PNAS RESEARCH ARTICLE | SYSTEMS BIOLOGY PHYSICS

# Epigenetic factor competition reshapes the EMT landscape

M. Ali Al-Radhawi[a,b], Shubham Tripathi[b,c,d], Yun Zhang[e,f], Eduardo D. Sontag[a,g,h,1], and Herbert Levine[b,d,g,1]

Contributed by Herbert Levine; received June 23, 2022; accepted September 2, 2022; reviewed by Sylvia Plevritis and Kim Sneppen

**The emergence of and transitions between distinct phenotypes in isogenic cells can be attributed to the intricate interplay of epigenetic marks, external signals, and gene-regulatory elements. These elements include chromatin remodelers, histone modifiers, transcription factors, and regulatory RNAs. Mathematical models known as gene-regulatory networks (GRNs) are an increasingly important tool to unravel the workings of such complex networks. In such models, epigenetic factors are usually proposed to act on the chromatin regions directly involved in the expression of relevant genes. However, it has been well-established that these factors operate globally and compete with each other for targets genome-wide. Therefore, a perturbation of the activity of a regulator can redistribute epigenetic marks across the genome and modulate the levels of competing regulators. In this paper, we propose a conceptual and mathematical modeling framework that incorporates both local and global competition effects between antagonistic epigenetic regulators, in addition to local transcription factors, and show the counterintuitive consequences of such interactions. We apply our approach to recent experimental findings on the epithelial–mesenchymal transition (EMT). We show that it can explain the puzzling experimental data, as well as provide verifiable predictions.**

EMT | epigenetic | gene network

## Significance

Epigenetics refers to the role of structural modifications in the genome that encode the context governing gene-expression dynamics. Although there have been attempts at creating mathematical models of these structural changes, how these changes couple to the transcriptional circuitry has remained unclear. Here, we study the role of histone modifications—namely, the adding or removing of methyl groups from histone H3 tails—in interacting with regulatory factors governing the transitions between epithelial and mesenchymal phenotypes. Our results show the critical importance of taking into account off-target global effects due to the competition between different DNA segments for the same histone-modifying enzymes.

Multicellular organisms start from the mitotic division of a single cell and then proliferate and differentiate into increasingly more specialized lineages. Within the central dogma of molecular biology, differences between cells in different lineages can be explained by different patterns of gene activity (1). Even within the same lineage, different genes can be activated, depending on the external signals, environmental factors, or internal stochasticity. Hence, precise mechanisms for gene regulation must exist within the cell. Such mechanisms are commonly ascribed to a multitude of processes involving epigenetic, transcriptional, translational, and posttranslational regulatory elements. However, the manner in which the different layers of regulation interact is far from being fully understood.

The complex interplay between transcriptional and epigenetic control is strikingly evident in the regulation of the epithelial–mesenchymal transition (EMT), a process that allows cells to lose cell–cell adhesion and apico-basal polarity and acquire migratory and invasive traits (2). EMT is essential for embryonic development and wound healing. However, dysregulated EMT is a key contributor to cancer mortality, playing crucial roles in metastasis and the emergence of drug resistance. Epigenetic processes have been known to act both upstream and downstream of the core EMT gene-regulatory circuit (2), and various epigenetic modifications are known to alter the expression of transcription factors (TFs) and micro-RNAs involved in EMT control (3–5). In turn, the global epigenetic state of cells can be altered by the induction of one or more EMT TFs (4).

The complexity of the epigenetic–transcriptional interplay in EMT has been highlighted in a recent study by Zhang et al. (6). This work has shown that the knockout of different histone methyltransferases in the human mammary epithelial cell line HMLER could induce two distinct trajectories of EMT, characterized by distinct and unexpected changes in gene-expression profiles. It would clearly be useful to interpret these results within the context of increasing powerful mathematical models of gene regulation and the EMT process.

A popular framework to model and predict the consequences of interactions between various biomolecular regulators uses the concept of a gene-regulatory network (GRN) (7). Such descriptions can recapitulate the phenotypic heterogeneity that may be exhibited by genotypically identical cells by taking into account the differences in the configuration of the network, initial conditions, and external factors (8, 9). Early GRN models mainly considered the interactions between TFs and promoters (7, 10, 11). GRNs were later expanded to include the activities of noncoding RNAs, such as micro-RNAs (12), DNA methylation (13, 14), and histone modifiers (15–18). Complementary mathematical models have focused on the effect of antagonistic epigenetic factors (EFs) acting on the histone tails (19, 20) or the interaction between repressive histone modifiers and TFs (21).

