PLOS COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Epigenetic cell memory: The gene's inner chromatin modification circuit

Simone Bruno[1], **Ruth J. Williams**[2], **Domitilla Del Vecchio**[1]*

1 Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America, **2** Department of Mathematics, University of California, San Diego, La Jolla, California, United States of America

* ddv@mit.edu

Citation: Bruno S, Williams RJ, Del Vecchio D (2022) Epigenetic cell memory: The gene's inner chromatin modification circuit. PLoS Comput Biol 18(4): e1009961. https://doi.org/10.1371/journal.pcbi.1009961

Editor: Carl Herrmann, Heidelberg University, GERMANY

Received: October 18, 2021

Accepted: February 24, 2022

Published: April 6, 2022

Data Availability Statement: Matlab and Mathematica codes used in the paper are available on a GitHub repository at https://github.com/simonbruno100/EpigeneticCellMemorypaper2021. We have also used Zenodo to assign a DOI to the repository: 10.5281/zenodo.5973843.

Funding: S.B. was supported by NIH/NIBIB Grant Number R01EB024591 (PI: D.D.V.) (https://www.nibib.nih.gov/), in part by NSF Collaborative Research grant MCB-2027949 (PI: D.D.V.) (https://www.nsf.gov/) and by the DoD Newton Award for Transformative Ideas during the COVID-19

Abstract

Epigenetic cell memory allows distinct gene expression patterns to persist in different cell types despite a common genotype. Although different patterns can be maintained by the concerted action of transcription factors (TFs), it was proposed that long-term persistence hinges on chromatin state. Here, we study how the dynamics of chromatin state affect memory, and focus on a biologically motivated circuit motif, among histones and DNA modifications, that mediates the action of TFs on gene expression. Memory arises from time-scale separation among three circuit's constituent processes: basal erasure, auto and cross-catalysis, and recruited erasure of modifications. When the two latter processes are sufficiently faster than the former, the circuit exhibits bistability and hysteresis, allowing active and repressed gene states to coexist and persist after TF stimulus removal. The duration of memory is stochastic with a mean value that increases as time-scale separation increases, but more so for the repressed state. This asymmetry stems from the cross-catalysis between repressive histone modifications and DNA methylation and is enhanced by the relatively slower decay rate of the latter. Nevertheless, TF-mediated positive autoregulation can rebalance this asymmetry and even confers robustness of active states to repressive stimuli. More generally, by wiring positively autoregulated chromatin modification circuits under time scale separation, long-term distinct gene expression patterns arise, which are also robust to failure in the regulatory links.

Author summary

Epigenetic cell memory ensures that cells are locked into specialized functions for the lifetime of an organism. Phenotype loss is often associated with disease, such as cancer, and also required for artificially reprogramming cells from one type to another. Chromatin state, determined by histone modifications and DNA methylation, has recently appeared as a key mediator of epigenetic cell memory. However, a mechanistic understanding of how the dynamics of chromatin state affect the temporal duration of this memory is lacking. Here, we developed and analyzed a theoretical framework that includes these dynamics in gene regulation. Our results show that when both recruited erasure and auto/cross-catalysis among histone modifications and DNA methylation are sufficiently slower than

Pandemic (https://basicresearch.defense.gov/COVID-19/Newton-Award-for-Transformative-Ideas/). R.J.W. was supported in part by NSF Collaborative Research grant MCB-2027947 (PI: R. J.W.) and by the Charles Lee Powell Foundation (https://clpf.org/). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

basal erasure of all modifications, loss of cell memory will occur. Our mathematical formulas show how the parameters capturing these time scales depend on the abundance of methyl-DNA-binding proteins, on writers, erasers, and readers of nucleosome modifications, and on cell division time. With this information, one may design experimental interventions to either enforce phenotypic plasticity or re-lock phenotypes in aberrant cells.

Introduction

The maintenance of distinct cell identities without any difference in genotype is a critical property of multicellular organisms. This property is captured by the word "epigenetic", coined by Conrad H. Waddington in 1942, to broadly indicate that different cell phenotypes can be associated with the same genotype [1]. A natural question is how multicellular organisms can ensure that distinct phenotypes are safely locked-in despite sharing the same genotype. This question has been extensively investigated in systems biology [2, 3] and has been recently revisited in light of the so-called "Epigenetic Revolution" [4]. Maintenance of distinct phenotypes requires the inheritance of different gene expression patterns through subsequent cell divisions [5]. During the natural process of cell differentiation, cells go through multiple fate decisions and select among different gene expression patterns that are mutually exclusive and are robust to both intrinsic noise and TF perturbations [6]. According to classical theory, these patterns are all stable attractors of the dynamics of gene regulation networks (GRNs) enabled by TF binding to DNA. In each of these attractors, lineage-specific genes become continuously expressed and often maintain their expression through positive autoregulation [7–9]. Therefore, the GRN-centric perspective proposes that TF-enabled regulatory networks are fully capable of maintaining epigenetic memory of cell-type-specific gene expression patterns in differentiated cells.

Following an explosive number of discoveries fueling the Epigenetic Revolution, a different and somewhat contrasting view has emerged on epigenetic cell memory [4]. A key argument that detracts to the role of TF-enabled regulatory networks in maintaining memory states is that maintenance is dictated by the binding of TFs to DNA. But, this binding is disrupted upon DNA replication and cell division. Allis specifically says "This mechanism is transient in dividing cells because progression of the DNA replication fork usually disrupts these protein-DNA interactions, which then need to be re-established in the resulting daughter cells." [10] (Chapter 22). In [11], the authors also argue that ". . .feedback loops can clearly enable heritable states of altered gene expression without any need to evoke chromatin. However, it is unlikely that such feedback loops alone would enable the propagation of states throughout the length of development and in the germline of complex organisms."

It has therefore been proposed that it is the chromatin state, i.e., the extent of DNA compaction mediated by histone and DNA modifications, that enables long-term transmission to successive cell generations [10](Chapter 3), [11]. This view is supported by the discovery of enzymatic processes that can copy histone modifications from modified histones to unmodified nearby histones, facilitating the maintenance of these modifications after DNA replication and cell division [10, 12]. It is also reinforced by the already established fact that DNA methylation is copied from the parental DNA strand to the progeny strand immediately after DNA replication [13, 14] and by the discovery of cooperative interactions between DNA methylation and histone modifications [15]. However, histone and DNA modifications result from dynamic enzymatic processes that continuously write and erase modifications and not from a

static read-and-copy process. It thus remains unclear how these dynamic processes may enable long-term maintenance of gene expression states. These considerations motivate the need for a theoretical framework that includes the dynamics of chromatin state in the regulation of gene expression by TFs and determines how these dynamics affect the temporal duration of memory. This question is even more pressing for engineering persistent synthetic genetic circuits in the chromosome of mammalian cells, in which silencing is a major obstacle to keeping proper circuit function for extended temporal durations [16–19].

In this paper, we create a biologically motivated chemical reaction model of the *gene's inner chromatin modification circuit* that mediates the effect of TF inputs on gene expression through histone modifications and DNA methylation. The circuit comprises auto-catalysis of histone modifications, competitive recruitment of erasure between activating and repressive modifications, cross-catalysis between repressive histone modifications and DNA methylation, and basal erasure.

In this model, the effect of TFs on gene expression is not determined by the binding of TFs to DNA alone, but by the recruitment of the enzymes for the *de novo* establishment of either activating or repressive modifications. We first determine how time scale separation among the circuit's constituent processes affects memory of a TF input stimulus for both active and repressed chromatin states, by analyzing the stability and probability of chromatin states, the mean temporal duration of a chromatin state, and the variability of this duration. Then, we study the relative contributions of TF-mediated regulation and chromatin dynamics to memory of active and repressed gene states in two ubiquitous GRN motifs: TF-mediated positive autoregulation and mutual inhibition. Specifically, we analyze the mean temporal duration of active and repressed chromatin states and their robustness to both repetitive disruptions of TF-enabled regulatory links and to undesired input stimuli, such as due to endogenous silencing.

Related work

Models of chromatin-mediated gene regulation, in which histone modifications, DNA methylation, and TF-mediated regulation are combined together, remain under-represented in the literature [20]. Most existing models that include chromatin modifications into gene expression regulation are based on phenomenological rules rather than on the molecular reactions that regulate chromatin state, and focus on specific biological processes, such as iPSC reprogramming [21–24] or epithelial-mesenchymal transition [25], and most of them are only suitable for computational simulations [21–23]. On the opposite side of the spectrum, highly detailed mechanistic models have appeared for histone modifications alone, in which each nucleosome within a gene is modeled and simulated as a distinct unit [26]. Different from these existing works, we explicitly address how the duration of memory of a chromatin state is modulated by the topological properties of the chromatin modification circuit, by relevant biochemical parameters, and by the interplay between TF-based regulation and chromatin dynamics.

Models

Model of the gene's inner chromatin modification circuit

In this paper, we focus on a circuit motif that captures the interactions among H3K4 methylation/acetylation, H3K9 methylation, and DNA methylation, and has the nucleosome as the basic unit that can carry these modifications (Fig 1A). H3K4 methylation (H3K4me3) typically co-exists with acetylation and is associated with active chromatin state [10](Chapter 3), [27], while H3K9 methylation (H3K9me2/3) and DNA methylation are a hallmark of repressed

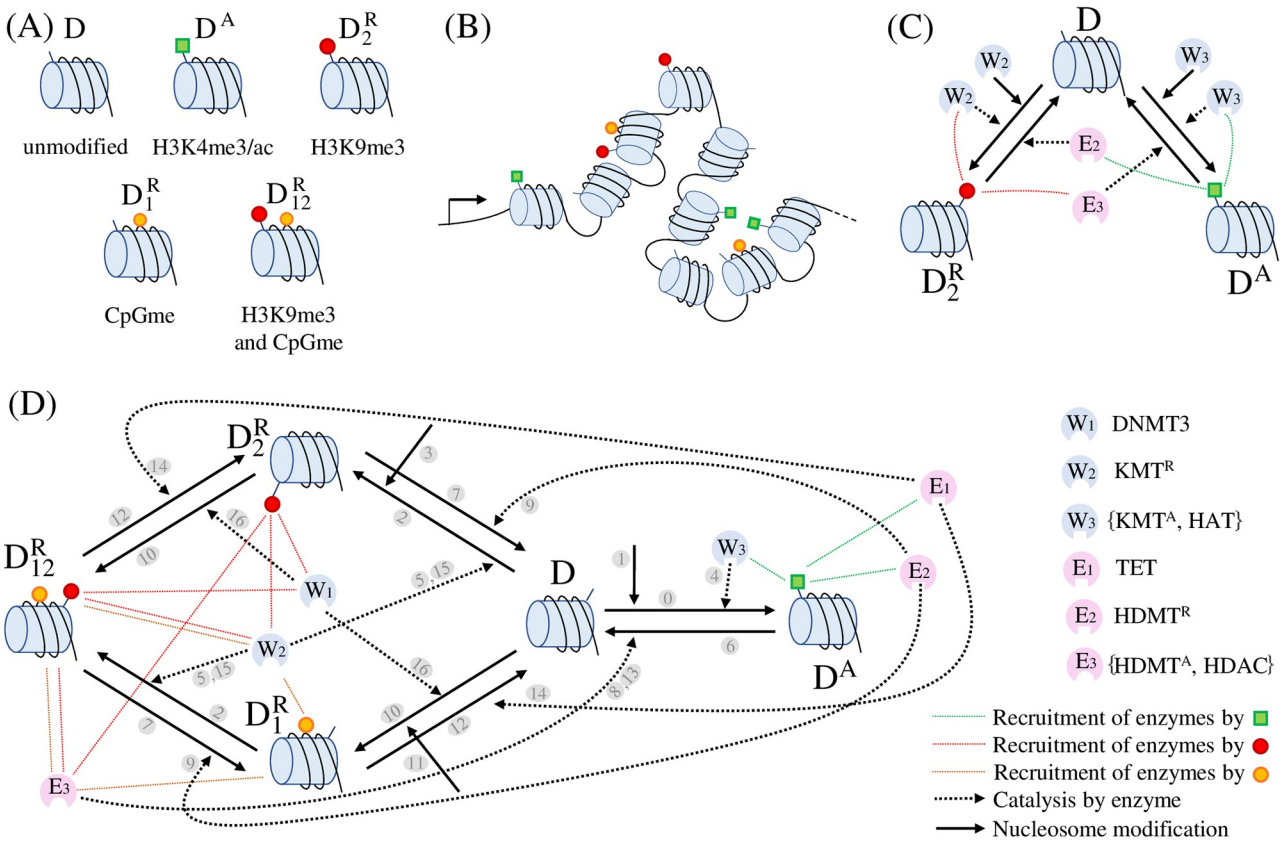

Fig 1. The gene's inner chromatin modification circuit. (A) Nucleosome modifications: D (unmodified nucleosome), D^A (nucleosome with a activating histone modification, H3K4me3/ac), D_1^R (nucleosome with only DNA methylation, CpGme), D_2^R (nucleosome with only a repressive histone modification, H3K9me3) and D_{12}^R (nucleosome with both H3K9me3 and CpGme). (B) Pictorial representation of a gene with its nucleosomes carrying various modifications. The left side arrow represents the promoter. (C) Competitive interactions between opposing histone modifications (activating D^A and repressive D_2^R), wherein each modification recruits writers of itself and erasers of the opposing modification. (D) Complete chromatin modification circuit that includes all the interactions described in Section "Models" (see also Figs B and C in S1 File). The numbers shaded in gray on the arrows correspond to the reactions associated with the arrows, described in the main text and in Fig 2. In (C) and (D), enzymes that write (writers) and erase (erasers) each modification are represented as W_1, W_2, W_3 and E_1, E_2, E_3, respectively. The socket on each enzyme represents a domain that binds to protein readers of the indicated (by the dashed lines) modification, enabling the process by which each modification recruits writers or erasers to nearby histones. To distinguish the KMTs for H3K9 and H3K4, we define the writers for H3K9 methylation (Suv39H1) as KMT^R and the writers for H3K4 methylation (SETs and MLL1/2) as KMT^A. Similarly, to distinguish the HDMTs for H3K9 methylation and H3K4 methylation, we define the erasers for H3K9 methylation (JMJD2A) as $HDMT^R$ and the erasers for H3K4 methylation (JARID) as $HDMT^A$. We use colored dotted lines to depict the recruitment process by H3K4me3/ac (green lines), H3K9me3 (red lines), and CpGme (orange lines) and we use dotted black arrows to depict the consequent effect on writing/erasing. The solid black arrow represents the nucleosome modification.

https://doi.org/10.1371/journal.pcbi.1009961.g001

chromatin state and, as such, are found at pluripotency genes and at lineage specific genes in terminally differentiated cells [5].

Specifically, H3K4 and H3K9 methylation are typically mutually exclusive since the enzymes that methylate H3K4 do not do so if the neighboring K9 residues are already methylated [28] and *viceversa* [29]. We therefore assume that each nucleosome carries either H3K9 methylation or H3K4 methylation but not both. We also assume that H3K4 methylation and acetylation co-exist and we represent them by an activating "A" nucleosome modification (Fig 1A). Furthermore, for H3K9 methylation, we lump together the two methylation states (me2 and me3) for simplicity and call them "me3" because both of them are associated with gene repression. Also, although nucleosomes carry two copies of each histone, we consider them as one unit that can carry only one histone modification.

