2007).

783 55. Nan, X. *et al.* Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone
784 deacetylase complex. *Nature* **393,** 383–389. https://doi.org/10.1038/30764 (1998).

785 56. Zeng, H. What is a cell type and how to define it? *Cell* **185,** 2739–2755. https://doi.org/10.1016/j.cell.2022.
786 06.031 (2022).

787 57. Liu, H. *et al.* DNA methylation atlas of the mouse brain at single-cell resolution. *Nature* **598,** 120–128.
788 https://doi.org/10.1038/s41586-020-03182-8 (2021).

58. Clark, S. J. *et al.* scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat Commun* **9,** 781. https://doi.org/10.1038/s41467-018-03149-4 (2018).

59. Waddington, C. H. *The strategy of the genes* (Routledge, 1957).

60. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596,** 583–589. ISSN: 1476-4687. https://doi.org/10.1038/s41586-021-03819-2 (Aug. 2021).

61. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nature Methods* **19,** 679–682. ISSN: 1548-7105. https://doi.org/10.1038/s41592-022-01488-1 (June 2022).

62. Jin, C. *et al.* TET1 is a maintenance DNA demethylase that prevents methylation spreading in differentiated cells. *Nucleic Acids Research* **42,** 6956–71. https://doi.org/10.1093/nar/gku372 (2014).

63. Huisinga, K. L., Brower-Toland, B. & Elgin, S. C. The contradictory definitions of heterochromatin: transcription and silencing. *Chromosoma* **115,** 110–22. https://doi.org/10.1007/s00412-006-0052-x (2006).

64. Del Vecchio, D. & Murray, R. M. *Biomolecular Feedback Systems* (Princeton University Press, 2014).

65. Bruno, S., Williams, R. J. & Del Vecchio, D. *Model reduction and stochastic analysis of the histone modification circuit* in *2022 European Control Conference (ECC)* (2022), 264–271.

66. Felsenfeld, G. & Groudine, M. Controlling the double helix. *Nature* **421,** 448–53. https://doi.org/10.1038/nature01411 (2003).

67. Friedman, N., Cai, L. & Xie, X. S. Linking Stochastic Dynamics to Population Distribution: An Analytical Framework of Gene Expression. *Phys. Rev. Lett* **97.** https://doi.org/10.1103/PhysRevLett.97.168302 (2006).

Supplementary Information (SI)

S.1 Model of chromatin modification circuit

The chromatin modification circuit reaction model considered in our study, shown in Fig. SM.1A, was constructed starting from the one in [42] and it is based on known molecular interactions from the literature. Here, we give a brief description of the reaction model (see [42] for additional details). The chromatin modifications considered in the circuit are H3K9 methylation (H3K9me3), H3K4 methylation (H3K4me3), and DNA methylation. The basic unit of the model is the nucleosome with DNA wrapped around it, D, that can be modified with H3K4me3, $\mathrm{D^A}$, DNA methylation, $\mathrm{D_1^R}$, H3K9me3, $\mathrm{D_2^R}$, or both H3K9me3 and DNA methylation, $\mathrm{D_{12}^R}$. The reaction model can be graphically represented by the circuit of Fig. SM.1A, whose associated reactions are given in SI Section S.6, Table SM.1. More precisely, writer enzymes have the ability to *de novo* establish chromatin marks. Additionally, histone modifications recruit writer enzymes for the same modification to nearby modifiable nucleosomes (*auto-catalysis*) [11, 19, 20, 38], and DNA methylation and repressive histone modification cooperate by recruiting each other's writer enzymes (*cross-catalysis*) [31]. Ultimately, these modifications can be passively removed through dilution during DNA replication or by the action of eraser enzymes (*basal erasure*). Both activating and repressive modifications are involved in the recruitment of each other's eraser enzymes (*recruited erasure*) [12, 39–41].

Now, defining the number of $\mathrm{D^A}, \mathrm{D_1^R}, \mathrm{D_2^R}, \mathrm{D_{12}^R}, \mathrm{D}$ as n^A, n_1^R, n_2^R, n_{12}^R, and n^D, let us derive the related ordinary differential equation (ODE) model in terms of the fractions $\bar{D}^A = n^A/\mathrm{D_{tot}}$, $\bar{D}_1^R = n_{D_1^R}/\mathrm{D_{tot}}$, $\bar{D}_2^R = n_{D_2^R}/\mathrm{D_{tot}}$, $\bar{D}_{12}^R = n_{D_{12}^R}/\mathrm{D_{tot}}$ and $\bar{D} = n_D/\mathrm{D_{tot}}$, with $\mathrm{D_{tot}}$ the total number of nucleosomes in a gene of interest. This can be done by assuming that $\mathrm{D_{tot}}$ is sufficiently large so that $\bar{D}^A$, $\bar{D}_1^R$, $\bar{D}_2^R$, $\bar{D}_{12}^R$ and $\bar{D}$ can be treated as real number. Now, let us introduce $D_{tot} = \mathrm{D_{tot}}/\Omega$ where Ω is the reaction volume, and define the normalized inputs: $u_{10}^R = k_{W0}^1/k_M^A D_{tot}$, $u_1^R = k_W^1/(k_M^A D_{tot})$, $u_{20}^R = k_{W0}^2/(k_M^A D_{tot})$, $u_2^R = k_W^2/(k_M^A D_{tot})$, $u_0^A = k_{W0}^A/(k_M^A D_{tot})$, and $u^A = k_W^A/(k_M^A D_{tot})$. Furthermore, let us introduce the following dimensionless parameters:

$$\alpha = \frac{k_M}{k_M^A} \quad \bar{\alpha} = \frac{\bar{k}_M}{k_M^A}, \quad \alpha' = \frac{k'_M}{k_M^A}, \tag{SM.1}$$

in which α ($\bar{\alpha}$, α') is the non-dimensional rate constant associated with auto-catalysis (cross-catalysis). Let

$$\varepsilon = \frac{\delta + \bar{k}_E^A}{k_M^A D_{tot}}, \quad \varepsilon' = \frac{k_E^A}{k_M^A}, \quad \mu = \frac{k_E^R}{k_E^A}, \quad \mu' = \frac{k_T'^*}{k_E^A}, \tag{SM.2}$$

with $b = O(1)$ such that $(\delta + \bar{k}_E^R)/(\delta + \bar{k}_E^A) = b\mu$ and $\beta = O(1)$ such that $(\delta' + k'_T)/(\delta + \bar{k}_E^A) = \beta\mu'$. Then,

μ (μ') encapsulates the asymmetry between the erasure rate of repressive histone modifications (the DNA demethylation rate) and the erasure rate of activating marks. Furthermore, given that $(\delta + \bar{k}_E^R)/(k_M^A D_{tot}) = b\varepsilon\mu$, $(\delta' + k_T')/(k_M^A D_{tot}) = \beta\varepsilon\mu'$, $k_E^R/k_M^A = \mu\varepsilon'$ and $k_T'^*/k_M^A = \mu'\varepsilon'$, ε (ε') is a parameter that scales the ratio between the basal (recruited) erasure rate and the auto-catalysis rate of each modification. Considering the normalized time $\tau = tk_M^A D_{tot}$, the ODEs associated with the chromatin modification circuit are

$$
\begin{aligned}
\frac{d\bar{D}_1^R}{d\tau} &= (u_{10}^R + u_1^R + \alpha'(\bar{D}_2^R + \bar{D}_{12}^R))\bar{D} + \mu(b\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{12}^R - (u_{20}^R + \alpha(\bar{D}_2^R + \bar{D}_{12}^R) + \bar{\alpha}(\bar{D}_1^R + \bar{D}_{12}^R))\bar{D}_1^R \\
&\quad - \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_1^R \\
\frac{d\bar{D}_{12}^R}{d\tau} &= (u_{10}^R + \alpha'(\bar{D}_2^R + \bar{D}_{12}^R))\bar{D}_2^R + (u_{20}^R + \alpha(\bar{D}_2^R + \bar{D}_{12}^R) + \bar{\alpha}(\bar{D}_1^R + \bar{D}_{12}^R))\bar{D}_1^R \\
&\quad - (\mu'(\beta\varepsilon + \varepsilon'\bar{D}^A) + \mu(b\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_{12}^R \qquad \text{(SM.3)} \\
\frac{d\bar{D}_2^R}{d\tau} &= (u_{20}^R + u_2^R + \alpha(\bar{D}_2^R + \bar{D}_{12}^R) + \bar{\alpha}(\bar{D}_1^R + \bar{D}_{12}^R))\bar{D} + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{12}^R - (u_{10}^R + \alpha'(\bar{D}_2^R + \bar{D}_{12}^R))\bar{D}_2^R \\
&\quad - \mu(b\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_2^R \\
\frac{d\bar{D}^A}{d\tau} &= (u_0^A + u^A + \bar{D}^A)\bar{D} - (\varepsilon + \varepsilon'(\bar{D}_2^R + \bar{D}_{12}^R) + \varepsilon'(\bar{D}_1^R + \bar{D}_{12}^R))\bar{D}^A,
\end{aligned}
$$

with $\bar{D} = 1 - \bar{D}_1^R - \bar{D}_{12}^R - \bar{D}_2^R - \bar{D}^A$ (see [42] for full derivation). It is important to note that in our reporter gene, the expression of fluorescent protein EBFP2 is driven by a constitutive promoter. This means that the promoter remains active continuously under normal conditions, without the need for external activators. This is the reason why, for $\mathrm{D^A}$, we introduce a constant non-basal *de-novo* establishment term ($u^A > 0$). As for the other species, we will describe in the following section how external inputs, such as KRAB, DNMT3A, and TET1, modulate the parameters of the chromatin modification circuit model.

S.1.1 How KRAB, DNMT3A and TET1 modulate model parameters

In our study, we exploit KRAB, DNMT3A, and TET1 to modulate chromatin modification state and they affect the model parameters as follows. KRAB is a well-known epigenetic effector that mediates the recruitment of H3K9me3 writers [33]. As we introduce KRAB through transient transfection, its concentration will gradually decrease due to dilution until it eventually dies out. Furthermore, KRAB is fused to the DNA binding domain (DBD) PhlF. This fusion not only allows us to obtain sequence-specific recruitment, but also to regulate the binding and unbinding of KRAB through DAPG, as DAPG inhibits the ability of PhlF-KRAB to bind to DNA. To incorporate these aspects into our model, we write the expression for k_W^2, i.e., reaction rate constant for H3K9me3

establishment (see reactions in SI Table SM.1), as follows:

$$k_W^2 = \tilde{k}_W^2 \frac{1}{1 + \frac{DAPG}{K_2}} W_2 e^{-\delta t}, \tag{SM.4}$$

in which W_2 is the total amount of PhlF-KRAB, $\tilde{k}_W^2$ is a parameter independent of PhlF-KRAB, K_2 is the dissociation constant related to the binding - unbinding reactions between DAPG and PhlF-KRAB, and δ is the dilution rate constant. Then, the normalized input u_2^R in the ODE model (SM.3) can be written as

$$u_2^R = \frac{k_W^2}{k_M^A D_{tot}} = \frac{\tilde{k}_W^2}{k_M^A} \frac{1}{1 + \frac{DAPG}{K_2}} \frac{W_2}{D_{tot}} e^{-\frac{\delta}{k_M^A D_{tot}} t k_M^A D_{tot}} = \tilde{u}_2^R \frac{1}{1 + \frac{DAPG}{K_2}} \bar{W}_2 e^{-\varepsilon\tau}, \tag{SM.5}$$

in which we define $\tilde{u}_2^R = \tilde{k}_W^2 / k_M^A$ and $\bar{W}_2 = W_2 / D_{tot}$.

The enzyme DNMT3A is a DNA methyltransferase that catalyzes *de novo* DNA methylation ([11], Chapter 5) and, as a consequence, it modulates the rate of *de novo* DNA methylation establishment k_W^1 (see reactions in SI Table SM.1). As for KRAB, we introduce DNMT3 through transient transfection, and therefore its concentration will gradually decrease. To incorporate this aspect into our model, we can modify the expression for k_W^1 as follows:

$$k_W^1 = \tilde{k}_W^1 W_1 e^{-\delta t}, \tag{SM.6}$$

in which W_1 is the total amount of DNMT3A and $\tilde{k}_W^1$ is a parameter independent of DNMT3A. Then, the normalized input u_1^R in the ODE model (SM.3) can be written as

$$u_1^R = \frac{k_W^1}{k_M^A D_{tot}} = \frac{\tilde{k}_W^1}{k_M^A} \frac{W_1}{D_{tot}} e^{-\frac{\delta}{k_M^A D_{tot}} t k_M^A D_{tot}} = \tilde{u}_1^R \bar{W}_1 e^{-\varepsilon\tau}, \tag{SM.7}$$

in which we define $\tilde{u}_1^R = \tilde{k}_W^1 / k_M^A$ and $\bar{W}_1 = W_1 / D_{tot}$.

Finally, TET1 is an enzyme that catalyzes the conversion of DNA methylation into modifications that are not recognized by the maintenance enzyme DNMT1, and hence are passively removed through dilution [12, 40, 62]. This enzyme then modulates the constants k_T' and $k_T'^*$ in the reactions of SI Table SM.1. As for KRAB, we introduce TET1 through transient transfection and we fuse it to a DBD. More specifically, we employ rTetR, a DBD that binds DNA exclusively in the presence of doxycycline (dox). To incorporate these aspects into our model, we can modify the expressions for k_T' and $k_T'^*$ as follows:

$$k_T' = k_{T1}' + \tilde{k}_T' \frac{\frac{dox}{K_1}}{1 + \frac{dox}{K_1}} T e^{-\delta t}, \quad k_T'^* = k_{T1}'^* + \tilde{k}_T'^* \frac{\frac{dox}{K_1}}{1 + \frac{dox}{K_1}} T e^{-\delta t}, \tag{SM.8}$$

in which k'_{T1} (k'^{*}_{T1}) is the component of the rate coefficient that does not depend on the external rTetR-TET1 transiently transfected, T is the total amount of rTetR-TET1 transiently transfected, $\tilde{k}'_T$ and $\tilde{k}'^{*}_T$ are parameters independent of rTetR-TET1, and K_1 is the dissociation constant related to the binding - unbinding reactions between dox and rTetR-TET1. Then, the non-dimensional parameter μ' in the ODE model (SM.3) can be written as

$$\mu' = \frac{k'^{*}_T}{k^A_E} = \frac{k'^{*}_{T1}}{k^A_E} + \frac{\tilde{k}'^{*}_T D_{tot}}{k^A_E} \frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}} \frac{T}{D_{tot}} e^{-\frac{\delta}{k^A_M D_{tot}} t k^A_M D_{tot}} = \mu'_1 + \tilde{\mu}' \frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}} \bar{T} e^{-\varepsilon\tau}, \tag{SM.9}$$

in which we define $\mu'_1 = \frac{k'^{*}_{T1}}{k^A_E}$, $\tilde{\mu}'$, and $\bar{T} = T/D_{tot}$. From (SM.9), it is possible to note that, without transfection of rTetR-TET1, $\mu' = \mu'_1$, indicating that μ' is non-zero. This is because of the presence of passive DNA demethylation. However, previous experimental studies suggest that passive DNA demethylation, although present, is a slow process [5, 43].

S.2 Model of transcriptional regulation

The gene expression process involves two main steps. The first step is transcription, during which the genetic information is transcribed from DNA into mRNA (m). The second step is translation, during which the mRNA is translated into the gene product (X). Chromatin state affects transcription by modulating nucleosome compaction and then gene expression [11, 63]. Thus, we assume that transcription is predominantly allowed by $\mathrm{D^A}$, while allowing a low level of transcription to D and to all the species characterized by repressive chromatin modifications. Furthermore, mRNA m is subject to dilution due to cell division and degradation, while the gene product X is subject to dilution only. This is because fluorescent reporters are highly stable and thus not degraded but only diluted [64]. The reactions associated with the gene expression model can be found in SI Table SM.3. Considering the normalized time $\tau = t k^A_M D_{tot}$ and introducing the non-dimensional parameters $\bar{\beta}^A_m = \beta^A_m / k^A_M D_{tot}$, $\bar{\beta}_m = \beta_m / k^A_M D_{tot}$, $\bar{\beta} = \beta / k^A_M D_{tot}$, $\bar{\gamma}_m = \gamma_m / k^A_M D_{tot}$, $\bar{\delta} = \delta / k^A_M D_{tot}$, the ODE model for X and m can be written as

$$\frac{d\bar{m}}{d\tau} = \bar{\beta}^A_m \bar{D}^A + \bar{\beta}_m (\bar{D} + \bar{D}^R_1 + \bar{D}^R_2 + \bar{D}^R_{12}) - \bar{\gamma}_m \bar{m}, \quad \frac{d\bar{X}}{d\tau} = \bar{\beta}\bar{m} - \bar{\delta}\bar{X}. \tag{SM.10}$$

In this paper, we will refer to the model that combines the complete chromatin modification circuit model (SI Section S.1, equations (SM.3)) along with the transcriptional regulation model outlined in this section as the "4D+X model".

S.3 Effect of α' and μ' on the behavior of the system

In this section, we analyze the stochastic behavior of the 4D+X model with a specific focus on understanding how the parameters α' (normalized rate of DNA methylation establishment through repressive histone modifications) and μ' (ratio between the DNA demethylation rate and the activating histone modification erasure rate) affect the probability distribution of gene expression levels and thus the type of memory that can be achieved.

When the rate of DNA demethylation is 0 ($\mu' = 0$), achieving analog memory is possible only when repressive histone modification (H3K9me3) does not catalyze the *de novo* establishment of DNA methylation ($\alpha' = 0$) (Fig. SM.1B). For $\alpha' > 0$, the gene expression level shifts either to a low or high level. Furthermore, the higher α', the more the distribution tends to shift towards a low gene expression level (Fig. SM.2 - top panel). When μ' is different from 0, only digital memory can be achieved (Fig. SM.2 - intermediate and bottom panels). Moreover, the higher μ' is, the more the gene expression level shifts towards a high level (Fig. SM.2 - bottom panel). Overall, these results suggest that analog memory can be achieved only when DNA demethylation rate is sufficiently small compared to histone modification erasure rate (that is, μ' can be approximated as zero on the time scales of interest) and DNA methylation is not catalyzed by repressive histone modifications ($\alpha' = 0$).

If these conditions are verified, and external inputs (KRAB, DNMT3A, and TET1) are not applied or are applied transiently (transient transfection), then the number of nucleosomes with methylated DNA stabilizes at a constant value. Introducing the variable $\bar{Y}_1 = \bar{D}_1^R + \bar{D}_{12}^R$, that is, the fraction of nucleosomes with methylated DNA in the gene, the behavior of the original model (SM.3) can be captured by a reduced 2D ordinary differential equation (ODE) model below:

$$\frac{d\bar{D}_2^R}{d\tau} = (\alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_1)\bar{D} - \mu(b\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_2^R, \qquad \frac{d\bar{D}^A}{d\tau} = (u^A + \bar{D}^A)\bar{D} - (\varepsilon + \varepsilon'(\bar{D}_2^R + \bar{Y}_1))\bar{D}^A, \qquad \text{(SM.11)}$$

with $\bar{D} = 1 - \bar{D}^A - \bar{D}_2^R - \bar{Y}_1$ and $\bar{Y}_1 = constant$ (full derivation below). A representative diagram of the reduced system is depicted in Fig. SM.1C. It is important to point out that in the original model of chromatin modification circuit, it is assumed that the basic unit is a nucleosome with DNA wrapped around it and with only one modifiable CpG [42]. There are cases in which this assumption is not verified. However, it is possible to demonstrate that removing this assumption yields a similar reduced model. This indicates that the assumption does not impact the qualitative results regarding the influence of the parameters α' and μ' on the probability distribution of gene expression levels (see SI Section S.7).

Derivation of model (SM.11)

Let us first merge the rates linked to the enhancement of H3K9me3 establishment by $\bar{D}_{12}^R$ into a single rate. We will assume that this rate is identical to the one associated with the enhancement of H3K9me3 establishment by $\bar{D}_1^R$. Similarly, let us merge the rates linked to the recruited erasure of H3K4me3 by $\bar{D}_{12}^R$ into a single rate and assume that this rate is identical to the one associated with the recruited erasure of H3K4me3 by $\bar{D}_1^R$. These simplifying assumptions do not affect the qualitative results related to the effect of the cooperative and competitive interactions among chromatin modifications on epigenetic cell memory. The ODE system (SM.3) can then be rewritten as

$$\begin{aligned}
\frac{d\bar{D}_1^R}{d\tau} &= (u_{10}^R + u_1^R + \alpha'(\bar{D}_2^R + \bar{D}_{12}^R))\bar{D} + \mu(b\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{12}^R - (u_{20}^R + \alpha(\bar{D}_2^R + \bar{D}_{12}^R) + \bar{\alpha}(\bar{D}_1^R + \bar{D}_{12}^R))\bar{D}_1^R \\
&\quad - \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_1^R \\
\frac{d\bar{D}_{12}^R}{d\tau} &= (u_{10}^R + \alpha'(\bar{D}_2^R + \bar{D}_{12}^R))\bar{D}_2^R + (u_{20}^R + \alpha\bar{D}_2^R + \bar{\alpha}(\bar{D}_1^R + \bar{D}_{12}^R))\bar{D}_1^R \\
&\quad - (\mu'(\beta\varepsilon + \varepsilon'\bar{D}^A) + \mu(b\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_{12}^R \qquad \text{(SM.12)} \\
\frac{d\bar{D}_2^R}{d\tau} &= (u_{20}^R + u_2^R + \alpha\bar{D}_2^R + \bar{\alpha}(\bar{D}_1^R + \bar{D}_{12}^R))\bar{D} + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{12}^R - (u_{10}^R + \alpha'(\bar{D}_2^R + \bar{D}_{12}^R))\bar{D}_2^R \\
&\quad - \mu(b\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_2^R \\
\frac{d\bar{D}^A}{d\tau} &= (u_0^A + u^A + \bar{D}^A)\bar{D} - (\varepsilon + \varepsilon'(\bar{D}_2^R) + \varepsilon'(\bar{D}_1^R + \bar{D}_{12}^R))\bar{D}^A,
\end{aligned}$$

Now, let us rewrite system (SM.12) by assuming negligible basal *de novo* establishment ($u_{10}^R = u_{20}^R = u_0^A = 0$) and introducing the variable $\bar{Y}_1 = \bar{D}_1^R + \bar{D}_{12}^R$, which corresponds to the fraction of nucleosomes with methylated DNA. We then obtain

$$\begin{aligned}
\frac{d\bar{Y}_1}{d\tau} &= (u_1^R)\bar{D} + (\alpha'(\bar{D}_2^R + \bar{D}_{12}^R))(\bar{D} + \bar{D}_2^R) - \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{Y}_1 \\
\frac{d\bar{D}_{12}^R}{d\tau} &= (\alpha'(\bar{D}_2^R + \bar{D}_{12}^R))\bar{D}_2^R + (\alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_1)\bar{D}_1^R - (\mu'(\beta\varepsilon + \varepsilon'\bar{D}^A) + \mu(b\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_{12}^R \qquad \text{(SM.13)} \\
\frac{d\bar{D}_2^R}{d\tau} &= (u_2^R + \alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_1)\bar{D} + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{12}^R - (\alpha'(\bar{D}_2^R + \bar{D}_{12}^R))\bar{D}_2^R - (\mu(b\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_2^R \\
\frac{d\bar{D}^A}{d\tau} &= (u^A + \bar{D}^A)\bar{D} - (\varepsilon + \varepsilon'\bar{D}_2^R + \varepsilon'\bar{Y}_1)\bar{D}^A,
\end{aligned}$$

with $\bar{D} = 1 - \bar{Y}_1 - \bar{D}_2^R - \bar{D}^A$. Based on previous experimental results [5] and our experimental findings (see Fig. 2H), we can further simplify our model by accounting for the fact that the enhancement of DNA methylation

establishment by H3K9me3 is negligible, i.e., $\alpha' = 0$. The ODE system (SM.13) can then be rewritten as

$$\begin{aligned}
\frac{d\bar{Y}_1}{d\tau} &= (u_1^R)\bar{D} - \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{Y}_1 \\
\frac{d\bar{D}_{12}^R}{d\tau} &= (\alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_1)\bar{D}_1^R - (\mu'(\beta\varepsilon + \varepsilon'\bar{D}^A) + \mu(b\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_{12}^R \\
\frac{d\bar{D}_2^R}{d\tau} &= (u_2^R + \alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_1)\bar{D} + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{12}^R - \mu(b\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_2^R \\
\frac{d\bar{D}^A}{d\tau} &= (u^A + \bar{D}^A)\bar{D} - (\varepsilon + \varepsilon'\bar{D}_2^R + \varepsilon'\bar{Y}_1)\bar{D}^A,
\end{aligned} \tag{SM.14}$$

with $\bar{D} = 1 - \bar{Y}_1 - \bar{D}_2^R - \bar{D}^A$. Now, let us rewrite the ODE system (SM.14) introducing the expressions for u_2^R, u_1^R, and μ' derived in SI Section S.1.1 (Expressions (SM.5), (SM.7), and (SM.9)):

$$\begin{aligned}
\frac{d\bar{Y}_1}{d\tau} &= \tilde{u}_1^R\bar{W}_1 e^{-\varepsilon\tau}\bar{D} - (\mu_1' + \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau})(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{Y}_1 \\
\frac{d\bar{D}_{12}^R}{d\tau} &= (\alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_1)\bar{D}_1^R - ((\mu_1' + \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau})(\beta\varepsilon + \varepsilon'\bar{D}^A) + \mu(b\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_{12}^R \\
\frac{d\bar{D}_2^R}{d\tau} &= (\tilde{u}_2^R\frac{1}{1+\frac{DAPG}{K_2}}\bar{W}_2 e^{-\varepsilon\tau} + \alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_1)\bar{D} + (\mu_1' + \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau})(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{12}^R - \mu(b\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_2^R \\
\frac{d\bar{D}^A}{d\tau} &= (u^A + \bar{D}^A)\bar{D} - (\varepsilon + \varepsilon'\bar{D}_2^R + \varepsilon'\bar{Y}_1)\bar{D}^A,
\end{aligned} \tag{SM.15}$$

with $\bar{D} = 1 - \bar{Y}_1 - \bar{D}_2^R - \bar{D}^A$ and $\bar{D}_1^R = \bar{Y}_1 - \bar{D}_{12}^R$. Let us now consider the parameter regime in which the rate of DNA demethylation, without transfection of rTetR-TET1, is sufficiently low compared to histone modification dynamics, i.e., $\mu_1' \approx 0$. This assumption is consistent with experimental data that suggest that the passive DNA demethylation is a slow process [5, 43]. The system (SM.15) can then be rewritten as follows:

$$\begin{aligned}
\frac{d\bar{Y}_1}{d\tau} &= \tilde{u}_1^R\bar{W}_1 e^{-\varepsilon\tau}\bar{D} - \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau}(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{Y}_1 \\
\frac{d\bar{D}_{12}^R}{d\tau} &= (\alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_1)\bar{D}_1^R - (\tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau}(\beta\varepsilon + \varepsilon'\bar{D}^A) + \mu(b\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_{12}^R \\
\frac{d\bar{D}_2^R}{d\tau} &= (\tilde{u}_2^R\frac{1}{1+\frac{DAPG}{K_2}}\bar{W}_2 e^{-\varepsilon\tau} + \alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_1)\bar{D} + \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau}(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{12}^R - \mu(b\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_2^R \\
\frac{d\bar{D}^A}{d\tau} &= (u^A + \bar{D}^A)\bar{D} - (\varepsilon + \varepsilon'\bar{D}_2^R + \varepsilon'\bar{Y}_1)\bar{D}^A,
\end{aligned} \tag{SM.16}$$

with $\bar{D} = 1 - \bar{Y}_1 - \bar{D}_2^R - \bar{D}^A$ and $\bar{D}_1^R = \bar{Y}_1 - \bar{D}_{12}^R$. From (SM.16), it follows that, after a temporary phase during which the externally transfected inputs decrease, $e^{-\varepsilon\tau} \approx 0$ and then $d\bar{Y}_1/d\tau \approx 0$, and the last two ODEs in (SM.16) depend only on $\bar{D}_2^R$, $\bar{D}^A$, and $\bar{Y}_1$. This implies that, after a transient change, $\bar{Y}_1$ stabilizes at a constant

value and the behavior of the original model can be captured by the reduced two-dimensional ODE model given by the last two equations in (SM.16), that corresponds to the ODE model in (SM.11).

