



## CENTRE DE RECERCA MATEMÀTICA

Title: *A multiscale model of epigenetic heterogeneity-driven cell fate decisionmaking*  
Journal Information: *PLoS Computational Biology*,  
Author(s): Folguera-Blasco N., Pérez-Carrasco R., Cuyàs E.,  
Menendez J.A., Alarcón T..  
Volume, pages: 15 1, DOI:[10.1371/journal.pcbi.1006592]

RESEARCH ARTICLE

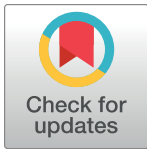
# A multiscale model of epigenetic heterogeneity-driven cell fate decision-making

Núria Folguera-Blasco<sup>1,2\*</sup>, Rubén Pérez-Carrasco<sup>3</sup>, Elisabet Cuyàs<sup>4,5</sup>, Javier A. Menendez<sup>4,5</sup>, Tomás Alarcón<sup>1,2,6,7</sup>

**1** Centre de Recerca Matemàtica, Edifici C, Campus de Bellaterra, 08193 Bellaterra, Barcelona, Spain, **2** Departament de Matemàtiques, Universitat Autònoma de Barcelona, 08193 Bellaterra, Barcelona, Spain, **3** Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK, **4** ProCURE (Program Against Cancer Therapeutic Resistance), Metabolism and Cancer Group, Catalan Institute of Oncology, Girona, Spain, **5** Girona Biomedical Research Institute (IDIBGI), Girona, Spain, **6** ICREA, Pg. Lluís Companys 23, 08010 Barcelona, Spain, **7** Barcelona Graduate School of Mathematics (BGSMath), Barcelona, Spain

✉ Current address: The Francis Crick Institute, London, United Kingdom

\* [nuria.folguerablasco@crick.ac.uk](mailto:nuria.folguerablasco@crick.ac.uk)



## OPEN ACCESS

**Citation:** Folguera-Blasco N, Pérez-Carrasco R, Cuyàs E, Menendez JA, Alarcón T (2019) A multiscale model of epigenetic heterogeneity-driven cell fate decision-making. *PLoS Comput Biol* 15(4): e1006592. <https://doi.org/10.1371/journal.pcbi.1006592>

**Editor:** Natalia L. Komarova, University of California Irvine, UNITED STATES

**Received:** October 18, 2018

**Accepted:** March 19, 2019

**Published:** April 30, 2019

**Copyright:** © 2019 Folguera-Blasco et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the manuscript and its Supporting Information files.

**Funding:** This work is supported by a grant of the Obra Social La Caixa Foundation on Collaborative Mathematics awarded to the Centre de Recerca Matemàtica. The authors have been partially funded by the CERCA Programme of the Generalitat de Catalunya. EC is the recipient of a Sara Borrell post-doctoral contract (CD15/00033, Ministerio de Sanidad y Consumo, Fondo de

## Abstract

The inherent capacity of somatic cells to switch their phenotypic status in response to damage stimuli *in vivo* might have a pivotal role in ageing and cancer. However, how the entry-exit mechanisms of phenotype reprogramming are established remains poorly understood. In an attempt to elucidate such mechanisms, we herein introduce a stochastic model of combined epigenetic regulation (ER)-gene regulatory network (GRN) to study the plastic phenotypic behaviours driven by ER heterogeneity. To deal with such complex system, we additionally formulate a multiscale asymptotic method for stochastic model reduction, from which we derive an efficient hybrid simulation scheme. Our analysis of the coupled system reveals a regime of tristability in which pluripotent stem-like and differentiated steady-states coexist with a third indecisive state, with ER driving transitions between these states. Crucially, ER heterogeneity of differentiation genes is for the most part responsible for conferring abnormal robustness to pluripotent stem-like states. We formulate epigenetic heterogeneity-based strategies capable of unlocking and facilitating the transit from differentiation-refractory (stem-like) to differentiation-primed epistates. The application of the hybrid numerical method validates the likelihood of such switching involving solely kinetic changes in epigenetic factors. Our results suggest that epigenetic heterogeneity regulates the mechanisms and kinetics of phenotypic robustness of cell fate reprogramming. The occurrence of tunable switches capable of modifying the nature of cell fate reprogramming might pave the way for new therapeutic strategies to regulate reparative reprogramming in ageing and cancer.

Investigación Sanitaria, Spain). NF-B and TA acknowledge MINECO and AGAUR for funding under grants MTM2015-71509-C2-1-R and 2014SGR1307. TA acknowledges support from MINECO for funding awarded to the Barcelona Graduate School of Mathematics under the “María de Maeztu” programme, grant number MDM-2014-0445. RP-C also acknowledges the UCL Mathematics Clifford Fellowship. This work was supported by grants from MINECO (SAF2016-80639-P) and AGAUR (2014 SGR229) to JAM. 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.

## Author summary

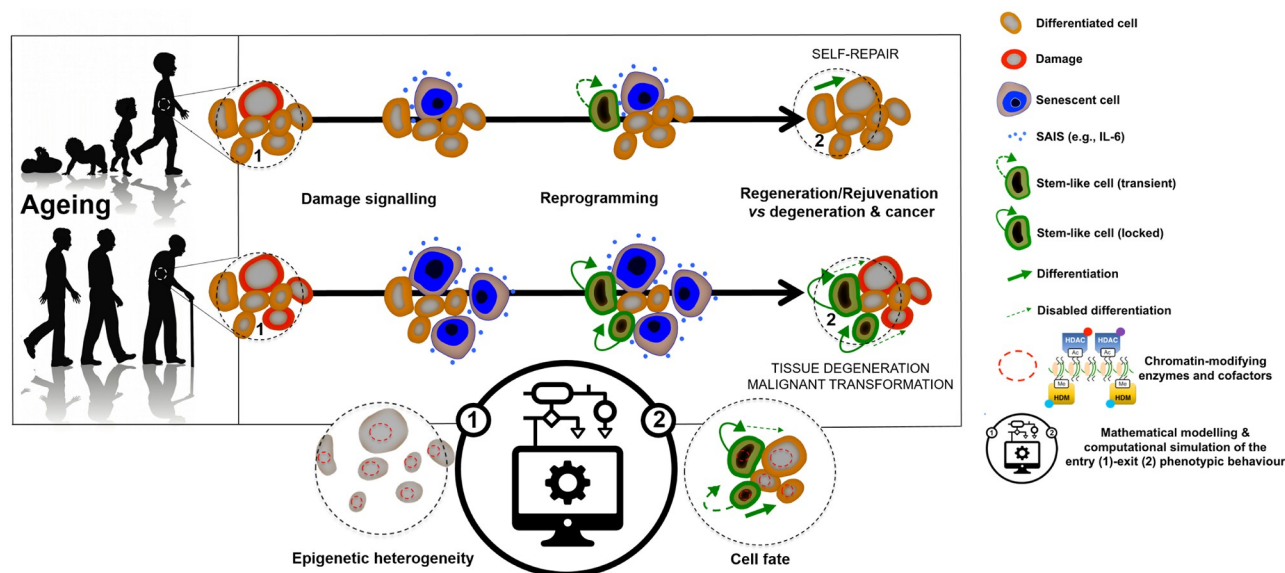
Certain modifications of the structure and functioning of the protein/DNA complex called chromatin can allow adult, fully differentiated, cells to adopt a stem cell-like pluripotent state in a purely epigenetic manner, not involving changes in the underlying DNA sequence. Such reprogramming-like phenomena may constitute an innate reparative route through which human tissues respond to injury and could also serve as a novel regenerative strategy in human pathological situations in which tissue or organ repair is impaired. However, it should be noted that *in vivo* reprogramming would be capable of maintaining tissue homeostasis provided the acquisition of pluripotency features is strictly transient and accompanied by an accurate replenishment of the specific cell types being lost. Crucially, an excessive reprogramming in the absence of controlled re-differentiation would impair the repair or the replacement of damaged cells, thereby promoting pathological alterations of cell fate. A mechanistic understanding of how the degree of chromatin *plasticity* dictates the reparative versus pathological behaviour of *in vivo* reprogramming to *rejuvenate* aged tissues while preventing tumorigenesis is urgently needed, including especially the intrinsic epigenetic heterogeneity of the tissue resident cells being reprogrammed. We here introduce a novel method that mathematically captures how epigenetic heterogeneity is actually the driving force that governs the routes and kinetics to entry into and exit from a pathological stem-like state. Moreover, our approach computationally validates the likelihood of unlocking chronic, unrestrained plastic states and drive their differentiation down the correct path by solely manipulating the intensity and direction of few epigenetic control switches. Our approach could inspire new therapeutic approaches based on *in vivo* cell reprogramming for efficient tissue regeneration and rejuvenation and cancer treatment.

## Introduction

The ability of the ageing process to interfere with the robustness of the epigenetic regulation (ER) of differentiated phenotypes might suffice to promote tissue dysfunction and malignization [1].

Fully committed somatic cells can spontaneously reprogram to pluripotent stem-like cells during the normal response to injury or damage *in vivo* [2]. Such cellular processes involving dedifferentiation and cell-fate switching might constitute a fundamental element of a tissue's capacity to self-repair and rejuvenate [3, 4]. However, such physiological/reparative cell reprogramming might have pathological consequences if the acquisition of epigenetic and phenotypic plasticity is not transient. In response to chronically permissive tissue environments for *in vivo* reprogramming, the occurrence of unrestrained epigenetic plasticity might permanently lock cells into self-renewing pluripotent cell states disabled for reparative differentiation and prone to tumorigenesis (see Fig 1) [1, 5–8].

Central to such so-called stem-lock model for ageing and cancer [6, 9, 10] is the sufficient capacity of ER to drive cell fate in the absence of *bona fide*, initiating events. ER refers to a series of modifications of the cell's DNA without modifying its genetic sequence. Such modifications can disrupt or allow expression of particular genes. By switching on or off different parts of the genome, ER is in fact responsible for the variety of phenotypes in complex multicellular organisms (where all somatic cells are genetically identical). Recent advances in experimental determination of ER mechanisms have triggered an ever-growing interest in developing mathematical models regarding both ER of gene expression [11–16] and epigenetic



**Fig 1. Physiological and pathological cell fate reprogramming: A mathematical approach.** Reprogramming-like phenomena in response to damage signalling may constitute a reparative route through which human tissues respond to injury, stress, and disease via induction of a transient acquisition of epigenetic plasticity and phenotypic malleability. However, tissue regeneration/rejuvenation should involve not only the transient epigenetic reprogramming of differentiated cells, but also the committed re-acquisition of the original or alternative committed cell fate. Chronic or unrestrained epigenetic plasticity would drive ageing/cancer phenotypes by impairing the repair or the replacement of damaged cells; such uncontrolled phenomena of *in vivo* reprogramming might also generate cancer-like cellular states. Accordingly, we now know that chronic senescence-associated inflammatory signalling (SAIS) might lock cells in highly plastic epigenetic states disabled for reparative differentiation and prone to malignant transformation. We herein introduce a first-in-class stochastic, multiscale reduction method of combined epigenetic regulation (ER)-gene regulatory network (GRN) to mathematically model and computationally simulate how ER heterogeneity regulates the entry-exit mechanisms and kinetics of physiological and pathological cell fate reprogramming.

<https://doi.org/10.1371/journal.pcbi.1006592.g001>

memory [12–14, 17–20]. Identification of the molecular culprits underlying the ER capacity to drive the transition between normal and highly restrictive/permissive chromatin states is expected to have major impact in the understanding and therapeutic management of ageing-related diseases including cancer [7, 21, 22]. Unfortunately, robust and standardised approaches capable of capturing such fundamental stochastic aspects of ageing biology are mostly lacking.

Here, we present a mathematical and computational systems biology approach capable of deconstructing, modelling and simulating the predictive power that ER may have on the susceptibility of cells to loss (and re-gain) their normal identity. By adding ER to the picture, our current work significantly extends previous approaches where phenotypes are associated with the attractors of complex gene regulatory systems and their robustness, with the resilience of such attractors in the presence of intrinsic noise, environmental fluctuations, and other disturbances [23–31]. Specifically, we develop a stochastic model of a coupled ER-gene regulatory network (GRN) system aimed at analysing the impact of ER heterogeneity on the causal relationship between epigenetic plasticity and cell-fate reprogramming and determination. Moreover, by introducing a stochastic model reduction analysis based on multiple scale asymptotics of the combined ER-GRN system [32–39], we provide our model with the capacity of evaluating a variety of phenotypic behaviours due solely to ER systems heterogeneity [40].

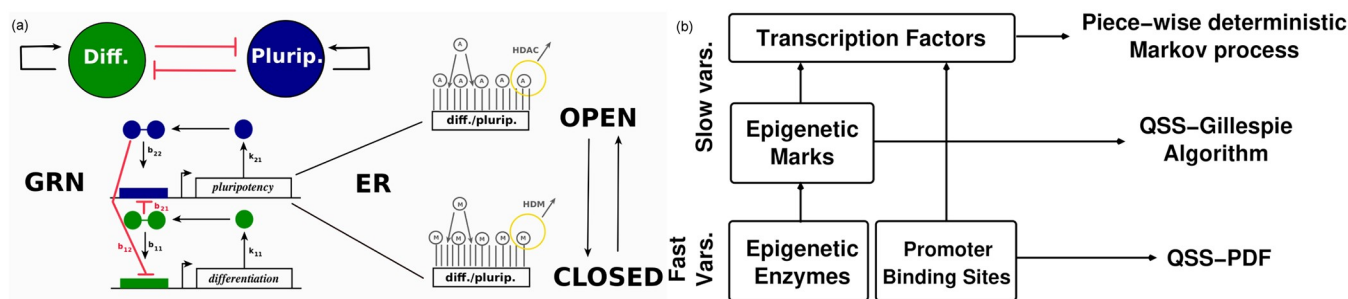
This work is organised as follows. In section *Materials and methods*, we present a summary of the formulation of our ER-GRN model and its analysis. First, we present the general description of an ER-GRN model, and later, we focus on the ER component of the model. We

detail how the transitions between stable states within the ER system are computed and how the ensemble of ER systems is generated. In order to analyse the consequences of the existing heterogeneity within the ER systems generated, we develop a multiscale asymptotic theory to study ER-GRN systems (for which additional details are given in the [S1 Text](#)). Such theory allows us to reduce the complexity of the model to a hybrid system, for which a numerical simulation method was implemented (see [S1 Appendix](#)). The *Results* section starts by presenting the findings regarding the effects of ER in the GRN. In particular, we initially evaluate the epigenetic parameters regulating the entry into robust epigenetic states throughout the entire ER-GRN system. Then, we focus on the role that the ER heterogeneity may have in giving rise to different system behaviours, namely, differentiation-primed and differentiation-resilient (pluripotency-locked) states. Once these two different behaviours are identified, we perform an analysis to identify the mechanisms regulating the phenotypic robustness of the pluripotency-locked and differentiation-primed states. We then formulate epigenetic heterogeneity-based strategies capable of directing the exit and transit from stem-locked to differentiation-primed epistates. We finally apply the hybrid numerical method derived from our theoretical analysis to determine the efficiency of the epigenetic strategies formulated to unlock a persistent state of pathological pluripotency. Finally, in the *Discussion* section, we summarise our findings and present our conclusions.