In addition to these histone modifications, DNA itself can be methylated in correspondence to the CpG dinucleotide (CpGme) [5, 10, 30]. DNA methylation is correlated with the absence of H3K4 methylation [10](Chapter 6), since the enzymes that methylate H3K4 recognize unmethylated DNA binding motifs [10](Chapter 1) and *viceversa* [10](Chapter 6), [31]. Therefore, we assume that each nucleosome can carry either one or the other but not both (Fig 1A). By contrast, DNA methylation can co-exist with H3K9 methylation [10](Chapters 6, 22) and hence we allow a nucleosome to carry both modifications (Fig 1A). As a result, each gene will include a number of nucleosomes with different modifications (Fig 1B).

In the next section, we introduce the molecular mechanisms and corresponding reactions by which these modifications can be written, erased, and copied to nearby nucleosomes, which will form the chromatin modification circuit (Fig 1C and 1D).

Histone modifications. Histone modifications can be modeled by enzymatic reactions wherein an unmodified histone can be *de novo* modified by the action of writer enzymes. These enzymes are called KMTs (lysine methyltransferases) for methylation and HATs (histone acetyltransferases) for acetylation and are recruited to DNA by, among others, sequence-specific TFs [32](Chapter 6). These modifications can be actively removed by the action of eraser enzymes. These enzymes are called HDMTs (histone demethylases) for de-methylation and HDACs (histone deacetylases) for de-acetylation. Additionally, histone modifications can be copied through a read-write process, wherein a modified histone is bound by a reader protein, which, in turn, recruits writers of the same modification [10]. We describe these in detail.

***De novo* establishment.** As described in [10](Chapter 21), sequence-specific transcriptional activators bind DNA and recruit HATs, such as the SAGA complex, to the promoter of a gene, which becomes acetylated. Examples of transcriptional activators that recruit HATs include Myc, GATA.1, and Gal4. The deposition of H3K4me3 then can occur co-transcriptionally as RNA Polymerase II recruits SETs, a KMT that methylates H3K4 [33], [10](Chapter 3), or through the recruitment of SETs and MLL1/2 (also KMTs for H3K4) to DNA by the CxxC binding domain that specifically recognizes unmethylated DNA [5](Chapter 7). Finally, MLLs can be recruited to specific promoters by transcriptional activators such as Oct4 [34]. H3K9me3 is established by the writer action of Suv39H1 (a KMT for H3K9), which can also be recruited to DNA by sequence-specific TFs. An example of this is the recruitment of this enzyme to GATA.1 targets by the PU.1 TF, as a means to silence GATA.1 targets and promote the myeloid lineage [35]. These modifications can be captured by the one-step enzymatic reactions ⓪, ①, ② and ③ in Fig 2. In the sequel, we will consider the k_W parameters as inputs to the chromatin modification circuit since they can be modulated by TFs external to the circuit (see Eqs (82), (83) in S1 File).

Auto-catalysis. Histone modifications can recruit more alike modifications to nearby unmodified histones through a read-write mechanism. Specifically, protein readers bind the modification and recruit writers of the same modification to nearby unmodified histones [10] (Chapter 22), [27, 37–39]. For example, protein HP1α contains a chromodomain that specifically recognizes H3K9me3 and recruits the Suv39H1 methyltransferase responsible for H3K9 methylation [37]. Similarly, WDR5 is a protein reader that recognizes specifically H3K4me3 and binds to SETs to recruit them for H3K4 methylation [27]. Similarly, for histone acetylation, the p300/CBP complex is a HAT that also has a bromodomain, so it is a reader and a writer at the same time [10](Chapters 4,21), [5](Chapter 2). We capture this auto-catalytic process by reactions ④ for activating histone modifications and by reactions ⑤ for repressive histone modifications in Fig 2.

Recruited and basal erasure. An activating (repressing) histone modification can recruit erasers for a repressing (activating) histone modification [40]. For example, JMJD2A is an eraser for H3K9me3/2 and de-methylates H3K9me3/2 through its Jumonji domain while

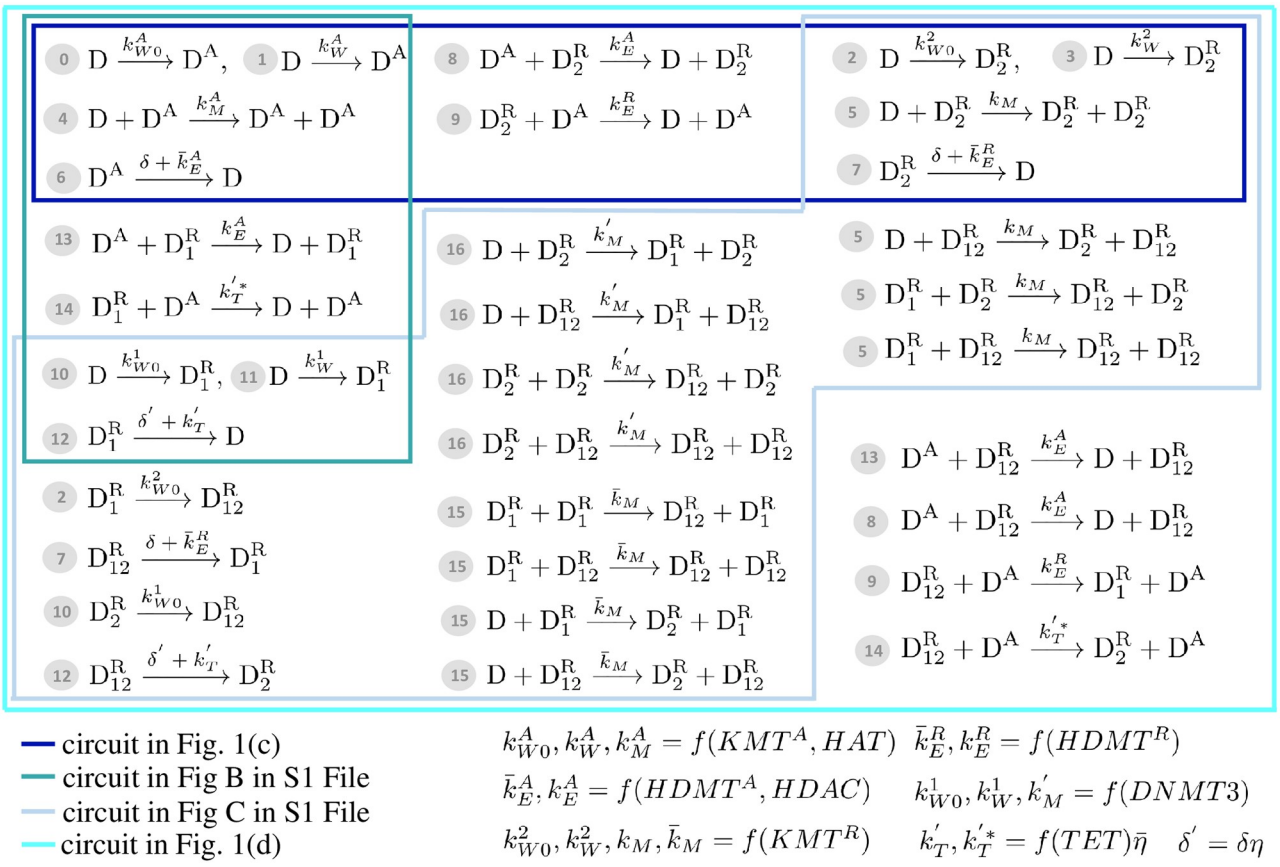

Fig 2. Reactions associated with the chromatin modification circuit motifs. Each reaction is associated with a number, which is referred to in the main text. Specifically, reactions ⓪, ①, ②, ③, ⑩ and ⑪ describe *de novo establishment*. Reactions ④ and ⑤ describe *auto-catalysis*, wherein a modification recruits writers of the same modification to nearby nucleosomes. Reactions ⑮ and ⑯ describe *cross-catalysis*, wherein DNA methylation recruits writers of repressive histone modifications and *viceversa*, respectively. Reactions ⑥, ⑦, and ⑫ represent *basal erasure* while reactions ⑧, ⑨, ⑬, and ⑭ represent *recruited erasure*, wherein competing modifications recruit erasers of each other. The different colored lines delimit the sets of reactions that take place for each of the circuit motifs shown in Fig 1 and Figs B and C in S1 File. Specifically, Fig B depicts the circuit between activating histone modifications and DNA methylation and Fig C depicts the circuit among repressive modifications. Here, we use one-step enzymatic reaction models [36] to capture histone and DNA modifications. In the bottom-right corner, we indicate how the reaction rate constants depend on the amount of writers and erasers, in which $f(\cdot)$ is a generic increasing function. Here, η and $\bar{\eta}$ are between 0 and 1 and defined in Section "Models". Specifically, $\eta = 1$ in the absence of DNMT1 and $\bar{\eta} = 1$ in the absence of MBD (see Section 1 in S1 File for the detailed form of the reaction constants and their derivation).

https://doi.org/10.1371/journal.pcbi.1009961.g002

being able to bind H3K4me3 through the Tudor domain. Therefore, H3K4me3 helps recruit this H3K9me3 eraser to DNA. In turn, JARID is an eraser of H3K4me3 and does so through one of its PHD domains. Through a different PHD domain, it binds H3K9me3. Therefore, H3K9me3 helps recruit this H3K4me3 eraser to DNA. Furthermore, the CHD4 subunit of the NuRD (nucleosome remodeling and de-acetylase) complex contains a domain (a PHD domain) that recognizes H3K9me3. This provides a mechanism through which H3K9me3 recruits HDACs. The corresponding reactions are number ⑧ and ⑨ in Fig 2. Furthermore, in addition to recruited erasure, histone modifications are subject to basal erasure, wherein a modification can be removed passively through dilution due to DNA replication during S phase or through non-specific de-methylation [10](Chapter 22). Basal erasure is captured by reactions ⑥ and ⑦ in Fig 2. The resulting circuit diagram is shown in Fig 1C, wherein each modification auto-catalyzes itself and inhibits the other one.

DNA methylation and crosstalk with histone modifications. Similar to histone modifications, DNA methylation can be captured by enzymatic reactions. Specifically, *de novo* establishment of CpG methylation is mediated by the DNMT3 enzyme (reactions ⑩ and ⑪ in Fig 2) and copy-write occurs through the DNMT1 enzyme, which quickly copies the methylation pattern on the nascent DNA strand at replication. De-methylation occurs passively through dilution due to DNA replication and by an active de-methylation mediated by TET enzymes [10](Chapter 15). In particular, TET enzymes recognize CpGme dinucleotides and have catalytic activity to convert methylated CpG to hydroxilmethylated CpG, then to formylcytosine, and finally to carbolxylcytosine [41, 42]. None of these modified forms are recognized by DNMT1 and therefore they are subject to dilution *via* DNA replication [5](Chapter 17). The ability of TET enzymes to convert methylated CpG to hydorxilmethylated CpG *in vivo* is hampered by the binding of MBD proteins to methylated CpG [43]. Reversely, MBD proteins cannot bind hydroxylmethylated DNA [42]. The combination of these processes can be captured by a basal erasure reaction (reaction ⑫ in Fig 2), where the expressions of the decay rate constants δ' and k_T' are derived in Section 1.3, Eqs (33) and (62) in S1 File. Specifically, we have that $\delta' = \delta\eta$, in which δ is the dilution rate constant, $\eta = 0$ if the efficiency of the maintenance process by DNMT1 is 100%, and $\eta = 1$ if DNMT1 is absent, consistent with earlier models [44]. Similarly, we have that $k_T' = k_T\bar{\eta}$, in which k_T is a constant proportional to the level of TET enzyme, $\bar{\eta} = 0$ if MBD proteins are highly abundant, and $\bar{\eta} = 1$ if MBD proteins are absent.

Recruited erasure between activating histone modifications and DNA methylation. DNA methylation and H3K4 methylation/acetylation have mechanisms to recruit erasers of each other. Specifically, methylated CpGs recruit MeCP2 proteins, which associate with HDACs to establish de-acetylation [10](Chapter 15), [45–47]. Similarly, methylated CpGs recruit MBD2, which interacts with the NuRD complex to also promote de-acetylation [10, 45]. These processes are captured by reaction ⑬ in Fig 2. On the other hand, TET1 enzyme has high propensity to bind to unmethylated CpGs through the CXXC domain [5](Chapter 17), [42]. This suggests a potential mechanism by which methylated H3K4 recruits TET1 to nearby methylated CpGs, enhancing their de-methylation. This is captured by reaction ⑭ in Fig 2, in which the reaction constant $k_T'^*$ has the same trend as k_T' with the abundance of MBD (Eq (76) in S1 File). A diagram representing the mutual inhibition between activating histone modifications and DNA methylation is shown in Fig B in S1 File.

Cross-catalysis between repressive chromatin modifications. DNA methylation and H3K9 methylation reinforce each other. Specifically, MBD proteins recognize methylated CpG dinucleotides and recruit both histone modifying complexes to the methylated sites. MBD1, in particular, binds to methylated CpG sites and recruits histone methyltransferases for H3K9, SETDB1 and Suv39H1, which bring H3K9me3 about [45]. Similarly, MeCP2 binds methylated CpGs and recruits histone methylases that lead to H3K9me3 [48]. We capture the process by which DNA methylation recruits H3K9 methyltransferases by reactions ⑮ in Fig 2. On the other hand, DNMT3 binds to HP1 protein (reader of H3K9me3), suggesting that H3K9me3 recruits DNA methylation enzymes through HP1 [49]. This is captured by reactions ⑯ in Fig 2. A diagram representing these synergistic interactions is shown in Fig C in S1 File. These interactions create positive feedback loops, wherein each modification enhances the creation of the other.

By combining the competitive interactions between activating and repressive modifications (Fig 1C and Fig B in S1 File), with the cross-catalysis among repressive modifications (Fig C in S1 File), we obtain the full chromatin modification circuit of Fig 1D, whose corresponding reactions are listed in Fig 2. Sections 1.1–1.7 in S1 File contain the derivation of the models

and of reaction constants. The expressions of reaction constants are function of the level of eraser enzymes, writer enzymes, methyl-DNA binding proteins, and of cell division rate. Fig 2 summarizes these dependencies in a qualitative way. Section 1.8 in S1 File contains the list of assumptions that were made to write these reactions. In particular, in this model we do not include spatial regulation for simplicity, assuming that the higher-order structure of chromatin allows any modified nucleosome to act on any other nucleosome within the gene, as in previous work [26]. Even if the model does not include spatial regulation, it still includes "beyond-neighbor" interactions that play a role in bistability [26].

Model of transcriptional regulation by chromatin state

Chromatin state affects transcription by modulating nucleosome compaction. DNA wrapped around highly compacted nucleosomes is difficult to access by TFs and therefore genes in such DNA regions are not transcribed. By contrast, DNA wrapped around less compacted nucleosomes is more likely to be accessible by TFs, and hence is susceptible to transcription. Repressive histone modifications and DNA methylation generally lead to increased compaction while activating histone modifications keep a more open chromatin structure [10](Chapter 3).

We therefore assume that transcription is allowed only by D^A, but not by D or any of D_1^R, D_2^R, and D_{12}^R. In principle, nucleosome state D may allow some low basal transcription, through non-specific targeting by chromatin remodelers [32, 50], but we assume this is sufficiently low to be negligible. A more general model could include this basal rate, accounting for the fact that transcription by RNA Pol II of D occurs concurrently with the deposition of H3K4me3 and hence with the conversion of D to D^A [33]. Thus, by lumping together transcription and translation, we obtain the following gene expression model (Section 4 in S1 File):

$$D^A \xrightarrow{\alpha_x} D^A + X, X \xrightarrow{\gamma_x} \emptyset, \quad (1)$$

in which X is the gene product, γ_x is the decay rate constant of X and $\alpha_x = F(A, R)$ is an increasing function of the concentration of TF activators A and a decreasing function of the concentration of TF repressors R. Here, for simplicity, we assume that α_x is a constant. The main conclusions of this paper are not affected by this assumption.