S.4 Effect of ε and ε' on the behavior of the system

In this section, we analyze the deterministic and stochastic behavior of our system, with a specific focus on understanding the effect of parameters ε (parameter that scales the ratio between the basal erasure rate and the auto-catalysis rate of each modification) and ε' (parameter that scales the ratio between the recruited erasure rate and the auto-catalysis rate of each modification) on the probability distribution of gene expression levels.

We first conduct a deterministic analysis by studying the ODE model associated with the simplified chromatin modification circuit, i.e., equations (SM.11), and determine how ε and ε' affect the value of $\bar{D}^A$ at the equilibrium for different fractions of methylated CpGs in the gene, i.e., $\bar{Y}_1$ (Fig. SM.3A). When ε is large, the system consistently exhibits a unique stable steady state characterized by low $\bar{D}^A$ (Fig. SM.3A). The value of $\bar{D}^A$ is smaller for higher $\bar{Y}_1$. As ε decreases, the steady-state value of D^A increases, particularly for low values of $\bar{Y}_1$, where $\bar{D}^A \approx 1$ (Fig. SM.3A). Further reduction in ε results in the system becoming bistable for intermediate values of $\bar{Y}_1$ (Fig. SM.3A). Changes in ε' do not significantly impact these trends, except for when ε is small. In such cases, larger values of ε' reduce the range of $\bar{Y}_1$ for bistability and diminish the difference in the values of $\bar{D}^A$ between the two steady states (Fig. SM.3A). We next study the impact of ε and ε' on the number of methylated CpGs at equilibrium for various initial levels of DNMT3A, represented by $\bar{W}_1$ as defined in SI Eq. (SM.6) (Fig. SM.3B). In general, $\bar{Y}_1$ increases with higher $\bar{W}_1$. However, for large values of ε, $\bar{Y}_1$ remains low over an extended range of $\bar{W}_1$ values. As ε decreases, $\bar{Y}_1$ reaches higher values. In the case of small ε, an increase in ε' results in a more ultrasensitive curve.

Overall, this analysis suggests that high fractions of nucleosomes with H3K4me3, and then a high level of gene expression, is achievable only when ε is sufficiently small (Fig. SM.3A). In this parameter regime, when ε' is sufficiently high, $\bar{Y}_1$ has an ultrasensitive response to DNMT3A transient dosage and then different ranges of DNMT3A pulse height would lead to either high or low gene expression levels only (Fig. SM.1D and Fig. SM.3B). To validate these findings, we use the Stochastic Simulation Algorithm (SSA) [45] to perform a computational study on the 4D+X model, that is, the model that incorporates the full chromatin modification circuit, whose reactions are listed in SI Tables SM.1 and SM.3 (Fig. SM.1E and Fig. SM.3C). Here, we consider several ranges of the initial amount of DNMT3A transfected ($\bar{W}_1$), which are uniformly distributed.

For different ranges of initial DNMT3A transfection levels, a bimodal probability distribution of gene expression

levels emerges when ε' is sufficiently large (Fig. SM.3C, left-hand side plots). As ε' decreases, the stationary distribution becomes unimodal (Fig. SM.3C, left-hand side plots). When considering ranges that include higher values of $\bar{W}_1$, it becomes more evident how the unique peak of the distribution shifts towards lower gene expression levels for sufficiently high values of $\bar{W}_1$ (Fig. SM.1E, left-hand side plots and Fig. SM.3C, right-hand side plots).

We then conduct a similar computational analysis to assess the impact of ε and ε' on number of methylated CpGs and gene expression level across various initial levels of TET1 enzyme ($\bar{T}$) (Fig. SM.3D,E). The obtained results mirror those described in the previous analysis (Fig. SM.3B,C).

By looking at our experimental results involving transfection of KRAB (Fig. 2F), it is possible to note that the probability distribution of gene expression levels is bimodal. Previous computational study suggests that this behavior can be obtained only for sufficiently small values of ε [42, 65]. Furthermore, our experimental results involving transfection of DNMT3A (Fig. 2C) also show a bimodal probability distribution for gene expression levels. Based on our mathematical analysis (Fig. SM.3B,C), these outcomes can be observed only when ε is sufficiently small and ε' is sufficiently large. We validate these hypotheses by successfully replicating our experimental data via simulations considering this parameter regime, i.e, small ε, large ε' (Fig. SM.4).

S.5 Refined chromatin modification circuit model to replicate the bisulfite sequencing data from the monoclonal experiment (Fig. 4)

In the study conducted in the previous sections, we observed that, in the absence of external stimuli and when α' and μ' can be assumed equal to zero, the chromatin modification circuit model, whose associated reactions are listed in Table SM.1, predicts a "frozen" level of DNA methylation, i.e., $\dot{\bar{Y}}_1 = 0$ (SI Section S.3). However, this model does not perfectly align with the experimental results obtained from the bisulfite sequencing data of the synthetic reporter system in the eight different clones, each once associated with a distinct EBFP2 expression level (Fig. 4A, C, D, and E). In fact, for each clone, we do not observe a single value for methylated CpGs across the entire cell population. Instead, we observe a distribution with mean and standard deviation that remain approximately constant throughout the observed time points (Fig. 4F). Notably, the distribution of methylated CpGs appears more spread for clones associated with an intermediate level of methylated CpGs.

In order to recapitulate this observed behavior, we incorporated in the full model known molecular mechanisms involved in the DNA methylation maintenance process that were omitted before. Specifically, in the original full

model we have

$$\mathrm{D}_1^{\mathrm{R}} \xrightarrow{\delta'} \mathrm{D}, \tag{SM.17}$$

in which δ' represents the net passive demethylation rate constant obtained from the balance between the dilution and the DNMT1 maintenance process. In order to recapitulate the experimental results, we added the following reactions:

$$\mathrm{D} \xrightarrow{c_0} \mathrm{D}_1^{\mathrm{R}}, \quad \mathrm{D}_1^{\mathrm{R}} \xrightarrow{\tilde{\delta}' \cdot \left(\frac{1}{1+D_2^R/K}\right)} \mathrm{D}, \quad \mathrm{D} \xrightarrow{c \cdot D_2^R} \mathrm{D}_1^{\mathrm{R}}. \tag{SM.18}$$

It is important to point out that, for the sake of simplicity in the description of the DNA methylation maintenance model above, we used, with abuse of notation, $\mathrm{D}_1^{\mathrm{R}}$ to denote any species with DNA methylation and $\mathrm{D}_2^{\mathrm{R}}$ to denote any species with H3K9me3. The first reaction in (SM.18) represents the possibility that the DNA methylation maintenance enzyme DNMT1 can cause *de novo* DNA methylation, although at a very low rate (low c_0) [47, 48]. The second reaction represents the fact that H3K9me3 helps recruit DNMT1 to the gene and then reinforces DNMT1-mediated maintenance [46]. Finally, we introduce the third reaction since H3K9me3 helps recruit DNMT1 and DNMT1 can, by mistake, establish DNA methylation on previously unmethylated sites.

Considering the normalized time $\tau = tk_M^A D_{tot}$ and introducing the non-dimensional parameters $\bar{\delta}' = \delta'/k_M^A D_{tot}$, $\bar{\tilde{\delta}}' = \tilde{\delta}'/k_M^A D_{tot}$, $\bar{c}_0 = c_0/k_M^A D_{tot}$, $\bar{c} = c/k_M^A$, the ODE describing the dynamics of fraction of nucleosomes with methylated DNA, that is, $\bar{Y}_1$, can be then rewritten as

$$\frac{d\bar{Y}_1}{d\tau} = (\bar{c}_0 + \bar{c}\bar{D}_2^R)\bar{D} - (\bar{\delta}' + \bar{\tilde{\delta}}'\frac{1}{1+\bar{D}_2^R/\bar{K}})\bar{D}_1^R. \tag{SM.19}$$

Then, the average level of DNA methylation remains approximately constant in time when $d\bar{Y}_1/d\tau \approx 0$, that is, if all the reaction rate constants are sufficiently small ($\bar{c}_0, \bar{c}, \bar{\delta}', \bar{\tilde{\delta}}' \ll 1$).

This is confirmed by the computational study conducted by simulating the reactions of the refined chromatin modification circuit model that incorporates the DNA methylation maintenance model described above, whose reactions are listed in SI Table SM.5, using the SSA (SI Fig. SM.6A). Additionally, this analysis shows that the reactions encapsulating the recruitment mechanism of DNMT1 by H3K9me3 (the second and the third in (SM.18)) are crucial for obtaining broader distributions in correspondence of intermediate levels of methylated CpGs (SI Fig. SM.6B). In fact, as illustrated in the figure, when the rate constants associated with these reactions are low, all distributions have approximately the same height and width. However, higher values of these rates result in broader distributions at intermediate levels of methylated CpGs.

S.6 4D+X model: reactions

S.6.1 Chromatin modification circuit

In this reaction model, it is assumed that the rate constants of the establishment, auto and cross-catalytic, and erasure processes of H3K9me3 (DNA methylation) do not change if the other repressive mark is present on the same modifiable unit. The chromatin modification circuit reaction model can then be written as in SI Table SM.1.

R_j	Reaction	Prop.Func.(a_j)	Param.	R_j	Reaction	Prop.Func.(a_j)	Param.
1	$\mathrm{D} \xrightarrow{k^A_{W0}} \mathrm{D^A}$	$a_1 = k^A_{W0} n^D$	k^A_{W0}	21	$\mathrm{D} + \mathrm{D^R_2} \xrightarrow{k_M} \mathrm{D^R_2} + \mathrm{D^R_2}$	$a_{21} = \frac{k_M}{\Omega} n^D n^R_2$	$\frac{k_M}{\Omega}$
2	$\mathrm{D} \xrightarrow{k^A_W} \mathrm{D^A}$	$a_2 = k^A_W n^D$	k^A_W	22	$\mathrm{D} + \mathrm{D^R_{12}} \xrightarrow{k_M} \mathrm{D^R_2} + \mathrm{D^R_{12}}$	$a_{22} = \frac{k_M}{\Omega} n^D n^R_{12}$	$\frac{k_M}{\Omega}$
3	$\mathrm{D^A} \xrightarrow{\bar{k}^A_E} \mathrm{D}$	$a_3 = \bar{k}^A_E n^A$	$\bar{k}^A_E$	23	$\mathrm{D} + \mathrm{D^R_1} \xrightarrow{\bar{k}_M} \mathrm{D^R_2} + \mathrm{D^R_1}$	$a_{23} = \frac{\bar{k}_M}{\Omega} n^D n^R_1$	$\frac{\bar{k}_M}{\Omega}$
4	$\mathrm{D^A} \xrightarrow{\delta} \mathrm{D}$	$a_4 = \delta n^A$	δ	24	$\mathrm{D} + \mathrm{D^R_{12}} \xrightarrow{\bar{k}_M} \mathrm{D^R_2} + \mathrm{D^R_{12}}$	$a_{24} = \frac{\bar{k}_M}{\Omega} n^D n^R_{12}$	$\frac{\bar{k}_M}{\Omega}$
5	$\mathrm{D} + \mathrm{D^A} \xrightarrow{k^A_M} \mathrm{D^A} + \mathrm{D^A}$	$a_5 = \frac{k^A_M}{\Omega} n^D n^A$	$\frac{k^A_M}{\Omega}$	25	$\mathrm{D^R_2} + \mathrm{D^A} \xrightarrow{k^R_E} \mathrm{D} + \mathrm{D^A}$	$a_{25} = \frac{k^R_E}{\Omega} n^R_2 n^A$	$\frac{k^R_E}{\Omega}$
6	$\mathrm{D^A} + \mathrm{D^R_1} \xrightarrow{k^A_E} \mathrm{D} + \mathrm{D^R_1}$	$a_6 = \frac{k^A_E}{\Omega} n^A n^R_1$	$\frac{k^A_E}{\Omega}$	26	$\mathrm{D^R_1} \xrightarrow{k^2_{W0}} \mathrm{D^R_{12}}$	$a_{26} = k^2_{W0} n^R_1$	k^2_{W0}
7	$\mathrm{D^A} + \mathrm{D^R_{12}} \xrightarrow{k^A_E} \mathrm{D} + \mathrm{D^R_{12}}$	$a_7 = \frac{k^A_E}{\Omega} n^A n^R_{12}$	$\frac{k^A_E}{\Omega}$	27	$\mathrm{D^R_{12}} \xrightarrow{\bar{k}^R_E} \mathrm{D^R_1}$	$a_{27} = \bar{k}^R_E n^R_{12}$	$\bar{k}^R_E$
8	$\mathrm{D^A} + \mathrm{D^R_2} \xrightarrow{k^A_E} \mathrm{D} + \mathrm{D^R_2}$	$a_8 = \frac{k^A_E}{\Omega} n^A n^R_2$	$\frac{k^A_E}{\Omega}$	28	$\mathrm{D^R_{12}} \xrightarrow{\delta} \mathrm{D^R_1}$	$a_{28} = \delta n^R_{12}$	δ
9	$\mathrm{D^A} + \mathrm{D^R_{12}} \xrightarrow{k^A_E} \mathrm{D} + \mathrm{D^R_{12}}$	$a_9 = \frac{k^A_E}{\Omega} n^A n^R_{12}$	$\frac{k^A_E}{\Omega}$	29	$\mathrm{D^R_1} + \mathrm{D^R_2} \xrightarrow{k_M} \mathrm{D^R_{12}} + \mathrm{D^R_2}$	$a_{29} = \frac{k_M}{\Omega} n^R_1 n^R_2$	$\frac{k_M}{\Omega}$
10	$\mathrm{D} \xrightarrow{k^1_{W0}} \mathrm{D^R_1}$	$a_{10} = k^1_{W0} n^D$	k^1_{W0}	30	$\mathrm{D^R_1} + \mathrm{D^R_{12}} \xrightarrow{k_M} \mathrm{D^R_{12}} + \mathrm{D^R_{12}}$	$a_{30} = \frac{k_M}{\Omega} n^R_1 n^R_{12}$	$\frac{k_M}{\Omega}$
11	$\mathrm{D} \xrightarrow{k^1_W} \mathrm{D^R_1}$	$a_{11} = k^1_W n^D$	k^1_W	31	$\mathrm{D^R_1} + \mathrm{D^R_1} \xrightarrow{\bar{k}_M} \mathrm{D^R_{12}} + \mathrm{D^R_1}$	$a_{31} = \frac{\bar{k}_M}{\Omega} \frac{n^R_1(n^R_1-1)}{2}$	$\frac{\bar{k}_M}{\Omega}$
12	$\mathrm{D^R_1} \xrightarrow{k'_T} \mathrm{D}$	$a_{12} = k'_T n^R_1$	k'_T	32	$\mathrm{D^R_1} + \mathrm{D^R_{12}} \xrightarrow{\bar{k}_M} \mathrm{D^R_{12}} + \mathrm{D^R_{12}}$	$a_{32} = \frac{\bar{k}_M}{\Omega} n^R_1 n^R_{12}$	$\frac{\bar{k}_M}{\Omega}$
13	$\mathrm{D^R_1} \xrightarrow{\delta'} \mathrm{D}$	$a_{13} = \delta' n^R_1$	δ'	33	$\mathrm{D^R_{12}} + \mathrm{D^A} \xrightarrow{k^R_E} \mathrm{D^R_1} + \mathrm{D^A}$	$a_{33} = \frac{k^R_E}{\Omega} n^R_{12} n^A$	$\frac{k^R_E}{\Omega}$
14	$\mathrm{D} + \mathrm{D^R_2} \xrightarrow{k'_M} \mathrm{D^R_1} + \mathrm{D^R_2}$	$a_{14} = \frac{k'_M}{\Omega} n^D n^R_2$	$\frac{k'_M}{\Omega}$	34	$\mathrm{D^R_2} \xrightarrow{k^1_{W0}} \mathrm{D^R_{12}}$	$a_{34} = k^1_{W0} n^R_2$	k^1_{W0}
15	$\mathrm{D} + \mathrm{D^R_{12}} \xrightarrow{k'_M} \mathrm{D^R_1} + \mathrm{D^R_{12}}$	$a_{15} = \frac{k'_M}{\Omega} n^D n^R_{12}$	$\frac{k'_M}{\Omega}$	35	$\mathrm{D^R_{12}} \xrightarrow{k'_T} \mathrm{D^R_2}$	$a_{35} = k'_T n^R_{12}$	k'_T
16	$\mathrm{D^R_1} + \mathrm{D^A} \xrightarrow{k'^*_T} \mathrm{D} + \mathrm{D^A}$	$a_{16} = \frac{k'^*_T}{\Omega} n^R_1 n^A$	$\frac{k'^*_T}{\Omega}$	36	$\mathrm{D^R_{12}} \xrightarrow{\delta'} \mathrm{D^R_2}$	$a_{36} = \delta' n^R_{12}$	δ'
17	$\mathrm{D} \xrightarrow{k^2_{W0}} \mathrm{D^R_2}$	$a_{17} = k^2_{W0} n^D$	k^2_{W0}	37	$\mathrm{D^R_2} + \mathrm{D^R_2} \xrightarrow{k'_M} \mathrm{D^R_{12}} + \mathrm{D^R_2}$	$a_{37} = \frac{k'_M}{\Omega} \frac{n^R_2(n^R_2-1)}{2}$	$\frac{k'_M}{\Omega}$
18	$\mathrm{D} \xrightarrow{k^2_W} \mathrm{D^R_2}$	$a_{18} = k^2_W n^D$	k^2_W	38	$\mathrm{D^R_2} + \mathrm{D^R_{12}} \xrightarrow{k'_M} \mathrm{D^R_{12}} + \mathrm{D^R_{12}}$	$a_{38} = \frac{k'_M}{\Omega} n^R_2 n^R_{12}$	$\frac{k'_M}{\Omega}$
19	$\mathrm{D^R_2} \xrightarrow{\bar{k}^R_E} \mathrm{D}$	$a_{19} = \bar{k}^R_E n^R_2$	$\bar{k}^R_E$	39	$\mathrm{D^R_{12}} + \mathrm{D^A} \xrightarrow{k'^*_T} \mathrm{D^R_2} + \mathrm{D^A}$	$a_{39} = \frac{k'^*_T}{\Omega} n^R_{12} n^A$	$\frac{k'^*_T}{\Omega}$
20	$\mathrm{D^R_2} \xrightarrow{\delta} \mathrm{D}$	$a_{20} = \delta n^R_2$	δ				

Table SM.1: Full chromatin modification circuit model: reactions

R_j	Reaction	Prop.Func.(a_j)	Param.	R_j	Reaction	Prop.Func.(a_j)	Param.
1	$\mathrm{D} \xrightarrow{k^A_{W0}} \mathrm{D^A}$	$a_1 = k^A_{W0} n^D$	k^A_{W0}	7	$\mathrm{D} \xrightarrow{k^2_{W0}} \mathrm{D^R_2}$	$a_7 = k^2_{W0} n^D$	k^2_{W0}
2	$\mathrm{D} \xrightarrow{k^A_W} \mathrm{D^A}$	$a_2 = k^A_W n^D$	k^A_W	8	$\mathrm{D} \xrightarrow{k^2_W} \mathrm{D^R_2}$	$a_8 = k^2_W n^D$	k^2_W
3	$\mathrm{D^A} \xrightarrow{\bar{k}^A_E} \mathrm{D}$	$a_3 = \bar{k}^A_E n^A$	$\bar{k}^A_E$	9	$\mathrm{D^R_2} \xrightarrow{\bar{k}^R_E} \mathrm{D}$	$a_9 = \bar{k}^R_E n^R_2$	$\bar{k}^R_E$
4	$\mathrm{D^A} \xrightarrow{\delta} \mathrm{D}$	$a_4 = \delta n^A$	δ	10	$\mathrm{D^R_2} \xrightarrow{\delta} \mathrm{D}$	$a_{10} = \delta n^R_2$	δ
5	$\mathrm{D} + \mathrm{D^A} \xrightarrow{k^A_M} \mathrm{D^A} + \mathrm{D^A}$	$a_5 = \frac{k^A_M}{\Omega} n^D n^A$	$\frac{k^A_M}{\Omega}$	11	$\mathrm{D} + \mathrm{D^R_2} \xrightarrow{k_M} \mathrm{D^R_2} + \mathrm{D^R_2}$	$a_{11} = \frac{k_M}{\Omega} n^D n^R_2$	$\frac{k_M}{\Omega}$
6	$\mathrm{D^A} + \mathrm{D^R_2} \xrightarrow{k^A_E} \mathrm{D} + \mathrm{D^R_2}$	$a_6 = \frac{k^A_E}{\Omega} n^A n^R_2$	$\frac{k^A_E}{\Omega}$	12	$\mathrm{D^R_2} + \mathrm{D^A} \xrightarrow{k^R_E} \mathrm{D} + \mathrm{D^A}$	$a_{12} = \frac{k^R_E}{\Omega} n^R_2 n^A$	$\frac{k^R_E}{\Omega}$

Table SM.2: Histone modification circuit model: reactions

When DNA methylation is not included in the system, the only species included in the model are D, $\mathrm{D^A}$, and $\mathrm{D^R_2}$, and the reactions included in the model become only the ones listed in SI Table SM.2.