## Materials and methods

### Model formulation and analysis

In this paper, we aim to study an ER-GRN model which can describe cell differentiation and cell reprogramming. One of the simplest GRNs which allows to study such situation consists of two genes, one promoting differentiation, and the other promoting pluripotency (see [Fig 2\(a\)](#)). Nevertheless, in this section, we formulate and analyse our model considering a generic case with  $N_G$  genes. By doing so, our theoretical analysis can be further applied to any ER-GRN model, which implies a wide applicability of the derived formulation. However, when possible, we try to relate the theory developed to our particular ER-GRN so as to keep track of our case study.



**Fig 2. Schematic representation of the ER-GRN model and its multiscale reduction.** (a): Gene regulatory network (GRN) of two self-activating, mutually-inhibitory genes with epigenetic regulation (ER). In the GRN model, the gene product (single circle, denoted by  $X_i$  in [S2 Table](#)) is its own transcription factor which, upon dimerisation (two joint circles), binds the promoter region of the gene thus triggering gene transcription. The transition rates corresponding to this GRN are given in [S2 Table](#). For simplicity, we use an effective model in which the formation of the dimer and binding to the promoter region is taken into account in a single reaction, and the resulting number of promoter sites bound by two transcription factors is denoted by  $X_{ij}$  (see [S2 Table](#)). Furthermore, depending on whether the epigenetic state is open (i.e. predominantly acetylated (A)) or closed (i.e. predominantly methylated (M)) the promoter region of the gene is accessible or inaccessible to the transcription factor, respectively. (b): Schematic representation of the time separation structure of the multiscale method developed to simulate the ER-GRN system. See text and [S1 Text](#) for more details.

<https://doi.org/10.1371/journal.pcbi.1006592.g002>

## General description of the stochastic model of an epigenetically-regulated gene regulatory network

Consider a gene regulatory network composed of  $N_G$  self-activating genes which can repress each other. In particular, we consider that the gene product of each of these genes forms homodimers, which act as a transcription factor (TF) for its own gene by binding to its own promoter. Furthermore, each gene within the network has a number of inhibitors, which operate via competitive inhibition: the homodimers of protein  $j$  bind to the promoter of gene  $i$ , and by doing so they impede access of the TF to the promoter of gene  $i$ . In Fig 2(a), an illustrative scheme of the simplified case of two mutually inhibiting genes, one promoting pluripotency (blue) and one promoting differentiation (green), is shown. The regulation topology of the network can be represented by the binding rates of homodimers of protein  $j$  to the promoter of gene  $i$ ,  $b_{ij} > 0$  (see Fig 2(a)). Moreover, the expression of gene  $i$  is induced at a constant basal production rate,  $\hat{R}_i$ , independent of the regulatory mechanism described above. Proteins (TF monomers) of type  $i$  are synthesised at a rate proportional to the number of bound promoter sites with rate constant  $k_{i1}$  and degraded with degradation rate  $k_{i2}$  (see Fig 2(a) and S2 Table).

In addition to TF regulation, we further consider that each gene is under ER. ER controls gene transcription by modulating access of TFs to the promoter regions of the genes. In other words, in our model, ER is associated with an upstream drive that regulates gene expression [41]. Such epigenetic control is often related to alternative covalent modifications of histones. To address the high complexity of ER, we focus on a simpler stochastic model of ER, based on that formulated in [11, 16] and [17]. Our model belongs to a wider class of models which consider that single unmodified (U) chromatin loci can be modified so as to acquire positive (A) or negative (M) marks. Of such modifications, we consider methylation (associated with negative marks) and acetylation (associated with positive marks) [17]. An illustrative example on how epigenetic modifications alter the accessibility of TFs to the promoter regions of the genes is shown in Fig 2(a). Depending on whether the promoter region of a gene is mainly acetylated (A) or methylated (M), that promoter region is accessible (*open*) or inaccessible (*closed*) to the TFs, respectively. Both modifications are mediated by epigenetic enzymes: histone methylases (HMs) and demethylases (HDMs), and histone acetylases (HACs) and deacetylases (HDACs), which add or remove methylation and acetylation marks, respectively. Following [16], we explicitly account for HDM and HDAC activity only (see Fig 2(a)). In our model, a positive feedback mechanism is introduced whereby M marks help to both add more M marks and remove A marks from neighbouring loci [17]. The positive marks are assumed to be under the effects of a similar positive reinforcement mechanism [14, 17]. In this framework, each ER state is defined by the vector  $(Y_{i1}, \dots, Y_{i7})$ , describing the abundance of epigenetic marks and epigenetic enzymes at a given time. A full description of the details of the ER model are given in Section *Stochastic model of epigenetic regulation* of the S1 Text (see also [16]) and S3 Table, where the transition rates for the ER model are provided.

Under suitable conditions, determined by the activity and abundance of histone-modifying enzymes and co-factors, the positive reinforcement mechanism produces robust bistable behaviour. In this bistable regime, the two possible ER stable states are an *open* epigenetic state and a *silenced* epigenetic state. In the *open* epigenetic state, the levels of positive (negative) marks are elevated (downregulated). In this case, the promoter of the gene is accessible to TFs and transcription can occur. By contrast, in the absence (abundance) of positive (negative) marks, the gene is considered to be *silenced*, as TFs cannot reach the promoter.

An essential part of the stochastic dynamics of the ER system is the noise-induced transitions between the open and silenced states. Escape from steady states is a well-established



phenomenon (see e.g. [42]) and thoroughly analysed within the theory of rate processes [43] and large deviation theory [28, 44, 45]. As we will illustrate below, these noise-induced dynamics are essential to classify the epiphenotypes of somatic cells [16] and stem cells and to unravel the mechanisms of reprogramming and locking.

### Transitions between ER states: Minimum action path approach

Noise-induced transitions are essential to understand ER dynamics and their effect on cell-fate determination [46]. Throughout the bistable regime, sufficiently large fluctuations in the stochastic ER system will induce switching between the open and silenced states. The rate at which such transitions occur can be described using reaction-rate theory [43] and large deviation theory [44], which show that the waiting time between transitions is exponentially distributed. The average switching time,  $\tau_s$ , increases exponentially with system size, which in this case is given by the scale of ER substrates,  $Y$  [44, 45, 47, 48]:

$$\tau_s = C e^{Y\mathcal{S}}, \quad (1)$$

where  $C$  is a constant and  $\mathcal{S}$  is the minimum action of the stochastic switch. Eq (1) is derived from considering the probability distribution of the so-called *fluctuation paths*,  $\varphi(\tau)$ , in the ER space  $(Y_{i1}, \dots, Y_{i7})$ , which connect the mean-field steady states in a time  $\tau$ . According to large deviation theory [44, 45], we have  $P(\varphi(\tau)) \sim e^{-Y\mathcal{A}_{FW}(\varphi(\tau))}$ , which implies that the probability of observing paths different from the optimal, i.e. the path  $\varphi_*$  that minimises the Freidlin-Wentzel (FW) action functional  $\mathcal{A}_{FW}$ , is exponentially suppressed as system size,  $Y$ , increases. This means that, for a large enough system size, the behaviour of the system regarding large fluctuations is characterised by the optimal path:

$$\mathcal{S} \equiv \mathcal{A}_{FW}(\varphi_*) = \min_{\tau, \varphi(\tau)} \mathcal{A}_{FW}(\varphi(\tau)). \quad (2)$$

An explicit form of the functional  $\mathcal{A}_{FW}(\varphi(\tau))$  (see S1 Text) can be given if the dynamics is described by the corresponding chemical Langevin equation [49]. In this case, the optimal value of the action,  $\mathcal{S}$ , can be found by numerical minimisation, which provides both the optimal or minimum action path (MAP),  $\varphi_*$ , and the rate at which the ER system switches state driven by intrinsic noise [28]. Details regarding implementation of the action-optimisation algorithm are given in Section *Summary of the minimum action path theory and numerical method* of the S1 Text. A complete description of  $\tau_s$  requires to estimate the pre-factor  $C$  (see Eq (1)), which is not provided by the FW theory, but can be easily estimated using stochastic simulation.

### ER-system ensemble generation and parameter sensitivity analysis

In order to compute the switching time between the open and closed states (and vice versa) of the ER system, we should consider a particular ER system for each gene within the network. In order to mimic the existing ER heterogeneity within a set of cells from a particular tissue, we generated an ensemble of ER systems for each gene, which can be used to identify properties at population level. Such an ensemble is generated using approximate Bayesian computation (ABC) [50, 51], whereby we generate an ensemble of parameter sets  $\{c_{ij}|i = 1, \dots, N_G; j = 1, \dots, 16\}$ , with  $c_{ij}$  the kinetic parameter of the  $j$ -th reaction of the ER model for the gene  $i$  (see S3 Table), compatible with simulated data for the epigenetic regulation systems. The generated kinetic rate constants are dimensionless, i.e. they are relative to a given rate scale [16]. Such a feature implies that there is an undetermined time scale in our system. This additional degree of freedom can be used to fit our model of epigenetic (de-)activation to particular data. Since

the global time scales associated with different ER systems may differ among them, our model has the capability of reproducing different systems characterised by different time scales as previously shown by Bintu et al. [12].

Our approach follows closely that of [16], to which readers are referred to for a detailed presentation of the implementation. To summarise, we start by generating synthetic (simulated) data (denoted as “raw data” in S1 Fig) for the ER system of a genetic network epigenetically poised for differentiation, i.e. open differentiation-promoting genes and silenced pluripotency-promoting genes (see the example shown in S1 Fig). These raw data will play the role of the experimental data,  $x_0$ , to which we wish to fit our model. The raw data set consists of 10 realisations and 25 time points per realisation for each of the  $N_G$  epigenetic regulatory systems. For each time point,  $t_i$ , we consider two summary statistics: the mean over realisations,  $\bar{x}(t_i)$ , and the associated standard deviation,  $\sigma(t_i)$ . We then run the ABC rejection sampler method until we reach an ensemble of 10000 parameter sets which fit the raw data,  $x_0$ , within prescribed tolerances for the mean and standard deviation. We generate a 10000 parameter set ensemble of the ER system for the pluripotency-promoting gene and another 10000 parameter set ensemble for the gene promoting differentiation. In order to illustrate the effectiveness of the ABC method, S1(a) & S1(b) Fig show results comparing the reference (raw simulated) data to a subensemble average consisting of the 100 sets that best fit the data, for the differentiation- and pluripotency-promoting genes, respectively.

The above procedure provides us with an ensemble of parameter sets  $\{c_{ij}\}$  that are compatible with our raw data, i.e. such that they fit the data within the prescribed tolerances. The heterogeneity within this ensemble is compatible with existing biological variability in the activity of the different enzymes that carry out the epigenetic-regulatory modifications (HDMs, HDACs, as well as, histone methylases (HMs) and histone acetylases (HACs)), so that variation in  $\{c_{ij}\}$  can be traced back to heterogeneity in the availability of cofactors, many of them of metabolic origin such as NAD<sup>+</sup>, which are necessary for these enzymes to perform their function [16].

Our ensemble method also allows parameter sensitivity analysis regarding robustness of different ER behaviours, as described in detail in [16]. Once we have generated the ensemble, we identify subsets exhibiting certain properties of potential biological interest (e.g. mono- vs bistability) and perform a comparison between the parameter sets belonging to each of the subensembles, as well as a comparison of each subensemble with the whole ensemble. Such approach allows an identification of the essential parameters (via comparison of their empirical cumulative distribution functions (CDFs)) required by the system to exhibit the dynamic behaviour associated with a particular subensemble. A parameter is deemed as significant for a given behaviour when statistically significant differences can be detected between the subensemble-specific CDF and the CDF of other subensembles (or the whole population). The CDFs' shapes provide also useful information as to how the significant parameter should be changed for the ER system to switch a given behaviour. More details are given later in the *Results* Section.

## Coupling ER-GRN models: Multiscale analysis and model reduction

One should acknowledge that the system that results from coupling the ER and GRN models becomes rather cumbersome and computationally intractable as the GRN grows. We therefore took advantage of the intrinsic separation of time scales to analyse the behaviour of the resulting stochastic model [32–39]. Such approach allowed us to achieve meaningful model reduction via stochastic quasi-steady state approximations (QSSA) involving asymptotic analysis of the stochastic ER-GRN system [34–36, 38] (see Fig 2(b)). We assumed that the characteristic



scale for the number of TF monomers ( $S$ ), the number of promoter binding sites ( $E$ ), the number of ER modification sites ( $Y$ ), and the number of ER enzymes ( $Z$ ), are such that  $S \gg E$ ,  $Y \gg Z$  and  $\mathcal{O}(E) = \mathcal{O}(Y)$  (see [S1 Table](#) for the definition of these variables). It is noteworthy that the assumption  $Y \gg Z$  is exactly the Briggs-Haldane hypothesis for enzyme kinetics [52] because the ER modification sites are the substrates for the ER enzymes (see Section *Stochastic model of epigenetic regulation* in the [S1 Text](#)). The multiscale analysis and its technical details are provided in Sections *Multiscale analysis of the GRN system: WKB approximation and multiscale optimal path theory* and *Stochastic model reduction method* of the [S1 Text](#). Furthermore, the corresponding numerical method derived from this stochastic model reduction is described in [S1 Appendix](#).