Results

We analyze both the deterministic and stochastic behavior of the chromatin modification circuit model, focusing on the question of temporal duration of memory of transient input stimuli. Therefore, we regard this circuit as a dynamical system that takes k_W^A and k_W^R, modulated by the binding of sequence-specific TFs to DNA, as inputs and gives as output the chromatin state, captured by the number of modified nucleosomes D^A, D_1^R, D_2^R, and D_{12}^R within the gene (Fig 3A and 3B).

We then consider two applications: TF-enabled positive autoregulation for robustness of the active gene state to undesirable repressive input stimuli, such as due to endogenous silencing, and mutually repressing positively autoregulated chromatin modification circuits for long-term persistence of multiple distinct gene expression patterns.

Time scale separation among constituent processes control bistability, hysteresis, and time to memory loss

We first analyze the ordinary differential equation (ODE) model of the chromatin modification circuit, focusing on how parameters affect bistability and hysteresis. Bistability is the coexistence of two stable steady states, specifically active and repressed chromatin states, in the

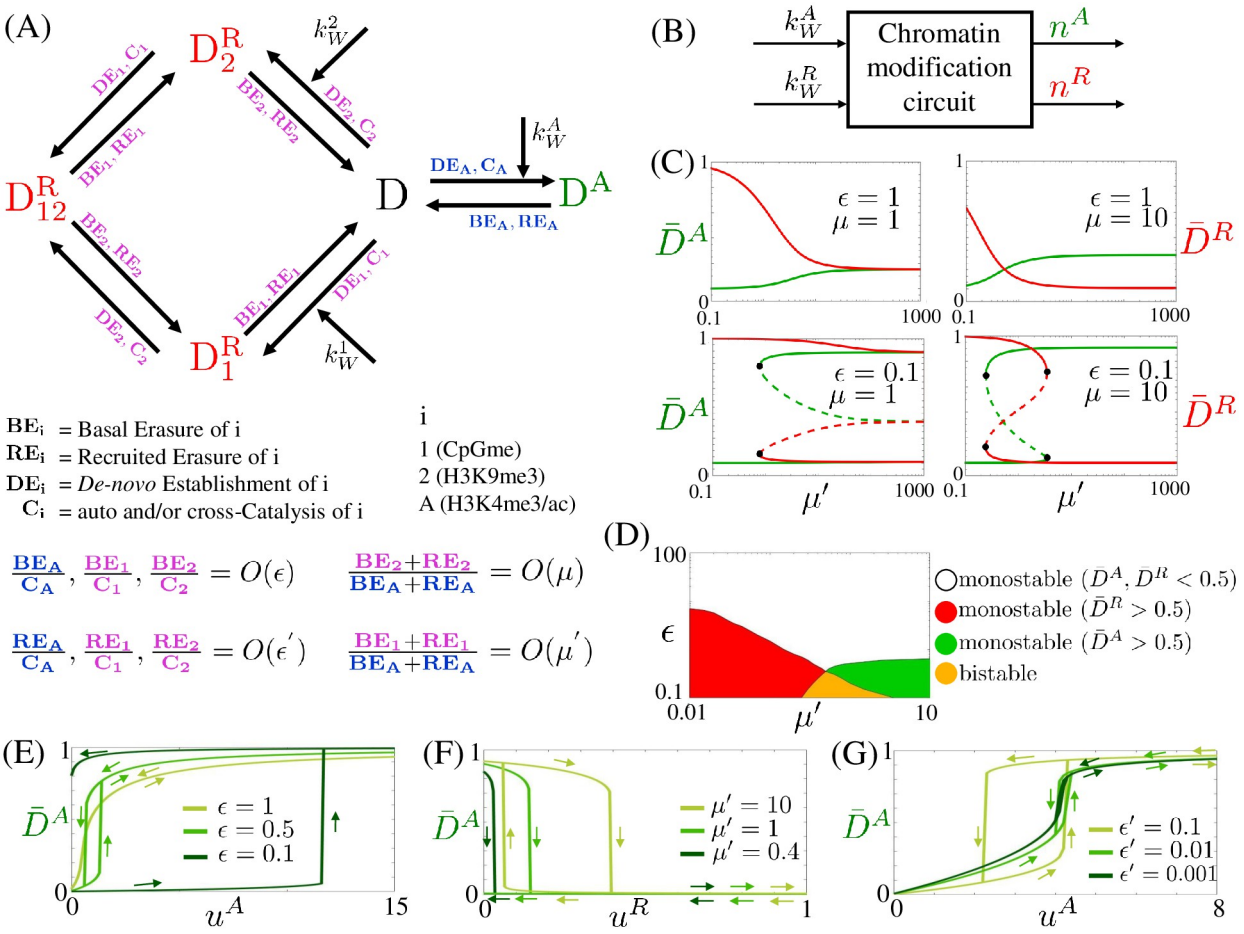

Fig 3. Time scale parameters ϵ, ϵ', μ and μ' control bistability and hysteresis in the chromatin modification circuit. (A) Diagram of the gene's inner chromatin modification circuit in which, compared to Fig 1D, we removed the lines and arrows indicating recruitment and catalysis. The labels on each arrow specify the processes enabling that nucleosome modification as indicated below the reaction diagram. We use purple labels for repressive modifications and blue labels for activating modifications. A visual representation of the relationships between the rates of these processes and the parameters ϵ, ϵ', μ and μ' is provided. In particular, ϵ and ϵ' quantify the time scales of basal and recruited erasure rates of all modifications relative to those of auto and cross-catalysis. Similarly, μ and μ' quantify the time scales of erasure rates (basal and recruited) of repressive histone modifications and DNA methylation, respectively, relative to those of activating histone modifications. For the mathematical definition, refer to Eq (2) and the related text. (B) Block diagram corresponding to the chromatin modification circuit. Here, n^A, n_1^R, n_2^R and n_{12}^R denote the numbers of modified nucleosomes D^A, D_1^R, D_2^R, and D_{12}^R within the gene, $k_W^R = (k_W^1, k_W^2)$ and $n^R = n_1^R + n_2^R + n_{12}^R$. The pair (k_W^R, k_W^A) are the inputs and (n^R, n^A) are the outputs. (C) Steady states of the system as a function of ϵ, μ and μ'. Here, $\bar{D}^A = n^A/D_{tot}$ and $\bar{D}^R = n^R/D_{tot}$ are the fractions of nucleosomes with activating or repressive modifications within the gene with a total of D_{tot} nucleosomes. Plots are obtained from system (3) with $u^A = u_1^R = u_2^R = 0$. The solid lines represent stable steady states, the dashed lines represent unstable steady states and the black circle represents the bifurcation point (saddle-node bifurcation). In these plots $u_0^A = u_{10}^R = u_{20}^R = u_0 = 0.1$ and all the other parameters are set equal to 1 (Fig K in S1 File shows different values). (D) Chart depicting the (ϵ, μ') combinations that result in a monostable (red, green or white) or bistable (yellow) system for $\mu = 10$ (Fig L in S1 File shows different values of μ). Here, $\epsilon = 1$. (E) Input/output steady state characteristics displaying hysteresis for the $(u^A, \bar{D}^A)$ and $(u^R, \bar{D}^R)$ pairs, with $u_1^R = u_2^R = u^R$, for different values of ϵ obtained from simulations of system (3). We consider $(\bar{D}^R, \bar{D}^A) = (1, 0)$ as initial conditions and we set $u^R = 0$, $\alpha = \bar{\alpha} = \alpha' = 1$, $u_0^A = u_{10}^R = u_{20}^R = 0.1$, $\epsilon = 1$, $\mu = 1$ and $\mu' = 0.8$ (Fig M in S1 File shows different values of μ, μ' and ϵ'). (F) Input/output steady state characteristics for the $(u^R, \bar{D}^A)$ pair, for different values of μ' obtained from simulations of system (3). We consider $(\bar{D}^R, \bar{D}^A) = (0, 1)$ as initial conditions and we set $u^A = 0$, $\epsilon = 0.07$, $u_0^A = u_{10}^R = u_{20}^R = 0.1$, $\alpha = \bar{\alpha} = \alpha' = 0.1$, and all the other parameters equal to 1 (Fig OA in S1 File shows the $(u^R, \bar{D}^R)$ steady state characteristics for the same parameter values). (G) Input/output steady state characteristics for the $(u^A, \bar{D}^A)$ pair, for different values of ϵ' obtained from simulations of system (3). We consider $(\bar{D}^R, \bar{D}^A) = (1, 0)$ as initial conditions and we set $u^R = 0$, $\epsilon = 0.15$, $\mu = 1$, $\mu' = 1$, $u_0^A = u_{10}^R = u_{20}^R = 0.1$ and all the other parameters equal to 1 (Fig OB in S1 File shows the $(u^R, \bar{D}^R)$ steady state characteristics for different values of ϵ'). In all plots $u^A = k_W^A/(k_M^A D_{tot})$, $u_i^R = k_W^i/(k_M^A D_{tot})$ for $i \in \{1, 2\}$, $(\bar{D}^A, \bar{D}^R) \approx (1, 0)$ corresponds to the active state and $(\bar{D}^A, \bar{D}^R) \approx (0, 1)$ corresponds to the repressed state. In the figure, we use green and red, respectively, to indicate the activating and repressive modifications and related quantities.

https://doi.org/10.1371/journal.pcbi.1009961.g003

absence of an input stimulus. Hysteresis occurs when the output of the system follows two different paths depending on whether the input is increased or decreased. In particular, the output remains in the proximity of the value reached for high input when this is removed that is, the system "remembers" a transient input stimulus. We then consider a stochastic model to address the question of temporal duration of this memory, given that noise intrinsic to molecular reactions can always displace the system's state.

For our analysis, we let n^A, n_1^R, n_2^R, n_{12}^R and n^D denote the number of nucleosomes with activating (D^A), repressive (D_1^R, D_2^R, and D_{12}^R), and no modifications (D), respectively, within a gene composed of a total of D_{tot} nucleosomes. All analyses are performed considering the unitless normalized time $\tau = tk_M^A D_{tot}$ with $D_{tot} = D_{tot}/\Omega$ and Ω the reaction volume. We define the following non-dimensional parameters, capturing the characteristic time scales of the circuit's constituent processes:

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

Here, α is the non-dimensional auto-catalysis rate parameter, while $\bar{\alpha}$, and α' are the non-dimensional cross-catalysis rate parameters, which we assume on the same order without loss of generality. We further let $(\delta + \bar{k}_E^R)/(\delta + \bar{k}_E^A) = b\mu$ with $b = O(1)$, implying that μ scales the ratio between the erasure rates of repressive and activating histone modifications. We similarly let $(\delta' + k_T')/(\delta + \bar{k}_E^A) = \beta\mu'$ with $\beta = O(1)$, implying that μ' scales the ratio between the erasure rates of DNA methylation and those of activating histone modifications. From these relationships, it also follows that $(\delta + \bar{k}_E^R)/(k_M^A D_{tot}) = b\epsilon\mu$ and that $(\delta' + k_T')/(k_M^A D_{tot}) = \beta\epsilon\mu'$. Therefore, the parameter ϵ scales the ratio between the basal rate at which all modifications are erased and the rate at which they are auto or cross-catalyzed. Finally, since $k_E^R/k_M^A = \mu\epsilon'$ and $k_T^*/k_M^A = \mu'\epsilon'$, it follows that ϵ' scales the ratio between the rate at which modifications are removed through mutual recruitment of erasers and that at which they are auto or cross-catalyzed. These parameter definitions are pictorially conveyed in Fig 3A.

Putting these definitions together, in our model, ϵ and ϵ' are non-dimensional parameters that quantify the time scales of basal and recruited erasure rates of all modifications relative to those of auto and cross-catalysis. Similarly, μ and μ' are non-dimensional parameters that quantify the time scales of erasure rates (basal and recruited) of repressive histone modifications and DNA methylation, respectively, relative to those of activating histone modifications. Since these normalized parameters are functions of the biochemical reaction rate constants (Eq (2)) and, in turn, the reaction rate constants are known functions of the level of writer and eraser enzymes, TET enzymes, and MBD proteins (see Fig 2 and Section 1.6 in S1 File), it is possible to experimentally vary these parameters by over-expressing, inhibiting, or by recruiting these molecular players.

Deterministic behavior

We first analyze the ordinary differential equation (ODE) model of the chromatin modification circuit, focusing on how the above parameters affect dynamics. To this end, we assume that the number of total modifiable units D_{tot} is sufficiently large such that the variables n^A, n_1^R, n_2^R, n_{12}^R and n^D can be considered real-valued and thus their temporal evolution can be described by ODEs. In particular, we describe the system in terms of fractions of modifiable units, that is $\bar{D}^A = n^A/D_{tot}$, $\bar{D}_1^R = n_1^R/D_{tot}$, $\bar{D}_2^R = n_2^R/D_{tot}$, $\bar{D}_{12}^R = n_{12}^R/D_{tot}$ and $\bar{D} = n^D/D_{tot}$. We further introduce the normalized inputs $\bar{u}^A = u_0^A + u^A$ with $u_0^A = k_{W0}^A/(k_M^A D_{tot})$ and $u^A = k_W^A/(k_M^A D_{tot})$, $\bar{u}_1^R = u_{10}^R + u_1^R$ with $u_{10}^R = k_{W0}^1/k_M^A D_{tot}$ and $u_1^R = k_W^1/(k_M^A D_{tot})$, $\bar{u}_2^R =$

$u^R_{20} + u^R_2$ with $u^R_{20} = k^2_{W0}/(k^A_M D_{tot})$ and $u^R_2 = k^2_W/(k^A_M D_{tot})$. With these definitions and letting $\dot{x} := dx/d\tau$, the ODEs describing the system in non-dimensional variables are given by

$$
\begin{aligned}
\dot{\bar{D}}^R_1 &= \overbrace{(\bar{u}^R_1 + \alpha'(\bar{D}^R_2 + \bar{D}^R_{12}))\bar{D}}^{\textcircled{10}+\textcircled{11}+\textcircled{16}} + \overbrace{\mu(b\epsilon + \epsilon'\bar{D}^A)\bar{D}^R_{12}}^{\textcircled{7}+\textcircled{9}} \\
&\quad \overbrace{-(u^R_{20} + \alpha(\bar{D}^R_2 + \bar{D}^R_{12}) + \bar{\alpha}(\bar{D}^R_1 + \bar{D}^R_{12}))\bar{D}^R_1}^{\textcircled{2}+\textcircled{5}+\textcircled{15}} \overbrace{-\mu'(\beta\epsilon + \epsilon'\bar{D}^A))\bar{D}^R_1}^{\textcircled{12}+\textcircled{14}} \\
\dot{\bar{D}}^R_2 &= \overbrace{(\bar{u}^R_2 + \alpha(\bar{D}^R_2 + \bar{D}^R_{12}) + \bar{\alpha}(\bar{D}^R_1 + \bar{D}^R_{12}))\bar{D}}^{\textcircled{2}+\textcircled{3}+\textcircled{5}+\textcircled{15}} + \overbrace{\mu'(\beta\epsilon + \epsilon'\bar{D}^A)\bar{D}^R_{12}}^{\textcircled{12}+\textcircled{14}} \\
&\quad \overbrace{-(u^R_{10} + \alpha'(\bar{D}^R_2 + \bar{D}^R_{12}))\bar{D}^R_2}^{\textcircled{10}+\textcircled{16}} \overbrace{-\mu(b\epsilon + \epsilon'\bar{D}^A))\bar{D}^R_2}^{\textcircled{7}+\textcircled{9}} \\
\dot{\bar{D}}^R_{12} &= \overbrace{(u^R_{10} + \alpha'(\bar{D}^R_2 + \bar{D}^R_{12}))\bar{D}^R_2}^{\textcircled{10}+\textcircled{16}} + \overbrace{(u^R_{20} + \alpha(\bar{D}^R_2 + \bar{D}^R_{12}) + \bar{\alpha}(\bar{D}^R_1 + \bar{D}^R_{12}))\bar{D}^R_1}^{\textcircled{2}+\textcircled{5}+\textcircled{15}} \\
&\quad \overbrace{-(\mu'(\beta\epsilon + \epsilon'\bar{D}^A) + \mu(b\epsilon + \epsilon'\bar{D}^A))\bar{D}^R_{12}}^{\textcircled{7}+\textcircled{9}+\textcircled{12}+\textcircled{14}} \\
\dot{\bar{D}}^A &= \overbrace{(\bar{u}^A + \bar{D}^A)\bar{D}}^{\textcircled{0}+\textcircled{1}+\textcircled{4}} \quad \overbrace{-(\epsilon + \epsilon'(\bar{D}^R_2 + \bar{D}^R_{12}) + \epsilon'(\bar{D}^R_1 + \bar{D}^R_{12}))\bar{D}^A}^{\textcircled{6}+\textcircled{8}+\textcircled{13}},
\end{aligned}
\tag{3}
$$

in which $\bar{D} = 1 - \bar{D}^R_1 - \bar{D}^R_2 - \bar{D}^R_{12} - \bar{D}^A$ and the circled numbers indicate the corresponding reactions from Fig 2 (see Section 1.6 and Eqs (80) in S1 File for derivation).