S.6.2 Transcriptional regulation

The reactions associated with the gene expression model described in SI Section S.2 are listed in SI Table SM.3. Here, m is the mRNA, X is the gene product, γ_m is the decay rate constant of m due to dilution and degradation, δ is the decay rate constant of X due to dilution, β_m^A, β_m are the transcription rate constants, with $\beta_m < \beta_m^A$, and β is the translation rate constant.

$\mathbf{R}_j$	Reaction	Prop.Func.(a_j)	Param.	$\mathbf{R}_j$	Reaction	Prop.Func.(a_j)	Param.
1	$\mathrm{D^A} \xrightarrow{\beta_m^A} \mathrm{D^A} + \mathrm{m}$	$a_1 = \beta_m^A n^A$	β_m^A	5	$\mathrm{D_{12}^R} \xrightarrow{\beta_m} \mathrm{D_{12}^R} + \mathrm{m}$	$a_5 = \beta_m n_{12}^R$	β_m
2	$\mathrm{D} \xrightarrow{\beta_m} \mathrm{D} + \mathrm{m}$	$a_2 = \beta_m n^D$	β_m	6	$\mathrm{m} \xrightarrow{\beta} \mathrm{m} + \mathrm{X}$	$a_6 = \beta n^m$	β
3	$\mathrm{D_1^R} \xrightarrow{\beta_m} \mathrm{D_1^R} + \mathrm{m}$	$a_3 = \beta_m n_1^R$	β_m	7	$\mathrm{m} \xrightarrow{\gamma_m} \emptyset$	$a_7 = \gamma_m n^m$	γ_m
4	$\mathrm{D_2^R} \xrightarrow{\beta_m} \mathrm{D_2^R} + \mathrm{m}$	$a_4 = \beta_m n_2^R$	β_m	8	$\mathrm{X} \xrightarrow{\delta} \emptyset$	$a_8 = \delta n^X$	δ

Table SM.3: Gene expression model: reactions

When DNA methylation, and then species $\mathrm{D_1^R}$ and $\mathrm{D_{12}^R}$ are not included in the chromatin modification circuit, then the gene expression model can be simplified as in SI Table SM.4.

$\mathbf{R}_j$	Reaction	Prop.Func.(a_j)	Param.	$\mathbf{R}_j$	Reaction	Prop.Func.(a_j)	Param.
1	$\mathrm{D^A} \xrightarrow{\beta_m^A} \mathrm{D^A} + \mathrm{m}$	$a_1 = \beta_m^A n^A$	β_m^A	4	$\mathrm{m} \xrightarrow{\beta} \mathrm{m} + \mathrm{X}$	$a_4 = \beta n^m$	β
2	$\mathrm{D} \xrightarrow{\beta_m} \mathrm{D} + \mathrm{m}$	$a_2 = \beta_m n^D$	β_m	5	$\mathrm{m} \xrightarrow{\gamma_m} \emptyset$	$a_5 = \gamma_m n^m$	γ_m
3	$\mathrm{D_2^R} \xrightarrow{\beta_m} \mathrm{D_2^R} + \mathrm{m}$	$a_3 = \beta_m n_2^R$	β_m	6	$\mathrm{X} \xrightarrow{\delta} \emptyset$	$a_6 = \delta n^X$	δ

Table SM.4: Gene expression model when we remove species $\mathrm{D_1^R}$, $\mathrm{D_{12}^R}$: reactions

S.6.3 DNA methylation maintenance and refined chromatin modification circuit

To recapitulate the experimental results shown in Fig. 4F, modifications need to be made to the reaction system that models the DNA methylation maintenance process, as explained in SI Section S.5. The reaction model for the chromatin modification circuit can then be revised and expressed as detailed in SI Table SM.5.

$\mathbf{R}_j$	Reaction	Prop.Func.(a_j)	Param.	$\mathbf{R}_j$	Reaction	Prop.Func.(a_j)	Param.
1	$\mathrm{D} \xrightarrow{k^A_{W0}} \mathrm{D^A}$	$a_1 = k^A_{W0} n^D$	k^A_{W0}	21	$\mathrm{D} + \mathrm{D}^\mathrm{R}_2 \xrightarrow{k_M} \mathrm{D}^\mathrm{R}_2 + \mathrm{D}^\mathrm{R}_2$	$a_{21} = \frac{k_M}{\Omega} n^D n^R_2$	$\frac{k_M}{\Omega}$
2	$\mathrm{D} \xrightarrow{k^A_W} \mathrm{D^A}$	$a_2 = k^A_W n^D$	k^A_W	22	$\mathrm{D} + \mathrm{D}^\mathrm{R}_{12} \xrightarrow{k_M} \mathrm{D}^\mathrm{R}_2 + \mathrm{D}^\mathrm{R}_{12}$	$a_{22} = \frac{k_M}{\Omega} n^D n^R_{12}$	$\frac{k_M}{\Omega}$
3	$\mathrm{D^A} \xrightarrow{\bar{k}^A_E} \mathrm{D}$	$a_3 = \bar{k}^A_E n^A$	$\bar{k}^A_E$	23	$\mathrm{D} + \mathrm{D}^\mathrm{R}_1 \xrightarrow{\bar{k}_M} \mathrm{D}^\mathrm{R}_2 + \mathrm{D}^\mathrm{R}_1$	$a_{23} = \frac{\bar{k}_M}{\Omega} n^D n^R_1$	$\frac{\bar{k}_M}{\Omega}$
4	$\mathrm{D^A} \xrightarrow{\delta} \mathrm{D}$	$a_4 = \delta n^A$	δ	24	$\mathrm{D} + \mathrm{D}^\mathrm{R}_{12} \xrightarrow{\bar{k}_M} \mathrm{D}^\mathrm{R}_2 + \mathrm{D}^\mathrm{R}_{12}$	$a_{24} = \frac{\bar{k}_M}{\Omega} n^D n^R_{12}$	$\frac{\bar{k}_M}{\Omega}$
5	$\mathrm{D} + \mathrm{D^A} \xrightarrow{k^A_M} \mathrm{D^A} + \mathrm{D^A}$	$a_5 = \frac{k^A_M}{\Omega} n^D n^A$	$\frac{k^A_M}{\Omega}$	25	$\mathrm{D}^\mathrm{R}_2 + \mathrm{D^A} \xrightarrow{k^R_E} \mathrm{D} + \mathrm{D^A}$	$a_{25} = \frac{k^R_E}{\Omega} n^R_2 n^A$	$\frac{k^R_E}{\Omega}$
6	$\mathrm{D^A} + \mathrm{D}^\mathrm{R}_1 \xrightarrow{k^A_E} \mathrm{D} + \mathrm{D}^\mathrm{R}_1$	$a_6 = \frac{k^A_E}{\Omega} n^A n^R_1$	$\frac{k^A_E}{\Omega}$	26	$\mathrm{D}^\mathrm{R}_1 \xrightarrow{k^2_{W0}} \mathrm{D}^\mathrm{R}_{12}$	$a_{26} = k^2_{W0} n^R_1$	k^2_{W0}
7	$\mathrm{D^A} + \mathrm{D}^\mathrm{R}_{12} \xrightarrow{k^A_E} \mathrm{D} + \mathrm{D}^\mathrm{R}_{12}$	$a_7 = \frac{k^A_E}{\Omega} n^A n^R_{12}$	$\frac{k^A_E}{\Omega}$	27	$\mathrm{D}^\mathrm{R}_{12} \xrightarrow{\bar{k}^R_E} \mathrm{D}^\mathrm{R}_1$	$a_{27} = \bar{k}^R_E n^R_{12}$	$\bar{k}^R_E$
8	$\mathrm{D^A} + \mathrm{D}^\mathrm{R}_2 \xrightarrow{k^A_E} \mathrm{D} + \mathrm{D}^\mathrm{R}_2$	$a_8 = \frac{k^A_E}{\Omega} n^A n^R_2$	$\frac{k^A_E}{\Omega}$	28	$\mathrm{D}^\mathrm{R}_{12} \xrightarrow{\delta} \mathrm{D}^\mathrm{R}_1$	$a_{28} = \delta n^R_{12}$	δ
9	$\mathrm{D^A} + \mathrm{D}^\mathrm{R}_{12} \xrightarrow{k^A_E} \mathrm{D} + \mathrm{D}^\mathrm{R}_{12}$	$a_9 = \frac{k^A_E}{\Omega} n^A n^R_{12}$	$\frac{k^A_E}{\Omega}$	29	$\mathrm{D}^\mathrm{R}_1 + \mathrm{D}^\mathrm{R}_2 \xrightarrow{k_M} \mathrm{D}^\mathrm{R}_{12} + \mathrm{D}^\mathrm{R}_2$	$a_{29} = \frac{k_M}{\Omega} n^R_1 n^R_2$	$\frac{k_M}{\Omega}$
10	$\mathrm{D} \xrightarrow{k^1_{W0}} \mathrm{D}^\mathrm{R}_1$	$a_{10} = k^1_{W0} n^D$	k^1_{W0}	30	$\mathrm{D}^\mathrm{R}_1 + \mathrm{D}^\mathrm{R}_{12} \xrightarrow{k_M} \mathrm{D}^\mathrm{R}_{12} + \mathrm{D}^\mathrm{R}_{12}$	$a_{30} = \frac{k_M}{\Omega} n^R_1 n^R_{12}$	$\frac{k_M}{\Omega}$
11	$\mathrm{D} \xrightarrow{k^1_W} \mathrm{D}^\mathrm{R}_1$	$a_{11} = k^1_W n^D$	k^1_W	31	$\mathrm{D}^\mathrm{R}_1 + \mathrm{D}^\mathrm{R}_1 \xrightarrow{\bar{k}_M} \mathrm{D}^\mathrm{R}_{12} + \mathrm{D}^\mathrm{R}_1$	$a_{31} = \frac{\bar{k}_M}{\Omega} \frac{n^R_1(n^R_1-1)}{2}$	$\frac{\bar{k}_M}{\Omega}$
12	$\mathrm{D}^\mathrm{R}_1 \xrightarrow{k'_T} \mathrm{D}$	$a_{12} = k'_T n^R_1$	k'_T	32	$\mathrm{D}^\mathrm{R}_1 + \mathrm{D}^\mathrm{R}_{12} \xrightarrow{\bar{k}_M} \mathrm{D}^\mathrm{R}_{12} + \mathrm{D}^\mathrm{R}_{12}$	$a_{32} = \frac{\bar{k}_M}{\Omega} n^R_1 n^R_{12}$	$\frac{\bar{k}_M}{\Omega}$
13	$\mathrm{D}^\mathrm{R}_1 \xrightarrow{\delta'} \mathrm{D}$	$a_{13} = \delta' n^R_1$	δ'	33	$\mathrm{D}^\mathrm{R}_{12} + \mathrm{D^A} \xrightarrow{k^R_E} \mathrm{D}^\mathrm{R}_1 + \mathrm{D^A}$	$a_{33} = \frac{k^R_E}{\Omega} n^R_{12} n^A$	$\frac{k^R_E}{\Omega}$
14	$\mathrm{D}^\mathrm{R}_1 \xrightarrow{\frac{\tilde{\delta}'}{1+(D^R_2+D^R_{12})/K}} \mathrm{D}$	$a_{14} = \frac{\tilde{\delta}'}{1+(D^R_2+D^R_{12})/K} n^R_1$	$\tilde{\delta}'$	34	$\mathrm{D}^\mathrm{R}_2 \xrightarrow{k^1_{W0}} \mathrm{D}^\mathrm{R}_{12}$	$a_{34} = k^1_{W0} n^R_2$	k^1_{W0}
15	$\mathrm{D} \xrightarrow{c_0} \mathrm{D}^\mathrm{R}_1$	$a_{15} = c_0 n^D$	c_0	35	$\mathrm{D}^\mathrm{R}_{12} \xrightarrow{k'_T} \mathrm{D}^\mathrm{R}_2$	$a_{35} = k'_T n^R_{12}$	k'_T
16	$\mathrm{D} + \mathrm{D}^\mathrm{R}_2 \xrightarrow{c} \mathrm{D}^\mathrm{R}_1 + \mathrm{D}^\mathrm{R}_2$	$a_{16} = \frac{c}{\Omega} n^D n^R_2$	$\frac{c}{\Omega}$	36	$\mathrm{D}^\mathrm{R}_{12} \xrightarrow{\delta'} \mathrm{D}^\mathrm{R}_2$	$a_{36} = \delta' n^R_{12}$	δ'
17	$\mathrm{D} + \mathrm{D}^\mathrm{R}_{12} \xrightarrow{c} \mathrm{D}^\mathrm{R}_1 + \mathrm{D}^\mathrm{R}_{12}$	$a_{17} = \frac{c}{\Omega} n^D n^R_{12}$	$\frac{c}{\Omega}$	37	$\mathrm{D}^\mathrm{R}_{12} \xrightarrow{\frac{\tilde{\delta}'}{1+(D^R_2+D^R_{12})/K}} \mathrm{D}^\mathrm{R}_2$	$a_{14} = \frac{\tilde{\delta}'}{1+(D^R_2+D^R_{12})/K} n^R_{12}$	$\tilde{\delta}'$
18	$\mathrm{D} + \mathrm{D}^\mathrm{R}_2 \xrightarrow{k'_M} \mathrm{D}^\mathrm{R}_1 + \mathrm{D}^\mathrm{R}_2$	$a_{18} = \frac{k'_M}{\Omega} n^D n^R_2$	$\frac{k'_M}{\Omega}$	38	$\mathrm{D}^\mathrm{R}_2 \xrightarrow{c_0} \mathrm{D}^\mathrm{R}_{12}$	$a_{38} = c_0 n^R_2$	c_0
19	$\mathrm{D} + \mathrm{D}^\mathrm{R}_{12} \xrightarrow{k'_M} \mathrm{D}^\mathrm{R}_1 + \mathrm{D}^\mathrm{R}_{12}$	$a_{19} = \frac{k'_M}{\Omega} n^D n^R_{12}$	$\frac{k'_M}{\Omega}$	39	$\mathrm{D}^\mathrm{R}_2 + \mathrm{D}^\mathrm{R}_2 \xrightarrow{c} \mathrm{D}^\mathrm{R}_{12} + \mathrm{D}^\mathrm{R}_2$	$a_{39} = \frac{c}{\Omega} \frac{n^R_2(n^R_2-1)}{2}$	$\frac{c}{\Omega}$
20	$\mathrm{D}^\mathrm{R}_1 + \mathrm{D^A} \xrightarrow{k'^*_T} \mathrm{D} + \mathrm{D^A}$	$a_{20} = \frac{k'^*_T}{\Omega} n^R_1 n^A$	$\frac{k'^*_T}{\Omega}$	40	$\mathrm{D}^\mathrm{R}_2 + \mathrm{D}^\mathrm{R}_{12} \xrightarrow{c} \mathrm{D}^\mathrm{R}_{12} + \mathrm{D}^\mathrm{R}_{12}$	$a_{40} = \frac{c}{\Omega} n^R_2 n^R_{12}$	$\frac{c}{\Omega}$
21	$\mathrm{D} \xrightarrow{k^2_{W0}} \mathrm{D}^\mathrm{R}_2$	$a_{21} = k^2_{W0} n^D$	k^2_{W0}	41	$\mathrm{D}^\mathrm{R}_2 + \mathrm{D}^\mathrm{R}_2 \xrightarrow{k'_M} \mathrm{D}^\mathrm{R}_{12} + \mathrm{D}^\mathrm{R}_2$	$a_{41} = \frac{k'_M}{\Omega} \frac{n^R_2(n^R_2-1)}{2}$	$\frac{k'_M}{\Omega}$
22	$\mathrm{D} \xrightarrow{k^2_W} \mathrm{D}^\mathrm{R}_2$	$a_{22} = k^2_W n^D$	k^2_W	42	$\mathrm{D}^\mathrm{R}_2 + \mathrm{D}^\mathrm{R}_{12} \xrightarrow{k'_M} \mathrm{D}^\mathrm{R}_{12} + \mathrm{D}^\mathrm{R}_{12}$	$a_{42} = \frac{k'_M}{\Omega} n^R_2 n^R_{12}$	$\frac{k'_M}{\Omega}$
23	$\mathrm{D}^\mathrm{R}_2 \xrightarrow{\bar{k}^R_E} \mathrm{D}$	$a_{23} = \bar{k}^R_E n^R_2$	$\bar{k}^R_E$	43	$\mathrm{D}^\mathrm{R}_{12} + \mathrm{D^A} \xrightarrow{k'^*_T} \mathrm{D}^\mathrm{R}_2 + \mathrm{D^A}$	$a_{43} = \frac{k'^*_T}{\Omega} n^R_{12} n^A$	$\frac{k'^*_T}{\Omega}$
24	$\mathrm{D}^\mathrm{R}_2 \xrightarrow{\delta} \mathrm{D}$	$a_{24} = \delta n^R_2$	δ				

Table SM.5: Full chromatin modification circuit model incorporating the refined DNA methylation maintenance model described in SI Section S.5: reactions

S.7 Modification to the chromatin modification circuit model to include more than one CpG per nucleosome

In the original model of the chromatin modification circuit, it is assumed that the basic unit, that is a nucleosome with DNA wrapped around, can have only one modifiable CpG [42]. There are cases in which this assumption is not verified. However, it is possible to show that this assumption does not affect the qualitative results related to the effect of the parameters α' (normalized rate of DNA methylation establishment through repressive histone modifications) and μ' (ratio between the DNA demethylation rate and the activating histone modification erasure rate) on the probability distribution of gene expression levels.

To this end, let us assume that the basic unit of the model, D, is the nucleosome with DNA wrapped around it with two modifiable CpGs. The conclusion obtained in this section would not change if we consider the DNA wrapped around the nucleosome having $n \in \mathbb{Z}^+$ modifiable CpGs. The basic unit of the model can then be modified with H3K4me3, $\mathrm{D^A}$, one DNA methylation, $\mathrm{D^R_1}$, H3K9me3, $\mathrm{D^R_2}$, one DNA methylation and H3K9me3, $\mathrm{D^R_{12}}$, two DNA methylations, $\mathrm{D^R_{11}}$, and two DNA methylation and H3K9me3, $\mathrm{D^R_{112}}$. Now, defining the number of $\mathrm{D^A}, \mathrm{D^R_1}, \mathrm{D^R_2}, \mathrm{D^R_{12}}, \mathrm{D^R_{11}}, \mathrm{D^R_{112}}, \mathrm{D}$ as n^A, n^R_1, n^R_2, n^R_{12}, $n_{D^R_{11}}$, n^R_{112} and n^D, let us derive the related ordinary differential equation (ODE) model in terms of the fractions $\bar{D}^A = n^A/\mathrm{D_{tot}}$, $\bar{D}^R_1 = n_{D^R_1}/\mathrm{D_{tot}}$, $\bar{D}^R_2 = n_{D^R_2}/\mathrm{D_{tot}}$, $\bar{D}^R_{12} = n_{D^R_{12}}/\mathrm{D_{tot}}$, $\bar{D}^R_{11} = n_{D^R_{11}}/\mathrm{D_{tot}}$, $\bar{D}^R_{112} = n_{D^R_{112}}/\mathrm{D_{tot}}$ and $\bar{D} = n_D/\mathrm{D_{tot}}$, with $\mathrm{D_{tot}}$ the total number of nucleosomes on a gene of interest. This can be done by assuming that $\mathrm{D_{tot}}$ is sufficiently large so that $\bar{D}^A$, $\bar{D}^R_1$, $\bar{D}^R_2$, $\bar{D}^R_{12}$ and $\bar{D}$ can be treated as real number. Now, considering the dimensionless parameters and normalized inputs introduced SI Section S.1 and the normalized time $\tau = tk^A_M D_{tot}$, the ODEs associated with the chromatin modification circuit are

$$
\begin{aligned}
\frac{d\bar{D}^R_1}{d\tau} &= (u^R_{10} + u^R_1 + \alpha'(\bar{D}^R_2 + \bar{D}^R_{12} + \bar{D}^R_{112}))\bar{D} + \mu(b\varepsilon + \varepsilon'\bar{D}^A)\bar{D}^R_{12} + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}^R_{11} \\
&\quad - (u^R_{20} + \alpha(\bar{D}^R_2 + \bar{D}^R_{12} + \bar{D}^R_{112}) + \bar{\alpha}(\bar{D}^R_1 + \bar{D}^R_{12} + 2\bar{D}^R_{11} + 2\bar{D}^R_{112}))\bar{D}^R_1 \\
&\quad - (u^R_{10} + u^R_1 + \alpha'(\bar{D}^R_2 + \bar{D}^R_{12} + \bar{D}^R_{112}) + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{D}^R_1 \\
\frac{d\bar{D}^R_{12}}{d\tau} &= (u^R_{10} + \alpha'(\bar{D}^R_2 + \bar{D}^R_{12} + \bar{D}^R_{112}))\bar{D}^R_2 \\
&\quad + (u^R_{20} + \alpha(\bar{D}^R_2 + \bar{D}^R_{12} + \bar{D}^R_{112}) + \bar{\alpha}(\bar{D}^R_1 + \bar{D}^R_{12} + 2\bar{D}^R_{11} + 2\bar{D}^R_{112}))\bar{D}^R_1 + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}^R_{112} \\
&\quad - (\mu'(\beta\varepsilon + \varepsilon'\bar{D}^A) + \mu(b\varepsilon + \varepsilon'\bar{D}^A) + u^R_{10} + u^R_1 + \alpha'(\bar{D}^R_2 + \bar{D}^R_{12} + \bar{D}^R_{112}))\bar{D}^R_{12}
\end{aligned}
$$

$$
\begin{aligned}
\frac{d\bar{D}_{11}^R}{d\tau} &= (u_{10}^R + u_1^R + \alpha'(\bar{D}_2^R + \bar{D}_{12}^R + \bar{D}_{112}^R))\bar{D}_1^R + \mu(b\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{112}^R \\
&\quad - (u_{20}^R + \alpha(\bar{D}_2^R + \bar{D}_{12}^R + \bar{D}_{112}^R) + \bar{\alpha}(\bar{D}_1^R + \bar{D}_{12}^R + 2\bar{D}_{11}^R + 2\bar{D}_{112}^R) + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_1^R \\
\frac{d\bar{D}_{112}^R}{d\tau} &= (u_{20}^R + \alpha(\bar{D}_2^R + \bar{D}_{12}^R + \bar{D}_{112}^R) + \bar{\alpha}(\bar{D}_1^R + \bar{D}_{12}^R + 2\bar{D}_{11}^R + 2\bar{D}_{112}^R))\bar{D}_{11}^R \\
&\quad + (u_{10}^R + \alpha'(\bar{D}_2^R + \bar{D}_{12}^R + \bar{D}_{112}^R))\bar{D}_{12}^R \\
&\quad - (\mu'(\beta\varepsilon + \varepsilon'\bar{D}^A) + \mu(b\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_{112}^R \\
\frac{d\bar{D}_2^R}{d\tau} &= (u_{20}^R + u_2^R + \alpha(\bar{D}_2^R + \bar{D}_{12}^R + \bar{D}_{112}^R) + \bar{\alpha}(\bar{D}_1^R + \bar{D}_{12}^R + 2\bar{D}_{11}^R + 2\bar{D}_{112}^R))\bar{D} + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{12}^R \\
&\quad - (u_{10}^R + \alpha'(\bar{D}_2^R + \bar{D}_{12}^R + \bar{D}_{112}^R) + \mu(b\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_2^R \\
\frac{d\bar{D}^A}{d\tau} &= (u_0^A + u^A + \bar{D}^A)\bar{D} - (\varepsilon + \varepsilon'(\bar{D}_2^R + \bar{D}_{12}^R + \bar{D}_{112}^R) + \varepsilon'(\bar{D}_1^R + \bar{D}_{12}^R + 2\bar{D}_{11}^R + 2\bar{D}_{112}^R))\bar{D}^A,
\end{aligned}
\tag{SM.20}
$$

with $\bar{D} = 1 - \bar{D}_1^R - \bar{D}_{12}^R - \bar{D}_{11}^R - \bar{D}_{112}^R - \bar{D}_2^R - \bar{D}^A$.