Under appropriate assumptions regarding the characteristic scales of the different molecular species, our model exhibits a hierarchy of time scales, thereby allowing a model simplification and its computational simulation. Since  $S \gg E$  and  $Y \gg Z$ , the number of bound-to-promoter TFs and ER enzyme-substrate complexes are fast variables that can be sampled from their quasi-equilibrium distribution with respect (or conditioned to) their associated slow variables. TFs and ER modification sites (i.e., ER substrates) are slow variables whose dynamics, which dominate the long-time behaviour of the system, are given by their associated stochastic dynamics with the fast variables sampled from their quasi-steady state approximation (QSSA) probability density functions (PDFs).

The assumption that  $S \gg Y$  allows for additional simplification of the model, as it enables to take the limit of  $S \gg 1$  in the stochastic equations for the TFs monomers, which leads to a *piece-wise deterministic* Markov description where the dynamics of the number of TFs monomers is given by an ordinary differential equation (ODE) perturbed at discrete times by a noise source [36]. See [Fig 2\(b\)](#) for a schematic representation of the different techniques applied in this multiscale method leading to the numerical method applied in the *Results* section. After applying the asymptotic model reduction presented in detail in Section *Stochastic model reduction method* of the [S1 Text](#), the resulting model equations of the full ER-GRN system reduce to:

$$\frac{dx_i}{d\tau} = R_i + \omega_{i1}x_{i1} - \omega_{i2}x_i - 2 \sum_{j=1}^{N_G} \left( \beta_{ij} \eta_i \left( \frac{e_j}{E} - \sum_{k=1}^{N_G} x_{jk} \right) x_i^2 - \delta_{ij} x_{ij} \right), \quad i = 1, \dots, N_G \quad (3)$$

$$y_{ij} = y_{ij}(0) + \sum_{k=1}^{R_E} r_{E_{ijk}} \frac{1}{E} \mathcal{P} \left( E \frac{1}{\epsilon_2} \int_0^\tau v_{ik}(y_i(\sigma)) d\sigma \right), \quad j = 1, 2, 3, \quad (4)$$

with  $\eta_i$  coupling the system, as it is defined as  $\eta_i = H(y_{i3} - y_0)$ , i.e. gene  $i$  is epigenetically open if the corresponding level of acetylation  $y_{i3}$  exceeds the threshold  $y_0$ . The remaining variables and parameters are in [S1–S3 Tables](#).

The resulting dynamics consists of a coupled hybrid system where the dynamics of the TF monomers,  $x_i(\tau)$  ([Eq \(3\)](#)) is described in terms of a piece-wise deterministic Markov process [53, 54], i.e., by a system of ODEs perturbed at discrete times by two random processes, one corresponding to stochastic ER ([Eq \(4\)](#)) and the other to TF dimers binding to the promoter regions,  $x_{jk}$ . The latter are sampled from their QSSA PDFs ([Eq. \(S.51\) in the S1 Text](#)). The stochastic dynamics of the slow ER variables, [Eq \(4\)](#), is in turn coupled to the random variation of the associated fast variables (ER enzymes, HDM and HDAC, and complexes). The number of complexes,  $Y_{i5}$  and  $Y_{i7}$ , are also sampled from their QSSA PDFs ([Eqs. \(S.53\)–\(S.54\) in the S1 Text](#)). We refer to this method as a hybrid method, since it involves coupling both types of mathematical descriptions and its numerical implementation, namely, the coupling between

ODE solvers and stochastic simulation methods. The corresponding numerical method used to simulate such system is described in detail in [S1 Appendix](#).

The computational time needed by the Gillespie algorithm (used to perform the stochastic simulations) increases linearly with the number of reactions in the system. Our model reduction method increases computational efficiency by reducing the number of (stochastic) reactions involved in the dynamics of the system (as illustrated in Fig. D in the [S1 Text](#)). Whilst in the example been considered (i.e. a two-gene regulatory system) this reduction is significant, such increase in efficiency might not be of enough significance when using larger GRNs. In such cases, the model reduction method should be supplemented with a Next Reaction Method implementation of the stochastic dynamics of the fast variables [55] or, alternatively, by approximating such stochastic dynamics using either the  $\tau$ -leaping method [49] or the Chemical Langevin Equation (CLE) method [56]. In the latter case, a careful error analysis will be required when coupling the ODE solver for the slow variables, [Eq \(4\)](#), with the  $\tau$ -leaping or CLE simulations.

## Results

In order to focus our discussion, we study a gene regulatory circuit with two genes, one that promotes differentiation and another one that induces pluripotency. These two genes are further assumed to interact through mutual competitive inhibition and to be under the effects of epigenetic regulation (see [Fig 2\(a\)](#)).

We proceed to analyse how ER sculpts the epigenetic landscape over the substrate of the phase space given by the GRN model. The latter provides the system with a variety of cell fates, corresponding to the stable steady states of the dynamical system underpinning the model of gene regulatory network [58]. The transitions between such cellular states, both deterministic and stochastic, depend upon the ability of the cell regulatory systems to elevate or lower the barriers between them. Epigenetic regulation is one of such mechanisms able to alter these barriers. Here, we show that the intrinsic ER heterogeneity within the ensemble generated—understood to be originated by variations in the availability of the co-factors necessary for histone modifying enzymes (HMEs) to carry out their chromatin-modifying functions—suffices to produce a variety of behaviours, including differentiation-primed and stem-locked states.

### The GRN model exhibits a complex phase space, including an undecided regulatory state

We start our analysis by studying the phase space of the dynamical system underlying our model of gene regulation, schematically illustrated in [Fig 2\(a\)](#). Using the methodology described in detail in Section *Multiscale analysis of the GRN system: WKB approximation and multiscale optimal path theory* of the [S1 Text](#), we have derived the (quasi-steady state approximation) equations for the optimal path theory of the stochastic model of the mutually inhibitory two-gene system [11]. Such equations describe the most likely relaxation trajectories towards their steady states [47, 48], under conditions of time scale separation:

$$\frac{dq_1}{d\tau} = R_1 + p_{\infty 1} p \frac{\omega_{11} \frac{\beta_{11}}{\delta_{11}} q_1^2}{1 + \frac{\beta_{11}}{\delta_{11}} q_1^2 + \frac{\beta_{12}}{\delta_{12}} q_2^2} - \omega_{12} q_1 \quad (5)$$

$$\frac{dq_2}{d\tau} = R_2 + p_{\infty 2} p \frac{\omega_{21} \frac{\beta_{22}}{\delta_{22}} q_2^2}{1 + \frac{\beta_{21}}{\delta_{21}} q_1^2 + \frac{\beta_{22}}{\delta_{22}} q_2^2} - \omega_{22} q_2 \quad (6)$$

where  $q_1$  and  $q_2$  are the variables (generalised coordinates) associated with the number of

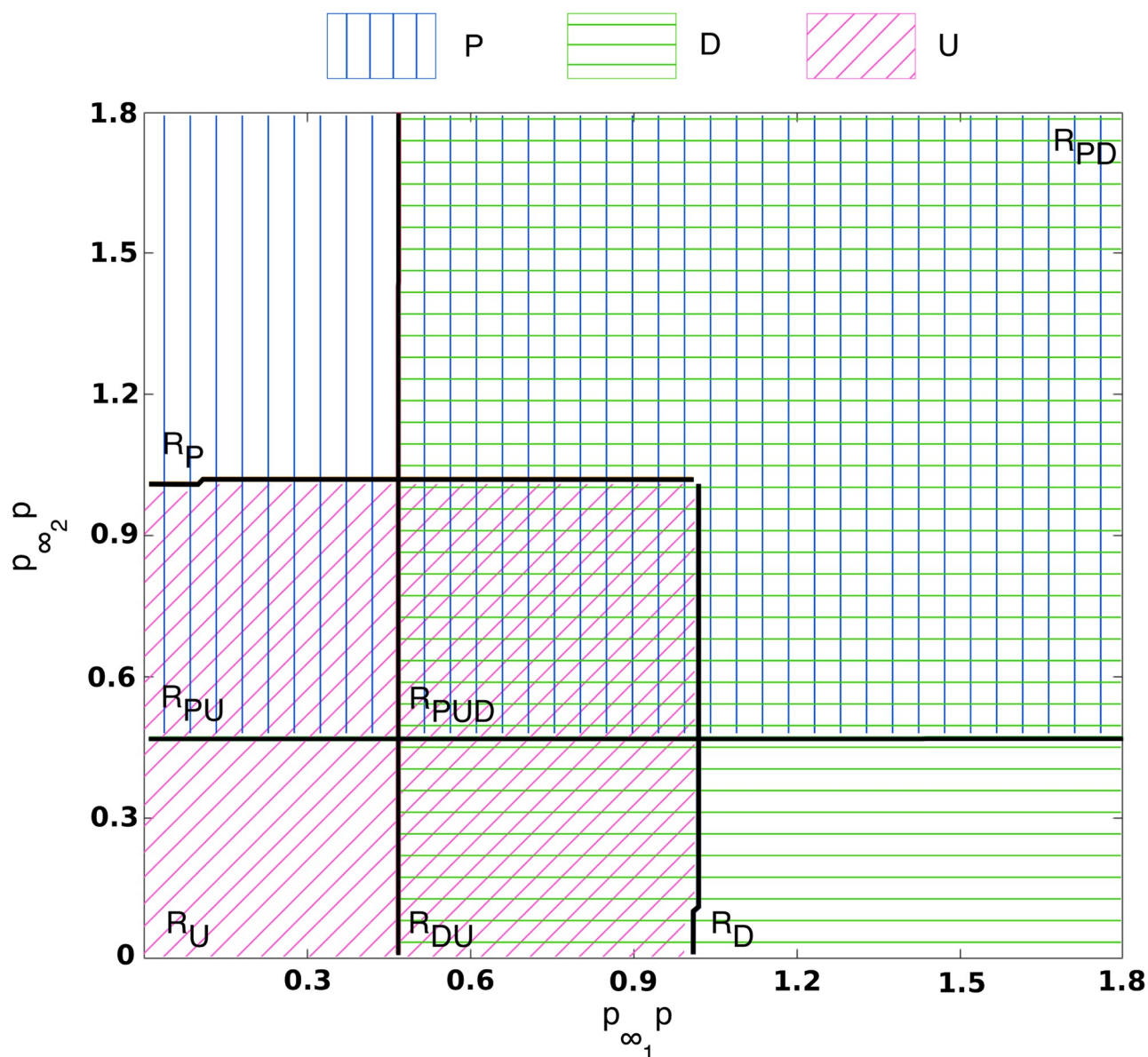
molecules of proteins,  $X_1$  and  $X_2$ , related to differentiation and pluripotent behaviours, respectively. The re-scaled variables,  $q_i$  and  $q_{ij}$ , and the re-scaled parameters,  $\omega_{ij}$ ,  $\beta_{ij}$ , and  $\delta_{ij}$ , are defined in [S2 Table](#).

The multiscale analysis carried out in Section *Multiscale analysis of the GRN system: WKB approximation and multiscale optimal path theory* of the [S1 Text](#) shows that the parameters  $p$ ,  $p_{\infty_1}$  and  $p_{\infty_2}$  are such that  $p_{\infty_1}p = \frac{e_1}{E}$  and  $p_{\infty_2}p = \frac{e_2}{E}$ , where  $e_1$  and  $e_2$  are the number of binding sites in the promoter region of the differentiation and pluripotency gene, respectively, which are exposed to and available for binding by TFs. This implies that  $p_{\infty_1}p$  and  $p_{\infty_2}p$  can be directly related to ER. Since  $E$  is a constant denoting the average number of binding sites at the gene promoter region, a change in the parameter value of  $p_{\infty_i}$  is due to a different value of the available number of binding sites,  $e_i$ . Hence,  $p_{\infty_i}p \rightarrow 0$  (i.e., few binding sites are available),  $i = 1, 2$ , corresponds to an epigenetically silenced gene, whereas  $p_{\infty_i}p \geq \mathcal{O}(1)$  (i.e., a large number of binding sites is available) associates with an epigenetically open gene. In this section, we study the phase space of the system when both  $p_{\infty_1}p$  and  $p_{\infty_2}p$  are varied. This allows us to understand how the behaviour of the GRN changes when its components are subject to ER. Our results are shown in [Fig 3](#).

The system described by Eqs (5) and (6) exhibits three types of biologically relevant stable steady states, namely, the *pluripotency* steady state (PSS), the *differentiation* steady state (DSS), and the *undecided* steady state (USS). The PSS (DSS) corresponds to a steady state with high (low) levels of expression of the pluripotency gene and low (high) levels of expression of the differentiation gene,  $q_1 \ll 1$  and  $q_2 = \mathcal{O}(1)$  ( $q_1 = \mathcal{O}(1)$  and  $q_2 \ll 1$ ), and the USS is associated with a state such that both genes are expressed at low levels, i.e.  $q_1 \ll 1$  and  $q_2 \ll 1$ . The existence of the latter state, the so-called “undecided”, is of particular interest because it closely relates to experimental results indicating that irreversible commitment to leave pluripotency and cell fate specification do not occur simultaneously [30, 59]. Therefore, such undecided state, which is characterised by low expression levels of pluripotency and differentiation genes, can be understood as a state where the cell has committed to leave pluripotency, indicated by the low expression levels of pluripotency related genes, but has not been committed yet to a particular differentiation fate, since differentiation-related genes are not up-regulated.