In the absence of inputs ($u^A = u^R_1 = u^R_2 = 0$), system (3) has a unique stable steady state, if ϵ is not sufficiently small. This steady state has $\bar{D}^R > \bar{D}^A$ for small μ'. When ϵ is sufficiently small, the system is bistable, with both active ($\bar{D}^A \approx 1$) and repressed ($\bar{D}^R \approx 1$) co-existing stable chromatin states. But when μ' is very small, even for small ϵ, the system returns monostable with a unique stable steady state corresponding to the repressed state (Fig 3C and 3D and Fig K in S1 File). Furthermore, as μ' increases (DNA methylation erasure rate increases), the system approximates well the histone modification system (Fig 1C), wherein DNA methylation is not present. This is because when μ' is sufficiently large, the DNA methylation is erased very quickly and then $\bar{D}^R = \bar{D}^R_1 + \bar{D}^R_2 + \bar{D}^R_{12} \approx \bar{D}^R_2$. As a consequence, when ϵ is sufficiently small the system is bistable for intermediate values of μ, monostable with $\bar{D}^A \approx 1$ when μ is large and monostable with $\bar{D}^R \approx 1$ when μ is small (Fig D in S1 File). Varying ϵ' does not change these trends, except for $\epsilon' \ll \epsilon$, in which case the system returns monostable (Fig K in S1 File).

Fig 3E shows the input/output steady state characteristics of the chromatin modification circuit. Specifically, as u^A is increased starting from $\bar{D}^A = 0$, the state switches to $\bar{D}^A \approx 1$ for u^A above a critical value, which increases as ϵ decreases. When u^A is decreased back to zero, the path that $\bar{D}^A$ follows is different and it can keep $\bar{D}^A \approx 1$ even when $u^A = 0$ if ϵ is sufficiently small. In this case, we say that the state transition becomes irreversible and *the system keeps memory* of its input stimulus. From a biological point of view, this implies that an activating

transient input stimulus allows to permanently re-activate a silenced gene. This memory vanishes as ϵ increases. When μ' decreases, the value of u^A required to flip the system's state to fully active increases until we reach a regime under which the memory is lost (Fig M in S1 File). A similar behavior is observed for the relationship between the repressing input u^R (with $u_1^R = u_2^R = u^R$) and $\bar{D}^R$ (Fig N in S1 File). Biologically, this implies that a repressive transient input stimulus can permanently silence an initially active gene. However, as opposed to the $(u^A, \bar{D}^A)$ characteristics, if we reduce μ', the minimum value of u^R required to silence the system ($\bar{D}^R \approx 1$, $\bar{D}^A \approx 0$) decreases (Fig 3F and Fig N in S1 File). This also shows that as μ' is decreased, the active state becomes less robust to repressive input stimuli.

A comparison of the plots of Figs M and N in S1 File for small ϵ and $\mu = 1$ reveals that, even when $\mu' = 1$, it takes smaller values of u^R to switch the state to $\bar{D}^R = 1$ than the values of u^A required to switch the state to $\bar{D}^A = 1$. This fact highlights that it is the structure itself of the chromatin modification circuit, rather than the rate constants alone, to "bias" the system towards a repressed chromatin state. In fact, the two-layer topology of the repressive modifications confers increased robustness to the repressed state with respect to activating stimuli k_W^A (Fig 3A). This is in contrast to what is observed for the system that does not include DNA methylation (Fig 1C), wherein once all parameters are equal to each other, the stability landscape is symmetric (Section 2.1 in S1 File). Varying ϵ' does not significantly affect these trends (Figs M and N in S1 File), unless $\epsilon' \ll \epsilon$, in which case hysteresis is lost (Fig 3G).

In summary, when recruited erasure (ϵ') and auto-catalysis/cross-catalysis ($\alpha, \bar{\alpha}, \alpha'$) are sufficiently faster than basal erasure (ϵ), the circuit displays bistability and hysteresis, except if μ and/or μ' are sufficiently small. In this case, the unique stable state is the fully repressed chromatin state, independent of all other parameters.

Stochastic behavior

We next evaluate how the stationary distribution of the chromatin state, the time to memory loss, and the ability to reactivate a repressed gene depend on ϵ, ϵ', μ and μ', and on the inputs u^A and u^R. To this end, we first perform a computational study by using Gillespie's Stochastic Simulation Algorithm (SSA) [51]. We then complement this with a mathematical derivation of the approximate expressions of both the stationary distribution and the time to memory loss as a function of the parameters and input stimuli.

Fig 4A and 4B shows how the stationary distribution is shaped by ϵ and μ' (A) and by inputs u^R and u^A (B). Specifically, ϵ has a most critical role. As it decreases, with all other parameters remaining fixed, the peaks of the distribution become more concentrated about the fully repressed and fully active chromatin states. Highly concentrated peaks imply that the probability of finding the chromatin state outside of these configurations is very small. This property is connected to the *time to memory loss*, defined as the expected value of the earliest time the chromatin state reaches the active state starting from the repressed state and *viceversa*. In fact, as ϵ is decreased, the chromatin state starting in either the repressed or active state takes longer, on average, to reach the active or repressed state for the first time, respectively (Fig 4C and 4D). In contrast to what found for the histone modification model of Fig 1C (Fig FA in S1 File), decreasing ϵ also has an effect in increasing the height of the peak in correspondence of the repressed state to the detriment of the peak corresponding to the active state (Fig 4A). When μ' is decreased, this bias increases, and the peak in correspondence of the repressed state becomes more concentrated (Fig 4A). Varying μ has a similar effect as varying μ' (Fig S in S1 File). Inputs u^R and u^A each increase the height of the peak corresponding to the repressed and active state, respectively (Fig 4B). Consistent with the deterministic analysis (Fig M in S1 File), for smaller μ', it is required a larger u^A to increase the probability of finding the state in

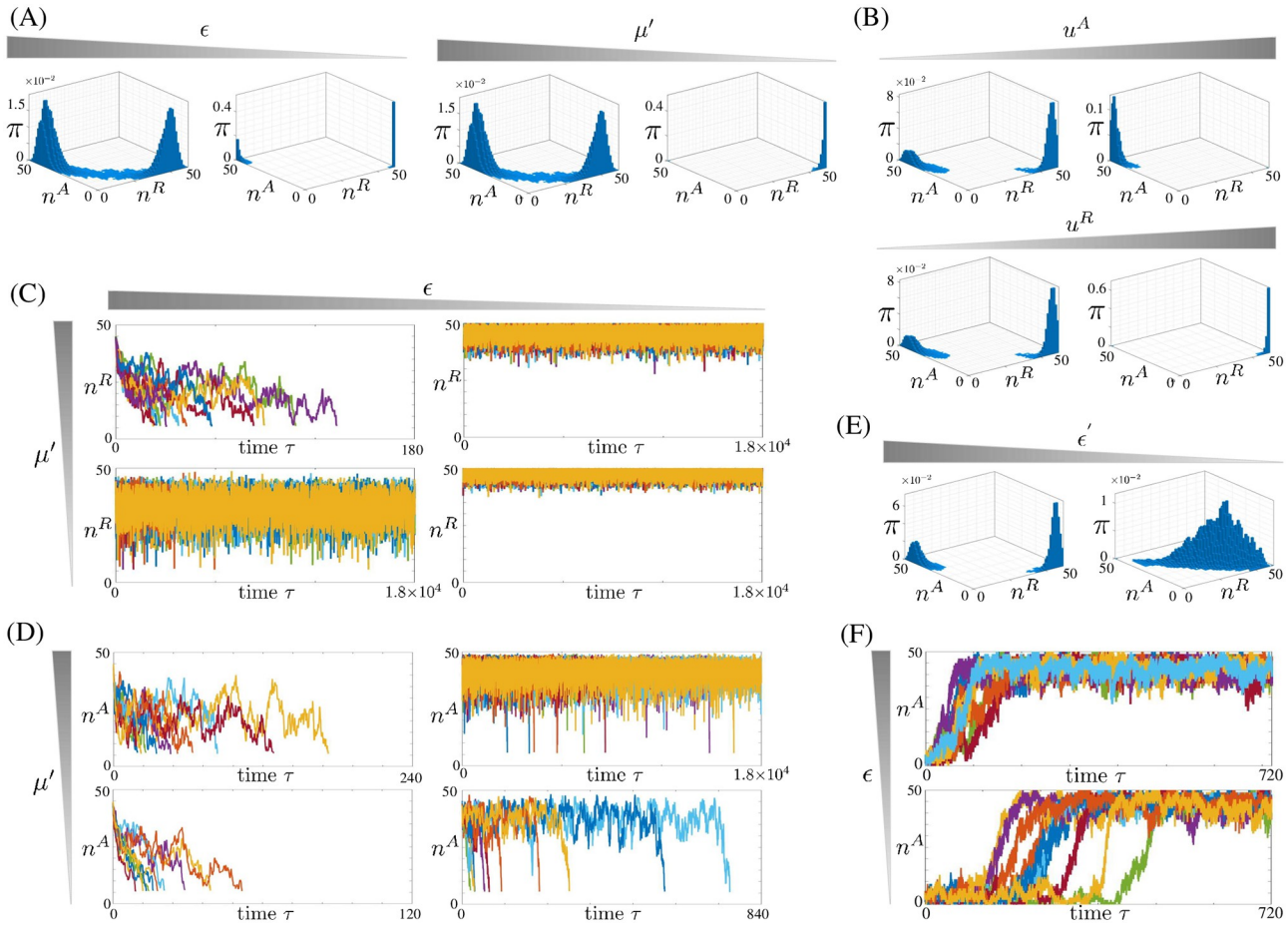

Fig 4. Time scale separation parameters ϵ, ϵ' and μ' control bimodality, time to memory loss, and reactivation time in the chromatin modification circuit. (A)-(B) Stationary probability distributions π of (n^A, n^R) for the chromatin modification circuit in Fig 3A obtained by simulating the reactions listed in Fig 2 with the SSA. (A) Effect of ϵ and μ' on the distribution: in the left side plots $\epsilon = 0.19$, 0.02 and $\mu' = 1$. In the right side plots, $\mu' = 1$, 0.1 and $\epsilon = 0.19$. (B) Effect of the input stimulus on the distribution: here, $\epsilon = 0.12$ and $\mu' = 0.8$. In the top plots, $u^A = 0$, 1, and in the bottom plots $u_1^R = u_2^R = u^R$ and $u^R = 0$, 1 (Fig R shows different values of μ' and ϵ). The parameter values for each panel are listed in Table A in S1 File. In all simulations, $\epsilon = 1$ and $\mu = 1$ (Figs S and T in S1 File show different values) and we decrease ϵ by decreasing $\delta + \bar{k}_E^A$ (similar results can be obtained if we vary k_M^A as shown in Fig Q in S1 File). (C) Time trajectories of n^R starting from a repressed chromatin state $n^A = 5$, $n_1^R = n_2^R = n_{12}^R = 15$, as ϵ and μ' are varied. Simulations are stopped when $n^R = 6$ for the first time. (D) Time trajectories of n^A starting from an active chromatin state $n^A = 45$, $n_{12}^R = 5$, as ϵ and μ' are varied. Simulations are stopped when $n^A = 6$ for the first time. In all plots of (C)-(D), time is normalized according to $\tau = t\frac{k_M^A}{\Omega}\mathrm{D}_{\mathrm{tot}}$, $\epsilon = 0.36$, 0.12, $\mu' = 1$, 0.5, $\mu = 1$, and $\epsilon' = 1$ (Fig V in S1 File shows different values of ϵ'). The parameter values are listed in Table B in S1 File. (E) Effect of ϵ' on the stationary probability distribution π. We set $\epsilon = 0.12$, $\mu' = 1$, $\mu = 1$ and $\epsilon' = 1$, 0.001 from left to right (Fig U in S1 File shows different values of ϵ'). The parameter values for each panel are listed in Table P in S1 File. (F) Time trajectories of the system starting from $n^R = 45$, $n^A = 5$ and applying an input u^A that, at steady state, leads to a unimodal distribution in the proximity of the active state $n^A = \mathrm{D}_{\mathrm{tot}}$. The parameter values are listed in Table C in S1 File. In particular, $u^A = 3.2$, $\mu' = 0.1$, $\epsilon = 0.24$, 0.16, $\mu = 1$ and $\epsilon' = 1$ (Fig W in S1 File shows different parameter values). In all plots, $n^R = n_1^R + n_2^R + n_{12}^R$ and simulations are obtained by implementing the set of reactions listed in Fig 2 with the SSA [51]. In all simulations, $\alpha = \alpha' = \bar{\alpha} = 0.2$. In our model, parameters ϵ and ϵ' quantify the time scales of basal and recruited erasure rates of all modifications relative to those of auto and cross-catalysis. Similarly, parameters μ and μ' quantify the time scales of erasure rates (basal and recruited) of repressive histone modifications and DNA methylation, respectively, relative to those of activating histone modifications. Mathematical definitions are found in Eq (2). In all plots $(n^A, n^R) \approx (50, 0)$ corresponds to the active state and $(n^A, n^R) \approx (0, 50)$ corresponds to the repressed state. In each panel of (C), (D), and (F), the number of trajectories plotted is 10.

https://doi.org/10.1371/journal.pcbi.1009961.g004

the fully active chromatin state (Fig RB in S1 File). By contrast, ϵ' does not significantly affect the trends observed for ϵ, μ', μ and the inputs, although it modulates some aspects of the distribution (Section 3.5 and Figs Q and T in S1 File). Specifically, when ϵ' decreases compared to ϵ, the peaks of the distribution become less concentrated and, in the extreme case where $\epsilon' \ll \epsilon$, the distribution becomes unimodal (Fig 4E).