As we did for the original ODE system (SM.3), let us merge the rates linked to the enhancement of H3K9me3 establishment by $\bar{D}_{12}^R$, $\bar{D}_{11}^R$, and $\bar{D}_{112}^R$ into a single rate. We will assume that this rate is identical to the one associated with the enhancement of H3K9me3 establishment by $\bar{D}_1^R$. Similarly, let us merge the rates linked to the recruited erasure of H3K4me3 by $\bar{D}_{12}^R$, $\bar{D}_{11}^R$, and $\bar{D}_{112}^R$ into a single rate and assume that this rate is identical to the one associated with the recruited erasure of H3K4me3 by $\bar{D}_1^R$. The ODE system (SM.20) can then be

rewritten as

$$
\begin{aligned}
\frac{d\bar{D}_1^R}{d\tau} &= (u_{10}^R + u_1^R + \alpha'(\bar{D}_2^R + \bar{D}_{12}^R + \bar{D}_{112}^R))\bar{D} + \mu(b\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{12}^R + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{11}^R \\
&\quad - (u_{20}^R + \alpha\bar{D}_2^R + \bar{\alpha}(\bar{D}_1^R + \bar{D}_{12}^R + \bar{D}_{11}^R + \bar{D}_{112}^R))\bar{D}_1^R \\
&\quad - (u_{10}^R + u_1^R + \alpha'(\bar{D}_2^R + \bar{D}_{12}^R + \bar{D}_{112}^R) + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_1^R \\
\frac{d\bar{D}_{12}^R}{d\tau} &= (u_{10}^R + \alpha'(\bar{D}_2^R + \bar{D}_{12}^R + \bar{D}_{112}^R))\bar{D}_2^R \\
&\quad + (u_{20}^R + \alpha\bar{D}_2^R + \bar{\alpha}(\bar{D}_1^R + \bar{D}_{12}^R + \bar{D}_{11}^R + \bar{D}_{112}^R))\bar{D}_1^R + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{112}^R \\
&\quad - (\mu'(\beta\varepsilon + \varepsilon'\bar{D}^A) + \mu(b\varepsilon + \varepsilon'\bar{D}^A) + u_{10}^R + u_1^R + \alpha'(\bar{D}_2^R + \bar{D}_{12}^R + \bar{D}_{112}^R))\bar{D}_{12}^R \qquad \text{(SM.21)} \\
\frac{d\bar{D}_{11}^R}{d\tau} &= (u_{10}^R + u_1^R + \alpha'(\bar{D}_2^R + \bar{D}_{12}^R + \bar{D}_{112}^R))\bar{D}_1^R + \mu(b\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{112}^R \\
&\quad - (u_{20}^R + \alpha\bar{D}_2^R + \bar{\alpha}(\bar{D}_1^R + \bar{D}_{12}^R + \bar{D}_{11}^R + \bar{D}_{112}^R) + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_1^R \\
\frac{d\bar{D}_{112}^R}{d\tau} &= (u_{20}^R + \alpha\bar{D}_2^R + \bar{\alpha}(\bar{D}_1^R + \bar{D}_{12}^R + \bar{D}_{11}^R + \bar{D}_{112}^R))\bar{D}_{11}^R \\
&\quad + (u_{10}^R + \alpha'(\bar{D}_2^R + \bar{D}_{12}^R + \bar{D}_{112}^R))\bar{D}_{12}^R - (\mu'(\beta\varepsilon + \varepsilon'\bar{D}^A) + \mu(b\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_{112}^R
\end{aligned}
$$

$$
\begin{aligned}
\frac{d\bar{D}_2^R}{d\tau} &= (u_{20}^R + u_2^R + \alpha\bar{D}_2^R + \bar{\alpha}(\bar{D}_1^R + \bar{D}_{12}^R + \bar{D}_{11}^R + \bar{D}_{112}^R))\bar{D} + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{12}^R \\
&\quad - (u_{10}^R + \alpha'(\bar{D}_2^R + \bar{D}_{12}^R + \bar{D}_{112}^R) + \mu(b\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_2^R \\
\frac{d\bar{D}^A}{d\tau} &= (u_0^A + u^A + \bar{D}^A)\bar{D} - (\varepsilon + \varepsilon'(\bar{D}_2^R + \bar{D}_1^R + \bar{D}_{12}^R + \bar{D}_{11}^R + \bar{D}_{112}^R))\bar{D}^A,
\end{aligned}
$$

Now, let us rewrite system (SM.21) by assuming negligible basal establishment ($u_{10}^R = u_{20}^R = u_0^A = 0$) and introducing the variable $\bar{Y}_1 = \bar{D}_1^R + \bar{D}_{12}^R$, which corresponds to the fraction of nucleosomes with one methylated CpG, and the variable $\bar{Y}_{1tot} = \bar{D}_1^R + \bar{D}_{12}^R + \bar{D}_{11}^R + \bar{D}_{112}^R$, which corresponds to the fraction of nucleosomes with

at least one methylated CpG. We then obtain

$$
\begin{aligned}
\frac{d\bar{D}_1^R}{d\tau} &= (u_1^R + \alpha'(\bar{D}_2^R + \bar{Y}_{1tot} - \bar{D}_1^R - \bar{D}_{11}^R))\bar{D} + \mu(b\varepsilon + \varepsilon'\bar{D}^A)(\bar{Y}_1 - \bar{D}_1^R) + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{11}^R \\
&\quad - (\alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_{tot} + u_1^R + \alpha'(\bar{D}_2^R + \bar{Y}_{1tot} - \bar{D}_1^R - \bar{D}_{11}^R) + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_1^R \\
\frac{d\bar{D}_{11}^R}{d\tau} &= (u_1^R + \alpha'(\bar{D}_2^R + \bar{Y}_{1tot} - \bar{D}_1^R - \bar{D}_{11}^R))\bar{D}_1^R + \mu(b\varepsilon + \varepsilon'\bar{D}^A)(\bar{Y}_{1tot} - \bar{Y}_1 - \bar{D}_{11}^R)) \\
&\quad - (\alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_{1tot} + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_1^R \\
\frac{d\bar{Y}_1}{d\tau} &= u_1^R(\bar{D} - \bar{D}_1^R) + (\alpha'(\bar{D}_2^R + \bar{Y}_{1tot} - \bar{D}_1^R - \bar{D}_{11}^R))(\bar{D} + \bar{D}_2^R) \\
&\quad + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)(\bar{Y}_{1tot} - \bar{Y}_1) - (\mu'(\beta\varepsilon + \varepsilon'\bar{D}^A) + \alpha'(\bar{D}_2^R + \bar{Y}_{1tot} - \bar{D}_1^R - \bar{D}_{11}^R))\bar{Y}_1 \\
\frac{d\bar{Y}_{1tot}}{d\tau} &= u_1^R\bar{D} + (\alpha'(\bar{D}_2^R + \bar{Y}_{1tot} - \bar{D}_1^R - \bar{D}_{11}^R))(\bar{D} + \bar{D}_2^R) - \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{Y}_1 \\
\frac{d\bar{D}_2^R}{d\tau} &= (u_2^R + \alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_{1tot})\bar{D} + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{12}^R - (\alpha'(\bar{D}_2^R + \bar{D}_{12}^R + \bar{D}_{112}^R) + \mu(b\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_2^R \\
\frac{d\bar{D}^A}{d\tau} &= (u^A + \bar{D}^A)\bar{D} - (\varepsilon + \varepsilon'(\bar{D}_2^R + \bar{Y}_{1tot}))\bar{D}^A,
\end{aligned}
\tag{SM.22}
$$

with $\bar{D} = 1 - \bar{Y}_{1tot} - \bar{D}_2^R - \bar{D}^A$. As we did for the ODE system associated with our original chromatin modification circuit in SI Section S.1, we can further simplify our model by assuming that the enhancement of DNA methylation establishment by H3K9me3 is negligible, i.e., $\alpha' = 0$. The ODE system (SM.22) can then be rewritten as

$$
\begin{aligned}
\frac{d\bar{D}_1^R}{d\tau} &= (u_1^R)\bar{D} + \mu(b\varepsilon + \varepsilon'\bar{D}^A)(\bar{Y}_1 - \bar{D}_1^R) + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{11}^R - (\alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_{tot} + u_1^R + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_1^R \\
\frac{d\bar{D}_{11}^R}{d\tau} &= (u_1^R)\bar{D}_1^R + \mu(b\varepsilon + \varepsilon'\bar{D}^A)(\bar{Y}_{1tot} - \bar{Y}_1 - \bar{D}_{11}^R)) - (\alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_{1tot} + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_1^R \\
\frac{d\bar{Y}_1}{d\tau} &= u_1^R(\bar{D} - \bar{D}_1^R) + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)(\bar{Y}_{1tot} - \bar{Y}_1) - \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{Y}_1 \\
\frac{d\bar{Y}_{1tot}}{d\tau} &= u_1^R\bar{D} - \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{Y}_1 \\
\frac{d\bar{D}_2^R}{d\tau} &= (u_2^R + \alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_{1tot})\bar{D} + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{12}^R - \mu(b\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_2^R \\
\frac{d\bar{D}^A}{d\tau} &= (u^A + \bar{D}^A)\bar{D} - (\varepsilon + \varepsilon'(\bar{D}_2^R + \bar{Y}_{1tot}))\bar{D}^A,
\end{aligned}
\tag{SM.23}
$$

with $\bar{D} = 1 - \bar{Y}_{1tot} - \bar{D}_2^R - \bar{D}^A$. Now, introducing in (SM.23) the expressions for u_2^R, u_1^R, and μ' derived in SI

Section S.1.1 (Expressions (SM.5), (SM.7), and (SM.9)), we obtain:

$$
\begin{aligned}
\frac{d\bar{D}_1^R}{d\tau} &= \tilde{u}_1^R \bar{W}_1 e^{-\varepsilon\tau}\bar{D} + \mu(b\varepsilon + \varepsilon'\bar{D}^A)(\bar{Y}_1 - \bar{D}_1^R) + (\mu_1' + \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau})(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{11}^R \\
&\quad - (\alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_{tot} + \tilde{u}_1^R\bar{W}_1 e^{-\varepsilon\tau} + (\mu_1' + \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau})(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_1^R \\
\frac{d\bar{D}_{11}^R}{d\tau} &= \tilde{u}_1^R \bar{W}_1 e^{-\varepsilon\tau}\bar{D}_1^R + \mu(b\varepsilon + \varepsilon'\bar{D}^A)(\bar{Y}_{1tot} - \bar{Y}_1 - \bar{D}_{11}^R)) \\
&\quad - (\alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_{1tot} + (\mu_1' + \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau})(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_1^R \\
\frac{d\bar{Y}_1}{d\tau} &= \tilde{u}_1^R \bar{W}_1 e^{-\varepsilon\tau}(\bar{D} - \bar{D}_1^R) + (\mu_1' + \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau})(\beta\varepsilon + \varepsilon'\bar{D}^A)(\bar{Y}_{1tot} - \bar{Y}_1) \\
&\quad - (\mu_1' + \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau})(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{Y}_1 \\
\frac{d\bar{Y}_{1tot}}{d\tau} &= \tilde{u}_1^R \bar{W}_1 e^{-\varepsilon\tau}\bar{D} - (\mu_1' + \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau})(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{Y}_1 \\
\frac{d\bar{D}_2^R}{d\tau} &= (\tilde{u}_2^R\frac{1}{1+\frac{DAPG}{K_2}}\bar{W}_2 e^{-\varepsilon\tau} + \alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_{1tot})\bar{D} + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{12}^R - \mu(b\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_2^R \\
\frac{d\bar{D}^A}{d\tau} &= (u^A + \bar{D}^A)\bar{D} - (\varepsilon + \varepsilon'(\bar{D}_2^R + \bar{Y}_{1tot}))\bar{D}^A,
\end{aligned}
\tag{SM.24}
$$

with $\bar{D} = 1 - \bar{Y}_{1tot} - \bar{D}_2^R - \bar{D}^A$. Furthermore, if we consider the parameter regime in which the rate of DNA demethylation, without transfection of rTetR-TET1, is sufficiently low compared to histone modification dynamics,i.e., $\mu_1' \to 0$, the system (SM.24) can then be rewritten as follows:

$$
\begin{aligned}
\frac{d\bar{D}_1^R}{d\tau} &= \tilde{u}_1^R \bar{W}_1 e^{-\varepsilon\tau}\bar{D} + \mu(b\varepsilon + \varepsilon'\bar{D}^A)(\bar{Y}_1 - \bar{D}_1^R) + \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau}(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{11}^R \\
&\quad - (\alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_{tot} + \tilde{u}_1^R\bar{W}_1 e^{-\varepsilon\tau} + \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau}(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_1^R \\
\frac{d\bar{D}_{11}^R}{d\tau} &= \tilde{u}_1^R \bar{W}_1 e^{-\varepsilon\tau}\bar{D}_1^R + \mu(b\varepsilon + \varepsilon'\bar{D}^A)(\bar{Y}_{1tot} - \bar{Y}_1 - \bar{D}_{11}^R)) \\
&\quad - (\alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_{1tot} + \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau}(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{D}_1^R \\
\frac{d\bar{Y}_1}{d\tau} &= \tilde{u}_1^R \bar{W}_1 e^{-\varepsilon\tau}(\bar{D} - \bar{D}_1^R) + \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau}(\beta\varepsilon + \varepsilon'\bar{D}^A)(\bar{Y}_{1tot} - \bar{Y}_1) \\
&\quad - \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau}(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{Y}_1 \\
\frac{d\bar{Y}_{1tot}}{d\tau} &= \tilde{u}_1^R \bar{W}_1 e^{-\varepsilon\tau}\bar{D} - \tilde{\mu}'\frac{\frac{dox}{K_1}}{1+\frac{dox}{K_1}}\bar{T}e^{-\varepsilon\tau}(\beta\varepsilon + \varepsilon'\bar{D}^A))\bar{Y}_1
\end{aligned}
\tag{SM.25}
$$

$$\frac{d\bar{D}_2^R}{d\tau} = (\tilde{u}_2^R \frac{1}{1+\frac{DAPG}{K_2}} \bar{W}_2 e^{-\varepsilon\tau} + \alpha \bar{D}_2^R + \bar{\alpha}\bar{Y}_{1tot})\bar{D} + \mu'(\beta\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_{12}^R - \mu(b\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_2^R$$
$$\frac{d\bar{D}^A}{d\tau} = (u^A + \bar{D}^A)\bar{D} - (\varepsilon + \varepsilon'(\bar{D}_2^R + \bar{Y}_{1tot}))\bar{D}^A,$$

with $\bar{D} = 1 - \bar{Y}_1 - \bar{D}_2^R - \bar{D}^A$. From the fourth equation in (SM.25), it is possible to notice that, following a temporary phase during which the externally transfected inputs decrease, $e^{-\varepsilon\tau} \approx 0$, and then $d\bar{Y}_{1tot}/d\tau \approx 0$. This implies that, once the transient phase concludes, $\bar{Y}_{1tot}$ stabilizes at a constant value and the behavior of the original model can be captured by a reduced 2D model described by the subsequent ODEs:

$$\frac{d\bar{D}_2^R}{d\tau} = (\alpha\bar{D}_2^R + \bar{\alpha}\bar{Y}_{1tot})\bar{D} - \mu(b\varepsilon + \varepsilon'\bar{D}^A)\bar{D}_2^R, \quad \frac{d\bar{D}^A}{d\tau} = (u^A + \bar{D}^A)\bar{D} - (\varepsilon + \varepsilon'(\bar{D}_2^R + \bar{Y}_{1tot}))\bar{D}^A, \quad \text{(SM.26)}$$

with $\bar{D} = 1 - \bar{D}^A - \bar{D}_2^R - \bar{Y}_{1tot}$ and $\bar{Y}_{1tot} = constant$. It is evident that the ODE model (SM.26) closely resembles the original model, in which we assume to have one CpG per nucleosome, i.e., (SM.11), with the only distinction being the presence of $\bar{Y}_{1tot}$ instead of $\bar{Y}_1$. That is, in this revised model, the representation of the fraction of the number of methylated CpGs in the gene is denoted by $\bar{Y}_{1tot}$. Then, the introduction of the simplifying assumption, having one CpG per nucleosome, does not affect the qualitative results of the analysis that focuses on identifying the impact of the parameters α' (normalized rate of DNA methylation establishment through repressive histone modifications) and μ' (ratio between the DNA demethylation rate and the activating histone modification erasure rate) on the probability distribution of gene expression levels.

S.8 Supplementary Modeling Figures

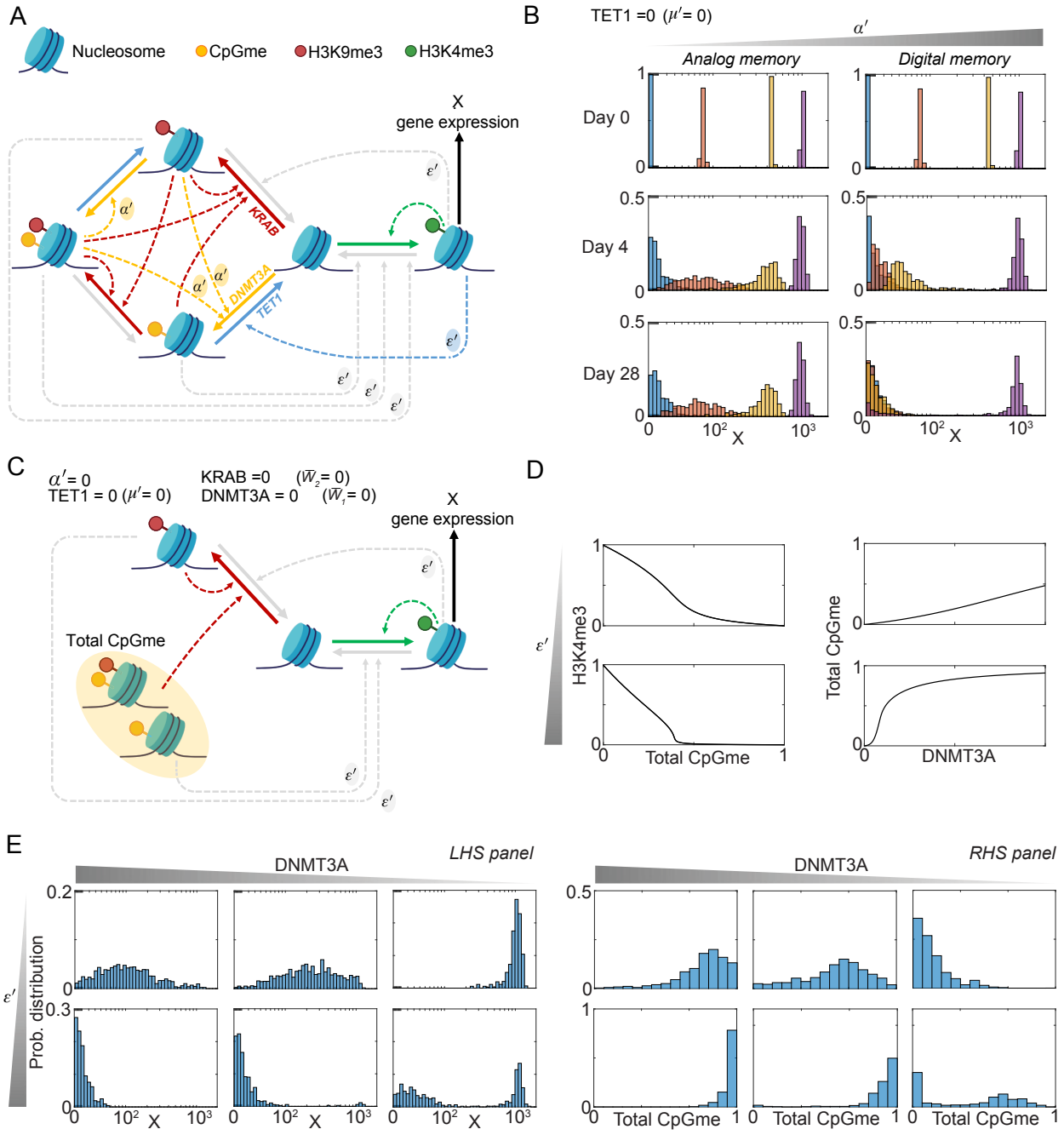

Fig. SM.1. Continued on the following page.

Fig. SM.1. The dynamics of the chromatin modification circuit can be captured by a reduced model, under specific parameter regimes. **A** Full chromatin modification circuit diagram. The species involved are unmodified nucleosome, nucleosome with H3K4me3, nucleosome with H3K9me3, nucleosome with CpGme, and nucleosome with both H3K9me3 and CpGme. We use solid arrows to represent the nucleosome modifications and gene expression. We use dashed arrows to represent the recruitment process and the consequent effect on establishment or erasure. For each arrow, we use red for processes involved in H3K9me3 establishment, orange for processes involved in DNA methylation establishment, green for processes involved in H3K4me3 establishment, light blue for processes involved in DNA demethylation, gray for processes involved in the erasure of histone modifications, and black for gene expression. In particular, $\alpha' = 0$ when H3K9me3 does not recruit writers of DNA methylation. In order to simplify the diagram without compromising the essential interactions among modifications, we removed some non-critical dashed arrows (refer to Fig. 1 in [42] for the complete diagram). **B** Probability distributions of the system represented by reactions listed in SI Tables SM.1 and SM.3. The distributions are obtained computationally using SSA [45] and we indicate with X the gene expression level (logicle scale). In particular, we determine how α' affects the probability distribution of the system when $\mu' = 0$. The parameter values used for these simulations are listed in SI Section S.9. In particular, we consider $\alpha' = 0, 0.1$ and four initial conditions (H3K9me3 and CpGme, CpGme only, H3K9me3 only, H3K4me3)= (n_{12}^R,n_1^R,n_2^R,n^A): (14,0,1,0) (blue), (6,0,8,1) (red), (4,0,5,6) (yellow), and (1,0,1,13) (purple). For each case, the initial value of X at time $t = 0$ was set to its steady state. For all the simulations we set $\alpha = \bar{\alpha} = 1$, $\varepsilon = 0.08$, $\varepsilon' = 25$, $\mu = 0.1$, $b = \beta = 1$. See SI Fig. SM.2 for how different values of α' and μ' influence the probability distribution of the system. **C** Simplified chromatin modification circuit diagram obtained when $\alpha' = 0$, TET1 = 0 (i.e., $\mu' = 0$), DNMT3A = 0 (i.e., $\bar{W}_1 = 0$), KRAB = 0 (i.e., $\bar{W}_2 = 0$). **D** Left-hand side plots: dose-response curve for the (Total CpGme,H3K4me3) pair, i.e, $(\bar{Y}_1, \bar{D}^A)$, for different values of ε' obtained from simulations of system (SM.11) with (Total CpGme,H3K9me3,H3K4me3) $= (\bar{Y}_1, \bar{D}_2^R, \bar{D}^A) = (k, 0, 1-k)$ as initial conditions, with $0 \leq k \leq 1$. Right hand-side plots: dose-response curve for the (DNMT3A,Total CpGme) pair, i.e, $(\bar{W}_1, \bar{Y}_1)$, for different values of ε' obtained from simulations of system (SM.16) with (Total CpGme,H3K9me3,H3K4me3) $= (\bar{Y}_1, \bar{D}_2^R, \bar{D}^A) = (0, 0, 1)$ as initial conditions. The parameter values used for these simulations are listed in SI Section S.9. In particular, we consider $\varepsilon = 0.1$ and $\varepsilon' = 1, 25$ (See SI Fig. SM.3 for how different values of ε and ε' influence the deterministic behavior of the system). Here, the DNMT3A dynamics is modeled as a pulse that exponentially decreases over time. The value on the x-axis corresponds to the DNMT3A value at time 0. **E** Probability distributions of the system, represented by reactions listed in SI Tables SM.1 and SM.3, after t = 28 days. The distributions are obtained computationally using SSA [45]. In the left-hand side panel we show the probability distributions of the gene expression level X (logicle scale) and in the right-hand side panel we show the probability distributions of the total CpGme level. The parameter values used for these simulations are listed in SI Section S.9. In particular, we consider $\varepsilon = 0.13$, $\varepsilon' = 1.5, 25$, DNMT3A $= \bar{W}_1 = [0, 0.4], [2.4, 2.8], [3.4, 3.8]$ and (H3K9me3 and CpGme, CpGme only, H3K9me3 only, H3K4me3) = (n_{12}^R,n_1^R,n_2^R,n^A) = $(0, 0, 0, \mathrm{D_{tot}})$ as initial condition (see SI Fig. SM.3 for how different values of ε and ε' influence the probability distribution of the system). In our model, α' is the normalized rate of DNA methylation establishment through repressive histone modifications, μ' is the ratio between the DNA demethylation rate and the activating histone modification erasure rate, and ε (ε') is the parameter that scales the ratio between the basal (recruited) erasure rate and the auto-catalysis rate of each modification. For panels B and E, we consider $\mathrm{D_{tot}} = 15$ and we consider $N = 1000$ samples to realize each distribution and, for each simulation, the value of $\bar{W}_1$ was randomly selected from a uniformly distributed range, whose extremes are listed above.