Author contributions: M.A.A.-R., S.T., E.D.S., and H.L. designed research; M.A.A.-R. performed research; M.A.A.-R., S.T., and Y.Z. analyzed data; and M.A.A.-R., S.T., and H.L. wrote the paper.

Reviewers: S.P., Stanford University; and K.S., Kobenhavns Universitet.

The authors declare no competing interest.



[1]To whom correspondence may be addressed. Email: e.sontag@northeastern.edu or h.levine@northeastern.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2210844119/-/DCSupplemental.

Published October 10, 2022.

In the context of modeling EMT, most of the systems-biology modeling effort has focused exclusively on the transcriptional and translational dynamics (22). Conversely, mathematical modeling of the underlying epigenetic processes has largely been limited to coarse-grained phenomenological approaches (23, 24). Thus, recent high-resolution characterizations of the epigenetic and transcriptional changes during EMT and their response to epigenetic and transcriptional perturbations (4, 6, 25) provide both a need and an opportunity for the development of new models that can shed light on the principles governing the complex transcriptional–epigenetic interplay.

In summary, the aforementioned formulations, both for general GRNs and for EMT, do not consider genome-wide effects. Instead, they focus on the genomically local interactions of the regulatory factors with a single gene or a small set of genes. Although such an assumption might be justified in many cases, this ignores the fact that many regulatory factors, and especially epigenetic ones, act globally and can have hundreds or thousands of targets. Furthermore, such factors compete with each other, and perturbations to the activity or expression level of one of them can have considerable off-target effects, as will be reviewed next.

## The Polycomb and Trithorax Groups of EFs

In this paper, we focus on the well-documented antagonism between the Polycomb (PcG) and Trithorax (TrX) protein groups. These protein families modulate histone tails that help maintain genes in silenced or active states and that act globally to regulate numerous cellular processes (26, 27). PcG and TrX act antagonistically, where the first is usually associated with silencing, while the latter is associated with activation (26, 27). For instance, PcG Repressive Complex 2 (PRC2), a PcG protein, is responsible for trimethylating Histone H3 lysine 27 (H3K27) to mark genes for silencing (28, 29). PRC2 has been reported to have more than 1,000 targets in a single human embryonic fibroblast cell (30), and it is estimated to target at least 10% of the genes in embryonic stem cells (ESCs) (31).

On the other hand, the COMPASS family of proteins (a subfamily of TrX proteins) is involved in depositing activating methylation marks. In particular, KMT2A/B (MLL1/2), KMT2C/D (MLL3/4), and SETD1A/B (KMT2F/G) deposit methylation marks at Histone H3 lysine 4 (H3K4) (32). These proteins differ in the genomic region targeted: SETD1A/B trimethylate H3K4 around transcription start sites (33), KMT2C/D monomethylate H3K4 around enhancer elements (34, 35), and KMT2A/B deposit H3K4me2 and H3K4me3 marks at developmental genes (36). It should be noted that these factors often have overlapping effects depending on the context. For instance, knockout of KMT2D causes genome-wide disappearance of H3K4me3 in mice B cells (37) and brain cells (38), and it has been shown to be essential for the maintenance of H3K4me2 marks in mice cardiomyocytes (39). In addition, KMT2D is required for acetylating H3K27 in conjunction with CREBBP and EP300 (40), where H3K27ac is an activating mark that is mutually exclusive with the silencing mark H3K27me3. Similar to PRC2, TrX group proteins act genome-wide. For instance, the transcription of 1,200 genes has been shown to be dependent on KMT2B in mouse ESCs (41), while KMT2D was shown to bind to 4,880 genes in mice cardiomyocytes (39). In summary, PcG and TrX proteins act genome-wide, deposit marks on similar histone sites, and have opposing functions.

## EF Competition

As reviewed above, antagonistic EFs deposit functionally opposing histone marks. Do they compete for the same (or nearby) genomic sites? PcG and TrX proteins are recruited to genes by regulatory sequences known as PcG response elements (PREs) and TrX response elements (TREs), respectively (42). Existing evidence in *Drosophila* shows that PREs are also TREs and that PcG/TrX proteins compete for them (19, 42). In addition, activating methylation marks (e.g., H3K27ac, H3K4me3, and H3K36me3) inhibit PRC2's ability to methylate H3K27 (43–45). On the other hand, PRC2's activity reduces the ability of CREBBP/EP300/KMT2D to deposit H3K27ac activating marks (46, 47). While it might be possible for H3K4me3 and H3K27me3 to exist in the same vicinity [a phenomenon known as bivalency (48)], they are mutually exclusive on the same histone tail (49). Alternative mechanisms of competition include PRC2 acting indirectly on nearby nucleosomes by recruiting other factors to remove activating marks (50, 51). Therefore, it is usually assumed that PcG and TrX counteract each other genome-wide (19, 27).