Different combinations of these states can coexist depending on the parameter values  $p_{\infty_i}p$  (see [Fig 3](#)), revealing a complex phase space with seven different regions: monostability regions ( $\mathcal{R}_U, \mathcal{R}_p, \mathcal{R}_D$ ), bistability regions ( $\mathcal{R}_{pU}, \mathcal{R}_{DU}, \mathcal{R}_{pD}$ ), and a tristability region ( $\mathcal{R}_{pUD}$ ). The lines shown in [Fig 3](#) correspond to the stability boundary of the different regimes. At such boundaries, saddle-node bifurcations occur, as illustrated in the example shown in [Fig B](#) (Section *Benchmark: stochastic model of a single self-activating gene* of the [S1 Text](#)). [S3 Fig](#) shows examples of trajectories illustrating the dynamics described by Eqs (5) and (6) for different values of the pair  $(p_{\infty_1}p, p_{\infty_2}p)$  corresponding to the different regions shown in [Fig 3](#). In particular, we show how the long term behaviour of different initial conditions differ as  $(p_{\infty_1}p, p_{\infty_2}p)$  varies, so that different cell fates (co)exist associated with different levels of TF accessibility.

The phase space shown in [Fig 3](#) illustrates the enormous relevance of ER, represented by the values of  $p_{\infty_i}p$ , which determine the pluripotent, differentiated or undecided cell fate. It is noteworthy that transitions between different cell fates can be achieved by simply altering the ER status, to which the GRN responds by jumping between different regions of the phase diagram ([Fig 3](#)). For example, by epigenetically silencing the differentiation gene (i.e., by decreasing  $p_{\infty_1}p$ ), the GRN system could move from a differentiated state ( $\mathcal{R}_D$  region) to an undecided one ( $\mathcal{R}_U$  region). If this epigenetically silencing is accompanied by an increase in the number of binding sites available at the promoter region of the pluripotency gene (i.e.,



**Fig 3. Phase diagram of the two-gene system, Eqs (5) and (6).** Vertical blue (horizontal green) hatching denotes regions where the pluripotency (differentiated) state is stable. Diagonal pink hatching denotes regions where the undecided state is stable. Regions of the phase diagram where different hatchings overlap correspond to regions of bistability or tristability. In the labels in the plot, *P* stands for pluripotency, *D* stands for differentiation and *U* for undecided. This phase diagram was obtained using the methodology formulated in [57]. Parameter values:  $\omega_{11} = \omega_{21} = 4.0$ . Other parameter values as per S7 Table.

<https://doi.org/10.1371/journal.pcbi.1006592.g003>

increase in  $p_{\infty 2}$ ), the GRN system can enter the  $\mathcal{R}_p$  region and reach a pluripotent cell fate. Indeed, there exist several possible transitions between cell fates by solely altering ER, thus highlighting its key role in switching fate. Its consequences for cell differentiation, cell reprogramming and cell locking is the focus of our study in the following sections.

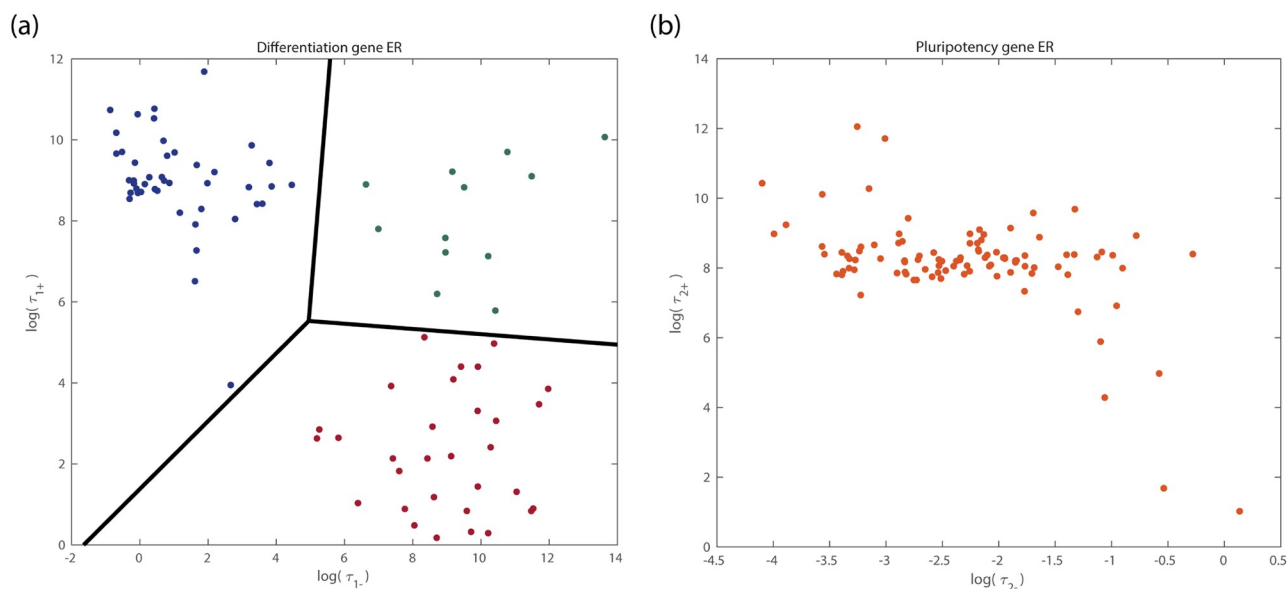
### Co-factor heterogeneity gives rise to both pluripotency-locked and differentiation-primed states

In the previous section, we have analysed the landscape (phase space) provided by the dynamical system describing the GRN. We now proceed to study the effect of ER on the robustness of

the different phases shown in Fig 3 (see also S3 Fig). ER is essential to the robustness of such phases and, consequently, to the stability of the associated cell fates. Stochastic transitions in bistable ER systems can induce (or facilitate) transitions between the GRN phases, which are associated with differentiation and reprogramming of cell fates. This phenomenon, so-called *epigenetic plasticity*, has been recently proposed as a major driver for disrupting cell-fate regulatory mechanisms in cancer and ageing [46]. We further focus on the role of heterogeneity within the ensemble of ER systems described in Section *Materials and methods* (see also [16]).

In order to characterise robustness of the different ER systems within the ensemble, we have focused on the analysis of the average transition times between the *open* and *closed* ER states. We perform this analysis for each differentiation ER system (DERS) and each pluripotency ER system (PERS). We define an ER system to be *open* (*closed*) when the acetylation levels are over (under) 90% (10%). We denote by  $\tau_{1+}$  ( $\tau_{1-}$ ) the average transition time for a DERS to switch from closed to open (open to closed). Similarly, the quantities  $\tau_{2+}$  and  $\tau_{2-}$  are analogously defined for the PERSs. The results are shown in Fig 4(a) and 4(b), where we present scatter plots of the average transition times within the ensemble of DERSs (Fig 4(a)) and PERSs (Fig 4(b)). Each point corresponds to an ER system (i.e. a given parameter set) within our ensemble. By inspection, we observe that the heterogeneity exhibited by the differentiation ER systems, showing large degrees of heterogeneity in both  $\tau_{1+}$  and  $\tau_{1-}$  (Fig 4(a)), is greater than the one corresponding to the pluripotency ER systems (Fig 4(b)). In particular, the dispersion in  $\tau_{2+}$  is much smaller than in  $\tau_{1+}$ , suggesting that DERSs have more variability in opening times than PERSs.

Heterogeneity in the differentiation ER systems exhibits an interesting pattern: DERSs organise themselves in three clusters obtained through *k*-means clustering, shown as blue,



**Fig 4. Scatter plots showing heterogeneity in the behaviour of bistable (a) differentiation ER systems (DERSs) and (b) pluripotency ER systems (PERSs).** The vertical axis corresponds to the average opening time and the horizontal axis, to the average closing time. Each dot in plot (a) represents a DERS within the ensemble (see Section *ER-system ensemble generation and parameter sensitivity analysis*). We analyse a total of 90 DERS parameter sets and 100 PERSs. The red cluster includes 31 sets, the green cluster contains 13 sets, and the blue cluster has 46 sets. Different colours and black lines show the three clusters resulting from a *k*-means analysis discussed in Sections *Co-factor heterogeneity gives rise to both pluripotency-locked and differentiation-primed states* and *Analysis of ensemble heterogeneity*. Dots in plot (b) represent PERSs within the ensemble defined in Section *ER-system ensemble generation and parameter sensitivity analysis*.

<https://doi.org/10.1371/journal.pcbi.1006592.g004>



green and red dots in Fig 4(a). DERSs within the *blue* cluster are characterised by long closed-to-open waiting times and short open-to-closed waiting times. DERSs belonging to the *red* cluster are the specular image of those within the blue cluster, i.e. they have short closed-to-open waiting times and long open-to-closed waiting times. Finally, DERSs in the *green* cluster are characterised by large values of both  $\tau_{1+}$  and  $\tau_{1-}$ .

Insight into the stochastic dynamics, particularly regarding transitions between open and silenced ER states due to intrinsic noise, can be gained by analysing the corresponding optimal escape paths,  $\phi$ , which by means of the MAP theory provide the transition times (see Eq (1)). To illustrate this, four examples of such paths, which were computed according to the MAP theory (see Section *Transitions between ER states: minimum action path approach*), for two DERSs and two PERSs of the corresponding generated ensembles are shown in S4 Fig. A comparison between the value of the minimum action,  $\mathcal{S}$  (see Eq (2)), associated with each of these systems, shows a tendency for DERSs to exhibit a larger degree of variability (see Table 1). Whilst the action value corresponding to the closed-to-open transition exhibits about a two-fold variability between PERSs, there is an over 8-fold increase when comparing the action values for this transition for DERSs. Similarly, when comparing the action  $\mathcal{S}$  for the open-to-closed optimal paths, we observe that the variability associated with the DERSs is also larger than the one in PERSs. This property partly explains the difference between Fig 4(a) and 4(b) regarding DERSs and PERSs heterogeneity, respectively. A similar argument can be put forward to help us explain the heterogeneity within the DERS ensemble (Fig 4(a)). Blue cluster DERSs (illustrated by DERS2 in Table 1) exhibit optimal closed-to-open paths with larger value of the optimal action than that found in their red cluster (depicted by DERS1 in Table 1) counterparts (see S4(a) Fig). This property has the consequence that the closed-to-open waiting time,  $\tau_{1+}$ , is longer for blue cluster DERSs.

To quantify the effects of bistable ER on the landscape related to the gene regulatory system (see Fig 3), we proceed to estimate the probability,  $Q$ , that the combined activity of each pair of DERS and PERS within our ensemble produces a global epigenetic regulatory state compatible with differentiation. DERS-PERS pairs with high values of  $Q$  are associated with *differentiation-primed* states. By contrast, those DERS-PERS combinations with low  $Q$  are identified with *pluripotency-locked* states.

Since escape times from a stable attractor in a stochastic multistable system are exponentially distributed [44, 45], the PDFs for the escape times for both DERSs and PERSs are fully determined by the corresponding values of  $\tau_{1\pm}$  and  $\tau_{2\pm}$ . We also assume that, for a given ER-GRN system, the DERS and the PERS evolve independently of each other.

We consider the PDF of the waiting time associated with a scenario of full remodelling of the epigenetic landscape,  $\tau_p$ . Such a scenario assumes that the system is initially in a pluripotency-locked ER state where the DERS is closed and the PERS is open. We denote such epigenetic state by  $D_- P_+$ . For the system to make its transit into the differentiation-primed state

**Table 1. Minimum action values,  $\mathcal{S}$ , corresponding to the optimal escape paths shown in S4 Fig (for details, see Section *Transitions between ER states: Minimum action path approach* and Section *Co-factor heterogeneity gives rise to both pluripotency-locked and differentiation-primed states* for details). Parameter values are given in S5 Table.**

ER system	Open to closed	Closed to open
DERS1	0.05387	0.007012
DERS2	0.09947	0.05813
PERS1	0.01502	0.1836
PERS2	0.02043	0.07645

<https://doi.org/10.1371/journal.pcbi.1006592.t001>

$D_+ P_-$ , corresponding to open DERS and closed PERS, there are two possible reprogramming routes:  $D_- P_+ \rightarrow D_- P_- \rightarrow D_+ P_-$  (route 1) and  $D_- P_+ \rightarrow D_+ P_+ \rightarrow D_+ P_-$  (route 2). Simultaneous switch of both ER systems is considered highly unlikely and therefore ignored. The PDF of the waiting time of the transition  $D_- P_+ \rightarrow D_+ P_-$ , denoted by  $P_{+,-,+}(\tau_p)$ , is given by:

$$P_{+,-,+}(\tau_p) = \mathcal{Z}^{-1}(P_1(\tau_p) + P_2(\tau_p)), \quad (7)$$

where

$$P_1(\tau) = \tau_{1-}^{-1} \tau_{2+}^{-1} e^{-\tau/\tau_{1-}} \left( \frac{e^{-\tau/\tau_{2+}} - e^{-\tau/\tau_{2-}}}{\tau_{2-}^{-1} - \tau_{2+}^{-1}} \right)$$

$$P_2(\tau) = \tau_{1-}^{-1} \tau_{2+}^{-1} e^{-\tau/\tau_{2+}} \left( \frac{e^{-\tau/\tau_{1+}} - e^{-\tau/\tau_{1-}}}{\tau_{1-}^{-1} - \tau_{1+}^{-1}} \right),$$

and

$$\mathcal{Z}^{-1} = \frac{(\tau_{1-} + \tau_{2+})(\tau_{1-}^{-1} + \tau_{2-}^{-1})(\tau_{1+}^{-1} + \tau_{2+}^{-1})}{\tau_{2+}^{-1} + \tau_{1+}^{-1} + \tau_{2-}^{-1} + \tau_{1-}^{-1}}.$$

$P_1(\tau_p)$  and  $P_2(\tau_p)$  are the probabilities related to each of the landscape reprogramming routes. The probability that the ER landscape has undergone reprogramming from pluripotency-locked to differentiation-primed state within the time interval  $(0, \tau_p]$ ,  $\mathcal{Q}$ , is thus given by:

$$\mathcal{Q} \equiv \int_0^{\tau_p} P_{+,-,+}(\tau) d\tau, \quad (8)$$

where in our case,  $\tau_p$  has been taken as the mean ensemble time for the differentiation ER systems (DERSs) to switch from the closed to the open state,  $\tau_{1+}$ . Furthermore,  $\tau_{1+}$  exhibits a greater range of variability than the time for the pluripotency ER systems to switch from its open to its closed state, which is also a necessary condition for the epigenetic remodelling to take place.