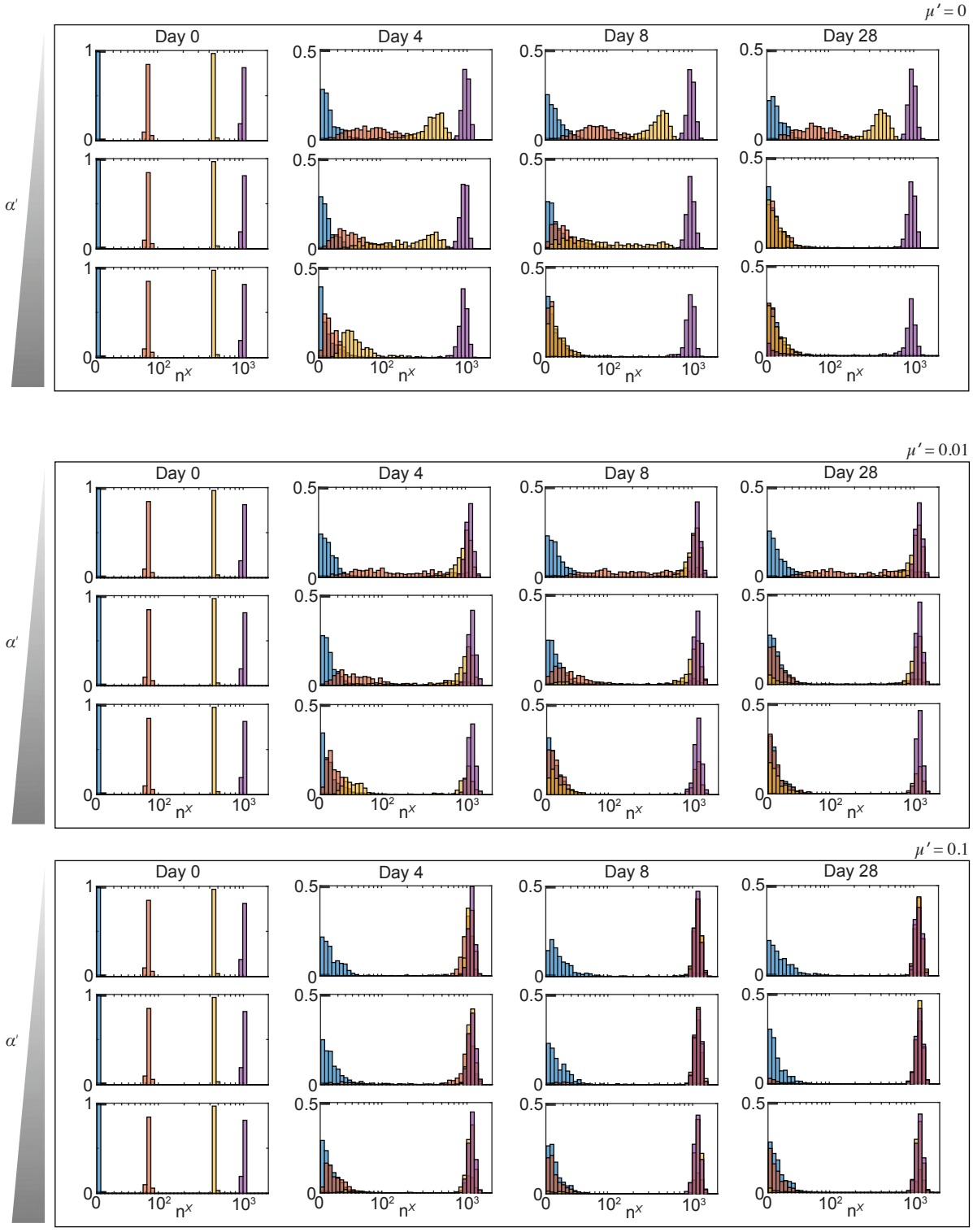

Fig. SM.2. Continued on the following page.

Fig. SM.2. Achieving analog memory is only possible when $\alpha' = 0$ and $\mu' = 0$. Probability distributions of the system represented by reactions listed in SI Tables SM.1 and SM.3. The distributions are obtained computationally using SSA [45] and we indicate with n^X the gene expression level (logicle scale). In particular, we determine how α' and μ' affect the probability distribution of the system. The parameter values used for these simulations are listed in SI Section S.9. In particular, we consider $\alpha' = 0, 0.01, 0.1$ and $\mu' = 0, 0.01, 0.1$ and four initial conditions: (n_{12}^R,n_1^R,n_2^R,n^A) = (14,0,1,0) (blue), (n_{12}^R,n_1^R,n_2^R,n^A) = (6,0,8,1) (red), (n_{12}^R,n_1^R,n_2^R,n^A) = (4,0,5,6) (yellow), (n_{12}^R,n_1^R,n_2^R,n^A) = (1,0,1,13) (purple). For each case, the initial values of n^X of n^m at time $t = 0$ were set to their steady states. For all the simulations we set $\alpha = \bar{\alpha} = 1$, $\varepsilon = 0.08$, $\varepsilon' = 25$, $\mu = 0.1$, $b = \beta = 1$. In our model, α' is the normalized rate of DNA methylation establishment through repressive histone modifications and μ' is the ratio between the DNA demethylation rate and the activating histone modification erasure rate. For all simulations, we consider $\mathrm{D_{tot}} = 15$ and we consider $N = 1000$ samples to realize each distribution.

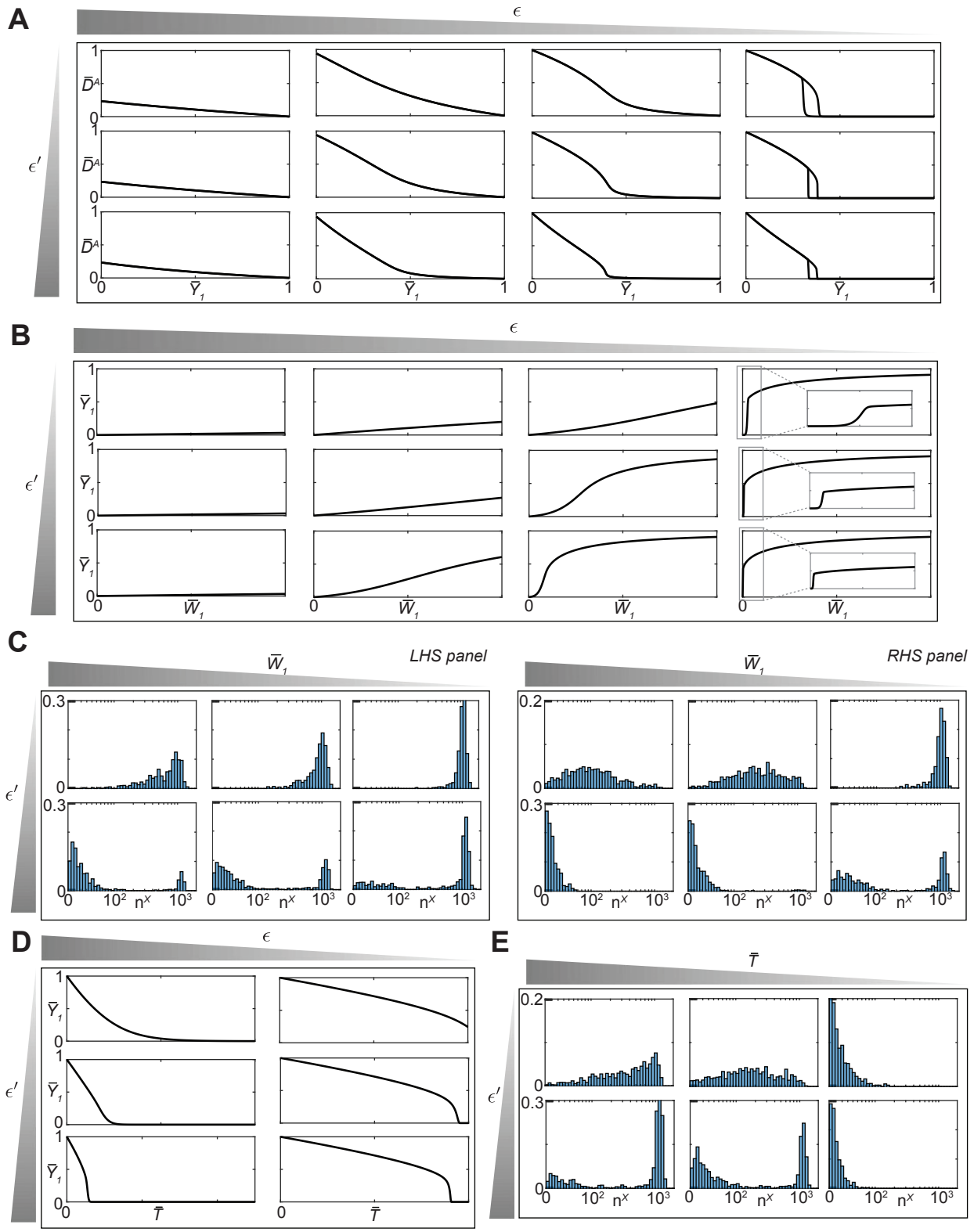

Fig. SM.3. Continued on the following page.

Fig. SM.3. The parameters ε and ε' affect the shape of the probability distribution for gene expression levels. A Input/output steady state characteristics for the $(\bar{Y}_1, \bar{D}^A)$ pair for different values of ε and ε' obtained from simulations of system (SM.11) with $(\bar{Y}_1, \bar{D}_2^R, \bar{D}^A) = (k, 0, 1-k)$ as initial conditions, with $0 \leq k \leq 1$. The parameter values used for these simulations are listed in SI Section S.9. In particular, $\varepsilon = 0.0001, 0.1, 1, 50$ and $\varepsilon' = 1, 5, 25$. **B** dose-response curve for the $(\bar{W}_1, \bar{Y}_1)$ pair for different values of ε and ε' obtained from simulations of system (SM.16) with $(\bar{Y}_1, \bar{D}_2^R, \bar{D}^A) = (0, 0, 1)$ as initial conditions. The parameter values used for these simulations are listed in SI Section S.9. In particular, $\varepsilon = 0.0001, 0.1, 1, 50$ and $\varepsilon' = 1, 5, 25$. **C** Probability distributions of the system represented by reactions listed in SI Tables SM.1 and SM.3 after t = 28 days. The distributions are obtained computationally using SSA [45] and we indicate with n^X the gene expression level (logicle scale). The parameter values used for these simulations are listed in SI Section S.9. In particular, in the left-hand side panel we consider $\varepsilon' = 1.5, 25$, $\bar{W}_1 = [0, 0.7], [0.5, 1.2], [1, 1.7]$ and (n_{12}^R,n_1^R,n_2^R,n^A) = $(0, 0, 0, \mathrm{D_{tot}})$ as initial conditions. The initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs. In the right-hand side panel, we consider $\bar{W}_1 = [0, 0.4], [2.4, 2.8], [3.4, 3.8]$ and all the other parameter values as the ones used for the simulations in the left-hand side panel. For all the simulations we set $\alpha = \bar{\alpha} = 1$, $\varepsilon = 0.13$, $\mu = 0.1$, $b = \beta = 1$, and we consider $\mathrm{D_{tot}} = 15$ and $N = 1000$ samples to realize each distribution. For each simulation, the value of $\bar{W}_1$ was randomly selected from a uniformly distributed range, whose extremes are listed above. **D** Dose-response curve for the $(\bar{Y}_1, \bar{T})$ pair for different values of ε and ε' obtained from simulations of system (SM.16) with $(\bar{Y}_1, \bar{D}_2^R, \bar{D}^A) = (1, 0, 0)$ as initial conditions. The parameter values used for these simulations are listed in SI Section S.9. In particular, $\varepsilon = 0.0001, 0.1$ and $\varepsilon' = 1, 5, 25$. **E** Probability distributions of the system represented by reactions listed in SI Tables SM.1 and SM.3 after t = 28 days. The distributions are obtained computationally using SSA [45] and we indicate with n^X the gene expression level (logicle scale). The parameter values used for these simulations are listed in SI Section S.9. In particular, we consider $\varepsilon' = 1.5, 25$, $\bar{T} = [0, 0.05], [0.24, 0.28], [0.36, 0.4]$ and (n_{12}^R,n_1^R,n_2^R,n^A) = $(\mathrm{D_{tot}}, 0, 0, 0)$ as initial conditions. The initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs. For all the simulations we set $\alpha = \bar{\alpha} = 1$, $\varepsilon = 0.13$, $\mu = 0.1$, $b = \beta = 1$, and we consider $\mathrm{D_{tot}} = 15$ and $N = 1000$ samples to realize each distribution. For each simulation, the value of $\bar{T}$ was randomly selected from a uniformly distributed range, whose extremes are listed above. In our model, ε (ε') is the parameter that scales the ratio between the basal (recruited) erasure rate and the auto-catalysis rate of each modification and $\bar{W}_1$ and $\bar{T}$ represent the normalized amounts of DNMT3A and TET1 at time 0, which exponentially decrease over time.

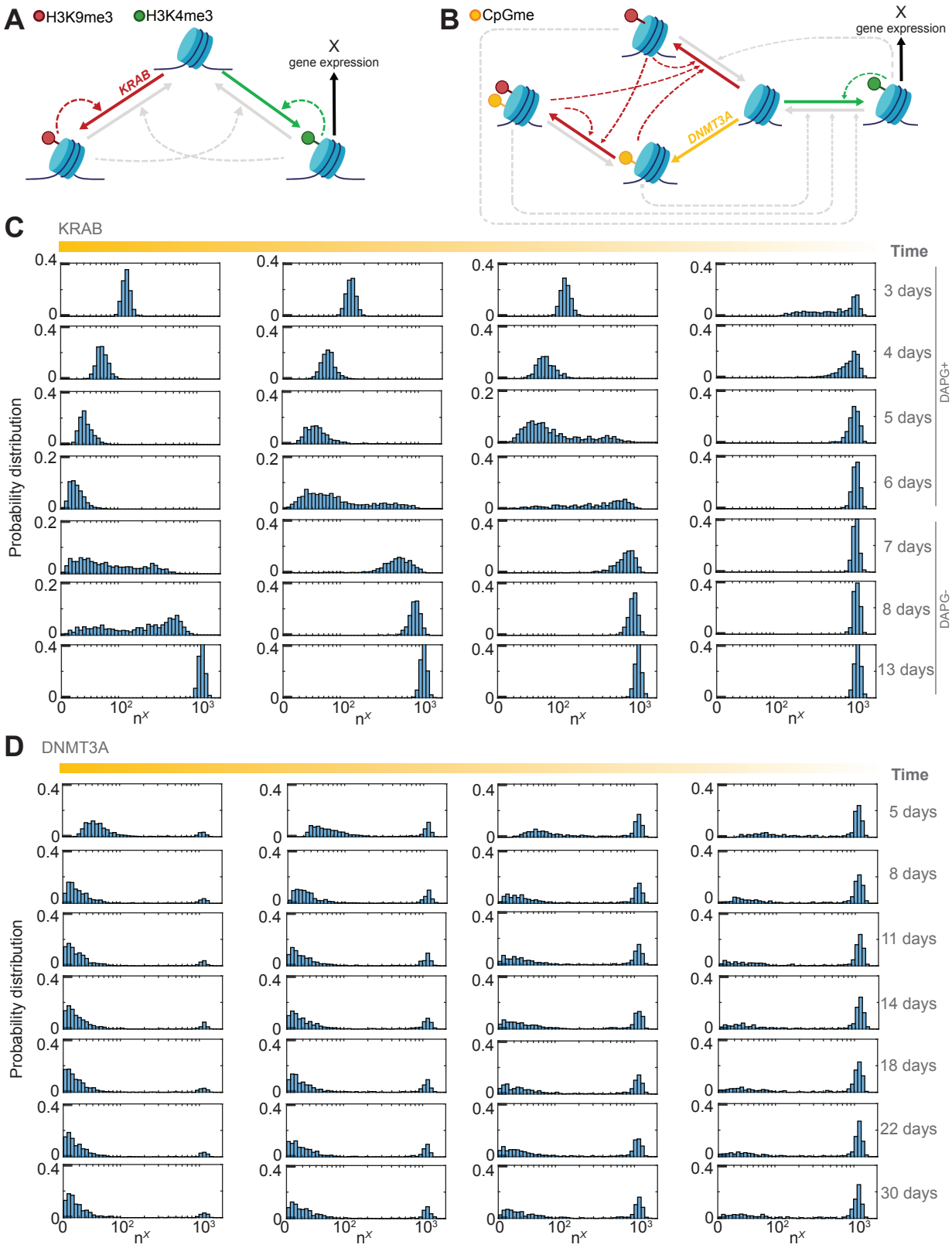

Fig. SM.4. Continued on the following page.

Fig. SM.4. Simulations recapitulating the effect of transient transfection of KRAB and DNMT3A on the gene expression probability distribution. A Diagram of the simplified chromatin modification circuit, including H3K9me3 and H3K4me3 only. We use solid arrows to represent the nucleosome modifications and gene expression. We use dashed arrows to represent the recruitment process and the consequent effect on establishment or erasure. For each arrow, we use red for processes involved in H3K9me3 establishment, green for processes involved in H3K4me3 establishment, gray for processes involved in the erasure of histone modifications, and black for gene expression. **B** Diagram of the chromatin modification circuit including H3K9me3, H3K4me3, and CpGme. We use solid arrows to represent the nucleosome modifications and gene expression. We use dashed arrows to represent the recruitment process and the consequent effect on establishment or erasure. For each arrow, we use red for processes involved in the H3K9me3 establishment, green for processes involved in H3K4me3 establishment, orange for processes involved in DNA methylation establishment, gray for processes involved in the erasure of histone modifications, and black for gene expression. **C** Probability distributions of the system represented in panel A, whose reactions are listed in SI Tables SM.2 and SM.4. The distributions are obtained computationally using SSA [45] and we indicate with n^X the gene expression level (logicle scale). The parameter values used for these simulations are listed in SI Section S.9. In particular, we consider $KRAB = \bar{W}_2 = [1.5, 7.5], [20, 100], [45, 225], [4000, 20000]$, and ($n_2^R$,$n^A$) = $(0, \mathrm{D_{tot}})$ as initial conditions. The initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs. **D** Probability distributions of the system represented in panel B, whose reactions are listed in SI Tables SM.1 and SM.3. The distributions are obtained computationally using SSA [45] and we indicate with n^X the gene expression level (logicle scale). The parameter values used for these simulations are listed in SI Section S.9. In particular, we consider $DNMT3A = \bar{W}_1 = [0.2, 0.7], [0.35, 1.23], [0.6, 2.1], [0.9, 3.1]$, $KRAB = 0$, and (n_{12}^R,n_1^R,n_2^R,n^A) = $(0, 0, 0, \mathrm{D_{tot}})$ as initial conditions. The initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs. For all the simulations we set $\alpha = 1$, $\varepsilon = 0.13$, $\varepsilon' = 25$, $\mu = 0.1$, $b = 1$, we model the dynamics of KRAB and DNMT3A as pulses that exponentially decrease over time, and we consider $\mathrm{D_{tot}} = 15$ and $N = 1000$ samples to realize each distribution. For each simulation, the value of $\bar{W}_1$ (or $\bar{W}_2$) was randomly selected from a uniformly distributed range, whose extremes are listed above.

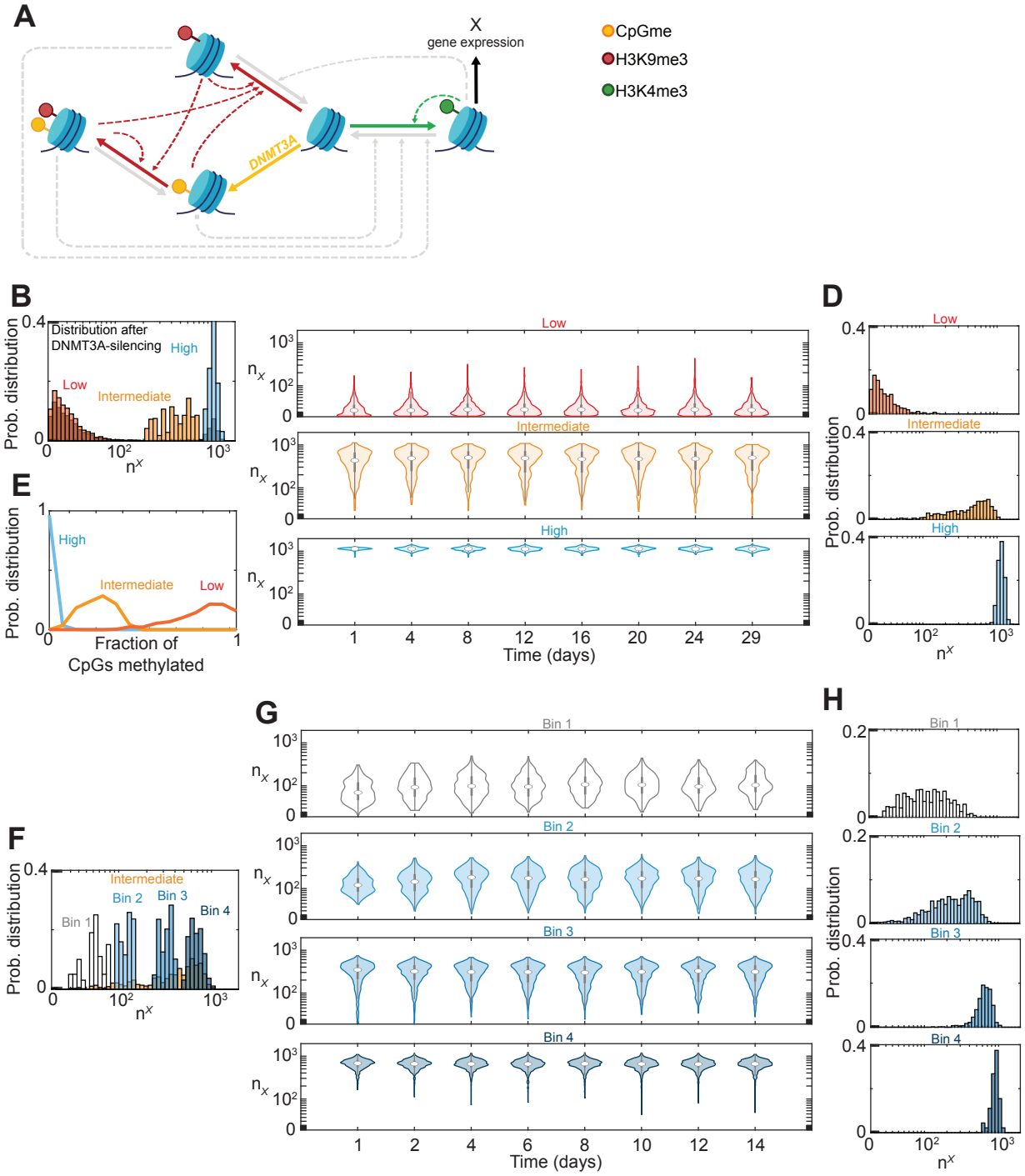

Fig. SM.5. Continued on the following page.

Fig. SM.5. Simulations recapitulating the experimental results shown in Fig. 3, showing that intermediate gene expression levels are stable and exhibit intermediate DNA methylation levels. A Diagram of the chromatin modification circuit including H3K9me3, H3K4me3, and CpGme. We use solid arrows to represent the nucleosome modifications and gene expression. We use dashed arrows to represent the recruitment process and the consequent effect on establishment or erasure. For each arrow, we use red for processes involved in H3K9me3 establishment, green for processes involved in H3K4me3 establishment, orange for processes involved in DNA methylation establishment, gray for processes involved in the erasure of histone modifications, and black for gene expression. **B** Probability distributions of the system represented in panel A, whose reactions are listed in SI Tables SM.1 and SM.3. The distributions are obtained computationally using SSA [45] and we indicate with n^X the gene expression level (logicle scale). The black distribution represents the probability distribution of the system at day 30 after DNMT3A-silencing. The parameter values used for these simulations are listed in SI Section S.9. In particular, we consider $DNMT3A = \bar{W}_1 = [0.6, 2.1]$, $KRAB = 0$, and (n^R_{12},n^R_1,n^R_2,n^A) = $(0, 0, 0, \mathrm{D_{tot}})$ as initial conditions. The initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs. The low (red), intermediate (orange), and high (blue) distributions are obtained by subdividing the black distribution with respect to the following intervals of n^X:[0,100], [200,800], [900,10000]. **C** Probability distributions of the system represented in panel A obtained as described in panel B. The parameter values used for these simulations are listed in SI Section S.9. In particular, for each case, we consider $DNMT3A = 0$, $KRAB = 0$, and, as initial conditions, points from the red, orange, and blue distributions shown in panel B, respectively. **D** Probability distributions of the system at day 29 from panel C. **E** Probability distribution of the fraction of CpGs methylated, i.e., $(n^R_1 + n^R_{12})/\mathrm{D_{tot}}$, at day 29 obtained from the simulations conducted to obtain panel C. **F** Probability distributions of the system represented in panel A, whose reactions are listed in SI Tables SM.1 and SM.3. The distributions are obtained computationally using SSA [45] and we indicate with n^X the gene expression level (logicle scale). The orange distribution is the one obtained and described in panel D. The other ones are obtained by subdividing the orange distribution with respect to the following intervals of n^X: [0,80] (Bin 1), [90,150] (Bin 2), [250,400] Bin 3, [500,10000] (Bin 4). **G** Probability distributions of the system represented in panel A, obtained as described in panel F. The parameter values used for these simulations are listed in SI Section S.9. In particular, for each case, we consider $DNMT3A = 0$, $KRAB = 0$, and, as initial conditions, points from the "Bin 1", "Bin 2", "Bin 3", and "Bin 4" distributions shown in panel F, respectively. **H** Probability distributions of the system at day 14 from panel G.