In the regime where both $\epsilon' \ll 1$ and $\epsilon \ll 1$, in which the system displays a bimodal distribution (Figs Q-T in S1 File), we can analytically derive a one-dimensional Markov chain approximation that allows analytical computation of the stationary distribution and of the time to memory loss as functions of the parameters (Sections 3.2–3.4 in S1 File). Biophysically, in this regime, the reactions represented by the labels BE_i and RE_i in Fig 3A become slower compared to those represented by the label C_i. In particular, letting $x = n_{12}^R$ with $x \in [0, D_{tot}]$, we have that, when $\epsilon \ll 1$, the stationary probability distribution $\pi(x)$ can be approximated by

$$\pi_{\epsilon \ll 1}(x) \approx \begin{cases} \frac{1}{1+P} & \text{if } x = 0 \\ 0 & \text{if } x \neq 0, D_{tot}\,, \\ \frac{P}{1+P} & \text{if } x = D_{tot} \end{cases} \tag{4}$$

with

$$\begin{aligned} P_1 &= \frac{(\bar{u}^A + \bar{u}_{12}^R + \alpha + \bar{\alpha} + \alpha')}{(\bar{u}^A + \bar{u}_{12}^R + 1)}, \\ P_2 &= \prod_{i=1}^{D_{tot}-1} \left(\frac{2\left(\bar{u}_{12}^R + (\alpha + \bar{\alpha} + \alpha')\frac{i}{D_{tot}}\right)}{\mu\mu'\epsilon'^{\frac{(D_{tot}-i)}{D_{tot}}} \bar{K}_i\left(\bar{u}^A + \frac{(D_{tot}-i)}{D_{tot}}\right)} \right) \cdot \frac{\bar{u}_{12}^R}{\mu\mu' b\beta\epsilon \bar{K}_{D_{tot}} \bar{u}^A}, \\ P &= P_1 \cdot P_2, \end{aligned} \tag{5}$$

in which $\bar{u}^A = u_0^A + u^A$, $\bar{u}_{12}^R = u_{10}^R + u_1^R + u_{20}^R + u_2^R$ and $\bar{K}_i$, defined in Eqs (161) and (171) in S1 File, independent of ϵ, μ, μ' and ϵ' (see Section 3.3 in S1 File for the mathematical derivations). From this expression, we conclude that if $\epsilon \ll 1$, $\pi(x) \approx 0$ for all x except for $x = D_{tot}$ (fully repressed state) and $x = 0$ (fully active state), implying that the probability of finding the system in one of the intermediate states is about zero. This is in accordance with the computational results of Fig 4A, which show higher concentration around the extreme states, and with Fig 4C and 4D, which show smaller likelihood of transitioning out of those extreme states as ϵ is decreased. When ϵ goes to zero, $P \to \infty$, which implies that $\pi(D_{tot}) \approx 1$. This is consistent with the structural bias of the chromatin modification circuit towards the repressed chromatin state (Fig 3A) and the computational results of Fig RA in S1 File. This bias is enhanced as $\mu\mu'$ is decreased since $P \to \infty$ as $\mu\mu' \to 0$, leading to $\pi(D_{tot}) \approx 1$. Furthermore, in agreement with what observed in Fig 4B, an increase of u^A decreases P and hence increases $\pi(0)$. By contrast, an increase of u^R increases P, which, in turn, increases $\pi(D_{tot})$. Also, larger u^A values are required to increase the probability of the active state to the same level when μ' is smaller.

Similarly, the normalized time to memory loss of the repressed state can be approximated as (Section 3.4 in S1 File)

$$\bar{\tau}_R \approx \frac{K_R}{\mu\mu'\epsilon^2} \left(1 + \sum_{i=1}^{D_{tot}-1} \frac{K_R^i}{h_1^i(\mu\mu')} \right), \tag{6}$$

with h_1^i an increasing function, $h_1^i(0) = 0$, and K_R and K_R^i functions independent of ϵ, μ' and μ, while the normalized time to memory loss of the active state can be approximated as

$$\bar{\tau}_A \approx \frac{K_A}{\epsilon} \left(1 + \sum_{i=1}^{D_{tot}-1} \frac{h_2^i(\mu\mu')}{K_A^i} \right), \tag{7}$$

with h_2^i an increasing function and $h_2^i(0) = 0$, and K_A and K_A^i functions independent of ϵ, μ'

and μ. In the limiting condition where $\epsilon \to 0$, we have that both $\bar{\tau}_R \to \infty$ and $\bar{\tau}_A \to \infty$. Therefore, decreased ϵ is the driver for increasing the extent of epigenetic memory of both the active and repressed chromatin states. Furthermore, because of the structural asymmetry of the chromatin modification circuit, the power of ϵ is different in the two above expressions, indicating that decreased ϵ increases the extent of memory much more for the repressed state than for the active state. This is in contrast with what obtained for the histone modification circuit, in which both $\bar{\tau}_R$ and $\bar{\tau}_A$ are $O(1/\epsilon)$ (Eqs (138) and (140) in S1 File). Finally, decreased μ and μ' lead to larger $\bar{\tau}_R$ but to lower $\bar{\tau}_A$. These trends are in agreement with the simulations in Fig 4C and 4D. In particular, in each panel, we plotted a number of trajectories equal to 10 to provide a comprehensive representation of the parameter effects on the trajectories and, at the same time, a clear reading of the figure.

We next determine *via* simulation how the parameters ϵ, μ' and ϵ' affect the time to reactivation, i.e., the time it takes to re-activate an initially repressed chromatin state by applying a sufficiently large activating input stimulus u^A (Fig 4F). The time trajectories show a switch-like behavior, which is more prominent for $\epsilon \ll 1$, $\mu' \ll 1$, and $\epsilon' > 1$ (Fig W in S1 File). Also, the latency, i.e., the time the trajectory takes to switch after an activating input is applied, is highly variable and its variability is mostly dictated by small ϵ and small μ', which, together, make gene reactivation a low probability event. Furthermore, the time to reactivation decreases by increasing μ' and/or increasing ϵ, consistent with results from the deterministic model (Fig X in S1 File).

Taken together, our results indicate that basal erasure rate sufficiently smaller than recruited erasure rate (ϵ sufficiently small and ϵ' sufficiently larger than ϵ) allows a bimodal distribution such that both active and repressed chromatin states each have high probabilities compared to intermediate states. As ϵ decreases, this distribution is more concentrated about the fully active and repressed states and the time to memory loss of either state increases. Even when there is no asymmetry between activating and repressive modification erasure rates ($\mu = 1$ and $\mu' = 1$), the time to memory loss is larger for the repressed chromatin state. This asymmetry is accentuated by $\mu' < 1$.

Relationship with published data

In [52], the authors study the kinetics of silencing and reactivation of chromatin modifications, including H3K9me3 and DNA methylation, at the single-cell level. In particular, they show that silencing and reactivation has a all-or-none behavior. Our model predicts that this all-or-none response is observed when the underlying chromatin modification circuit has a bimodal stationary distribution, which we proved requires $\epsilon \ll 1$ and ϵ' sufficiently larger than ϵ (Fig W in S1 File). This suggests that in practice we should expect $\epsilon \ll 1$ and ϵ' not too small. Furthermore, data on temporal dynamics of DNA de-methylation *in vivo* indicate that $\mu' \ll 1$ (Section 3.1 in S1 File). Early elegant experimental studies performed by [53] further show that the latency of reactivation of a silenced gene is highly variable. In our model, a highly variable latency of reactivation events is predicted under the parameter regime where both ϵ and μ' are small (Fig 4F and Fig W in S1 File), suggesting that $\mu' \ll 1$ may in fact be responsible for the experimentally observed latency of reactivation [53]. Furthermore, in [53] it is shown that increased proliferation rate leads to a faster reactivation kinetics. In our model, increased proliferation is captured by increased ϵ and in fact higher ϵ leads to faster kinetics of reactivation events (Fig YA-YC in S1 File), in agreement with the observations in [53].

Finally, a basic DNA methylation model that does not include MBD proteins cannot reflect the low effective active erasure rates of DNA methylation through the action of TET enzymes, k_T, encountered *in vivo* (reactions (45) in S1 File). In our model, the ability of TET enzymes to

convert methylated DNA to hydorxilmethylated DNA is hampered by the binding of MBD proteins to methylated DNA (Section 1.3 in S1 File). As a result, the effective active erasure rate constant, $k_T^{'}$, derived in Eq (62) in S1 File, can be written as $k_T^{'} = k_T\bar{\eta}$ in which k_T is a constant proportional to the level of TET enzyme, where $\bar{\eta} = 0$ if MBD proteins are highly abundant, and $\bar{\eta} = 1$ if MBD proteins are absent (Eq (62) in S1 File). This implies that the parameter $\mu^{'} = k_T^{*}/k_E^A$ is proportional to the amount of TET enzymes and decreases with the amount of MBD proteins. Lower levels of MBD proteins result in higher μ' and in our model this leads to faster reactivation kinetics and decreased variability of reactivation time (Fig YD-YF in S1 File). This is consistent with the experimental observations of [54], which show that knocking down of MBD proteins not only makes the reprogramming process faster, with almost the entire cell population converted into iPS cells within 7 days, but makes it also almost deterministic.

Positive TF-enabled autoregulation extends memory of active chromatin state robustly to perturbations

The study of the chromatin modification circuit reported in the previous section shows that the two-layer topology of the repressive modifications biases the system towards a repressed chromatin state, and that this bias becomes even more pronounced for low values of μ'. It follows that, when the chromatin state is in the active configuration, without any activating input stimulus, even a small perturbation can cause a quick silencing of the gene. For an engineered gene expression system, this silencing will disrupt the intended engineered function of the cell line, as observed in experiments with different cell types [16–19]. Here, we study how TF-enabled positive autoregulation can alleviate this problem by restoring memory of the active gene state and by making it more robust to endogenous silencing. A possible form of positive autoregulation is when the gene product X acts as a TF activator that recruits writers of activating histone modifications to the gene. This is the case of many TFs involved in cell fate determination, such as Oct4 and Nanog [55]. The analysis of this design thus not only provides a potential engineering solution to the silencing problem of chromosomally integrated gene expression systems, but it also allows for a different understanding of the role of positive autoregulation in natural networks.

In this case, the input k_W^A of the chromatin modification circuit is a monotonically increasing function of n^X (Section 1.7 in S1 File). The positive autoregulation model is thus given by combining the chromatin modification circuit reactions (Fig 2) with the gene expression reactions (1), and the expression for k_W^A (Eq (82) in S1 File). The corresponding interaction diagram is shown in Fig 5A and 5B shows the block diagram representation with input k_W^R and output n^X.

Deterministic behavior

Considering again the normalized time $\tau = tk_M^A D_{tot}$, and letting $\bar{X} := n^X/\mathrm{D_{tot}}$, with n^X the number of molecules of X, $\bar{\alpha}_x = \alpha_x/(k_M^A D_{tot})$, and $\bar{\gamma}_x = \gamma_x/(k_M^A D_{tot})$, $\dot{x} := dx/d\tau$, the ODE associated with gene expression in non-dimensional variables can be written as

$$\dot{\bar{X}} \quad = \bar{\alpha}_x \bar{D}^A - \bar{\gamma}_x \bar{X}. \tag{8}$$

By assuming molecular counts sufficiently high to use ODEs and by combining the ODEs of the chromatin modification circuit (3) with those of gene expression (8), we obtain the ODE model of the positive autoregulation system (Eqs (1) in S2 File). We first realize bifurcation plots with $p_x = \alpha_x/\gamma_x$ the bifurcation parameter (Fig A in S2 File). As p_x increases, the system transitions from monostable with one stable equilibrium point corresponding to the fully