In addition to direct competition with other EFs, PRC2 is antagonistic to active transcription. PRC2 activity leads to chromatin compaction (52), which makes it harder for activating TFs to access their target sites. On the other hand, PRC2 can read the epigenetic context to avoid acting on genes that are transcriptionally active (53–55). Possible mechanisms include PRC2 binding to nascent RNA (56), relative aversion to open chromatin (57), likely competition with TFs (55), or enhancer–PRE communication (58), among others (21).

## PcG Dilution and Redistribution upon Suppression of Competitors

Since PcG and its competitors vie for similar genomic sites, knockout of one factor can have far-reaching off-target effects via redistribution or dilution of its competing factors. Therefore, new genes might get activated or silenced. Such knockout experiments have been conducted in the literature with a particular attention to PcG proteins and the corresponding histone mark (H3K27me3). Below, we review experiments that provide evidence supporting the sequestration and redistribution of PcG proteins.

The protein MES4 is an H3K36 methyltransferase that is antagonistic to PcG. In ref. 59, loss of MES4 in *Caenorhabditis elegans* caused a reduction of H3K27me3 levels (deposited by PcG) at its target sites. Meanwhile, genomic sites that lost the antagonistic mark H3K36me3 gained H3K27me3. Similarly, in mice (60), an H3K36M mutant inhibits the activity of H3K36 methyltransferases. The authors observed sequestration of PRC2, as evidenced by increased levels of chromatin-bound EZH2 and SUZ12 (subcomponents of PRC2). In addition, many genes lost H3K27me3, and their expression levels increased. This indicated that the loss of H3K36 methylation can provide new substrates for PRC2. Redistribution of H3K27me3 was also observed in H3K36M mutants in *Drosophila* (61). In a recent investigation, rapid depletion of the BAF complex (a chromatin remodeler that is antagonistic to PcG) redistributed PcG from highly occupied domains to new genomic sites in mouse ESCs (62). Additional pieces of evidence are included in *Discussion*.

Overall, PcG proteins are highly sensitive to perturbations to other EFs and epigenetic marks. Such interventions reshape the global epigenetic landscape leading to aberrant changes in transcription. This, in turn, has been identified as a contributing factor to various malignancies (63, 64).

## Organization of the Paper

We first propose a general modeling framework and outline its underlying assumptions based on the experimental literature. We

then show how the epigenetic competition model can explain paradoxical knockout results in single- and multiple-knockout experiments in ref. 6. We then propose a combined model encompassing epigenetic competition and gene regulation to account for different patterns of gene activity in the observed EMTs. Then, we use our model to offer several predictions, some of which have already been verified. We conclude with a discussion that complements our literature review in the introduction and points to new directions. The mathematical details are included in *SI Appendix*.

## Results

**The Proposed Modeling Framework Accounts for Global Effects.** Consider a GRN consisting of a number of genes interacting with each other via TF binding and/or micro-RNA–mediated posttranscriptional regulation. The interactions can either be inhibitory or activating. A subset of genes in the GRN are also subject to the influence of global EFs, such as histone modifiers (e.g., PcG or TrX). Fig. 1 depicts the schematics of the network. We describe elements of the model next. The mathematical details are included in *SI Appendix*.

***EFs.*** The EFs can generally be histone modifiers, chromatin remodelers, and/or DNA methyltransferases. We focus here on PcG and TrX EFs and assume the presence of at least one repressing PcG protein and one activating TrX protein. Based on our review of the literature on the antagonism and redistribution of EFs and epigenetic marks, we make the following four assumptions:

1. *Competition*: The EFs compete for binding to similar genomic targets, which could be involved in the regulation of the same gene. For example, if a PcG protein binds to a PRE/TRE, or if it deposits a repressive mark, then another TrX protein cannot bind to the same gene, and vice versa. This assumption is justified by our previous review of the literature documenting the antagonism between the two groups of proteins across the genome.
2. *Global targeting*: The EFs have targets genome-wide and are not limited to the local GRN. This is justified by the fact that common EFs are known to have hundreds or even thousands of genomic targets.
3. *Scarcity*: The levels of the EFs are limited. More formally, we assume that the total levels are constant in the time scale of interest. This is justified by our review of the knockout experiments that observe the dilution of the EFs at their original targets and their redistribution to new targets.
4. *Localization*: If an EF is bound to a target on a specific gene, then it cannot simultaneously bind to targets on other genes. This is justified by the observation that EFs (e.g., PRC2) are physically localized to their targets. In addition, the aforementioned dilution and redistribution effects imply that a specific EF complex cannot affect two genes simultaneously.