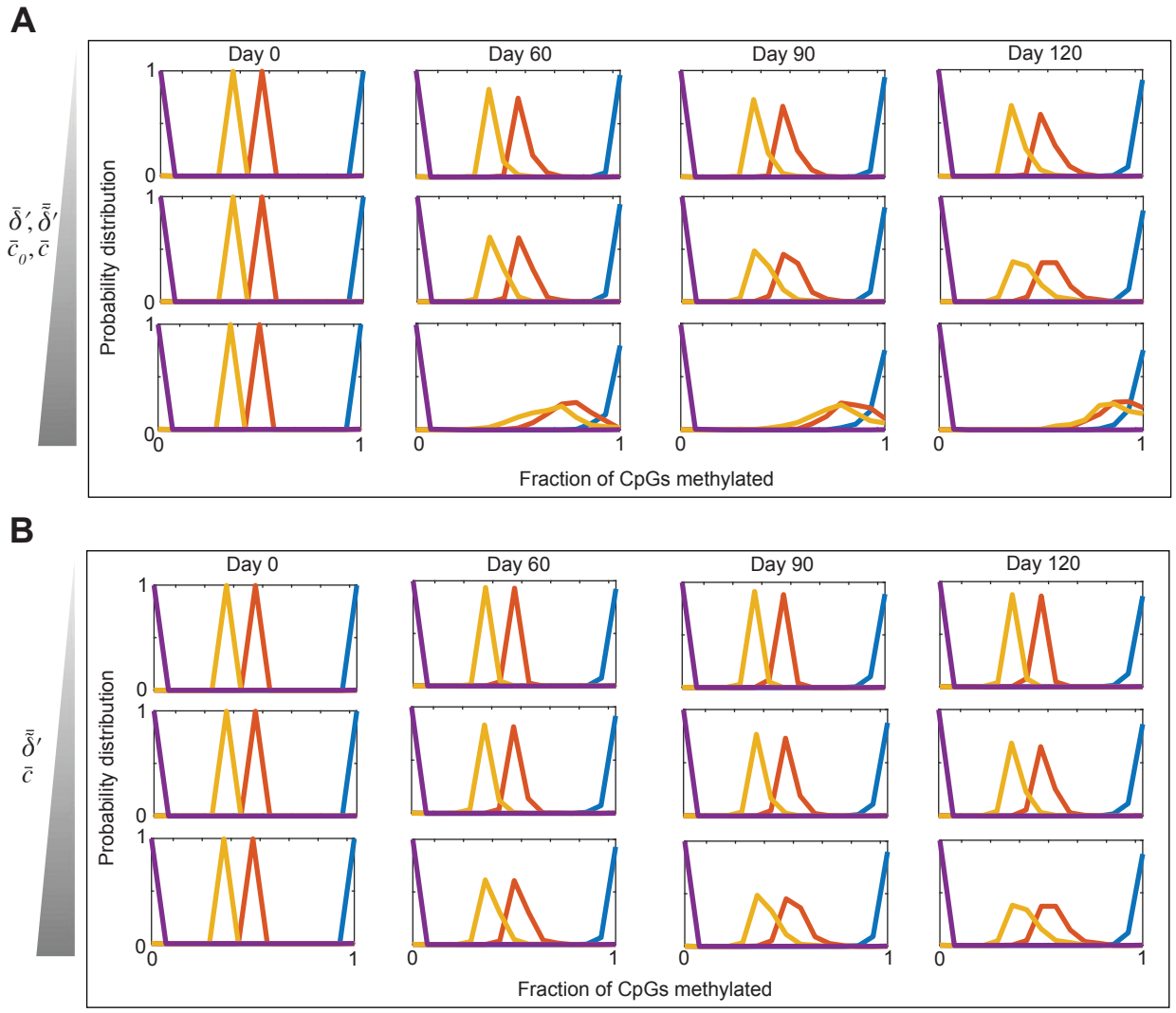

Fig. SM.6. Continued on the following page.

Fig. SM.6. Stochastic analysis of the full chromatin modification circuit model including the refined DNA methylation maintenance model. A Probability distributions of the system represented by reactions listed in SI Tables SM.5 and SM.3. The distributions are obtained computationally using SSA [45] and, on the x we have the fraction of CpGs methylated. In particular, we determine how lower parameter values affect the probability distribution of the system. The parameter values used for these simulations are listed in SI Section S.9. In particular, we consider $\bar{c}_0 = 3.3 \cdot 10^{-6}, 6.6 \cdot 10^{-6}, 6.6 \cdot 10^{-5}$, $\bar{\delta}' = 3.3 \cdot 10^{-6}, 6.6 \cdot 10^{-6}, 6.6 \cdot 10^{-5}$, $\bar{c} = 5 \cdot 10^{-5}, 1 \cdot 10^{-4}, 1 \cdot 10^{-3}$, $\bar{\tilde{\delta}}' = 3.3 \cdot 10^{-6}, 6.6 \cdot 10^{-6}, 6.6 \cdot 10^{-5}$, and four initial fractions of CpGs methylated, i.e., (0.93,0.4,0.27,0). Specifically, as initial conditions, we set (n_{12}^R,n_1^R,n_2^R,n^A) = (14,0,0,1) (blue), (n_{12}^R,n_1^R,n_2^R,n^A) = (6,0,0,9) (red), (n_{12}^R,n_1^R,n_2^R,n^A) = (4,0,0,11) (yellow), (n_{12}^R,n_1^R,n_2^R,n^A) = (0,0,0,15) (purple). For each case, The initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs. For all the simulations we set $\alpha = \bar{\alpha} = 1$, $\varepsilon = 0.13$, $\varepsilon' = 25$, $\mu = 0.1$, $b = \beta = 1$. **B** Probability distributions of the system represented by reactions listed in SI Tables SM.5 and SM.3. The distributions are obtained computationally using SSA [45] and, on the x we have the fraction of CpGs methylated. In particular, we determine how $\bar{c}$ and $\bar{\tilde{\delta}}'$, i.e., normalized rate constants associated with the reactions representing the effect of DNMT1 recruitment by H3K9me3 on the maintenance process, affect the probability distribution of the system. The parameter values used for these simulations are listed in SI Section S.9. In particular, we consider $\bar{c}_0 = 6.6 \cdot 10^{-6}$, $\bar{\delta}' = 6.6 \cdot 10^{-6}$, $\bar{c} = 3.47 \cdot 10^{-5}$, $3 \cdot 10^{-5}$, $3.47 \cdot 10^{-4}$, $\bar{\tilde{\delta}}' = 6.6 \cdot 10^{-7}, 1.94 \cdot 10^{-6}, 6.6 \cdot 10^{-6}$.The values of the other parameters and the initial conditions are the same as those used for panel A. For all simulations, we consider $\mathrm{D_{tot}} = 15$ and we consider $N = 1000$ samples to realize each distribution.

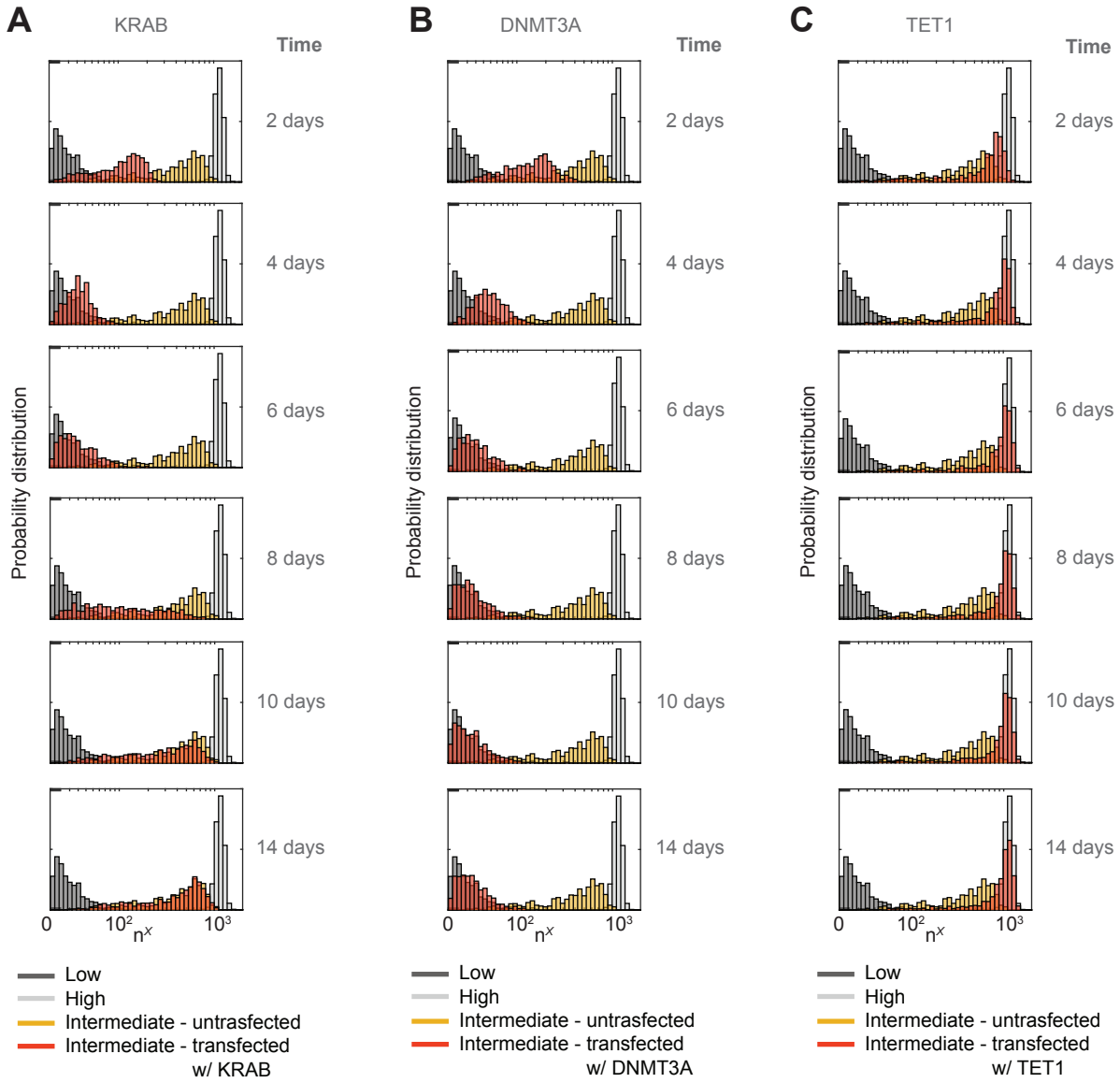

Fig. SM.7. Continued on the following page.

Fig. SM.7. Simulations recapitulating the effect of KRAB, DNMT3A, and TET1 on the gene expression probability distribution, starting from an intermediate level of gene expression. A Probability distributions of the system represented in Fig. SM.4A, whose reactions are listed in SI Tables SM.2 and SM.4. The distributions are obtained computationally using SSA [45] and we indicate with n^X the gene expression level (logicle scale). The parameter values used for these simulations are listed in SI Section S.9. In particular, we consider $KRAB = \bar{W}_2 = [50, 250]$, $DNMT3A = \bar{W}_1 = 0, TET1 = \bar{T} = 0$. As initial conditions, we consider values of (n^R_{12},n^R_1,n^R_2,n^A,n^X,n^m) from the simulations associated with the orange distribution in SI Fig. SM.5D. **B** Probability distributions of the system represented in Fig. SM.4B, whose reactions are listed in SI Tables SM.1 and SM.3. The distributions are obtained computationally using SSA [45]. The parameter values used for these simulations are listed in SI Section S.9. In particular, we consider $DNMT3A = \bar{W}_1 = [0.2, 0.7]$, $KRAB = \bar{W}_2 = 0$, and $TET1 = \bar{T} = 0$. **C** Probability distributions of the system represented in Fig. SM.4B, whose reactions are listed in SI Tables SM.1 and SM.3. The parameter values used for these simulations are listed in SI Section S.9. In particular, we consider $DNMT3A = \bar{W}_1 = 0$, $KRAB = \bar{W}_2 = 0$, and $TET1 = \bar{T} = [0.004, 0.014]$. For all the simulations we set $\alpha = \bar{\alpha} = 1$, $\varepsilon = 0.13$, $\varepsilon' = 25$, $\mu = 0.1$, $b = 1$, we model the dynamics of KRAB, DNMT3A, and TET1 as pulses that exponentially decrease over time, and we consider $\mathrm{D_{tot}} = 15$ and $N = 1000$ samples to realize each distribution. For each simulation, the value of $\bar{W}_1$, $\bar{W}_2$, or $\bar{T}$ was randomly selected from a uniformly distributed range, whose extremes are listed above.

S.9 Values of parameters and rationale for parameter selections used in generating figures in the main paper and SI

In our model, we defined the total number of nucleosomes within a gene of interest as $\mathrm{D_{tot}}$. Assuming approximately one nucleosome per 200 bp [66] and given that the length of our gene is ≈ 3000bp, then in our computational study we set the value of $\mathrm{D_{tot}}$ to 15.

With regard to the effect of dilution due to cell growth, we use a standard approach by modeling it as a first-order decay reaction [67]. The decay rate constant can then be expressed as $\delta = ln(2)/T$, where T corresponds to the cell cycle length. Given that in our experiments we use CHO cells, whose cycle length has been estimated to be $\approx$ 20 hr [5], in the simulations aimed at replicating our experimental data, we set $\delta = 0.035$ (corresponding to $T = 19.8$ h).

Additionally, our mathematical study in SI Section S.4 suggests that our 4D+X model can replicate the experimental data related to the transfection of KRAB and DNMT3A (Fig. 2) only within a parameter regime characterized by small ε and large ε'. The expressions defining these parameters are given in (SM.2). Therefore, we set $k_M^A/\Omega = 0.0347$ h^{-1}, $\bar{k}_E^A = 0.0176$ h^{-1}, and $k_E^A/\Omega = 0.8675$ h^{-1}, since these parameter values, along with the previously listed values of δ and $\mathrm{D_{tot}}$, allow us to be within the desired parameter regime and accurately replicate our experimental results.

Regarding the other parameters, we considered multiple values and selected those that allow us to accurately replicate all experimental results presented in this paper.

S.9.1 Parameter values used to realize the plots in Fig. 2J - top panel

$\alpha = 1$, $\bar{\alpha} = 1$, $\mu = 0.1$, $b = 1$, $\varepsilon = 0.1$, $\varepsilon' = 25$, $u^A = 15$. As initial conditions, we set $(\bar{Y}_1, \bar{D}_2^R, \bar{D}^A) = (k, 0, 1-k)$, with $0 \leq k \leq 1$.

S.9.2 Parameter values used to realize the plots in Fig. 2J - bottom panel

$\tilde{u}_2^R = 0$, $\tilde{\mu}' = 0$, $\beta = 1$, $\tilde{u}_1^R = 1$, $\bar{W}_1 = [0, 4]$ $\alpha = 1$, $\bar{\alpha} = 1$, $\mu = 0.1$, $b = 1$, $\varepsilon = 0.1$, $\varepsilon' = 25$, $u^A = 15$. As initial conditions, we set $(\bar{Y}_1, \bar{D}_2^R, \bar{D}^A) = (0, 0, 1)$.

S.9.3 Parameter values used to realize the plots in Fig. 2K

$k^A_{W0} = 0$, $k^A_W = 7.8075$ h^{-1}, $\bar{k}^A_E = 0.0118$ h^{-1}, $\delta = 0.0291$ h^{-1}, $\frac{k^A_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^A_E}{\Omega} = 0.8675$ h^{-1}, $k^1_{W0} = 0$, $k^1_W = 0$, $k'_T = 0$, $\delta' = 0$, $\frac{k'_M}{\Omega} = 0$, $\frac{k'^*_T}{\Omega} = 0$, $k^2_{W0} = 0$, $k^2_W = 0$, $\bar{k}^R_E = 0.0012$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^R_E}{\Omega} = 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta^A_m = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. We consider four initial conditions: (n^R_{12},n^R_1,n^R_2,n^A) = (14,0,1,0) (blue), (n^R_{12},n^R_1,n^R_2,n^A) = (6,0,8,1) (red), (n^R_{12},n^R_1,n^R_2,n^A) = (4,0,5,6) (yellow), (n^R_{12},n^R_1,n^R_2,n^A) = (1,0,1,13) (purple). For each case, the initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs.

S.9.4 Parameter values used to realize the plots in Fig. 2L

$k^A_{W0} = 0$, $k^A_W = 7.8075$ h^{-1}, $\bar{k}^A_E = 0.0315$ h^{-1}, $\delta = 0.035$ h^{-1}, $\frac{k^A_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^A_E}{\Omega} = 0.8675$ h^{-1}, $k^1_{W0} = 0$, $k^1_W \in [0, 0.3643]e^{-\delta t}, [0.2602, 0.6246]e^{-\delta t}, [0.5205, 0.8848]e^{-\delta t}$ h^{-1}, $k'_T = 0$, $\delta' = 0$, $\frac{k'_M}{\Omega} = 0$, $\frac{k'^*_T}{\Omega} = 0$, $k^2_{W0} = 0$, $k^2_W = 0$, $\bar{k}^R_E = 0.0032$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^R_E}{\Omega} = 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta^A_m = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. As initial condition, we consider (n^R_{12},n^R_1,n^R_2,n^A) = (0,0,0,15) and the initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs.

S.9.5 Parameter values used to realize the plots in Fig. 4G and I

$k^A_{W0} = 0$, $k^A_W = 7.8075$ h^{-1}, $\bar{k}^A_E = 0.0176$ h^{-1}, $\delta = 0.035$ h^{-1}, $\frac{k^A_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^A_E}{\Omega} = 0.8675$ h^{-1}, $k^1_{W0} = 0$, $k^1_W = 0$, $c = 1.11 \cdot 10^{-6}$ h^{-1}, $c_0 = 8.37 \cdot 10^{-5}$ h^{-1}, $\delta' = 2.09 \cdot 10^{-5}$ h^{-1}, $\tilde{\delta}' = 4.15 \cdot 10^{-5}$ h^{-1}, $K = 1$, $\frac{k'_M}{\Omega} = 0$, $\frac{k'^*_T}{\Omega} =$ h^{-1}, $k^2_{W0} = 0$, $k^2_W = 0$, $\bar{k}^R_E = 0.0018$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^R_E}{\Omega} = 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta^A_m = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. We consider eight initial conditions: (n^R_{12},n^R_1,n^R_2,n^A) = (13,0,0,2) (Bin 1), (12,0,0,3) (Bin 2), (8,0,0,7) (Bin 3), (6,0,0,9) (Bin 4), (5,0,0,10) (Bin 5), (4,0,0,11) (Bin 6), (1,0,0,14) (Bin 7), (0,0,0,15) (Bin 8). For each case, the initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs.

S.9.6 Parameter values used to realize the plots in SI Fig. SM.1B

$k^A_{W0} = 0$, $k^A_W = 7.8075$ h^{-1}, $\bar{k}^A_E = 0.0118$ h^{-1}, $\delta = 0.0291$ h^{-1}, $\frac{k^A_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^A_E}{\Omega} = 0.8675$ h^{-1}, $k^1_{W0} = 0$, $k^1_W = 0$, $k'_T = 0$, $\delta' = 0$, $\frac{k'_M}{\Omega} = 0, 0.00347$ h^{-1}, $\frac{k'^*_T}{\Omega} = 0$, $k^2_{W0} = 0$, $k^2_W = 0$, $\bar{k}^R_E = 0.0012$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^R_E}{\Omega} = 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta^A_m = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. We consider four initial conditions: (n^R_{12},n^R_1,n^R_2,n^A) = (14,0,1,0) (blue), (n^R_{12},n^R_1,n^R_2,n^A) = (6,0,8,1) (red),

$(n_{12}^R,n_1^R,n_2^R,n^A)$ = (4,0,5,6) (yellow), $(n_{12}^R,n_1^R,n_2^R,n^A)$ = (1,0,1,13) (purple). For each case, the initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs.

S.9.7 Parameter values used to realize the plots in SI Fig. SM.1D - left hand-side panel

$\alpha = 1$, $\bar{\alpha} = 1$, $\mu = 0.1$, $b = 1$, $\varepsilon = 0.1$, $\varepsilon' = 1,25$, $u^A = 15$. As initial conditions, we set $(\bar{Y}_1, \bar{D}_2^R, \bar{D}^A) = (k, 0, 1-k)$, with $0 \leq k \leq 1$.

S.9.8 Parameter values used to realize the plots in SI Fig. SM.1D - right hand-side panel

$\tilde{u}_2^R = 0$, $\tilde{\mu}' = 0$, $\beta = 1$, $\tilde{u}_1^R = 1$, $\bar{W}_1 = [0,4]$ $\alpha = 1$, $\bar{\alpha} = 1$, $\mu = 0.1$, $b = 1$, $\varepsilon = 0.1$, $\varepsilon' = 1,25$, $u^A = 15$. As initial conditions, we set $(\bar{Y}_1, \bar{D}_2^R, \bar{D}^A) = (0, 0, 1)$.

S.9.9 Parameter values used to realize the plots in SI Fig. SM.1E

$k_{W0}^A = 0$, $k_W^A = 7.8075$ h^{-1}, $\bar{k}_E^A = 0.0315$ h^{-1}, $\delta = 0.035$ h^{-1}, $\frac{k_M^A}{\Omega} = 0.0347$ h^{-1}, $\frac{k_E^A}{\Omega} = 0.0520, 0.8675$ h^{-1}, $k_{W0}^1 = 0$, $k_W^1 \in [0, 0.3643]e^{-\delta t}, [0.2602, 0.6246]e^{-\delta t}, [0.5205, 0.8848]e^{-\delta t}$ h^{-1}, $k_T' = 0$, $\delta' = 0$, $\frac{k_M'}{\Omega} = 0$, $\frac{k_T'^*}{\Omega} = 0$, $k_{W0}^2 = 0$, $k_W^2 = 0$, $\bar{k}_E^R = 0.0032$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k_E^R}{\Omega} = 0.0052, 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta_m^A = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. As initial condition, we consider $(n_{12}^R,n_1^R,n_2^R,n^A)$ = (0,0,0,15) and the initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs.

S.9.10 Parameter values used to realize the plots in SI Fig. SM.2

$k_{W0}^A = 0$, $k_W^A = 7.8075$ h^{-1}, $\bar{k}_E^A = 0.0118$ h^{-1}, $\delta = 0.0291$ h^{-1}, $\frac{k_M^A}{\Omega} = 0.0347$ h^{-1}, $\frac{k_E^A}{\Omega} = 0.8675$ h^{-1}, $k_{W0}^1 = 0$, $k_W^1 = 0$, $k_T' = 0, 0.00012, 0.0012$ h^{-1}, $\delta' = 0$, $\frac{k_M'}{\Omega} = 0, 3.47 \cdot 10^{-4}, 3.47 \cdot 10^{-3}$ h^{-1}, $\frac{k_T'^*}{\Omega} = 0, 0.0087, 0.087$ h^{-1}, $k_{W0}^2 = 0$, $k_W^2 = 0$, $\bar{k}_E^R = 0.0012$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k_E^R}{\Omega} = 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta_m^A = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. We consider four initial conditions: $(n_{12}^R,n_1^R,n_2^R,n^A)$ = (14,0,1,0) (blue), $(n_{12}^R,n_1^R,n_2^R,n^A)$ = (6,0,8,1) (red), $(n_{12}^R,n_1^R,n_2^R,n^A)$ = (4,0,5,6) (yellow), $(n_{12}^R,n_1^R,n_2^R,n^A)$ = (1,0,1,13) (purple). For each case, the initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs.

S.9.11 Parameter values used to realize the plots in SI Fig. SM.3A

$\alpha = 1$, $\bar{\alpha} = 1$, $\mu = 0.1$, $b = 1$, $\varepsilon = 0.0001, 0.1, 1, 50$, $\varepsilon' = 1, 5, 25$, $u^A = 15$. As initial conditions, we set $(\bar{Y}_1, \bar{D}_2^R, \bar{D}^A) = (k, 0, 1-k)$, with $0 \leq k \leq 1$.