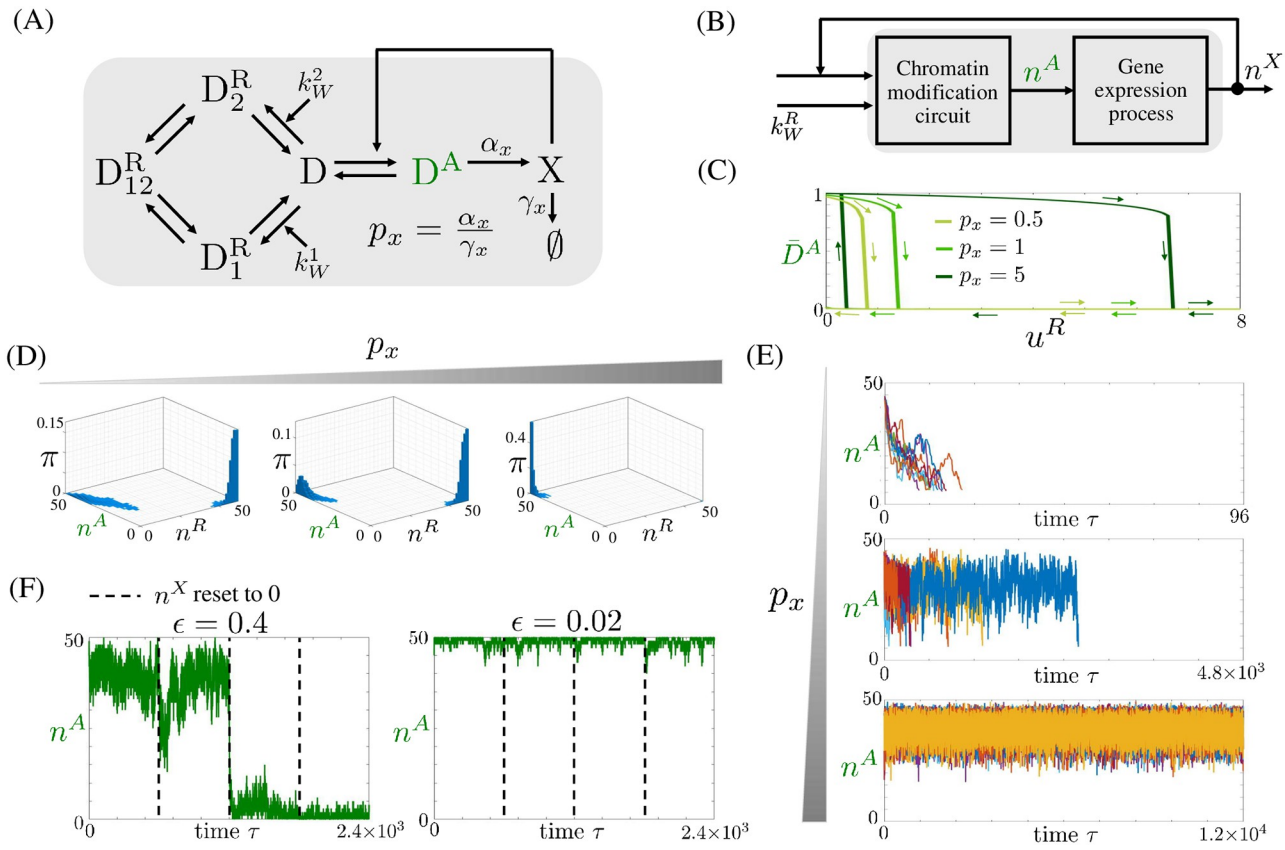

Fig 5. Positive autoregulation allows robust temporal extension of memory of the active chromatin state. (A) Diagram of a positively autoregulated gene where the gene's product X recruits writers of D^A. A simplified representation of the chromatin modification circuit is introduced, in which the the labels are not represented. (B) Block diagram corresponding to the circuit in panel (A). Here, $k_W^R = (k_W^1, k_W^2)$. Furthermore, k_W^R is the input and n^X, the number of molecules of X, is the output, which feeds back on k_W^A by increasing its value. (C) Input/output steady state characteristics for the $(u^R, \bar{D}^A)$ pair, with $u_1^R = u_2^R = u^R$, for different values of p_x obtained from simulations of system Eqs (1) in S2 File with $(\bar{D}^R, \bar{D}^A, \bar{X}) = (0, 1, p_x)$ as initial conditions. The parameter values are listed in Table A in S2 File. In particular, $\epsilon' = 1$, $\mu = 1$, $\mu' = 0.7$, $\epsilon = 0.1$. (D) Stationary probability distribution π obtained by simulating the reactions listed in Table B in S2 File with the SSA. As before, $n^R = n_1^R + n_2^R + n_{12}^R$. The parameter values of each plot are listed in Table B in S2 File. In particular, $p_x = 0$, 0.1, 10, $\epsilon' = 1$, $\mu = 1$, $\mu' = 0.5$, $\epsilon = 0.12$ (Figs B and C in S2 File show different parameter values). (E) Time trajectories obtained by simulating the reactions listed in Table C in S2 File with the SSA with no inputs and starting with initial conditions $n^A = 45$, $n_{12}^R = 5$ and $n^X = p_x\, n^A$. Simulations are stopped the first time at which $n^A = 6$. In all plots, time is normalized according to $\tau = t\frac{k_M^A}{\Omega} D_{tot}$, $\epsilon = 0.36$, $\mu' = 0.5$, $\mu = 1$, $\epsilon' = 1$, and $p_x = 0$, 0.2, 5. The parameter values of each panel are listed in Table C in S2 File. In each panel, the number of trajectories plotted is 10. (F) Time trajectories of system of reactions listed in Table D in S2 File with the SSA starting from the active chromatin state ($n^A = 45$, $n^D = 5$ and $n^X = 45$). At the indicated times, n^X is reset artificially to zero (dashed lines). The parameter values for each simulation are listed in Table D in S2 File. In particular, $p_x = 1$, $\mu' = 0.1$, $\mu = 1$, $\epsilon' = 1$ and $\epsilon = 0.4$, 0.02. For both values of ϵ, the system is bistable. In our model, ϵ, defined in Eq (2), is a non-dimensional parameter that quantifies the time scales of basal erasure rate of all modifications relative to those of auto and cross-catalysis. In the figure, we use green to indicate the activating modification and related quantities.

https://doi.org/10.1371/journal.pcbi.1009961.g005

repressed chromatin state to bistable in which also a fully active chromatin state is a stable equilibrium. For lower values of μ', higher values of p_x are required to make the fully active state appear as a stable equilibrium. Therefore, increased p_x can compensate, to some extent, for the asymmetry of the system, thus restoring stability of the active state. Varying μ has a similar effect as varying μ'. In order to gain a qualitative understanding of the way positive autoregulation affects the stability of the system, we also conducted a mathematical analysis assuming that protein dynamics are much faster than chromatin modification dynamics. This analysis shows that positive autoregulation has an effect equivalent to increasing the autocatalysis rate of the activating histone modification. Based on the bifurcation plots realized for

the chromatin modification circuit (Fig K in S1 File), a higher auto-catalysis rate constant k_M^A can restore the active state stability, consistent with the results in Fig A in S2 File.

The strength of positive autoregulation p_x also affects the robustness of the active gene state's memory to endogenous silencing. Specifically, we introduce in our model possible perturbations that can lead to silencing, due to for example chromatin spreading from surrounding gene loci, by considering $u_1^R, u_2^R > 0$. The input/output steady state characteristics $(u^R, \bar{D}^A)$ with $u_1^R = u_2^R = u^R$, shows persistence of the active state to increasing values of u^R as p_x is increased (Fig 5C). These results imply enhanced robustness of the active state to repressive input perturbations.

Stochastic behavior

Here, we determine how positive autoregulation modulates the stationary probability distribution of the chromatin state and the time to memory loss of the active state. To this end, we first perform a computational study by simulating the full set of reactions (Table B in S2 File) with the SSA [51]. This study shows that, increasing p_x, the peak in correspondence to the active state becomes more concentrated and its height increases to the detriment of the peak corresponding to the repressed state (Fig 5D). However, for lower values of μ', the height of the peak in correspondence to the active state is also lower, implying that a higher value of p_x is required to increase the peak to the same level (Fig C in S2 File). Reducing μ has a similar effect as reducing μ' (Fig C in S2 File). These results are in agreement with the temporal trajectories of n^A (Fig 5E), which indicate that the time to memory loss of the active state increases when p_x is increased. The trends with which p_x affects the time to memory loss of the active state do not depend on ϵ' (Fig D in S2 File).

To gain an analytical understanding of how p_x affects the stationary distribution and the time to memory loss, we also computed analytically these quantities as a function of the system parameters in the regime $\epsilon' \ll 1$ and $\epsilon \ll 1$ and by assuming that protein dynamics are fast compared to chromatin dynamics. The latter assumption may be satisfied if the protein is quickly degrading due to proteolysis [56]. In this regime, we obtain a one-dimensional Markov chain reduction of the system (Section 1.3 in S2 File), which leads to the stationary probability distribution for $\epsilon \ll 1$ as

$$\pi_{\epsilon \ll 1}(x) \approx \begin{cases} \frac{1}{1+P} & \text{if } x = 0 \\ 0 & \text{if } x \neq 0, \mathrm{D_{tot}} \\ \frac{P}{1+P} & \text{if } x = \mathrm{D_{tot}} \end{cases} \tag{9}$$

with

$$\begin{aligned} P_1 &= \frac{(u_{tot} + \alpha + \bar{\alpha} + \alpha')}{(u_{tot} + \tilde{u}^A p_x + 1)}, \\ P_2 &= \prod_{i=1}^{\mathrm{D_{tot}}-1} \left(\frac{2\left(\bar{u}_{12}^R + (\alpha + \bar{\alpha} + \alpha')\frac{i}{\mathrm{D_{tot}}}\right)}{\mu\mu'\epsilon' \frac{(\mathrm{D_{tot}} - i)}{\mathrm{D_{tot}}} \bar{K}_i \left(u_0^A + (\tilde{u}^A p_x + 1)\frac{(\mathrm{D_{tot}} - i)}{\mathrm{D_{tot}}} \right)} \right) \cdot \frac{(\bar{u}_{12}^R)}{\mu\mu' b\beta\epsilon \bar{K}_{\mathrm{D_{tot}}}(u_0^A)}, \\ P &= P_1 \cdot P_2, \end{aligned} \tag{10}$$

in which $x = n^R$, $u_{tot} = u_0^A + \bar{u}_{12}^R$, $\bar{u}_{12}^R = u_1^R + u_{10}^R + u_2^R + u_{20}^R$ and $\bar{K}_i$ defined in Eq (161) in S1 File. From these expressions, we conclude that by increasing p_x, P decreases and, in turn, the height of the peak in correspondence to $x = 0$ (active chromatin state) increases to the

detriment of the height of the peak at $x = D_{tot}$ (repressed chromatin state). Furthermore, the smaller μ and μ' are, the higher p_x has to be in order to raise the peak corresponding to the active state. These results are in agreement with the computational results (Fig 5D and Fig C in S2 File).

Using the same model reduction, we derive the formula of the normalized time to memory loss of the active chromatin state, which, for $\epsilon \ll 1$, can be approximated as

$$\bar{\tau}_A \approx \frac{f_2(p_x)}{\epsilon}\left(1 + \sum_{i=1}^{D_{tot}-1} \frac{h_2^i(p_x, \mu\mu')}{K_A^i}\right), \tag{11}$$

in which f_2 and h_2^i are increasing functions of their arguments, $h_2^i(p_x, 0) = 0$, and K_A^i are functions independent of ϵ, μ', μ and p_x (Section 1.4 in S2 File). From this expression, it follows that if p_x increases, $\bar{\tau}_A$ increases. This confirms that positive autoregulation helps extend the memory of the active state.

Taken together, these findings indicate that positive TF-enabled autoregulation contributes to extend the duration of memory of an active chromatin state, thus re-balancing the asymmetry that naturally biases the chromatin modification circuit towards a repressed state. This is also manifested in the effect of positive autoregulation on the time to reactivation of a gene, wherein the time to reactivation decreases by increasing p_x (Fig E in S2 File).

We also determined how the chromatin modification circuit contributes to the robustness of the active state's memory to perturbations that reduce the availability of X at the gene (Fig 5F). For low ϵ the robustness is improved, implying that the system is able to keep the active chromatin state even when X is transiently removed. Then, small ϵ may be a way in which the chromatin modification circuit enables robustness of the active state's memory to disruption of the binding of the TF X to DNA, such as due to DNA replication and cell division.

Wiring positively autoregulated chromatin modification circuits allows long-term persistence of gene expression patterns

Long-term and reconfigurable memory of multiple gene expression patterns is desirable for a number of applications, such as programmable cell therapies and *in vivo* transdifferentiation, wherein TFs need to be activated in a specific temporal sequence and in an exclusive fashion [57, 58]. We thus analyze a candidate circuit to implement this function: two positively autoregulated chromatin modification circuits that mutually recruit writers of repressive modifications to each other (Fig 6A). This is also a common motif in GRNs involved in cell fate determination, such as the PU.1/GATA.1 antagonism [35, 59], the Nanog/GATA interaction [6], and the Cdx2/Oct4 circuit motif [6, 60].

For each gene, the chromatin modification circuit is described by the reactions in Fig 2 and the gene expression reactions are listed in (1). In each gene's chromatin modification circuit, k_W^R is a monotonically increasing function of the other gene's product level and k_W^A is a monotonically increasing function of the gene's self product level (Section 2.1, Eqs (10) in S2 File). The system is then obtained by input/output connecting two chromatin modification circuits through the TFs that they are expressing (Fig 6B).

Deterministic behavior

By combining the ODEs of the chromatin modification circuit and those of the gene expression circuit (Eqs (3) and (8), respectively) for each gene, by defining $X := n^X/D_{tot}$, $Z := n^Z/D_{tot}$ and by properly setting the inputs according to Eqs (13) in S2 File, we obtain the ODEs of the mutual repression system (Eqs (14) in S2 File). To determine the number and stability of

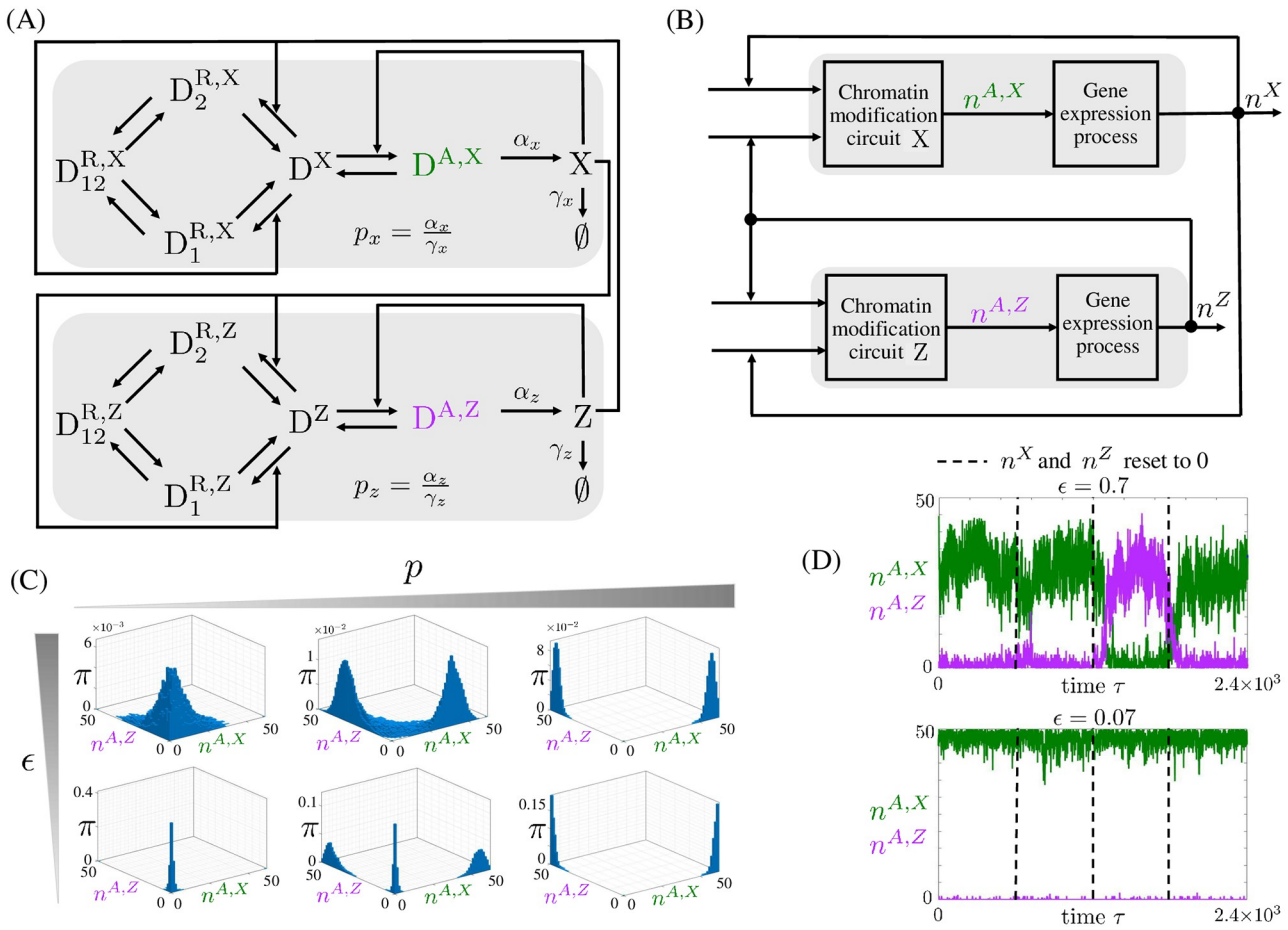

Fig 6. Mutual repression circuit: Robust memory of multiple co-existing gene expression patterns. (A) Interaction diagram of two mutually repressing and positively autoregulated genes, wherein the product of each gene recruits writers of repressive chromatin modifications to the other gene. Here, X and Z represent the products of the two genes. (B) Block diagram corresponding to the circuit in panel (A). Here, n^X and n^Z correspond to the number of molecules of X and Z, respectively. (C) Stationary probability distribution π of the system obtained by simulating the reactions listed in Tables E and F in S2 File with the SSA, in which $n^{A,\ell}$ with $\ell = X, Z$ represents the number of nucleosomes in each gene with activating histone modifications. In (C), $p_x = p_z = p$ with $p = 0$, 0.1, 10 and $\epsilon = 0.48$, 0.2. The parameter values of each plot are listed in Tables E and F in S2 File. For all simulations we have $\mu = 1$, $\mu' = 0.6$, and $\epsilon' = 1$ (Figs G-I in S2 File show different parameter values). (D) Time trajectories of $n^{A,X}$ and $n^{A,Z}$ starting from $n^{A,X} = \mathrm{D}_{\mathrm{tot}}$ and $n^{A,Z} = 0$, in which n^X and n^Z are reset to zero at the indicated times (dashed line). Time is normalized with respect to $\frac{k_M^A}{\Omega}\mathrm{D}_{\mathrm{tot}}$. The parameter values for each panel are listed in Tables G and H in S2 File. In particular, $p = 0.15$, $\mu' = 0.6$, $\mu = 1$, $\epsilon = 1$ and ϵ as indicated. In all plots, we assume equal parameters for both chromatin modification circuits. In our model, ϵ, defined in Eq (2), is a non-dimensional parameter that quantifies the time scales of basal erasure rate of all modifications relative to those of auto and cross-catalysis. In all plots $(n^{A,\ell}, n^{R,\ell}) \approx (50, 0)$ and $(n^{A,\ell}, n^{R,\ell}) \approx (0, 50)$ correspond to the active and repressed state of gene ℓ, with $\ell = X, Z$. In the figure, we use green and purple, respectively, to indicate $\mathrm{D}^{A,X}$ and $\mathrm{D}^{A,Z}$ and related quantities.

https://doi.org/10.1371/journal.pcbi.1009961.g006

equilibria, we exploit the results obtained for the positive autoregulation circuit. In fact, the block diagram in Fig 6B makes it explicit that the mutual repression circuit is the input/output composition of two positively autoregulated genes, in which the output of one gene, n^X or n^Z, is used as an input to the other gene by increasing its k_W^R. This analysis, described in Section 2.3 in S2 File, shows that, for low p values and large ϵ, the system has a unique stable steady state about the origin, where both genes are "off". By increasing p, the system acquires two stable steady states, in which one gene is "on" and the other is "off". That is, one equilibrium with $\bar{D}^{A,X} \gg \bar{D}^{A,Z}$ and the other equilibrium with $\bar{D}^{A,X} \ll \bar{D}^{A,Z}$. When ϵ is small, also the steady state about the origin is stable, while when ϵ is large, it is not. Thus, when p is large (high

expression rate) and ϵ is small, the system is tri-stable. When μ' is increased, the system can have four co-existing stable steady states, acquiring a new stable steady state in which both genes are "on" (Fig F in S2 File). Since we expect μ' to be much smaller than one, this last scenario may not be observed in practice.