We investigate the DERSs belonging to the different clusters of Fig 4(a) regarding their likelihood to produce pluripotency-locked epigenetic landscapes (results shown in S2 Fig). The analysis shows that when a DERS within the red cluster is paired with any PERS (S2(c) Fig), the resulting system corresponds to a differentiation-primed epigenetic landscape ( $\mathcal{Q} = 1$ ). By contrast, when a PERS is paired with DERSs from the blue cluster (S2(a) Fig) and the green cluster (S2(b) Fig), both differentiation-primed (large  $\mathcal{Q}$ ) and pluripotency-locked (small  $\mathcal{Q}$ ) epigenetic landscapes are obtained. As discussed in the next section, the latter are more likely within the blue cluster.

## Analysis of ensemble heterogeneity

We now proceed to analyse the patterns observed in our ensemble of ER systems regarding both the differences between the three clusters observed in the ensemble of DERSs (Fig 4(a)) and the distinctive features that characterise pluripotency-locked DERS-PERS pairs. We followed the methodology put forward in [16], whereby statistics (in our case, cumulative distribution functions (CDFs)) of a subensemble of systems exhibiting a particular behaviour are analysed. We focused on the study of the CDFs of the kinetic parameters of the ER reactions,  $c_{ij}$  (see S3 Table), belonging to the DERSs/PERSs associated with the relevant behaviour aiming

to characterise. In our case, such subensembles are either the different clusters, or they are determined by whether they exhibit low or high value of  $Q$ . By comparing such CDFs to either the general population (i.e. whole ensemble) or to different subensembles, we can detect statistically significant biases, which allows us to identify key parameters (and their biases) associated with the behaviour displayed by the subensemble under consideration.

**Significant differences within the ensemble of DERSs.** We start this analysis by studying the patterns emerging in the ensemble of DERSs, Fig 4(a). As discussed in the previous section, DERSs organise themselves in three clusters, which differ regarding their capability to trigger differentiation-primed epigenetic landscapes (see Fig 4(a) and S2 Fig). Our results are shown in Fig 5, where we depict the empirical CDFs for the relevant kinetic parameters of the ER reactions for the differentiation gene,  $c_{1_j}$  (see S3 Table), i.e. those  $c_{1_j}$  exhibiting statistically significant differences when comparing the CDFs of the clusters (red, green, blue) among them (see Fig 5). Each of these two-sample comparison is carried out by means of the Kolmogorov-Smirnov (KS) test. Statistically significant differences were found in the cases we comment below. The  $p$ -values are reported in Section *Analysis of ensemble heterogeneity: significant differences* of the S1 Text. The remaining CDFs for the ER reactions of the differentiation gene are given in S5 Fig.

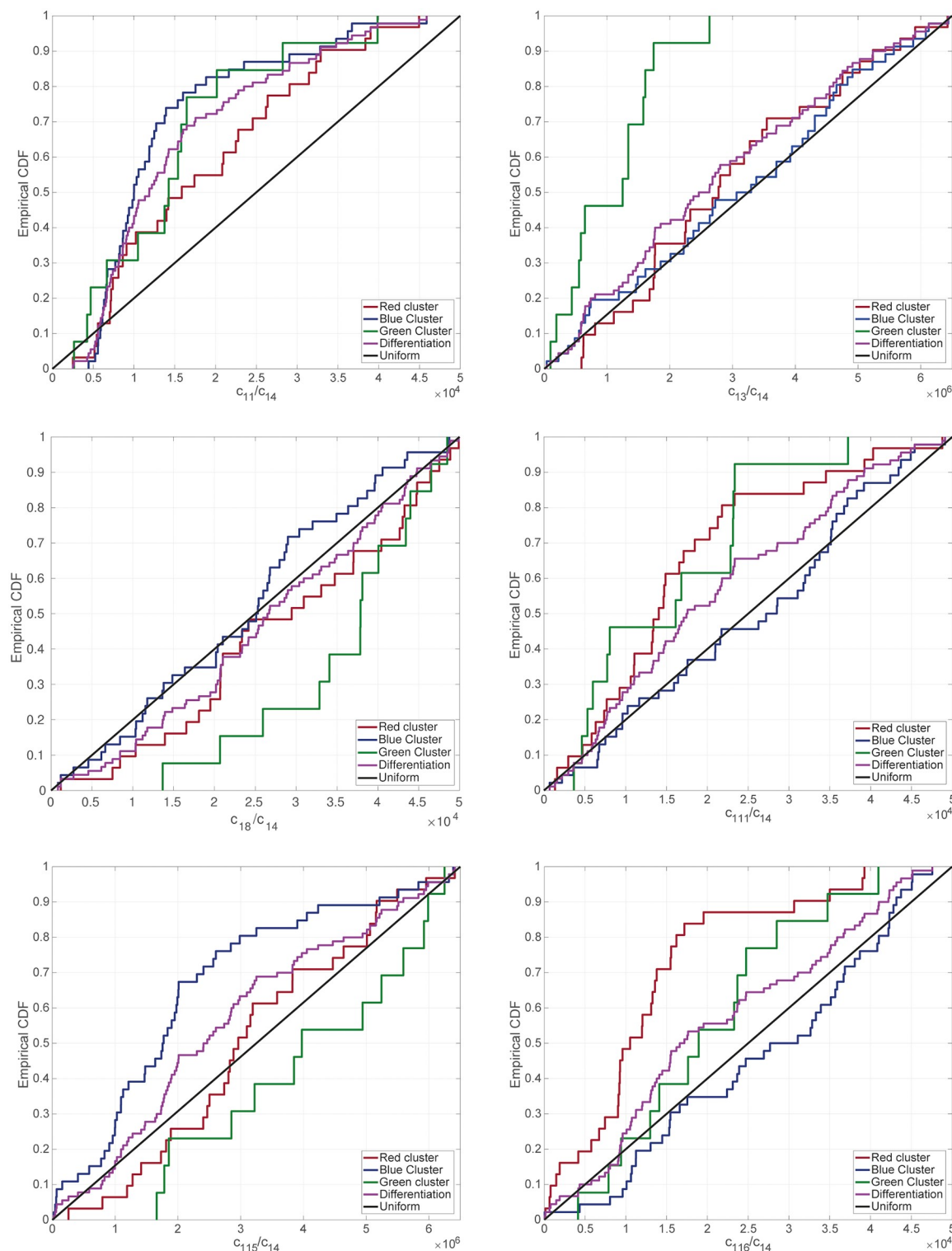
*Red cluster versus blue cluster.* As discussed in the previous section, the differences between DERSs within the blue and red clusters are essential to ascertain the main features that distinguish differentiation-primed and pluripotency-locked systems. The bias detected within the red (blue) cluster in the corresponding CDFs (see Fig 5) is towards bigger (smaller) values for  $c_{1_1}$  (unrecruited demethylation) and  $c_{1_{15}}$  (unrecruited acetylation) and towards smaller (larger) values for  $c_{1_{11}}$  (unrecruited deacetylation) and  $c_{1_{16}}$  (recruited acetylation). The behaviour of  $c_{1_1}$ ,  $c_{1_{11}}$ , and  $c_{1_{15}}$  is straightforward to interpret. The trends observed in the data are consistent with the DERSs within red cluster being more prone to differentiation-primed ER landscapes, as they promote removal of negative (methylation) marks and addition of positive (acetylation) marks.

*Red cluster versus green cluster.* In this case, the bias detected within the red (green) cluster in the corresponding CDFs (see Fig 5) is to larger (smaller) values for  $c_{1_3}$  (unrecruited demethylation) and to smaller (bigger) values for  $c_{1_{16}}$  (recruited acetylation). The tendency in the data corresponding to  $c_{1_3}$  is compatible with the features of the red cluster DERSs, as it involves an increase in the removal of negative marks.

*Blue cluster versus green cluster.* Fig 5 shows that DERSs within the green cluster have smaller values of  $c_{1_3}$  (unrecruited demethylation) and larger values of  $c_{1_8}$  (recruited methylation) than their blue cluster counterparts. Both of such effects stimulate addition of negative marks. However, DERSs in the green cluster also exhibit lower  $c_{1_{11}}$  (unrecruited deacetylation) and bigger  $c_{1_{15}}$  (unrecruited acetylation), which both encourage addition of positive marks. This can explain why the green cluster DERSs exhibit both long  $\tau_{1-}$  and  $\tau_{1+}$  (see Fig 4(a)).

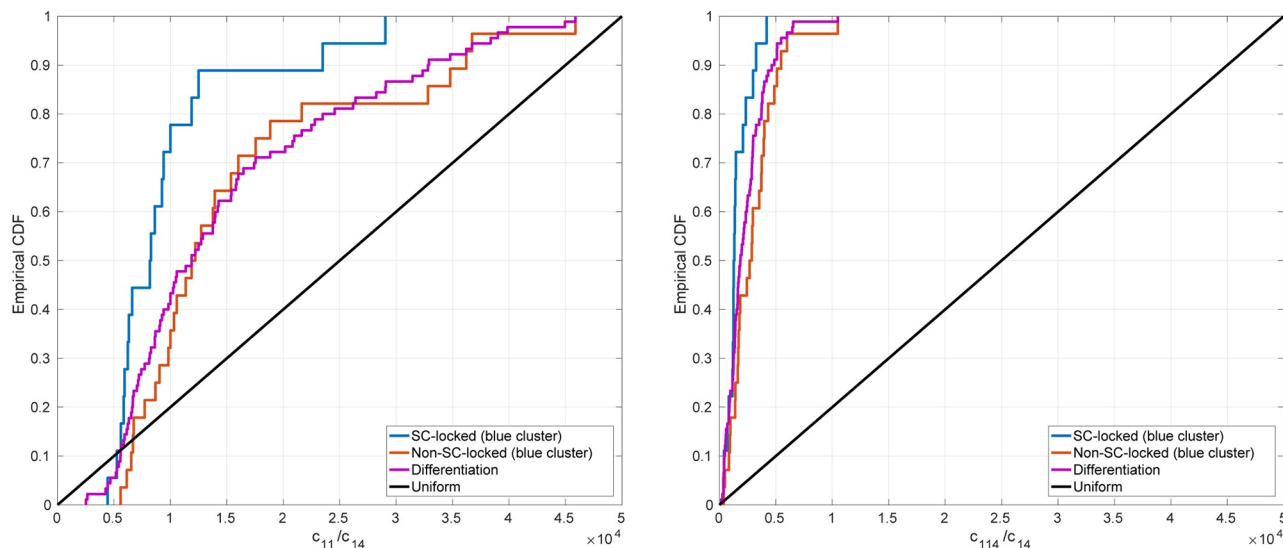
**Significant differences between differentiation-primed and pluripotency-locked ER landscapes.** The quantity  $Q$  allows us to classify each pair DERS-PERS drawn from our ensemble regarding their degree of resilience to switch into a state prone to differentiation. If  $Q$  is larger than a threshold value  $T$ , the corresponding DERS-PERS pair is categorised as differentiation-primed. By contrast, when  $Q < T$ , the DERS-PERS pair is classified as pluripotency-locked.

We first proceed to compare within the whole population (without discriminating among clusters) those DERSs such that  $Q \geq T$  (differentiation-primed ER landscapes) against those with  $Q < T$  (pluripotency-locked ER landscapes). We take  $T = 0.7$ . The results are



**Fig 5. Empirical CDFs for the whole ensemble of DERS parameter sets (magenta lines) for those  $c_{1j}$  significantly different between clusters.** This ensemble has been generated according to the methodology explained in Section *ER-system ensemble generation and parameter sensitivity analysis* (see also [16]). We also show the partial empirical CDFs corresponding to each of the clusters from Fig 4(a) (red, green, and blue lines). For reference, we also show the CDF for a uniform distribution (black line).

<https://doi.org/10.1371/journal.pcbi.1006592.g005>



**Fig 6. Empirical CDFs for the DERS parameter sets within the blue cluster for those  $c_{ij}$  significantly different.** This ensemble has been generated according to the methodology explained in Section *ER-system ensemble generation and parameter sensitivity analysis* (see also [16]). The DERSs within the blue cluster have been divided into two subsets: those such that  $Q < T$  (SC-locked, blue lines) and those such that  $Q \geq T$  (non-SC-locked, orange lines), with  $T = 0.7$ . For comparison, we plot the CDFs of the whole DERS ensemble (magenta lines), and, for guidance the CDF corresponding to a uniformly distributed random variable (black lines).

<https://doi.org/10.1371/journal.pcbi.1006592.g006>

shown in S6 Fig. The CDFs of the parameters  $c_{11}$  (unrecruited demethylation),  $c_{114}$  (recruited deacetylation), and  $c_{115}$  (unrecruited acetylation) are biased towards higher values for the sub-ensemble associated with differentiation-primed ER landscapes ( $Q \geq T$ ). The requirement for  $Q$  to be  $Q \geq T$  biases the CDF of  $c_{116}$  (recruited acetylation) towards lower values than in the general population. The interpretation of the results regarding  $c_{11}$  and  $c_{115}$  is clear, since they encourage the removal of negative marks and the addition of positive marks, respectively, and thus promote expression of the differentiation gene. The CDFs of  $c_{114}$  and  $c_{116}$  corresponding to differentiation-primed ER landscapes are virtually identical to the CDFs associated with the general population (see S6 Fig). These features are therefore inherent in bistable behaviour (see Section *General description of the stochastic model of an epigenetically-regulated gene regulatory network*), rather than being specific to differentiation-primed DERSs.

If we now restrict our analysis to those DERSs within the blue cluster (see Fig 6), we observe that the parameters whose CDFs differ significantly when splitted into differentiation-primed and pluripotency-locked are  $c_{11}$  (unrecruited demethylation) and  $c_{114}$  (recruited deacetylation). As in the analysis in the whole ensemble, only the result regarding  $c_{11}$  is relevant for the analysis of the features yielding differentiation-primed ER landscapes. The remaining CDFs for those  $c_{ij}$  not exhibiting significant differences when comparing those DERSs within the blue cluster giving rise to differentiation-primed behaviour to those behaving like pluripotency-locked, are given in S7 Fig.