S.9.12 Parameter values used to realize the plots in SI Fig. SM.3B

$\tilde{u}_2^R = 0$, $\tilde{\mu}' = 0$, $\beta = 1$, $\tilde{u}_1^R = 1$, $\bar{W}_1 = [0, 4]$ $\alpha = 1$, $\bar{\alpha} = 1$, $\mu = 0.1$, $b = 1$, $\varepsilon = 0.0001, 0.1, 1, 50$, $\varepsilon' = 1, 5, 25$, $u^A = 15$. As initial conditions, we set $(\bar{Y}_1, \bar{D}_2^R, \bar{D}^A) = (0, 0, 1)$.

S.9.13 Parameter values used to realize the plots in SI Fig. SM.3C

$k_{W0}^A = 0$, $k_W^A = 7.8075$ h^{-1}, $\bar{k}_E^A = 0.0315$ h^{-1}, $\delta = 0.035$ h^{-1}, $\frac{k_M^A}{\Omega} = 0.0347$ h^{-1}, $\frac{k_E^A}{\Omega} = 0.0520, 0.8675$ h^{-1}, $k_{W0}^1 = 0$, $k_W^1 \in [0, 0.3643]e^{-\delta t}, [0.2602, 0.6246]e^{-\delta t}, [0.5205, 0.8848]e^{-\delta t}$ h^{-1} (left hand-side panel) and $k_W^1 \in [0, 0.2082]e^{-\delta t}, [1.2492, 1.4574]e^{-\delta t}, [1.7697, 1.9779]e^{-\delta t}$ h^{-1} (right hand-side panel), $k_T' = 0$, $\delta' = 0$, $\frac{k_M'}{\Omega} = 0$, $\frac{k_T'^*}{\Omega} = 0$, $k_{W0}^2 = 0$, $k_W^2 = 0$, $\bar{k}_E^R = 0.0032$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k_E^R}{\Omega} = 0.0052, 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta_m^A = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. As initial condition, we consider (n_{12}^R,n_1^R,n_2^R,n^A) = (0,0,0,15) and the initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs.

S.9.14 Parameter values used to realize the plots in SI Fig. SM.3D

$\tilde{u}_2^R = 0$, $\tilde{\mu}' = 1$, $\beta = 1$, $\tilde{u}_1^R = 0$, $\bar{T} = [0, 4]$ $\alpha = 1$, $\bar{\alpha} = 1$, $\mu = 0.1$, $b = 1$, $\varepsilon = 0.0001, 0.1, 1, 50$, $\varepsilon' = 1, 5, 25$, $u^A = 15$. As initial conditions, we set $(\bar{Y}_1, \bar{D}_2^R, \bar{D}^A) = (1, 0, 0)$.

S.9.15 Parameter values used to realize the plots in SI Fig. SM.3E

$k_{W0}^A = 0$, $k_W^A = 7.8075$ h^{-1}, $\bar{k}_E^A = 0.0315$ h^{-1}, $\delta = 0.035$ h^{-1}, $\frac{k_M^A}{\Omega} = 0.0347$ h^{-1}, $\frac{k_E^A}{\Omega} = 0.0520, 0.8675$ h^{-1}, $k_{W0}^1 = 0$, $k_W^1 = 0$, $k_T' \in [0, 0.0016]e^{-\delta t}, [0.0076, 0.0088]e^{-\delta t}, [0.0113, 0.0126]e^{-\delta t}$ h^{-1}, $\delta' = 0$, $\frac{k_M'}{\Omega} = 0$, $\frac{k_T'^*}{\Omega} \in [0, 0.0434]e^{-\delta t}, [0.2082, 0.2429]e^{-\delta t}, [0.3123, 0.3470]e^{-\delta t}$ h^{-1}, $k_{W0}^2 = 0$, $k_W^2 = 0$, $\bar{k}_E^R = 0.0032$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k_E^R}{\Omega} = 0.0052, 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta_m^A = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. As initial condition, we consider (n_{12}^R,n_1^R,n_2^R,n^A) = (15,0,0,0) and the initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs.

S.9.16 Parameter values used to realize the plots in SI Fig. SM.4C

$k^A_{W0} = 0$, $k^A_W = 7.8075$ h^{-1}, $\bar{k}^A_E = 0.0176$ h^{-1}, $\delta = 0.035$ h^{-1}, $\frac{k^A_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^A_E}{\Omega} = 0.8675$ h^{-1}, $k^2_{W0} = 0$, $k^2_W \in [0.78, 3.90]e^{-\delta t}, [10.41, 52.05]e^{-\delta t}, [23.42, 117.11]e^{-\delta t}, [2082, 10410]e^{-\delta t}$ h^{-1}, $\bar{k}^R_E = 0.0018$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^R_E}{\Omega} = 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta^A_m = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. As initial condition, we consider (n^R_2,n^A) = (0,15) and the initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs.

S.9.17 Parameter values used to realize the plots in SI Fig. SM.4D

$k^A_{W0} = 0$, $k^A_W = 7.8075$ h^{-1}, $\bar{k}^A_E = 0.0176$ h^{-1}, $\delta = 0.035$ h^{-1}, $\frac{k^A_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^A_E}{\Omega} = 0.8675$ h^{-1}, $k^1_{W0} = 0$, $k^1_W \in [0.10, 0.36]e^{-\delta t}, [0.18, 0.64]e^{-\delta t}, [0.31, 1.09]e^{-\delta t}, [0.47, 1.61]e^{-\delta t}$ h^{-1}, $k'_T = 0$, $\delta' = 0$, $\frac{k'_M}{\Omega} = 0$, $\frac{k'^*_T}{\Omega} = 0$, $k^2_{W0} = 0$, $k^2_W = 0$ $\bar{k}^R_E = 0.0018$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^R_E}{\Omega} = 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta^A_m = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. As initial condition, we consider (n^R_{12},n^R_1,n^R_2,n^A) = (0,0,0,15) and the initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs.

S.9.18 Parameter values used to realize the plots in SI Fig. SM.5B

$k^A_{W0} = 0$, $k^A_W = 7.8075$ h^{-1}, $\bar{k}^A_E = 0.0176$ h^{-1}, $\delta = 0.035$ h^{-1}, $\frac{k^A_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^A_E}{\Omega} = 0.8675$ h^{-1}, $k^1_{W0} = 0$, $k^1_W \in [0.31, 1.09]e^{-\delta t}$ h^{-1}, $k'_T = 0$, $\delta' = 0$, $\frac{k'_M}{\Omega} = 0$, $\frac{k'^*_T}{\Omega} = 0$, $k^2_{W0} = 0$, $k^2_W = 0$ $\bar{k}^R_E = 0.0018$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^R_E}{\Omega} = 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta^A_m = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. As initial condition, we consider (n^R_{12},n^R_1,n^R_2,n^A) = (0,0,0,15) and the initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs. The distribution shown is after 30 days.

S.9.19 Parameter values used to realize the plots in SI Fig. SM.5C

$k^A_{W0} = 0$, $k^A_W = 7.8075$ h^{-1}, $\bar{k}^A_E = 0.0176$ h^{-1}, $\delta = 0.035$ h^{-1}, $\frac{k^A_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^A_E}{\Omega} = 0.8675$ h^{-1}, $k^1_{W0} = 0$, $k^1_W 0$, $k'_T = 0$, $\delta' = 0$, $\frac{k'_M}{\Omega} = 0$, $\frac{k'^*_T}{\Omega} = 0$, $k^2_{W0} = 0$, $k^2_W = 0$ $\bar{k}^R_E = 0.0018$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^R_E}{\Omega} = 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta^A_m = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. As initial condition, we consider values of (n^R_{12},n^R_1,n^R_2,n^A,n^X,n^m) from the simulations shown in SI Fig. SM.5B and divide them with respect to their value of n^X, following the following intervals: $[0, 100]$ (red), $[200, 800]$ (orange), $[900, 10000]$ (blue).

S.9.20 Parameter values used to realize the plots in SI Fig. SM.5G

$k^A_{W0} = 0$, $k^A_W = 7.8075$ h^{-1}, $\bar{k}^A_E = 0.0176$ h^{-1}, $\delta = 0.035$ h^{-1}, $\frac{k^A_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^A_E}{\Omega} = 0.8675$ h^{-1}, $k^1_{W0} = 0$, $k^1_W 0$, $k'_T = 0$, $\delta' = 0$, $\frac{k'_M}{\Omega} = 0$, $\frac{k'^*_T}{\Omega} = 0$, $k^2_{W0} = 0$, $k^2_W = 0$ $\bar{k}^R_E = 0.0018$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^R_E}{\Omega} = 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta^A_m = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. As initial condition, we consider values of (n^R_{12},n^R_1,n^R_2,n^A,n^X,n^m) from the simulations associated with the orange distribution in SI Fig. SM.5D and divide them with respect to their value of n^X, following the following intervals: $[0, 80]$ (Bin 1), $[90, 150]$ (Bin 2), $[250, 400]$ Bin 3, $[500, 10000]$ (Bin 4).

S.9.21 Parameter values used to realize the plots in SI Fig. SM.6A

$k^A_{W0} = 0$, $k^A_W = 7.8075$ h^{-1}, $\bar{k}^A_E = 0.0176$ h^{-1}, $\delta = 0.035$ h^{-1}, $\frac{k^A_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^A_E}{\Omega} = 0.8675$ h^{-1}, $k^1_{W0} = 0$, $k^1_W = 0$, $c = 1.74 \cdot 10^{-6}$, $3.47 \cdot 10^{-6}$, $3.47 \cdot 10^{-5}$ h^{-1}, $c_0 = 1.74 \cdot 10^{-6}$, $3.47 \cdot 10^{-6}$, $3.47 \cdot 10^{-5}$ h^{-1}, $\delta' = 1.74 \cdot 10^{-6}$, $3.47 \cdot 10^{-6}$, $3.47 \cdot 10^{-5}$ h^{-1}, $\tilde{\delta}' = 1.74 \cdot 10^{-6}$, $3.47 \cdot 10^{-6}$, $3.47 \cdot 10^{-5}$ h^{-1}, $K = 1$, $\frac{k'_M}{\Omega} = 0$, $\frac{k'^*_T}{\Omega} =$ h^{-1}, $k^2_{W0} = 0$, $k^2_W = 0$, $\bar{k}^R_E = 0.0018$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^R_E}{\Omega} = 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta^A_m = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. We consider four initial conditions: (n^R_{12},n^R_1,n^R_2,n^A) = (14,0,0,1) (blue), (n^R_{12},n^R_1,n^R_2,n^A) = (6,0,0,9) (red), (n^R_{12},n^R_1,n^R_2,n^A) = (4,0,0,11) (yellow), (n^R_{12},n^R_1,n^R_2,n^A) = (0,0,0,15) (purple). For each case, the initial values of n^X of n^m at time $t = 0$ were set to their steady states.

S.9.22 Parameter values used to realize the plots in SI Fig. SM.6B

$k^A_{W0} = 0$, $k^A_W = 7.8075$ h^{-1}, $\bar{k}^A_E = 0.0176$ h^{-1}, $\delta = 0.035$ h^{-1}, $\frac{k^A_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^A_E}{\Omega} = 0.8675$ h^{-1}, $k^1_{W0} = 0$, $k^1_W = 0$, $c = 3.47 \cdot 10^{-7}$, $1.04 \cdot 10^{-6}$, $3.47 \cdot 10^{-6}$ h^{-1}, $c_0 = 3.47 \cdot 10^{-6}$ h^{-1}, $\delta' = 3.47 \cdot 10^{-6}$ h^{-1}, $\tilde{\delta}' = 3.47 \cdot 10^{-7}$, $1.04 \cdot 10^{-6}$, $3.47 \cdot 10^{-6}$ h^{-1}, $K = 1$, $\frac{k'_M}{\Omega} = 0$, $\frac{k'^*_T}{\Omega} =$ h^{-1}, $k^2_{W0} = 0$, $k^2_W = 0$, $\bar{k}^R_E = 0.0018$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^R_E}{\Omega} = 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta^A_m = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. We consider four initial conditions: (n^R_{12},n^R_1,n^R_2,n^A) = (14,0,0,1) (blue), (n^R_{12},n^R_1,n^R_2,n^A) = (6,0,0,9) (red), (n^R_{12},n^R_1,n^R_2,n^A) = (4,0,0,11) (yellow), (n^R_{12},n^R_1,n^R_2,n^A) = (0,0,0,15) (purple). For each case, the initial values of n^X of n^m at time $t = 0$ were set to their steady states of the ODEs.

S.9.23 Parameter values used to realize the plots in SI Fig. SM.7A

$k^A_{W0} = 0$, $k^A_W = 7.8075$ h^{-1}, $\bar{k}^A_E = 0.0176$ h^{-1}, $\delta = 0.035$ h^{-1}, $\frac{k^A_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k^A_E}{\Omega} = 0.8675$ h^{-1}, $k^1_{W0} = 0$, $k^1_W = 0$, $k'_T = 0$, $\delta' = 0$, $\frac{k'_M}{\Omega} = 0$, $\frac{k'^*_T}{\Omega} = 0$, $k^2_{W0} = 0$, $k^2_W \in [26.02130.13]e^{-\delta t}$ h^{-1}, $\bar{k}^R_E = 0.0018$

h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k_E^R}{\Omega} = 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta_m^A = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. As initial condition, we consider values of (n_{12}^R,n_1^R,n_2^R,n^A,n^X,n^m) from the simulations associated with the orange distribution in SI Fig. SM.5C.

S.9.24 Parameter values used to realize the plots in SI Fig. SM.7B

$k_{W0}^A = 0$, $k_W^A = 7.8075$ h^{-1}, $\bar{k}_E^A = 0.0176$ h^{-1}, $\delta = 0.035$ h^{-1}, $\frac{k_M^A}{\Omega} = 0.0347$ h^{-1}, $\frac{k_E^A}{\Omega} = 0.8675$ h^{-1}, $k_{W0}^1 = 0$, $k_W^1 \in [0.1041, 0.36]e^{-\delta t}$ h^{-1}, $k_T' = 0$, $\delta' = 0$, $\frac{k_M'}{\Omega} = 0$, $\frac{k_T'^*}{\Omega} = 0$, $k_{W0}^2 = 0$, $k_W^2 = 0$, $\bar{k}_E^R = 0.0018$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k_E^R}{\Omega} = 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta_m^A = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. As initial condition, we consider values of (n_{12}^R,n_1^R,n_2^R,n^A,n^X,n^m) from the simulations associated with the orange distribution in SI Fig. SM.5C.

S.9.25 Parameter values used to realize the plots in SI Fig. SM.7C

$k_{W0}^A = 0$, $k_W^A = 7.8075$ h^{-1}, $\bar{k}_E^A = 0.0176$ h^{-1}, $\delta = 0.035$ h^{-1}, $\frac{k_M^A}{\Omega} = 0.0347$ h^{-1}, $\frac{k_E^A}{\Omega} = 0.8675$ h^{-1}, $k_{W0}^1 = 0$, $k_W^1 = 0$, $k_T' \in [0.7 \cdot 10^{-4}, 0.24 \cdot 10^{-3}]e^{-\delta t}$ h^{-1}, $\delta' = 0$, $\frac{k_M'}{\Omega} = 0$, $\frac{k_T'^*}{\Omega} in[3.5 \cdot 10^{-3}, 1.2 \cdot 10^{-2}]e^{-\delta t}$ h^{-1}, $k_{W0}^2 = 0$, $k_W^2 = 0$, $\bar{k}_E^R = 0.0018$ h^{-1}, $\frac{k_M}{\Omega} = 0.0347$ h^{-1}, $\frac{\bar{k}_M}{\Omega} = 0.0347$ h^{-1}, $\frac{k_E^R}{\Omega} = 0.0868$ h^{-1}, $\beta_m = 0.2556$ h^{-1}, $\beta_m^A = 0.0021$ h^{-1}, $\beta = 2.52$ h^{-1}, $\gamma_m = 0.24$ h^{-1}. As initial condition, we consider values of (n_{12}^R,n_1^R,n_2^R,n^A,n^X,n^m) from the simulations associated with the orange distribution in SI Fig. SM.5C.

S.10 Supplementary Experimental Figures

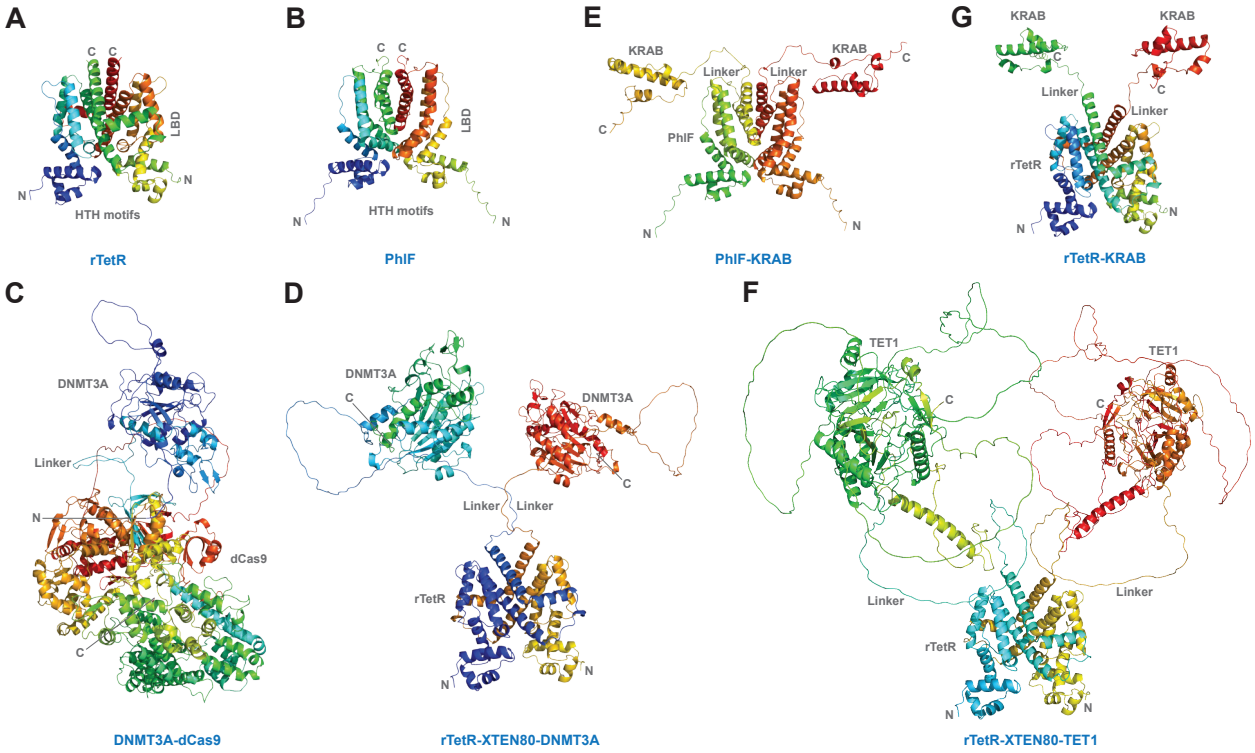

Fig. SE.1. Continued on the following page.

Fig. ES.1. Protein structures of epigenetic regulators used in this study. A Homodimer structure of DNA-binding protein rTetR. **B** Homodimer structure DNA-binding protein PhlF. **C** DNMT3A, writer of DNA methylation, fused to programmable DNA-binding protein dCas9. **D** DNMT3A fused to DNA-binding protein rTetR. **E** KRAB, a regulator of H3K9me3, fused to DNA-binding protein PhlF. **F** TET1, an an eraser of DNA methylation, fused to rTetR. **G** KRAB fused to rTeR. Predictions performed using AlphaFold2 and PyMol (Methods). For rTetR and PhlF TetR-family proteins, the epigenetic effectors KRAB, DNMT3A, and TET1 are fused proximal to alpha helix 9 at the C-terminus using a flexible linker. Helix-turn-helix (HTH) motifs bind DNA. The ligand binding domain (LBD) binds dox in the case or rTetR and DAPG in the case of PhlF. For dCas9, DNMT3A is fused at the N terminus using a flexible linker.

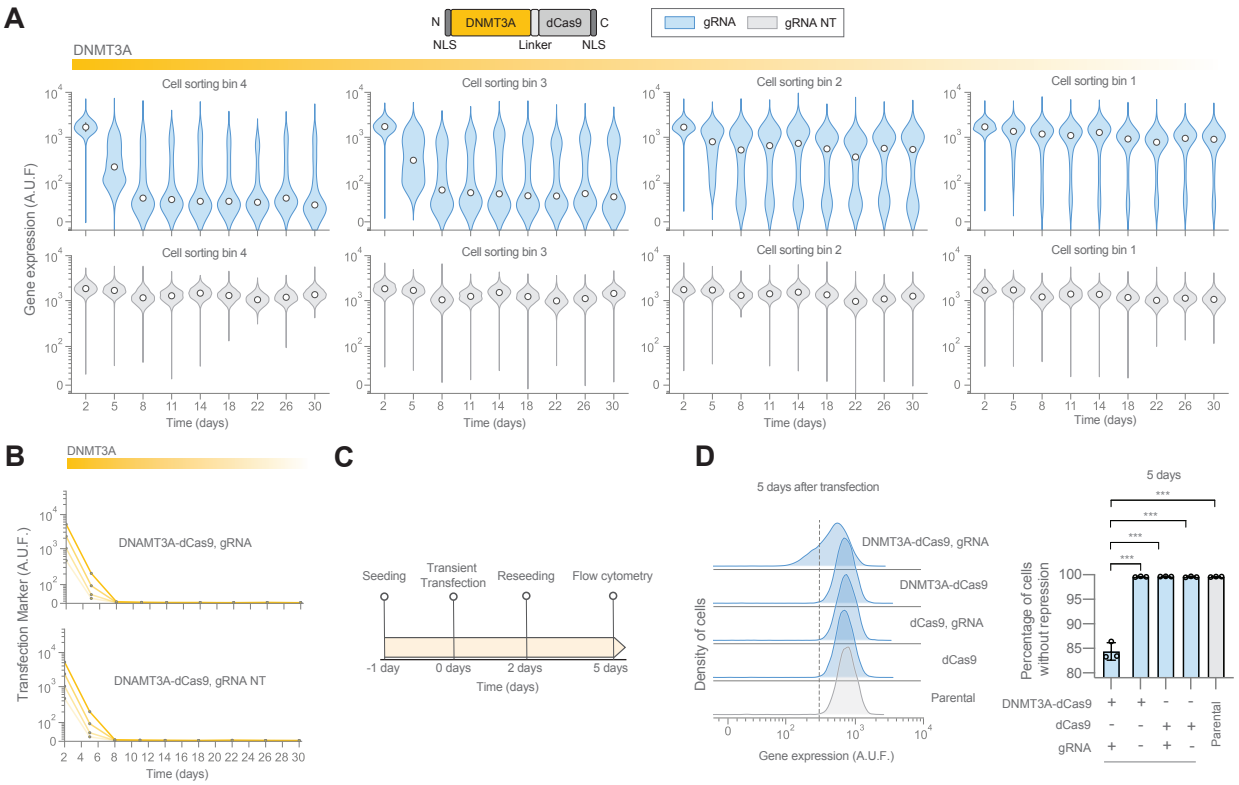

Fig. SE.2. Continued on the following page.

Fig. ES.2. Stationary distributions of DNMT3A-silenced cells exhibit continuous gene expression. A Gene expression dynamics after epigenetic modulation using DNMT3A-dCas9 corresponding to the cells shown in Fig.2C, D. Data shown are from a representative replicate from 3 independent replicates. DNMT3A-dCas9 was targeted to 5 target sites upstream of the promoter (gRNA). A gRNA with a scrambled target sequence (gRNA NT) was used as a control.**B** Transfection marker (EYFP) after epigenetic modulation using DNMT3A-dCas9 corresponding to the cells shown in Fig.2C, D and Panel A. Data shown are from a representative replicate from 3 independent replicates. **C** Experimental overview of DNMT3A-dCas9 for various experimental conditions. **D** Experimental analysis of DNMT3A-dCas9 using various experimental conditions for repression of the reporter gene according to the experimental overview shown in Panel C. Cell densities are from a representative replicate from 3 independent replicates. Percentages and mean percentages are shown from three independent replicates ($300 - 10^5$ gene expression A.U.F.). Statistical tests performed on all pairs and showing statistically significant results. Error bars are s.d. of mean percentage.