Stochastic behavior

Here, denoting the number of nucleosomes in each gene with activating histone modifications as $n^{A,X}$ and $n^{A,Z}$, we first determine how p_x, p_z, and ϵ shape the stationary distribution of ($n^{A,X}$, $n^{A,Z}$) and hence the admissible gene expression patterns. Specifically, we perform a computational study of the distribution by simulating the full set of reactions (Tables E and F in S2 File) with the SSA [51]. If the production rate constant p of both proteins is low, independent of ϵ, the distribution has one peak only in correspondence to both genes being "off", consistent with the deterministic analysis. Therefore, the only possible gene expression pattern in this regime is both genes "off". When p is sufficiently increased, the peak in correspondence of both genes "off" disappears and two peaks arise in correspondence of one gene "on" and the other "off", thus enabling two symmetric gene expression patterns. There is an intermediate protein production rate constant regime, in which a third gene expression pattern with both genes "off" is also possible. In all cases, decreased ϵ makes the distribution more concentrated about each of the peaks, thereby also reducing the probability of transitioning out of that expression pattern (Fig 6C). The admissible gene expression patterns are largely shaped by the choice of the two protein production rate constants (Fig I in S2 File). Reducing μ', as before, biases the distribution towards the repressed chromatin state ($n^{A,X}$, $n^{A,Z}$) $\approx$ (0, 0) (Fig H in S2 File) while varying ϵ' does not change the trends with which the key parameters affect the distribution (Fig G in S2 File). The computational results of Fig 6C are in agreement with those obtained from the analytical study (Section 2.4 in S2 File).

Furthermore, we determine computationally how ϵ affects the robustness of the memory of an admissible gene expression pattern with respect to resetting the amount of the gene's protein products X and Z to 0 (Fig 6D). Lower ϵ values result into enhanced robustness of the pattern's memory to resetting n^X and n^Z to zero, indicating that the chromatin modification circuit with lower ϵ helps memory of co-existing gene expression patterns persist through changes in the availability of TFs to the promoters. Finally, we analyzed whether overexpression of TF can transition the system from one gene expression pattern to the other one. Indeed, even when a gene expression pattern displays high robustness to TFs variability (low ϵ), the system allows a transition from one admissible expression pattern to the other one by TF overexpression (Fig J in S2 File). The speed of transition increases with the level of overexpression and decreases with ϵ.

Discussion

We analyzed the dynamics of a biologically motivated circuit motif among histone modifications and DNA methylation within a gene, which we called the gene's inner chromatin modification circuit (Fig 1D). We found that separation of time scales among three constituent processes, basal erasure, recruited erasure, and auto/cross-catalysis controls memory of input stimuli. Specifically, when basal erasure rate ϵ is sufficiently lower than both recruited erasure rate ϵ' and auto/cross-catalysis rates (α, $\bar{\alpha}$, α'), then the chromatin modification circuit shows bistability, hysteresis, and the probability distribution of chromatin state is highly concentrated about the active and repressed states (Figs 3C, 3D and 4A). Furthermore, the time to memory loss of either active or repressed states increases with ϵ, but significantly more so for the repressed chromatin state due to the cross-catalysis among repressive histone modifications

and DNA methylation. This asymmetry is enhanced by smaller μ', which is the ratio between the decay rate of DNA methylation and that of activating histone modifications. Since the decay rate of DNA methylation is very small due to the maintenance activity of DNMT1 enzyme, we estimated $\mu' \ll 1$ (Section 3.1 in S1 File). In this parameter regime, the active chromatin state becomes poorly robust to repressive input stimuli (Fig 3F) and the height of the peak of the distribution corresponding to the active chromatin state decreases (Fig 4A). TF-enabled positive autoregulation re-balances this asymmetry, thereby allowing enhanced robustness of the active chromatin state to repressive input stimuli (Fig 5C). At the same time, under time scale separation, the chromatin modification circuit enables robustness of the active chromatin state to repetitive disruptions of the positive autoregulation loop (Fig 5F). By wiring positively autoregulated chromatin modification circuits, we can thus obtain concurrent mutually exclusive gene expression patterns, which are robust to repeated disruptions of the regulatory links (Fig 6).

In addition to TF-enabled positive autoregulation, other forms of regulation are also possible, such as negative autoregulation and combinations of the two. For example, a form of TF-enabled negative autoregulation is obtained when the protein X recruits writers of repressive chromatin modifications. An instance of a TF that does so is the LANA protein, which recruits DNMT3 to DNA [61]. Furthermore, different forms of autoregulation motifs can be obtained when X regulates the expression of chromatin modifications' erasers instead of writers. This is observed, for example, for TF Oct4, which increases the expression of H3K9me3 eraser JMJD2A and the expression of TET enzymes [62, 63] or in the regulatory system of olfactory receptor activation, in which the expressed OR protein induces the expression of enzyme Adcy3, which removes the histone demethylase LSD1 [64]. Future work will be devoted to analyzing the dynamical properties of these biologically found regulatory motifs.

The non-dimensional parameters ϵ, ϵ', μ, and μ', which control the relative time scales among the circuit's constituent processes, are known functions of the concentrations of erasers, writers, and readers of nucleosome modifications as well as of cell proliferation rate. As a consequence, our model can predict how experimental interventions on these molecular players and cell division rate will affect plasticity, captured by less concentrated distributions of chromatin states, the kinetics of gene reactivation, and the robustness of active states to repressive stimuli. Indeed, our model predicts that in the absence of time scale separation, which can be obtained by increasing ϵ, the cell should be in a more plastic state, and gene reactivation should be faster and less variable. This is consistent with the experiments of [53], in which ϵ was artificially increased by increasing proliferation rate. Similarly, our model predicts that increasing μ' should make gene re-activation faster and less stochastic. Our expression $\mu' = k_T'^{*}/k_E^A$ also suggests that increasing the level of TET enzymes effectively increases μ' only if the level of MBD proteins is sufficiently low (Eq (76) in S1 File). Indeed, reprogramming experiments where TET was increased did not show dramatic changes on reprogramming kinetics [65], but knocking down MBD proteins showed fast and almost deterministic reactivation kinetics [54]. Future experiments on semi-synthetic chromosomally integrated reporter systems will be required to directly validate some of the model's predictions. For example, experiments in which ϵ is artificially increased by concurrently recruiting erasers for both H3K9 and H3K4 methylation, such as JMJD2A and JARID, will demonstrate the morphing of the distribution from bimodal to unimodal. Experiments where only erasers for H3K4 methylation (JARID) will be recruited can validate the prediction that the memory of the repressed state is extended through the decrease of μ and μ'. Finally, experiments in which H3K9 methylation is completely erased through recruitment of JMJD2A but DNA methylation is present should show loss of bimodality, since this is not present in a model that only

includes DNA methylation and H3K4 methylation (Fig B in S1 File). In addition, we will revise the model in order to remove some of the biological simplifications, such as the well-mixed model simplification and the assumption that a nucleosome cannot be characterized by more than one modified histone simultaneously.

Our results suggest that reduced time scale separation (larger ϵ and/or reduced ϵ') may be implicated in phenotypic plasticity, wherein larger ϵ allows for more probable transitions among multiple admissible gene expression patterns and thus among different phenotypes (Fig 6C). While small ϵ embodies a gene expression pattern with high robustness to loss of TF binding, larger ϵ allows for more probable transitions following TF binding disruptions (Fig 6D). It is plausible that physiological requirements for high degree of cellular plasticity at the initial stages of cell fate specification may be satisfied by a transiently larger ϵ at cell fate specific genes. Locking-in of the fate may then be aided by a decreased ϵ at these cell fate specific genes in terminally differentiated cells [8]. On the other hand, accidental changes of phenotype often result in disease such as cancer [66, 67]. These accidents may correlate with increased ϵ and/or μ' at those genes that become disregulated. Global increases in ϵ may then be caused by increased proliferation rates since these increase the basal decay rate of all modifications. Indeed, cellular states with high proliferation rates are associated with larger ϵ and hence with increased plasticity. These states include fastly proliferating cancer cells, pluripotent stem cells, and, most generally, progenitor cells.

Finally, genomic integration of engineered genetic systems has appeared as a promising approach for the creation of organoids [68] and for *in vivo* transdifferentiation [57, 58]. In particular, we found that positive autoregulation, by recruitment of writers of activating modifications or of erasers of repressive modifications, enables robustness of active chromatin states to endogenous silencing factors. This provides design guidelines for chromosomally integrated gene expression cassettes that are resilient to silencing. Similarly, mutual repression circuits obtained by chromatin state regulation, as opposed to traditional TF-enabled regulation, can be used for future cell fate decision circuits that remain locked-in with high probability but yet are reconfigurable by external interventions.

Supporting information

S1 File. Supporting information file with derivation of the models, detailed analysis of the histone and chromatin modification models, additional simulations and parameter values used in the simulations. Fig A: establishment, erasure, and maintenance of DNA methylation. Fig B: competitive interactions between activating histone modifications D^A and DNA methylation D_1^R. Fig C: cooperative interactions between DNA methylation D_1^R and repressive histone modification D_2^R. Fig D: competition between activating and repressive histone modifications. Fig E: trajectories of $\bar{D}^R$ and $\bar{D}^A$ of the full and the reduced system. Figs F-I: effect of system parameters on the stochastic behavior of the histone modification circuit. Fig J: trajectories of $\bar{D}_{12}^R$ and $\bar{D}^A$ of the full and the reduced system. Figs K-O: effect of system parameters on the deterministic behavior of the chromatin modification circuit. Figs P-W: effect of system parameters on the stochastic behavior of the chromatin modification circuit. Fig X: time trajectories of $\bar{D}^A$. Fig Y: effect of ϵ and μ' on the silencing and reactivation processes. Table A: reactions and parameter values used to generate the plots in Fig 4A, 4B. Table B: reactions and parameter values used to generate the plots in Fig 4C, 4D. Table C: reactions and parameter values used to generate the plots in Fig 4F. Table D: parameter values relative to Fig D. Table E: parameter values relative to Fig E. Table F: reactions and parameter values used to generate the plots in Fig F. Table G: reactions and parameter values used to generate the plots in Fig G. Table H: reactions and parameter values used to generate the plots in Fig H. Table I:

reactions and parameter values used to generate the plots in Fig I. Table J: parameter values relative to Fig J. Table K: reactions and parameter values used to generate the plots in Fig P. Table L: reactions and parameter values used to generate the plots in Fig Q. Table M: reactions and parameter values used to generate the plots in Fig R. Table N: reactions and parameter values used to generate the plots in Fig S. Table O: reactions and parameter values used to generate the plots in Fig T. Table P: reactions and parameter values used to generate the plots in Figs 4E and U. Table Q: reactions and parameter values used to generate the plots in Fig V. Table R: reaction and parameter values used to generate the plots in Fig W. Table S: reactions and parameter values used to generate the plots in Fig Y.
(PDF)

S2 File. Supporting information file with detailed analysis of the positively autoregulated gene model and of the mutual repression circuit model, additional simulations and parameter values used in the simulations. Fig A: effect of system parameters on the deterministic behavior of the positively autoregulated circuit. Figs B-E: effect of system parameters on the stochastic behavior of the positively autoregulated circuit. Fig F: effect of system parameters on the deterministic behavior of the mutual repression circuit. Figs: G-J: effect of system parameters on the stochastic behavior of the mutual repression circuit. Table A: parameter values relative to Fig 5C. Table B: reactions and parameter values used to generate the plots in Figs 5D and B. Table C: reactions and parameter values used to generate the plots in Figs 5E and D. Table D: reactions and parameter values used to generate the plots in Fig 5F. Tables E and F: reactions and parameter values used to generate the plots in Figs 6C and I. Tables G and H: reactions and parameter values used to generate the plots in Figs 6D and J. Table I: reactions and parameter values used to generate the plots in Fig C. Table J: reactions and parameter values used to generate the plots in Fig E. Tables K and L: reactions and parameter values used to generate the plots in Figs G and H.
(PDF)

Acknowledgments

R.J.W. gratefully acknowledges the hospitality of the Center of Mathematical Sciences and Applications at Harvard University for her sabbatical visit, during which preliminary discussions were conducted with S.B. and D.D.V. on the research reported here.

Author Contributions

Conceptualization: Domitilla Del Vecchio.

Data curation: Simone Bruno.

Formal analysis: Simone Bruno, Domitilla Del Vecchio.

Funding acquisition: Ruth J. Williams, Domitilla Del Vecchio.

Methodology: Simone Bruno, Ruth J. Williams, Domitilla Del Vecchio.

Project administration: Domitilla Del Vecchio.

Resources: Domitilla Del Vecchio.

Software: Simone Bruno.

Supervision: Domitilla Del Vecchio.

Visualization: Simone Bruno, Domitilla Del Vecchio.

Writing – original draft: Simone Bruno, Domitilla Del Vecchio.

Writing – review & editing: Simone Bruno, Domitilla Del Vecchio.