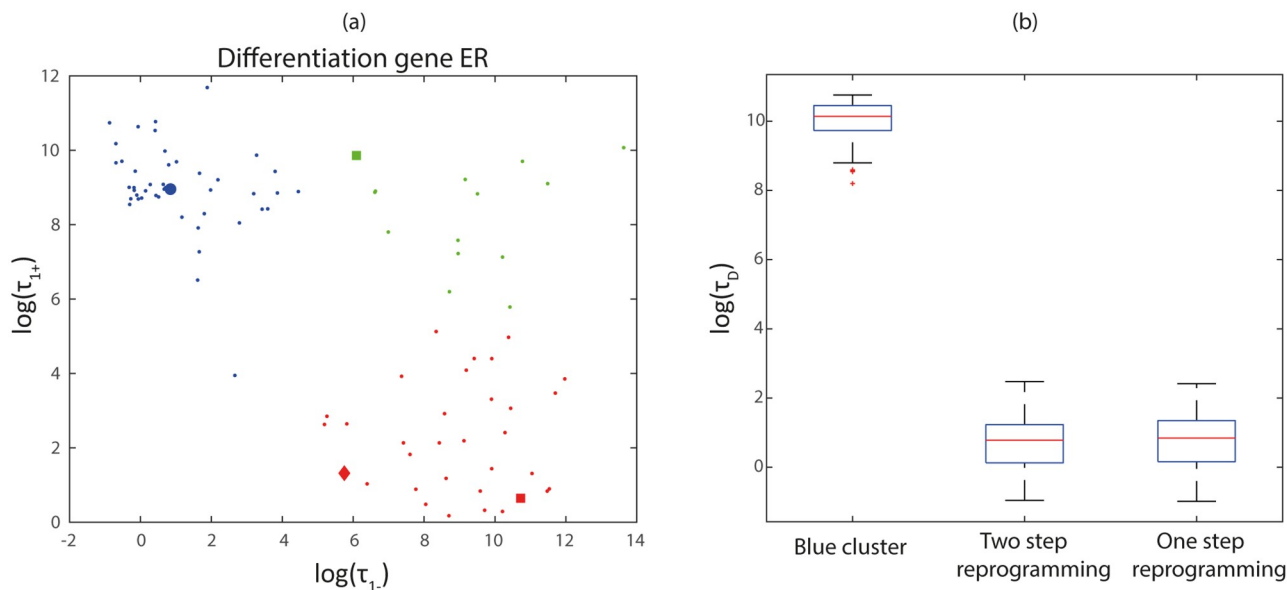
Regarding the PERSs, the results are less compelling. The results are shown in S8 Fig. Our analysis shows that significant differences can be found between the empirical distributions of three parameter values:  $c_{13}$  (unrecruited demethylation),  $c_{18}$  (recruited methylation), and  $c_{115}$  (unrecruited acetylation). PERSs such that  $Q \geq T$  exhibit larger values of all three parameters.



## Ensemble-based strategies for unlocking resilient pluripotency

Our analysis has illustrated which are the key  $c_{ij}$  giving rise to either pluripotency-locked or differentiation-primed landscapes, immediately suggesting a number of strategies to unlock resilient pluripotency states with hindered differentiation. One of our main conclusions is that states of resilient pluripotency are mostly vinculated to DERS-PERS combinations such that the DERS belongs to either the blue or the green cluster, since they are the ones exhibiting pluripotency-locking (low  $Q$ ). In view of this, a possible strategy in order to encourage differentiation-primed ER landscape consists in changing a selected combination of parameter values according to a rationale provided by the analysis carried out in the previous sections. Our results are shown in Fig 7.

One possible strategy consists in first transforming a blue cluster DERS into a green cluster one, and then completing the DERS *reprogramming* by transforming the resulting set into a red cluster DERS, where all DERSs exhibit differentiation-priming. A candidate strategy involves first changing a parameter whose CDF is significantly different when the blue cluster is compared with the green cluster. The second step is then to change a parameter that exhibits significant difference between the green and red cluster. Taking the results of the previous section into consideration, we consider the reduction of  $c_{111}$  (unrecruited deacetylation) and the increase of  $c_{13}$  (unrecruited demethylation). The result of this reprogramming strategy is shown in Fig 7(a), where we show that a blue cluster DERS is first transformed into a green cluster one (green square in Fig 7(a)), and then, finally, into a red cluster DERS (red square in Fig 7(a)). The initial blue cluster DERS has been chosen as the set with the largest value of  $c_{111}$ , which has been shown to be a significant difference when comparing the blue cluster to the red one, and the blue cluster to the green one, leading to the idea that this property is linked to



**Fig 7. Effect of the different reprogramming strategies of blue cluster DERSs, as evaluated in terms of the statistics of the differentiation time ( $\tau_D$ ).** (a) Two step reprogramming is illustrated by the green square (first step), which finally becomes the red square (second step). One step reprogramming is depicted as the red diamond (see Section *Ensemble-based strategies for unlocking resilient pluripotency* for details). (b) Comparison of  $\tau_D$  for the original DERS and the ones resulting from the reprogramming strategies. We consider a base-line scenario where the number of HMEs is exactly equal to average, i.e.  $e_{HDM} = e_{HDAC} = Z$ . We then compare the simulation results obtained for different scenarios regarding the different strategies to the base-line scenario. Parameter values:  $Z = 5$  and  $Y = 15$ . Other parameter values given in S6 Table.

<https://doi.org/10.1371/journal.pcbi.1006592.g007>

the blue cluster (idea which is reinforced because  $c_{111}$  is not significantly different when comparing the red and the green cluster).

The efficiency of such a strategy to unlock resilient pluripotency is shown in Fig 7(b) where we present statistics of the differentiation time,  $\tau_D$ , for the original blue cluster DERS and for the corresponding reprogrammed one (two step reprogramming, *red cluster-like*). These simulations have been done for the full ER-GRN, using the hybrid multiscale simulation algorithm described in S1 Appendix (and fully developed in Sections *Multiscale analysis of the GRN system: WKB approximation and multiscale optimal path theory* and *Stochastic model reduction method* of the S1 Text). The resulting differentiation times for the ER-GRN with reprogrammed ER landscape are orders of magnitude smaller than those with original ER-GRN within the blue cluster DERS, therefore confirming the success of our strategy.

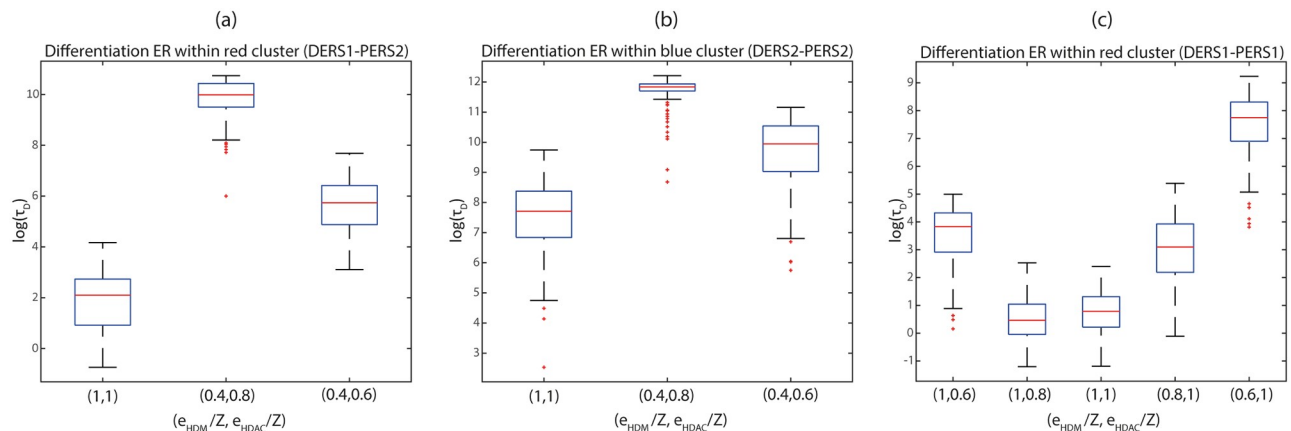
An alternative strategy, that involves changing the value of one parameter only, consists in increasing the value of  $c_{13}$  (unrecruited demethylation). Such a strategy is not obvious, since  $c_{13}$  is not one of the parameters whose empirical CDF has significant differences when DERS in the red cluster are directly compared with those in the blue cluster. However, since the CDF of  $c_{13}$  is significantly different when both the blue cluster and the red cluster are compared to the green cluster, it is conceivable that increasing  $c_{13}$  without further intervention could reprogram blue cluster DERSs. The result of this reprogramming strategy is shown in Fig 7(a) (red diamond). Simulation results shown in Fig 7(b) (two step reprogramming) confirm the viability of this approach. In fact, based on the statistics of the differentiation time, both strategies are virtually indistinguishable.

### Loss of HDAC activity hinders differentiation in our ER-GRN model

Besides variability associated with cofactor heterogeneity, represented in the ensemble of  $c_{ij}$  values, our model allows us to address the issue of variability regarding HME activity. HMEs are needed for the acetylation and methylation epigenetic modifications to take place, and their activities are known to be affected by physiological and pathological processes, including ageing and cancer. Here, we analyse the impact of HDM and HDAC loss of activity on the dynamics of differentiation. In particular, we simulate differentiation in our ER-GRN model to obtain statistics of the differentiation time to assess the effect of loss of HME activity. The simulations shown in this section have all been carried out using the hybrid multiscale simulation algorithm described in the S1 Appendix.

In order to clarify the effect of loss of HME activity on the ER model, we first consider the phase diagram of its mean-field limit in different situations (see [16] for details). This phase diagram depicts the closed, bistable and open region for a given ER system, using HMEs activity as parameters (HDM in the  $x$ -axis, HDAC in the  $y$ -axis). The results are shown in S9 Fig. The surface occupied by the bistable region (shaded blue region in S9 Fig) is much larger in blue cluster than in red cluster DERSs, because of the displacement of the boundary separating the closed and bistable behaviour. By comparison, the bistability region of the PERSs is narrower than that of the DERSs (see S9(b) and S9(d) Fig). In particular the boundary that separates the bistable phase from the closed phase (area at the left of the blue shaded region) is displaced towards smaller HDM activity (i.e., to the left) in the DERSs phase diagrams.

This property suggests that a possible strategy to promote a differentiation-primed landscape ( $D_- P_+ \rightarrow D_+ P_-$ ) would be to decrease HDM activity, as this would drive the PERSs into its closed phase whilst allowing the DERSs to remain within its bistability region. In order to assess this, we consider a base-line scenario where the number of HMEs is exactly equal to average, i.e.  $e_{HDM} = e_{HDAC} = Z$ . We then compare different scenarios regarding the abundance of HDM and HDAC to the base-line scenario.



**Fig 8. Plots showing the effect of the variation of HDM and HDAC on the statistics of the differentiation time ( $\tau_D$ ).** We consider a base-line scenario where the number of HMEs is exactly equal to average, i.e.  $e_{HDM} = e_{HDAC} = Z$ . We then compare the simulation results obtained for different scenarios regarding the abundance of HDM and HDAC to the base-line scenario, i.e. by changing the values of  $(e_{HDM}, e_{HDAC})$ . Parameter values:  $Z = 5$  and  $Y = 15$ . Other parameter values given in S5 Table.

<https://doi.org/10.1371/journal.pcbi.1006592.g008>

Contrary to what could be expected, simulation results show that the strategy of reducing HDM activity alone beyond the PERSs closing boundary further hinders differentiation. As can be seen in Fig 8(c), a decrease in HDM activity actually leads to longer differentiation time (see also [11]). Similarly, Fig 8(a) and 8(b), which show statistics of the differentiation time, reveal that a decrease in both HDM and HDAC activity also leads to an increment in differentiation times, that is, this strategy fails to decrease the differentiation time below the base-line scenario. In both cases, such hindrance of differentiation is the product of the increase in the opening times ( $\tau_{1+}$ ) of the DERSs. This effect occurs because, as HDM and HDAC activity is reduced, the DERSs are driven towards their closed-bistability boundary. Close to such a region, the DERSs closed state becomes more stable and thus the corresponding  $\tau_{1+}$  increases. By contrast, further reduction of HDAC activity moves the DERSs system closer to their bistable-open boundary, resulting in a reduction of the differentiation time. However, since the differentiation times remain above those corresponding to the base line HDM and HDAC activity scenario, we conclude that loss of both HDM and HDAC activity contributes towards hindering differentiation. Therefore, these results suggest that downregulation of HDM and HDAC activities, which has been observed in cancer and ageing, respectively, locks the ER landscapes in states more resilient to differentiation, as the differentiation time increases. This, in turn, is consistent with the theory that postulates that ageing and cancer may affect the ER control of cell fate, by locking cells into states disabled to differentiate and consequently, prone to malignant transformation.