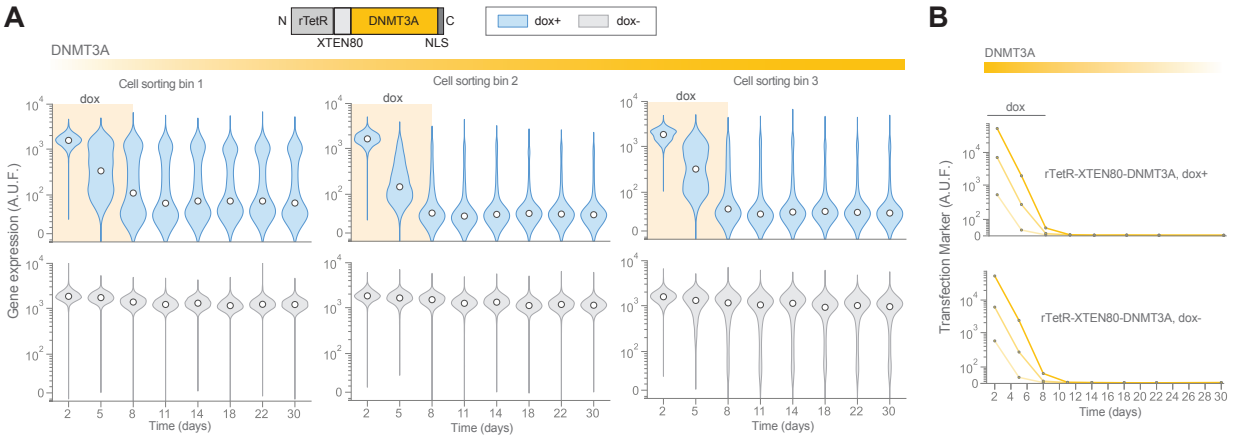

Fig. SE.3. Continued on the following page.

Fig. ES.3. Stationary distributions of cells silenced using rTetR-XTEN80-DNMT3A exhibit intermediate levels of gene expression. A Gene expression dynamics after DNMT3A-mediated silencing with rTetR-XTEN80-DNMT3A. Shown are flow cytometry measurements of the reporter gene (EBFP2) for three different levels of transfection that were sorted using fluorescence-activated cell sorting. Data is from a representative replicate from 3 independent replicates. **B** Transfection markers after epigenetic modulation using DNMT3A-dCas9 corresponding to the cells shown in Panel A. Data shown are from a representative replicate from 3 independent replicates.

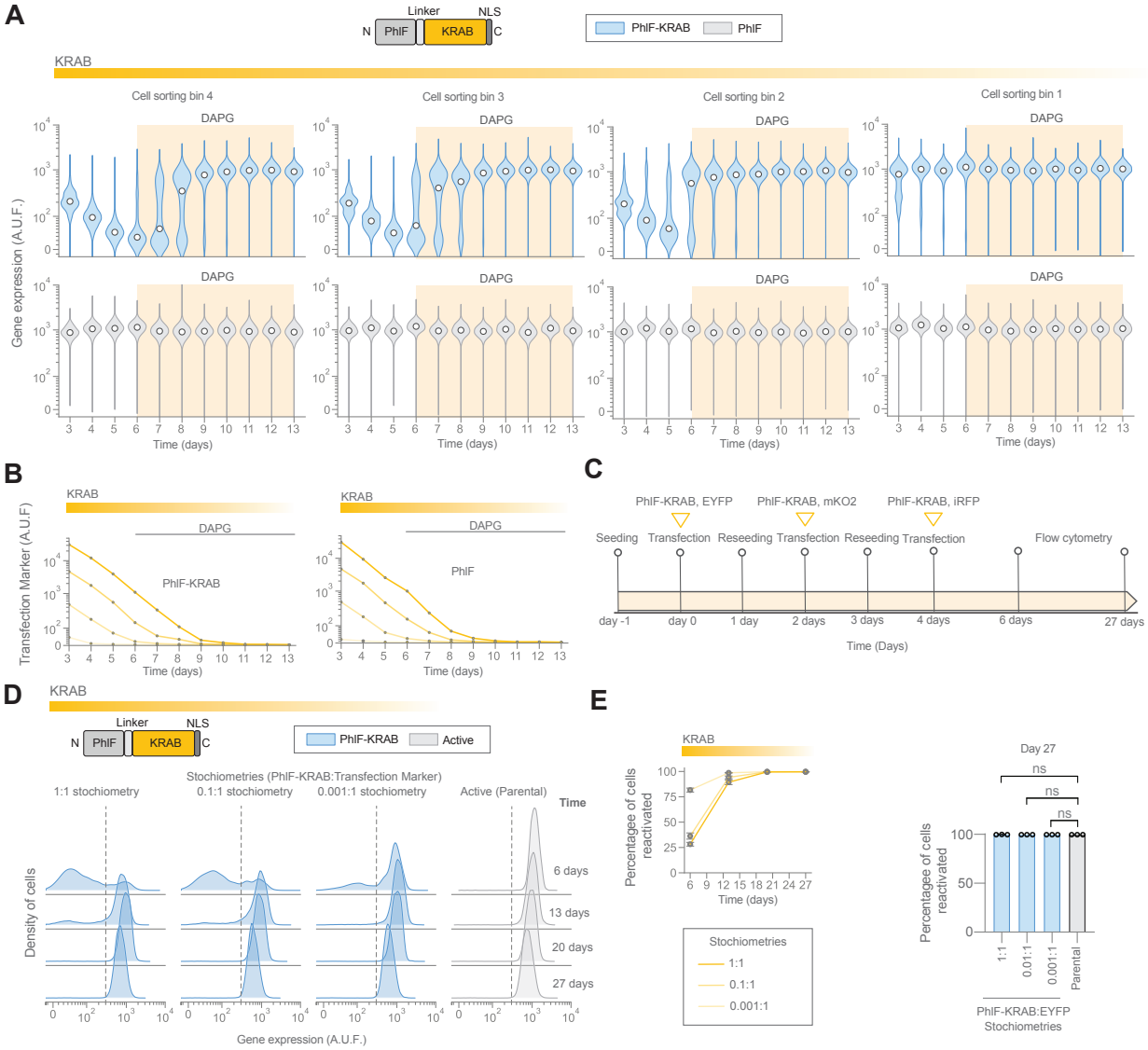

Fig. SE.4. Continued on the following page.

Fig. ES.4. KRAB-mediated silencing does to confer memory.

A Gene expression dynamics after PhlF-KRAB recruitment to the reporter gene. Shown are flow cytometry measurements of the reporter gene (EBFP2) for four different levels of transfection sorted using fluorescence-activated cell sorting. Data is from a representative replicate from 3 independent replicates. **B** Transfection markers of epigenetic modulation using PhlF-KRAB corresponding to Fig. 2F,G and Panel A. Data shown are from a representative replicate of three independent replicates. **C** Experimental overview of KRAB-mediated silencing using PhlF-KRAB. **D** Density of cells for KRAB-mediated silencing using PhlF-KRAB according to Panel C. Data shown is from a representative replicate from 3 independent replicates. **E** Percentage of cell reactivation $(300 - 10^5$ gene expression A.U.F.) in Panel D. Shown is the mean percentage from three independent replicates. Error bars are the s.d. of the mean.

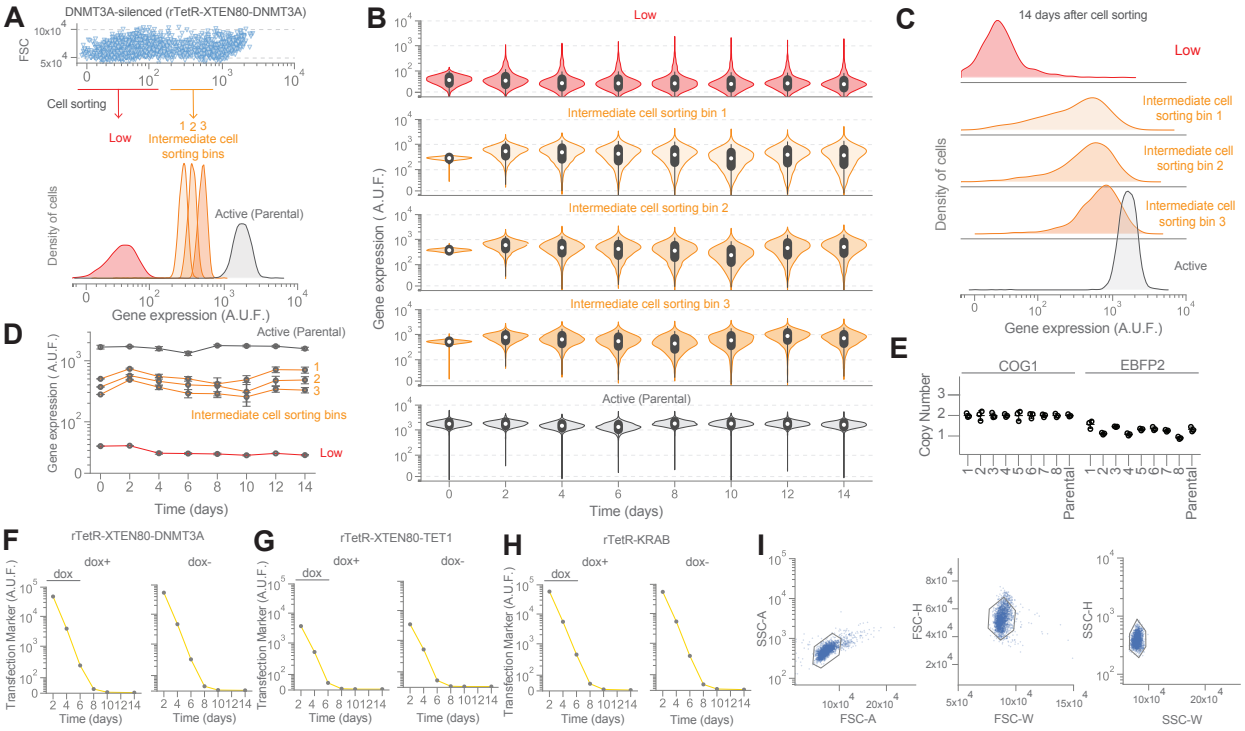

Fig. SE.5. Continued on the following page.

Fig. ES.5. Intermediate levels of gene expression are stable in cell silenced using rTetR instead of dCas9 for targeted recruitment of DNMT3A. A Fluorescence-activated cell sorting of cells exhibiting low, intermediate, and high levels of gene expression in DNMT3A-silenced cells (rTetR-XTEN80-DNMT3A). Data shown is from a representative replicate from three independent replicates. **B** Time-course flow cytometry analysis of the dynamics of gene expression exhibited by cells obtained in panel A. Data is from one representative replicate from three independent replicates. **C** Density of cells from flow cytometry measurements obtained on the last time point in panel B (14 days after cell sorting). Data shown is from one representative replicate from three independent replicates. **D** Dynamics of gene expression exhibited by cells shown in panels A-C. Shown is the mean of geometric mean fluorescence from three independent replicates. Error bars are s.d. of the mean. **E** Digital PCR corresponding to the parental cell line and monoclonal populations corresponding to Fig. 4 For each cell population, the average COG1 value was calculated (diploid). All data was then normalized to the COG1 average for each cell population separately. **F** Transfection markers corresponding to Fig. 5A. Data shown are from a representative replicate from 3 independent replicates. **G** Transfection markers corresponding to Fig. 5B. Data shown are from a representative replicate from 3 independent replicates. **H** Transfection markers corresponding to Fig. 5C. Data shown are from a representative replicate from 3 independent replicates. **I** Representative gating of cells used in this study. Data shown are from a representative replicate gating the cells bearing the reporter gene.

References

1. Felsenfeld, G. A brief history of epigenetics. *Cold Spring Harb Perspect Biol.* **6,** a018200. https://doi.org/10.1101/cshperspect.a018200 (2012).
2. Nishiyama, A. & Nakanishi, M. Navigating the DNA Methylation Landscape of Cancer. *Trends in Genetics* **37,** 1012–1027. ISSN: 0168-9525. https://www.sciencedirect.com/science/article/pii/S016895252100130X (2021).
3. Jones, M., Goodman, S. & Kobor, M. DNA Methylation and Healthy Human Aging. *Aging Cell* **14,** 924–932 (Dec. 2015).
4. Saxton, D. S. & Rine, J. Distinct silencer states generate epigenetic states of heterochromatin. *Mol Cell* **82,** 3566–3579. https://doi.org/10.1016/j.molcel.2022.08.002 (2022).
5. Bintu, L. *et al.* Dynamics of epigenetic regulation at the single-cell level. *Science*. https://doi.org/10.1126/science.aab2956 (2016).
6. Holmberg, J. & Perlmann, T. Maintaining differentiated cellular identity. *Nat Rev Genet* **13,** 429–39. https://doi.org/10.1038/nrg3209 (2012).
7. Zhang, Q. & Cao, X. Epigenetic regulation of the innate immune response to infection. *Nat Rev Immunol* **19,** 417–432. https://doi.org/10.1038/s41577-019-0151-6 (2019).
8. Day, J. J. & Sweatt, J. D. DNA methylation and memory formation. *Nat Neurosci* **13,** 1319–23. https://doi.org/10.1038/nn.2666 (2010).
9. Thakore, P., Black, J., Hilton, I. & Gersbach, C. Editing the epigenome: technologies for programmable transcription and epigenetic modulation. *Nat Methods* **13,** 127–137 (Feb. 2016).
10. Kouzarides, T. Chromatin modifications and their function. *Cell* **128,** 693–705. https://doi.org/10.1016/j.cell.2007.02.005 (2007).
11. Allis, C. D., Caparros, M.-L., Jenuwein, T. & Reinberg, D. *Epigenetics* (Cold Spring Harbor Laboratory Press, Second Edition, 2015).
12. Huang, S., Litt, M. & Blakey, C. A. *Epigenetic Gene Expression and Regulation* (Academic Press, 2015).
13. Allis, C. D. & Jenuwein, T. The molecular hallmarks of epigenetic control. *Nat Rev Genet* **17,** 487–500. https://doi.org/10.1038/nrg.2016.59 (2016).
14. Li, E. & Zhang, Y. DNA methylation in mammals. *Cold Spring Harb Perspect Biol* **6.** https://doi.org/10.1101/cshperspect.a019133 (2014).

1494 15. Cheng, X. Structural and functional coordination of DNA and histone methylation. *Cold Spring Harb Perspect*
1495 *Biol.* **6,** 693–705. https://doi.org/10.1101/cshperspect.a018747 (2014).

1496 16. Bannister, A. J. *et al.* Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain.
1497 *Nature* **410,** 120–4. https://doi.org/10.1038/35065138 (2001).

1498 17. Jones, P. L. *et al.* Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat*
1499 *Genet* **19,** 187–191. https://doi.org/10.1038/561 (1998).

1500 18. Schneider, R. *et al.* Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. *Nat Cell Biol* **6,**
1501 73–7. https://doi.org/10.1038/ncb1076 (2003).

1502 19. Hathaway, N. A. *et al.* Dynamics and memory of heterochromatin in living cells. *Cell* **149,** 1447–1460.
1503 https://doi.org/10.1016/j.cell.2012.03.052 (2012).

1504 20. Zhang, T., Cooper, S. & Brockdorff, N. The interplay of histone modifications - writers that read. *EMBO Rep*
1505 **16,** 1467–1481. https://doi.org/10.15252/embr.201540945 (2015).

1506 21. Angel, A., Song, J., Dean, C. & Howard, M. A Polycomb-based switch underlying quantitative epigenetic
1507 memory. *Nature* **476,** 105–108. https://doi.org/10.1038/nature10241 (2011).

1508 22. Dodd, I. B., Micheelsen, M. A., Sneppen, K. & Thon, G. Theoretical Analysis of Epigenetic Cell Memory by
1509 Nucleosome Modification. *Cell* **129,** 813–822. https://doi.org/10.1016/j.cell.2007.02.053 (2007).

1510 23. Ulm, E. A., Sleiman, S. F. & Chamberlin, H. M. Developmental functions for the Caenorhabditis elegans Sp
1511 protein SPTF-3. *Mech Dev* **128,** 428–41. https://doi.org/10.1016/j.mod.2011.08.003 (2011).

1512 24. Yao, Z. *et al.* A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell*
1513 **184,** 3222–3241. https://doi.org/10.1016/j.cell.2021.04.021 (2021).

1514 25. Stanley, G., Gokce, O., Malenka, R. C., Südhof, T. C. & Quake, S. R. Continuous and Discrete Neuron Types
1515 of the Adult Murine Striatum. *Neuron* **105,** 688–699. https://doi.org/10.1016/j.neuron.2019.11.004 (2020).

1516 26. Phillips, J. W. *et al.* A repeated molecular architecture across thalamic pathways. *Nat Neurosci* **22,** 1925–1935.
1517 https://doi.org/10.1038/s41593-019-0483-3 (2019).

1518 27. Gaidukov, L. *et al.* A multi-landing pad DNA integration platform for mammalian cell engineering. *Nucleic*
1519 *Acids Res.* **46,** 4072–4086 (May 2018).

1520 28. Recillas-Targa, F. *et al.* Position-effect protection and enhancer blocking by the chicken beta-globin insulator
1521 are separable activities. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 6883–6888 (May 2002).

1522 29. Stanton, B. C. *et al.* Systematic transfer of prokaryotic sensors and circuits to mammalian cells. *ACS Synth.*
1523 *Biol.* **3,** 880–891 (Dec. 2014).

1524 30. O'Geen, H. *et al.* dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient
1525 for target gene repression. *Nucleic Acids Res.* **45,** 9901–9916 (Sept. 2017).

1526 31. Fuks, F. *et al.* The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. *J Biol*
1527 *Chem* **278,** 4035–4040. https://doi.org/10.1074/jbc.M210256200 (2003).

1528 32. Margolin, J. F. *et al.* Krüppel-associated boxes are potent transcriptional repression domains. *Proc. Natl.*
1529 *Acad. Sci.* **91,** 4509–4513. 10.1073/pnas.91.10.4509 (1994).

1530 33. Schultz, D. C., Ayyanathan, K., Negorev, D., Maul, G. G. & Rauscher, F. J. SETDB1: a novel KAP-1-
1531 associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of
1532 euchromatic genes by KRAB zinc-finger proteins. *Annual Review of Physical Chemistry* **16,** 919–32. https:
1533 //doi.org/10.1101/gad.973302 (2002).

1534 34. Ecco, G., Imbeault, M. & Trono, D. KRAB zinc finger proteins. *Development* **144,** 2719–2729. https://doi.
1535 org/10.1242/dev.132605 (2017).

1536 35. Nuñez, J. K. *et al.* Genome-wide programmable transcriptional memory by CRISPR-based epigenome
1537 editing. *Cell* **184,** 2503–2519 (2021).

1538 36. Hervouet, E., Peixoto, P., Delage-Mourroux, R., Boyer-Guittaut, M. & Cartron, P.-F. Specific or not specific
1539 recruitment of DNMTs for DNA methylation, an epigenetic dilemma. *Clin Epigenetics,* 17. https://doi.org/10.
1540 1186/s13148-018-0450-y (2018).

1541 37. Lopez-Serra, L. & Esteller, M. Proteins that bind methylated DNA and human cancer: reading the wrong
1542 words. *Br J Cancer* **98,** 1881–1885. https://doi.org/10.1038/sj.bjc.6604374 (2008).

1543 38. Trerotola, M., Relli, V., Simeone, P. & Alberti, S. Epigenetic inheritance and the missing heritability. *Hum*
1544 *Genomics* **9,** 17. https://doi.org/10.1186/s40246-015-0041-3 (2015).

1545 39. Upadhyay, A. K., Horton, J. R., Zhang, X. & Cheng, X. Coordinated methyl-lysine erasure: structural and
1546 functional linkage of a Jumonji demethylase domain and a reader domain. *Curr Opin Struct Biol* **21,** 750–60.
1547 https://doi.org/10.1016/j.sbi.2011.08.003 (2011).

1548 40. Rasmussen, K. D. & Helin, K. Role of TET enzymes in DNA methylation, development, and cancer. *Genes*
1549 *Dev.* **30,** 733–50. https://doi.org/10.1101/gad.276568.115 (2016).

1550 41. Brown, D. A. *et al.* The SET1 Complex Selects Actively Transcribed Target Genes via Multivalent Interaction
1551 with CpG Island Chromatin. *Cell Rep* **20,** 2313–2327. https://doi.org/10.1016/j.celrep.2017.08.030 (2017).

1552 42. Bruno, S., Williams, R. J. & Del Vecchio, D. Epigenetic cell memory: The gene's inner chromatin modification
1553 circuit. *PLOS Computational Biology* **18,** 1–27. https://doi.org/10.1371/journal.pcbi.1009961 (2022).

1554 43. Jackson, M. *et al.* Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation
1555 in embryonic stem cells. *Mol Cell Biol* **24.** https://doi.org/10.1128/MCB.24.20.8862-8871 (2004).

1556 44. Bannister, A. J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **21,** 381–395.

1557 45. Gillespie, D. T. Stochastic Simulation of Chemical Kinetics. *Annual Review of Physical Chemistry* **58,** 35–55.
1558 https://doi.org/10.1146/annurev.physchem.58.032806.104637 (2007).

1559 46. Ren, W. *et al.* Direct readout of heterochromatic H3K9me3 regulates DNMT1-mediated maintenance DNA
1560 methylation. *Proceedings of the National Academy of Sciences of the United States of America* **117,**
1561 18439–18447. https://doi.org/10.1073/pnas.2009316117 (2020).

1562 47. Laird, C. D. *et al.* Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary
1563 strands of individual DNA molecules. *Proceedings of the National Academy of Sciences of the United States*
1564 *of America* **101,** 204–209. https://doi.org/10.1073/pnas.2536758100 (2004).

1565 48. Riggs, A. & Xiong, Z. Methylation and epigenetic fidelity. *Proceedings of the National Academy of Sciences*
1566 *of the United States of America* **101,** 4–5. https://doi.org/10.1073/pnas.2536758100 (2004).

1567 49. Ito, S. *et al.* Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*
1568 **333,** 1300–3. https://doi.org/10.1126/science.1210597 (2011).

1569 50. Liu, X. S. *et al.* Editing DNA methylation in the mammalian genome. *Cell* **167,** 233–247 (Sept. 2016).

1570 51. Mattei, A. L., Bailly, N. & Meissner, A. DNA methylation: a historical perspective. *Trends Genet* **38,** 676–707.
1571 https://doi.org/10.1016/j.tig.2022.03.010 (2022).

1572 52. Hong, T. *et al.* TET2 modulates spatial relocalization of heterochromatin in aged hematopoietic stem and
1573 progenitor cells. *Nat Aging* **3,** 1387–1400. https://doi.org/10.1038/s43587-023-00505-y (2023).

1574 53. Du, J., Johnson, L. M., Jacobsen, S. E. & Patel, D. J. DNA methylation pathways and their crosstalk with
1575 histone methylation. *Nat Rev Mol Cell Biol* **16,** 519–32. https://doi.org/10.1038/nrm4043 (2015).

1576 54. Ooi, S. K. *et al.* DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA.
1577 *Nature* **448,** 714–717. https://doi.org/10.1038/nature05987 (2007).

1578 55. Nan, X. *et al.* Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone
1579 deacetylase complex. *Nature* **393,** 383–389. https://doi.org/10.1038/30764 (1998).

1580 56. Zeng, H. What is a cell type and how to define it? *Cell* **185,** 2739–2755. https://doi.org/10.1016/j.cell.2022.
1581 06.031 (2022).

1582 57. Liu, H. *et al.* DNA methylation atlas of the mouse brain at single-cell resolution. *Nature* **598,** 120–128.
1583 https://doi.org/10.1038/s41586-020-03182-8 (2021).

58. Clark, S. J. *et al.* scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat Commun* **9,** 781. https://doi.org/10.1038/s41467-018-03149-4 (2018).

59. Waddington, C. H. *The strategy of the genes* (Routledge, 1957).

60. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596,** 583–589. ISSN: 1476-4687. https://doi.org/10.1038/s41586-021-03819-2 (Aug. 2021).

61. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nature Methods* **19,** 679–682. ISSN: 1548-7105. https://doi.org/10.1038/s41592-022-01488-1 (June 2022).

62. Jin, C. *et al.* TET1 is a maintenance DNA demethylase that prevents methylation spreading in differentiated cells. *Nucleic Acids Research* **42,** 6956–71. https://doi.org/10.1093/nar/gku372 (2014).

63. Huisinga, K. L., Brower-Toland, B. & Elgin, S. C. The contradictory definitions of heterochromatin: transcription and silencing. *Chromosoma* **115,** 110–22. https://doi.org/10.1007/s00412-006-0052-x (2006).

64. Del Vecchio, D. & Murray, R. M. *Biomolecular Feedback Systems* (Princeton University Press, 2014).

65. Bruno, S., Williams, R. J. & Del Vecchio, D. *Model reduction and stochastic analysis of the histone modification circuit* in *2022 European Control Conference (ECC)* (2022), 264–271.

66. Felsenfeld, G. & Groudine, M. Controlling the double helix. *Nature* **421,** 448–53. https://doi.org/10.1038/nature01411 (2003).

67. Friedman, N., Cai, L. & Xie, X. S. Linking Stochastic Dynamics to Population Distribution: An Analytical Framework of Gene Expression. *Phys. Rev. Lett* **97.** https://doi.org/10.1103/PhysRevLett.97.168302 (2006).