References

1. Waddington CH. The epigenotype. Endeavour. 1942.
2. Ptashne M. Epigenetics: Core misconcept. Proc. Natl. Acad. Sci. 2013. 110. https://doi.org/10.1073/pnas.1305399110 PMID: 23584020
3. Xiong W, Ferrell JE Jr. A positive-feedback-based bistable 'memory module' that governs a cell fate decision. Nature. 2003. 426. https://doi.org/10.1038/nature02089 PMID: 14647386
4. Carey N. The epigenetic revolution. Columbia University Press. 2013.
5. Huang S, Litt M, Blakey CA. Epigenetic Gene Expression and Regulation. Academic Press. 2015.
6. Lanza R, Atala A. Essentials of stem cell biology. Elsevier. 2014.
7. Huang S. Reprogramming cell fates: reconciling rarity with robustness. BioEssays. 2009. 31. https://doi.org/10.1002/bies.200800189 PMID: 19319911
8. Holmberg J, Perlmann T. Maintaining differentiated cellular identity. Nature Reviews. 2012. 13. https://doi.org/10.1038/nrg3209 PMID: 22596319
9. Huang S, Guo Y, May G, Enver T. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. Developmental Biology. 2007. 305(2): 695–713. https://doi.org/10.1016/j.ydbio.2007.02.036 PMID: 17412320
10. Allis CD, Caparros M-L, Jenuwein T, Reinberg D. Epigenetics. Cold Spring Harbor Laboratory Press, Second Edition. 2015.
11. Heard E, Martienssen RA. Transgenerational Epigenetic Inheritance: Myths and Mechanisms. Cell. 2014. 157. https://doi.org/10.1016/j.cell.2014.02.045 PMID: 24679529
12. Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. Nature Reviews. 2016. 17. https://doi.org/10.1038/nrg.2016.59 PMID: 27346641
13. Holliday R, Pugh JE. DNA Modification Mechanisms and Gene Activity during Development. American Association for the Advancement of Science. 1975. 187. https://doi.org/10.1126/science.187.4173.226 PMID: 1111098
14. Riggs AD. X inactivation, differentiation, and DNA methylation. Cytogenet Cell Genet. 1975. 14. https://doi.org/10.1159/000130315 PMID: 1093816
15. Du J, Johnson LM, Jacobsen SE, Patel DJ. DNA methylation pathways and their crosstalk with histone methylation. Nat Rev Mol Cell Biol. 2015. 16. https://doi.org/10.1038/nrm4043 PMID: 26296162
16. Ordovas L, Boon R, Pistoni M, Chen Y, Wolfs E, Guo W, et al. Efficient recombinase-mediated cassette exchange in hpscs to study the hepatocyte lineage reveals aavs1 locus-mediated transgene inhibition. Stem Cell Reports. 2015. 5. https://doi.org/10.1016/j.stemcr.2015.09.004 PMID: 26455413
17. Alhaji SY, Ngai SC, Abdullah S. Silencing of transgene expression in mammalian cells by DNA methylation and histone modifications in gene therapy perspective. Biotechnology and Genetic Engineering Reviews. 2019. 35. https://doi.org/10.1080/02648725.2018.1551594 PMID: 30514178
18. Fitzgerald M, Livingston M, Gibbs C, Deans TL. Rosa26 docking sites for investigating genetic circuit silencing in stem cells. Synthetic Biology. 2020. 5. https://doi.org/10.1093/synbio/ysaa014 PMID: 33195816
19. Zimak J, Wagoner ZW, Nelson N, Waechtler B, Schlosser H, Kopecky M, et al. Epigenetic silencing directs expression heterogeneity of stably integrated multi-transcript unit genetic circuits. Sci Rep. 2021. 11. https://doi.org/10.1038/s41598-021-81975-1 PMID: 33510302
20. Ringrose L, Howard M. Dissecting chromatin-mediated gene regulation and epigenetic memory through mathematical modelling. Current Opinion in Systems Biology. 2017. 3. https://doi.org/10.1016/j.coisb.2017.02.003
21. Artyomov MN, Meissner A, Chakraborty AK. A Model for Genetic and Epigenetic Regulatory Networks Identifies Rare Pathways for Transcription Factor Induced Pluripotency. PLOS Comput. Biol. 2010. 6. https://doi.org/10.1371/journal.pcbi.1000785 PMID: 20485562
22. Flöttmann M, Scharp T, Klipp E. A stochastic model of epigenetic dynamics in somatic cell reprogramming. PLOS Comput. Biol. 2012. https://doi.org/10.3389/fphys.2012.00216 PMID: 22754535
23. Ashwin SS, Sasai M. Effects of Collective Histone State Dynamics on Epigenetic Landscape and Kinetics of Cell Reprogramming. Nature Reports. 2015. https://doi.org/10.1038/srep16746 PMID: 26581803
24. Chen T, Al-Radhawi MA, Sontag ED. A mathematical model exhibiting the effect of DNA methylation on the stability boundary in cell-fate networks. Epigenetics. 2021. 16(4):436–457. https://doi.org/10.1080/15592294.2020.1805686 PMID: 32842865

25. Jia W, Mani ASA, Jolly MK, Levine H. A possible role for epigenetic feedback regulation in the dynamics of the epithelial-mesenchymal transition (EMT). Phys Biol. 2019. 16(6). https://doi.org/10.1088/1478-3975/ab34df

26. Dodd IB, Micheelsen MA, Sneppen K, Thon G. Theoretical Analysis of Epigenetic Cell Memory by Nucleosome Modification. Cell. 2007. 129. https://doi.org/10.1016/j.cell.2007.02.053 PMID: 17512413

27. Zhang T, Cooper S, Brockdorff N. The interplay of histone modifications—writers that read. EMBO Reports. 2015. https://doi.org/10.15252/embr.201540945 PMID: 26474904

28. Wysocka J, Myers MP, Laherty CD, Eisenman RN, Herr W. Human Sin3 deacetylase and trithorax-related Set1/Ash2 histone H3-K4 methyltransferase are tethered together selectively by the cell-proliferation factor HCF-1. GENES & DEVELOPMENT. 2003. 17. https://doi.org/10.1101/gad.252103 PMID: 12670868

29. Binda O, LeRoy G, Bua DJ, Garcia BA, Gozani O, Richard S. Trimethylation of histone H3 lysine 4 impairs methylation of histone H3 lysine 9. Epigenetics. 2010. 5. https://doi.org/10.4161/epi.5.8.13278 PMID: 21124070

30. Jin B, Li Y, Robertson KD. DNA Methylation. Superior or Subordinate in the Epigenetic Hierarchy?. Genes Cancer. 2011. 2(5):607–617. https://doi.org/10.1177/1947601910393957 PMID: 21941617

31. Ooi SKT, Qiu C, Bernstein E, Li K, Jia D, Yang Z, et al. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature. 2007. 448. https://doi.org/10.1038/nature05987 PMID: 17687327

32. Courey AJ. Mechanisms in Trancriptional Regulation. Blackwell Publishing. 2008.

33. Kim T, Buratowski S. Dimethylation of H3K4 by Set1 Recruits the Set3 Histone Deacetylase Complex to 50 Transcribed Regions. Cell. 2009. 137. https://doi.org/10.1016/j.cell.2009.02.045

34. Ang Y, Tsai S-Y, Lee D-F, Monk J, Su J, Ratnakumar K, et al. Wdr5 Mediates Self-Renewal and Reprogramming via the Embryonic Stem Cell Core Transcriptional Network. Cell. 2011. 145. https://doi.org/10.1016/j.cell.2011.03.003 PMID: 21477851

35. Stopka T, Amanatullah DF, Papetti M, Skoultchi AI. PU.1 inhibits the erythroid program by binding to GATA-1 on DNA and creating a repressive chromatin structure. The EMBO Journal. 2005. 24. https://doi.org/10.1038/sj.emboj.7600834 PMID: 16222338

36. Del Vecchio D, Murray RM. Biomolecular Feedback Systems. Princeton University Press. 2014.

37. Hathaway NA, Bell O, Hodges C, Miller EL, Neel DS, Crabtree GR. Dynamics and Memory of Heterochromatin in Living Cells. Cell. 2012. 149. https://doi.org/10.1016/j.cell.2012.03.052 PMID: 22704655

38. Trerotola M, Relli V, Simeone P, Alberti S. Epigenetic inheritance and the missing heritability. Human Genomics. 2015. 9. https://doi.org/10.1186/s40246-015-0041-3 PMID: 26216216

39. Alabert C, Barth TK, Reveron-Gomez N, Sidoli S, Schmidt A, Jensen ON, et al. Two distinct modes for propagation of histone PTMs across the cell cycle. Research Communication, CSHL. 2015. https://doi.org/10.1101/gad.256354.114 PMID: 25792596

40. Upadhyay AK, Horton JR, Zhang X, Cheng X. Coordinated Methyl-lysine Erasure: Structural and Functional Linkage of a Jumonji demethylasedomain and a Reader domain. Curr Opin Struct Biol. 2011. 21. https://doi.org/10.1016/j.sbi.2011.08.003 PMID: 21872465

41. Jin C, Lu Y, Jelinek J, Liang S, Estecio MRH, Craig Barton M, et al. TET1 is a maintenance DNA demethylase that prevents methylation spreading in differentiated cells. Nucleic Acids Research. 2014. 42. https://doi.org/10.1093/nar/gku372 PMID: 24875481

42. Rasmussen KD, Helin K. Role of TET enzymes in DNA methylation, development, and cancer. GENES & DEVELOPMENT. 2016. 30. https://doi.org/10.1101/gad.276568.115 PMID: 27036965

43. Ludwig AK, Zhang P, Hastert FD, Meyer S, Rausch C, Herce HD, et al. Binding of MBD proteins to DNA blocks Tet1 function thereby modulating transcriptional noise. Nucleic Acid Research. 2017. 45. https://doi.org/10.1093/nar/gkw1197 PMID: 27923996

44. Genereux DP, Miner BE, Bergstrom CT, Laird CD. A population-epigenetic model to infer site-specific methylation rates from double-stranded DNA methylation patterns. Proc. Natl. Acad. Sci. 2005. 102. https://doi.org/10.1073/pnas.0502036102 PMID: 15827124

45. Lopez-Serra L, Esteller M. Proteins that bind methylated DNA and human cancer: reading the wrong words. British Journal of Cancer. 2008. 98. https://doi.org/10.1038/sj.bjc.6604374 PMID: 18542062

46. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature. 1998. 393. https://doi.org/10.1038/30764 PMID: 9620804

47. Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, et al. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nature Genetics. 1998. 19. https://doi.org/10.1038/561 PMID: 9620779

48. Fuks F, Hurd PJ, Wolf D, Nan X, Bird AP, Kouzarides T. The Methyl-CpG-binding Protein MeCP2 Links DNA Methylation to Histone Methylation. The J. of Biological Chemistry. 2003. 278. https://doi.org/10.1074/jbc.M210256200 PMID: 12427740

49. Fuks F, Hurd PJ, Deplus R, Kouzarides T. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. Nucleic Acids Research. 2003. 31. https://doi.org/10.1093/nar/gkg332 PMID: 12711675

50. Tang L, Nogales E, Ciferri C. Structure and Function of SWI/SNF Chromatin Remodeling Complexes and Mechanistic Implications for Transcription. Prog Biophys Mol Biol. 2010. 102. https://doi.org/10.1016/j.pbiomolbio.2010.05.001 PMID: 20493208

51. Gillespie DT. Stochastic Simulation of Chemical Kinetics. Annual Review of Physical Chemistry. 2007. https://doi.org/10.1146/annurev.physchem.58.032806.104637 PMID: 17037977

52. Bintu L, Yong J, Antebi YE, McCue K, Kazuki Y, Uno N, et al. Dynamics of epigenetic regulation at the single-cell level. Science. 2016. https://doi.org/10.1126/science.aab2956 PMID: 26912859

53. Hanna J, Saha K, Pando B, van Zon J, Lengner CJ, Creyghton MP, et al. Direct cell reprogramming is a stochastic process amenable to acceleration. Nature. 2009. 462. https://doi.org/10.1038/nature08592 PMID: 19898493

54. Rais Y, Zviran A, Geula S, Gafni O, Chomsky E, Viukov S, et al. Deterministic direct reprogramming of somatic cells to pluripotency. Nature. 2013. https://doi.org/10.1038/nature12587 PMID: 24048479

55. Boyer LA, Ihn Lee T, Cole MF, Johnstone SE, Levine SS, Zucker JP, et al. Core Transcriptional Regulatory Circuitry in Human Embryonic Stem Cells. Cell. 2005. 122:947–956. https://doi.org/10.1016/j.cell.2005.08.020 PMID: 16153702

56. Lecker SH, Goldberg AL, Mitch WE. Protein Degradation by the Ubiquitin–Proteasome Pathway in Normal and Disease States. Journal of the American Society of Nephrology. 2006. 17. https://doi.org/10.1681/ASN.2006010083 PMID: 16738015

57. Jopling C, Boue S, Belmonte JCI. Dedifferentiation, transdifferentiation and reprogramming: three routes to regeneration. Nature Reviews Molecular Cell Biology. 2011. 12. https://doi.org/10.1038/nrm3043 PMID: 21252997

58. Cie´slar-Pobuda A, Knoflach V, Ringh MV, Stark j, Likus W, Siemianowicz K, et al. Transdifferentiation and reprogramming: Overview of the processes, their similarities and differences. Biochimica et Biophysica Acta (BBA)—Molecular Cell Research. 2017.

59. Burda P, Vargova J, Curik N, Salek C, Papadopoulos GL, Strouboulis J, et al. GATA-1 Inhibits PU.1 Gene via DNA and Histone H3K9 Methylation of Its Distal Enhancer in Erythroleukemia. PLOS ONE. 2015. https://doi.org/10.1371/journal.pone.0116840 PMID: 25679502

60. Ralston A, Rossant J. Genetic regulation of stem cell origins in the mouse embryo. Clinical Genetics. 2005. 68(2):106–112. https://doi.org/10.1111/j.1399-0004.2005.00478.x PMID: 15996204

61. Shamay M, Krithivas A, Zhang J, Hayward SD. Recruitment of the de novo DNA methyltransferase Dnmt3a by Kaposi's sarcoma-associated herpesvirus LANA. Proceedings of the National Academy of Sciences. 2006. 103(39):14554–14559. https://doi.org/10.1073/pnas.0604469103

62. Loh Y-H, Zhang W, Chen X, George J, Ng H-H. Jmjd1a and Jmjd2c histone H3 Lys 9 demethylases regulate self-renewal in embryonic stem cells. GENES & DEVELOPMENT. 2007. 21. https://doi.org/10.1101/gad.1588207 PMID: 17938240

63. Wu Y, Guo Z, Liu Y, Tang B, Wang Y, Yang L, et al. Oct4 and the small molecule inhibitor, SC1, regulates Tet2 expression in mouse embryonic stem cells. Mol Biol Rep. 2013. 40. https://doi.org/10.1007/s11033-012-2305-5 PMID: 23254757

64. Tian XJ, Zhang H, Sannerud J, Xing J. Achieving diverse and monoallelic olfactory receptor selection through dual-objective optimization design. Proc Natl Acad Sci U S A. 2016. 113(21):E2889–E2898. https://doi.org/10.1073/pnas.1601722113 PMID: 27162367

65. Gao Y, Chen J, Li K, Wu T, Huang B, Liu W, et al. Replacement of Oct4 by Tet1 during iPSC Induction Reveals an Important Role of DNA Methylation and Hydroxymethylation in Reprogramming. Cell Stem Cell. 2013. https://doi.org/10.1016/j.stem.2013.02.005 PMID: 23499384

66. Slack JMW. Metaplasia and transdifferentiation: from pure biology to the clinic. Nature Rev. Mol. Cell Bio. 2007. 8. PMID: 17377526

67. Holmberg J, He X, Peredo I, Orrego A, Hesselager G, Ericsson C, et al. Activation of neural and pluripotent stem cell signatures correlates with increased malignancy in human glioma. PLoS ONE. 2011. 6. https://doi.org/10.1371/journal.pone.0018454 PMID: 21483788

68. Garreta E, Kamm RD, Chuva de Sousa Lopes SM, Lancaster MA, Weiss R, Trepat X, et al. Rethinking organoid technology through bioengineering. Nat. Mater. 2021. 20. https://doi.org/10.1038/s41563-020-00804-4 PMID: 33199860