## Discussion

Epigenomic remodelling in response to cellular reprogramming can be viewed as a paradigmatic strategy capable of erasing the hallmarks of ageing at the molecular and cellular level [1, 3, 4, 8]. However, undesirable *trade-off* constraining phenotypes such as impairment of tissue repair/wound-healing, tissue dysfunction due to loss of cell identity, and tumorigenesis could also occur if such epigenetic remodelling is not accompanied by an adequate self-repair of injury or disease [7]. In this regard, our study provides mathematical and computational answers to one of the ageing research field's biggest challenges, namely, the understanding of

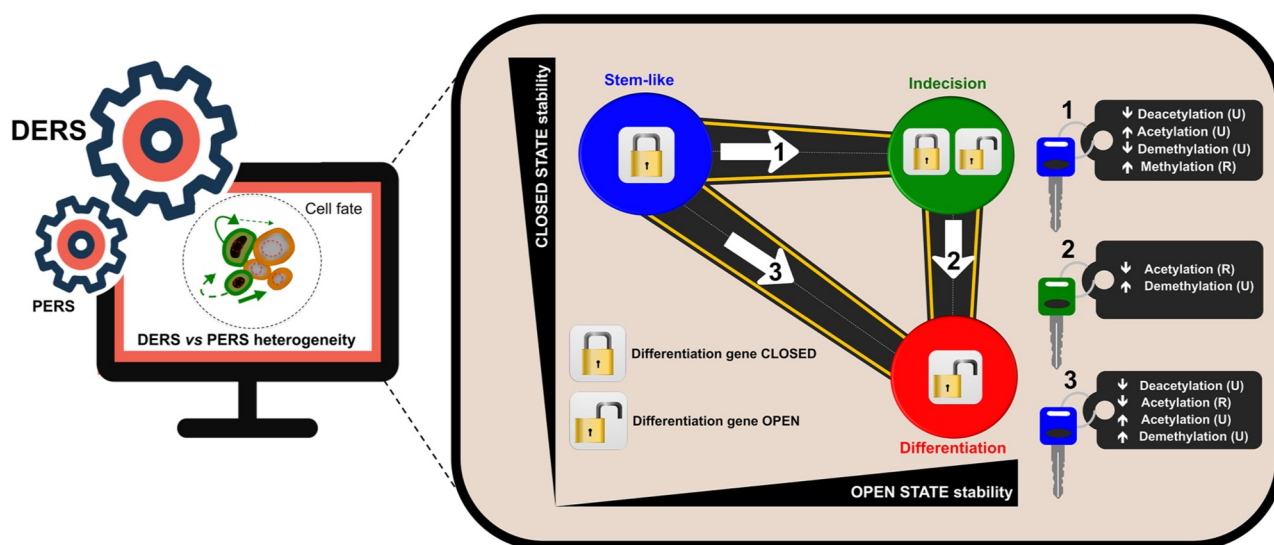
how epigenetic heterogeneity could operate as the fundamental driver of the beneficial versus deleterious effects of cellular reprogramming.

We are rapidly amassing evidence that, beyond *bona fide* genetic alterations, non-genetic stimuli such as inflammation, hypoxia, cell stress, and developmental and metabolic cues, can promote overly restrictive epigenetic states—capable of preventing the induction of tumour suppression programmes or blocking normal differentiation—or overly plastic epigenetic states—capable of stochastically activating oncogenic programmes and non-physiological cell fate transitions including those leading to the acquisition of stem cell-like states [16, 46]. Indeed, resilience and plasticity begin to be considered as the fundamental epigenetic dimensions that ultimately dictate the capacity of cells, tissues, and organs to undergo successful repair, degeneration, or malignization via phenotypic variation. In our model, the robustness or resilience of the cell phenotype attractors throughout the epigenetic landscape was determined by ER; then, a framework for the generation of the ensemble of ER systems allowed an ulterior analysis, whereas a multiscale asymptotic analysis-based method for model reduction enabled us to formulate an efficient numerical scheme to study the behaviour of the stochastic ER-GRN system. Our approximation, which is closely related to the notion of neutral networks formulated to analyse systems with genotype-phenotype maps [60–62], is applicable to broader scenarios because, by reducing a rather complex stochastic system into a hybrid, piece-wise deterministic Markov one, it is capable of providing an efficient and scalable, hybrid numerical method able to simulate more complex ER-GRN systems.

In order to determine the key mechanisms underlying epigenetic plasticity and its connections with aberrant stem cell-like locked states, we have considered a gene network model of two mutually-inhibiting genes regulating the phenotypic switch between differentiated and pluripotent states. Each gene within this regulatory system was acted upon by ER to restrict/enable its expression capability. Although it might be argued that such a system is too simplistic to describe realistic frameworks, one should acknowledge that mutual inhibition between two key transcription factors has been shown to control binary cell fate decisions in a number of biologically relevant situations [30, 63]. In addition, this system serves as a general tool to understand generic features of the role of multistability in more complex cell fate decision systems. Specific examples include lateral stabilization during early patterning in the pancreas (Ngn3–Ptf1a) [64], promotion of differentiation to trophectoderm in mammalian angiogenesis (Cdx2–Oct3/4) [65], cellular reprogramming (Oct4–Sox2) [11], and haematopoiesis (GATA1–PU.1) [66]. Furthermore, our approach might serve as a general tool that can be applied in a straightforward manner to adequately evaluate the epigenetic-regulatory features involved in the multistability of larger, more complex cell fate decision systems. In this regard, it should be noted that the robustness of the open/closed epigenetic states was assessed in terms of the average transition times and that the structure of DERSs/PERSs clusters was independent of the GRN dimensionality. Therefore, robustness analyses of more complex ER systems would be carried out by merely sampling DERSs/PERSs parameters from the generated ensemble. Nonetheless, even in our relatively simple case of a gene regulatory circuit involving solely two genes, the behaviour of the mean-field limit of the GRN exhibited a complex space with a tristability regime, which included not only the expected stem-locked and differentiated steady-states but also the so-called indecision state. From a developmental perspective, the latter state could serve the purpose of priming cells for differentiation. Perhaps more importantly, the transitions between the different phases could be triggered by changes related to ER (i.e., cofactors of chromatin-modifying enzymes), which thereby operate as *bona fide* molecular bridges that directly connect epigenetic and phenotypic plasticity via translation of changes in ER states into variations of GRN states.

Earlier studies provided evidence that the *de novo* reprogramming potential is higher within select subpopulations of cells and that such pre-existing epigenetic heterogeneity can be tuned to make cells more responsive to reprogramming stimuli [16, 40]. Along this line, intra- and inter-individual variability driven by the local interpretation of metabolic, epigenetic, and inflammatory regulators might not only reflect the occurrence of different ageing trajectories in different tissue cell subpopulations but might determine also the *de novo* responsiveness to therapeutic strategies aimed to remodel the organismal self-repair capacity for resistance to damage, stress, and disease [16, 67, 68]. In this regard, by uncovering the regulatory details of the phenotypic robustness of stem-like epi-states, we have been able to mathematically capture how epigenetic heterogeneity governs the routes and kinetics to entry and exit from unrestrained epigenetic plastic states. Thus, a sub-ensemble of ER systems with higher reprogramming potential was found to pre-exist within the ensemble of ER systems compatible with a terminally differentiated cell state; moreover, such a sub-ensemble could be harnessed to fine-tune the cellular response to reprogramming-to-stemness stimuli by solely targeting chromatin-modifying enzymes such as HDMs and HDACs, thus confirming and extending earlier experimental approaches [69]. It is reasonable to propose that epigenetic heterogeneity is a/the central regulator through which epigenetic plasticity allows cells to stochastically activate alternative regulatory programs and undergo distinct cell fate transitions. As such, epigenetic heterogeneity is largely responsible for the mechanistic dynamics determining the phenotypic robustness of cell fate reprogramming (see Fig 9).

A disruption of the homeostatic resilience of chromatin, causing it to become aberrantly restricted or permissive, has the potential to give rise to each classic cancer hallmark [46]. Intriguingly, a similar disruption of the entry-exit paths and kinetics of the endogenous injury-repair mechanisms appears to be also the convergent *trade-off* of a variety of strategies (e.g., metabolic manipulation, ablation of senescent cells, and cellular reprogramming)



**Fig 9. Strategies to unlock pluripotent stem-like states in ageing and cancer.** Epigenetic regulation heterogeneity of differentiation genes (DERSs), but not that of pluripotency genes (PERSs), was predominantly in charge of the entry and exit decisions of the pluripotent stem-like states (blue). The application of the hybrid numerical method validated the likelihood of epigenetic heterogeneity-based strategies capable of unlocking and directing the transit from differentiation-refractory to differentiation-primed (red) epistates via kinetics changes in epigenetic factors. (Note: The epigenetic parameters regulating the entry into robust epi-states throughout the entire ER-GRN system revealed a regime of tristability in which pluripotent stem-like (blue) and differentiated (red) steady-states coexisted with a third indecisive (green) state). (R: Recruited; U: Unrecruited).

<https://doi.org/10.1371/journal.pcbi.1006592.g009>



beginning to be recognised as valuable interventions directed against the ageing hallmarks [67]. Our current approach reconciles such apparently counterintuitive scenario by illuminating the occurrence of tunable switches in terms of epigenetic cofactor levels, which are capable of modifying the nature and direction of cell fate reprogramming. On the one hand, our mathematical deconstruction of epigenetic plasticity reveals that epigenetic heterogeneity may underlie the predisposition of cell populations to pathological reprogramming processes that cause a permanent, locked stem-like state disabled for reparative differentiation and prone to malignant transformation. On the other hand, we have computationally validated the likelihood of unlocking aberrant stem-like states disabled for reparative differentiation and drive them to a correct repair function by manipulating solely the intensity and direction of such epigenetic control switches. Therefore, we now propose that an ideal ageing-/cancer-targeted therapeutic approach must be able to correct chronic epigenetic plasticity of damaged/diseased tissues, but additionally, to “unlock” stem cell-like states to drive tissue regeneration, thereby preventing the occurrence of the abovementioned constraining phenotypes.

In summary, upon unearthing key regulatory dimensions of epigenetic plasticity in an unbiased manner, we here offer a conceptual and methodological re-orientation of how therapeutically approach pathological cellular reprogramming. As we enter a new era of therapeutic approaches to target ageing *per se* [67], our stochastic biomathematical modelling and computational simulation strategy might be incorporated as a valuable tool for assessing the benefit/risk ratio of new therapeutic approaches aimed to target and correct the ageing-/cancer-related perturbations of the epigenome.

## Supporting information

### S1 Text. Supplemental material.

(PDF)

### S1 Appendix. Numerical method.

(PDF)

### S1 Fig. Comparison of raw data and simulated data for the ER system.

(TIF)

### S2 Fig. Differentiation probability $Q$ within the DERSs and PERSs corresponding to the three clusters.

(TIF)

### S3 Fig. Dynamics of the different regions shown in Fig 3.

(TIF)

### S4 Fig. Optimal paths for DERS1, DERS2, PERS1 and PERS2.

(TIF)

### S5 Fig. Empirical CDFs for those $c_{ij}$ of the DERSs where significant differences between the clusters have not been found.

(TIF)

### S6 Fig. Empirical CDFs for the (whole) ensemble of DERS parameter sets classified according to $Q$ .

(TIF)

### S7 Fig. Empirical CDFs for those $c_{ij}$ of DERSs within the blue cluster where significant differences between differentiation-primed and pluripotency-locked behaviours have not

been found.

(TIF)

**S8 Fig. Empirical CDFs for the (whole) ensemble of PERS parameter sets classified according to  $Q$ .**

(TIF)

**S9 Fig. Phase diagrams depending on HME activity for DERS1, DERS2, PERS1 and PERS2.**

(TIF)

**S1 Table. General parameters involved in our ER-GRN model formulation.**

(PDF)

**S2 Table. Variables, rescaled variables and transition rates associated with the stochastic dynamics of the GRN submodel.**

(PDF)

**S3 Table. Variables, rescaled variables and transition rates associated with the stochastic dynamics of the ER submodel.**

(PDF)

**S4 Table. Reference parameter values used to generate the ensemble of DERSs and PERSs.**

(PDF)

**S5 Table. Parameter values for DERS1, DERS2, PERS1, PERS2.**

(PDF)

**S6 Table. Parameter values for the pluripotency-locked DERS being unlocked by the ER strategy.**

(PDF)

**S7 Table. Parameter values for the two-gene regulatory network.**

(PDF)

## Acknowledgments

The authors would like to thank Heather Harrington and Mariano Beguerisse for helpful discussions and advise.

## Author Contributions

**Conceptualization:** Núria Folguera-Blasco, Javier A. Menendez, Tomás Alarcón.

**Data curation:** Núria Folguera-Blasco.

**Formal analysis:** Núria Folguera-Blasco, Tomás Alarcón.

**Funding acquisition:** Javier A. Menendez, Tomás Alarcón.

**Investigation:** Núria Folguera-Blasco, Tomás Alarcón.

**Methodology:** Núria Folguera-Blasco, Rubén Pérez-Carrasco, Tomás Alarcón.

**Software:** Rubén Pérez-Carrasco.

**Supervision:** Tomás Alarcón.

**Writing – original draft:** Núria Folguera-Blasco, Rubén Pérez-Carrasco, Elisabet Cuyàs, Javier A. Menendez, Tomás Alarcón.

**Writing – review & editing:** Núria Folguera-Blasco, Rubén Pérez-Carrasco, Elisabet Cuyàs, Javier A. Menendez, Tomás Alarcón.

## References

- Mosteiro L, Pantoja C, Alcazar N, Marion RM, Chondronasin D, Rovira M, et al. Tissue damage and senescence provide critical signals for cellular reprogramming in vivo. *Science*. 2016; 354:aaf4445. <https://doi.org/10.1126/science.aaf4445> PMID: 27884981
- Yanger K, Zong Y, Maggs LR, Shapira SN, Maddipati R, Aiello NM, et al. Robust cellular reprogramming occurs spontaneously during liver regeneration. *Genes Dev*. 2013; 27:19–24. <https://doi.org/10.1101/gad.207803.112>
- Ocampo A, Reddy P, Martinez-Redondo P, Platero-Luengo A, Hatanaka F, Hishida T, et al. In Vivo Amelioration of Age-Associated Hallmarks by Partial Reprogramming. *Cell*. 2016; 167:1719–1733. <https://doi.org/10.1016/j.cell.2016.11.052> PMID: 27984723
- Ocampo A, Reddy P, Izpisua-Belmonte JC. Anti-aging strategies based on cellular reprogramming. *Trends Mol Med*. 2016; 22:725–738. <https://doi.org/10.1016/j.molmed.2016.06.005> PMID: 27426043
- Abad M, Mosteiro L, Pantoja C, Canamero M, Rayon T, Ors I, et al. Reprogramming in vivo produces teratomas and iPS cells with totipotency features. *Nature*. 2013; 502:340–345. <https://doi.org/10.1038/nature12586> PMID: 24025773
- Marion RM, Lopez de Silanes I, Mosteiro L, Gamache B, Abad M, Guerra C, et al. Common Telomere Changes during In Vivo Reprogramming and Early Stages of Tumorigenesis. *Stem Cell Reports*. 2017; 8:460–475. <https://doi.org/10.1016/j.stemcr.2017.01.001> PMID: 28162998
- Menendez JA, Alarcon T. Senescence-inflammatory regulation of reparative cellular reprogramming in aging and cancer. *Front Cell Dev Biol*. 2017; 5:49. <https://doi.org/10.3389/fcell.2017.00049> PMID: 28529938
- Mosteiro L, Pantoja C, de Martino A, Serrano M. Senescence promotes in vivo reprogramming through p16INK4a and IL-6. *Aging Cell*. 2018; 17(2). <https://doi.org/10.1111/acer.12711> PMID: 29280266
- Pal S, Tyler JK. Epigenetics and aging. *Sci Adv*. 2016; 2:e1600584. <https://doi.org/10.1126/sciadv.1600584> PMID: 27482540
- de Keizer PLJ. The fountain of youth by targeting senescent cells? *Trends Mol Med*. 2017; 23:6–17. <https://doi.org/10.1016/j.molmed.2016.11.006> PMID: 28041565
- Menéndez JA, Corominas-Faja B, Cuyàs E, García MG, Fernández-Arroyo S, Fernández AF, et al. Oncometabolic nuclear reprogramming of cancer stemness. *Stem Cell Reports*. 2016; 6:273–283. <https://doi.org/10.1016/j.stemcr.2015.12.012> PMID: 26876667
- Bintu L, Young J, Antebi YE, McCue K, Kazuki Y, Uno N, et al. Dynamics of epigenetic regulation at the single-cell level. *Science*. 2016; 351:720–724. <https://doi.org/10.1126/science.aab2956> PMID: 26912859
- Cortini R, Barbi M, Care BR, Lavelle C, Lesne A, Mozziconacci J, et al. The physics of epigenetics. *Rev Mod Phys*. 2016; 88:025002. <https://doi.org/10.1103/RevModPhys.88.025002>
- Ringrose L, Howard M. Dissecting chromatin-mediated gene regulation and epigenetic memory through mathematical modelling. *Current Opinion in Systems Biology*. 2017; 3:7–14. <https://doi.org/10.1016/j.coisb.2017.02.003>
- Tian XJ, Zhang H, Sannerud J, Xing J. Achieving diverse and monoallelic olfactory receptor selection through dual-objective optimization design. *Proceedings of the National Academy of Sciences*. 2016; 113:E2889–E2898. <https://doi.org/10.1073/pnas.1601722113>
- Folguera-Blasco N, Cuyàs E, Menéndez JA, Alarcón T. Epigenetic regulation of cell fate reprogramming in aging and disease: A predictive computational model. *PLoS Comp Biol*. 2018; 14:e1006052. <https://doi.org/10.1371/journal.pcbi.1006052>
- Dodd IB, Micheelsen MA, Sneppen K, Thon G. Theoretical analysis of epigenetic cell memory by nucleosome modification. *Cell*. 2007; 129:813–822. <https://doi.org/10.1016/j.cell.2007.02.053> PMID: 17512413
- David-Rus D, Mukhopadhyay S, Lebowitz JL, Sengupta AM. Inheritance of epigenetic chromatin silencing. *J Theor Biol*. 2009; 258:112–120. <https://doi.org/10.1016/j.jtbi.2008.12.021> PMID: 19174167
- Sneppen K, Dodd IB. A simple histone code opens many paths to epigenetics. *PLoS Comp Biol*. 2012; 8:e1002643. <https://doi.org/10.1371/journal.pcbi.1002643>

20. Sneppen K, Dodd IB. Nucleosome dynamics and maintenance of epigenetic states of CpG islands. *Phys Rev E*. 2016; 93:062417. <https://doi.org/10.1103/PhysRevE.93.062417> PMID: 27415308
21. Friedmann-Morvinski D, Verma IM. Dedifferentiation and reprogramming: origins of cancer stem cells. *EMBO Rep*. 2014; 15:244–253. <https://doi.org/10.1002/embr.201338254> PMID: 24531722
22. Wainwright EN, Scaffidi P. Epigenetics and Cancer Stem Cells: Unleashing, Hijacking, and Restricting Cellular Plasticity. *Trends Cancer*. 2017; 3:372–386. <https://doi.org/10.1016/j.trecan.2017.04.004> PMID: 28718414
23. Huang S, Eichler G, Bar-Yam Y, Ingber DE. Cell fates as high-dimensional attractors of a complex gene regulatory network. *Phys Rev Lett*. 2005; 94:128701. <https://doi.org/10.1103/PhysRevLett.94.128701> PMID: 15903968
24. Thomas P, Popović N, Grima R. Phenotypic switching in gene regulatory networks. *Proceedings of the National Academy of Sciences*. 2014; 111:6994–6999. <https://doi.org/10.1073/pnas.1400049111>
25. Zhang B, Wolynes PG. Stem cell differentiation as a many-body problem. *Proc Natl Acad Sci*. 2014; 111:10185–10190. <https://doi.org/10.1073/pnas.1408561111> PMID: 24946805
26. Ge H, Qian H, Xie XS. Stochastic phenotype transition of a single cell in an intermediate region of gene state switching. *Phys Rev Lett*. 2015; 114:078101. <https://doi.org/10.1103/PhysRevLett.114.078101> PMID: 25763973
27. Wells DK, Kath WL, Motter AE. Control of stochastic and induced switching in biophysical networks. *Phys Rev X*. 2015; 5:031036. <https://doi.org/10.1103/PhysRevX.5.031036> PMID: 26451275
28. Perez-Carrasco R, Guerrero P, Briscoe J, Page KM. Intrinsic noise profoundly alters the dynamics and steady state of the morphogen controlled bistable genetic switches. *PLoS Comp Bio*. 2016; 12:e1005154. <https://doi.org/10.1371/journal.pcbi.1005154>
29. Li YT, Hufton PG, Lee EJ, Potoyan DA. A stochastic and dynamical view of pluripotency in mouse embryonic stem cells. *PLoS Comp Biol*. 2018; 14:e1006000. <https://doi.org/10.1371/journal.pcbi.1006000>
30. Guantes R, Poyatos JF. Multistable Decision Switches for Flexible Control of Epigenetic Differentiation. *PLoS Computational Biology*. 2008; 4(11). <https://doi.org/10.1371/journal.pcbi.1000235> PMID: 19043543
31. Wu F, Su RQ, Lai YC, Wang X. Engineering of a synthetic quadrastable gene network to approach Waddington landscape and cell fate determination. *eLife*. 2017; 6(e23702).
32. Cao Y, Gillespie DT, Petzold LR. Multiscale stochastic simulation algorithm with stochastic partial equilibrium assumption for chemically reacting systems. *J Comp Phys*. 2005; 206:395–411. <https://doi.org/10.1016/j.jcp.2004.12.014>
33. Cao Y, Gillespie DT, Petzold LR. The slow-scale stochastic simulation algorithm. *J Chem Phys*. 2005; 122:014116. <https://doi.org/10.1063/1.1824902>
34. Ball K, Kurtz TG, Popovic L, Rempala G. Asymptotic analysis of multi-scale approximations to reaction networks. *Ann App Prob*. 2006; 16:1925–1961. <https://doi.org/10.1214/105051606000000420>
35. Anderson DF, Kurtz TG. *Stochastic analysis of biochemical systems*. Springer, New York, NY, USA; 2010.
36. Kang HW, Kurtz TG. Separation of time-scales and model reduction for stochastic reaction networks. *The Annals of Applied Probability*. 2013; 23(2):529–583. <https://doi.org/10.1214/12-AAP841>
37. Alarcón T. Stochastic quasi-steady state approximations for asymptotic solutions of the Chemical Master Equation. *J Chem Phys*. 2014; 140:184109. <https://doi.org/10.1063/1.4874653> PMID: 24832255
38. Kang HW, Kurtz TG, Popovic L. Central limit theorems and diffusion approximation for multiscale Markov chain models. *Ann Prob*. 2014; 24:721–759. <https://doi.org/10.1214/13-AAP934>
39. de la Cruz R, Guerrero P, Spill F, Alarcón T. The effects of intrinsic noise on the behaviour of bistable systems in quasi-steady state conditions. *J Chem Phys*. 2015; 143:074105. <https://doi.org/10.1063/1.4928575> PMID: 26298113
40. Pour M, Pilzer I, Rosner R, Smith ZD, Meissner A, Nachman I. Epigenetic predisposition to reprogramming fates. *EMBO Reports*. 2015; 18:370–378. <https://doi.org/10.15252/embr.201439264>
41. Dattani J, Barahona M. Stochastic models of gene transcription with upstream drives: exact solution and sample path characterization. *Journal of The Royal Society Interface*. 2017; 14. <https://doi.org/10.1098/rsif.2016.0833>
42. Gardiner CW. *Stochastic methods*. Springer-Verlag, Berlin, Germany; 2009.
43. Hanggi P, Talkner P, Borkovec M. Reaction-rate theory: fifty years after Kramers. *Rev Mod Phys*. 1990; 62:251. <https://doi.org/10.1103/RevModPhys.62.251>
44. Freidlin MI, Wentzell AD. *Random perturbations of dynamical systems*. Springer-Verlag, New York, NY, USA; 1998.

45. Touchette H. The large deviation approach to statistical mechanics. *Phys Rep.* 2009; 479:1–69. <https://doi.org/10.1016/j.physrep.2009.05.002>
46. Flavahan WA, Gaskell E, Bernstein BE. Epigenetic plasticity and the hallmarks of cancer. *Science.* 2017; 357:eaal2380. <https://doi.org/10.1126/science.aal2380> PMID: 28729483
47. Bressloff PC. *Stochastic processes in cell biology.* Springer-Verlag, Berlin, Germany.; 2014.
48. Bressloff PC, Newby JM. Path integrals and large deviations in stochastic hybrid systems. *Phys Rev E.* 2014; 89:042701. <https://doi.org/10.1103/PhysRevE.89.042701>
49. Gillespie DT. Approximate accelerated stochastic simulation of chemically reacting systems. *J Chem Phys.* 2001; 115:1716–1733. <https://doi.org/10.1063/1.1378322>
50. Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf MPH. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J R Soc Interface.* 2009; 6:187–202. <https://doi.org/10.1098/rsif.2008.0172> PMID: 19205079
51. Wilkinson DJ. *Stochastic modelling for systems biology.* Chapman & Hall/CRC, Boca Raton, FL, USA; 2012.
52. Keener J, Sneyd J. *Mathematical physiology.* Springer-Verlag, New York, NY, USA; 1998.
53. Davis MHA. Piecewise-Deterministic Markov Processes: A General Class of Non-Diffusion Stochastic Models. *Journal of the Royal Statistical Society Series B (Methodological).* 1984; 46:353–388. <https://doi.org/10.1111/j.2517-6161.1984.tb01308.x>
54. Bressloff PC, Lawley SD. Mean first passage times for piecewise deterministic Markov processes and the effects of critical points. *Journal of Statistical Mechanics: Theory and Experiment.* 2017; 2017:063202. <https://doi.org/10.1088/1742-5468/aa71df>
55. Gibson MA, Bruck J. Efficient exact stochastic simulation of chemical systems with many species and many channels. *J Phys Chem A.* 2000; 104:1876–1889. <https://doi.org/10.1021/jp993732q>
56. Gillespie DT. The chemical Langevin equation. *The Journal of Chemical Physics.* 2000; 113(1):297–306. <https://doi.org/10.1063/1.481811>
57. Harrington H, Feliu E, Wiuf C, Stumpf MH. Cellular Compartments Cause Multistability and Allow Cells to Process More Information. *Biophysical Journal.* 2013; 104:1824–1831. <https://doi.org/10.1016/j.bpj.2013.02.028> PMID: 23601329
58. Kauffman SA. *The origins of order.* Oxford University Press, New York, U.S.A.; 1993.
59. Semrau S, Goldmann JE, Soumilion M, Mikkelsen TS, Jaenisch R, van Oudenaarden A. Dynamics of lineage commitment revealed by single-cell transcriptomics of differentiating embryonic stem cells. *Nature Communications.* 2017; 8. <https://doi.org/10.1038/s41467-017-01076-4> PMID: 29061959
60. Wagner A. *Robustness and evolvability in living systems.* Princeton University Press, Princeton, NJ, U.S.A; 2007.
61. Wagner A. Neutralism and selectionism: A network-based reconciliation. *Nature Reviews Genetics.* 2008; 9:965–974. <https://doi.org/10.1038/nrg2473> PMID: 18957969
62. Wagner A. Robustness and evolvability: A paradox resolved. *Proc Roy Soc B.* 2008; 275:91–100. <https://doi.org/10.1098/rspb.2007.1137>
63. Huang S. Reprogramming cell fates: reconciling rarity with robustness. *BioEssays.* 2009; 31:546–560. <https://doi.org/10.1002/bies.200800189> PMID: 19319911
64. de Back W, Zhou JX, Brusch L. On the role of lateral stabilization during early patterning in the pancreas. *J R Soc Interface.* 2012; 10:20120766. <https://doi.org/10.1098/rsif.2012.0766>
65. Niwa H, Toyooka Y, Shimosato D, Strumpf D, Takahashi K, Yagi R, et al. Interaction between Oct3/4 and Cdx2 Determines Trophectoderm Differentiation. *Cell.* 2005; 123(5):917–929. <https://doi.org/10.1016/j.cell.2005.08.040> PMID: 16325584
66. Reikhtman N, Radparvar F, Evans T, Skoultschi AI. Direct interaction of hematopoietic transcription factors PU.1 and GATA-1: functional antagonism in erythroid cells. *Genes Dev.* 1999; 13:1389–1411. <https://doi.org/10.1101/gad.13.11.1398>
67. Mahmoudi S, Xu L, Brunet A. Turning back time with emerging rejuvenation strategies. *Nat Cell Biol.* 2019; 21:32–43. <https://doi.org/10.1038/s41556-018-0206-0> PMID: 30602763
68. Mahmoudi S, Mancini E, Moore A, Xu L, Jahanbani F, Hebestreit K, et al. Old fibroblasts secrete inflammatory cytokines that drive variability in reprogramming efficiency and may affect wound healing between old individuals. 2018;.
69. Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, et al. The Polycomb group protein EZH2 directly controls DNA methylation. *Nature.* 2006; 439:871–874. <https://doi.org/10.1038/nature04431> PMID: 16357870