Thus, when an EF is knocked out, there are sufficient binding sites to sequester the available EFs of another type and make them localize to other genomic loci.

***Genes.*** We use a coarse-grained model of the genes. The individual nucleosomes (that serve as substrates for the various histone-modifying enzymes) are not explicitly modeled. Instead, each gene is modeled as a collection of states that account for the possible histone marks, PRE/TRE occupancies, and TF binding sites. Hence, a given gene can be quantified as distribution of the aforementioned states. The rest of the genome is modeled as a single "mega-gene" with a very large copy number compared to the local GRN genes.

***Modeling the PREs/TREs.*** The PRE/TRE component allows the EFs to affect the target gene. For a given gene, the PRE/TRE can be in one the following states (shown in Fig. 1*B*):

1. *Neutral*: There is no PcG or TrX bound to it, and there are no histone marks.
2. *Active*: Either 1) an activating EF is bound to it (e.g., a TrX protein), or 2) it has an activating histone mark (e.g., H3K27ac or H3K4me3).

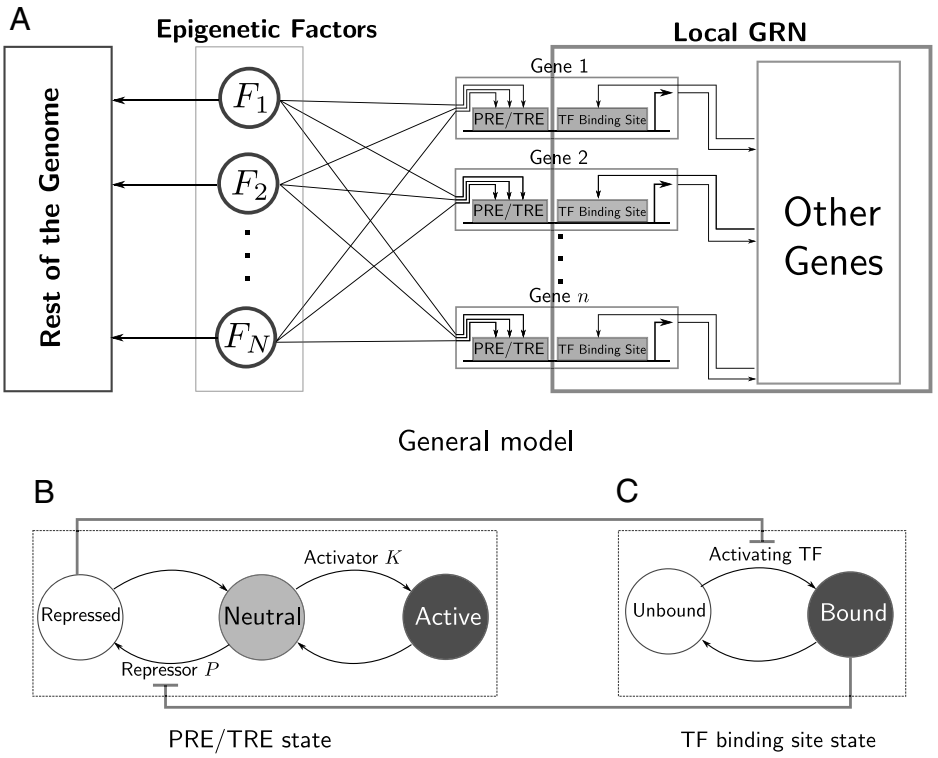



**Fig. 1.** (*A*) The proposed model. It consists of three compartments: the local GRN, the EFs, and the rest of the genome. The behavior of the local GRN can be "tuned" by global EFs that compete for similar targets and act genome-wide. On the other hand, the local context of a given GRN affects the binding of global EFs and, hence, affects their localization. (*B* and *C*) Modeling a self-activating gene. (*B*) The PRE/TRE. (*C*) The TF binding state. The interaction between EF binding and TF binding is also illustrated.

3. *Repressed*: Either 1) a repressive EF is bound to it (e.g., a PcG protein), or 2) it has a silencing histone mark (e.g., H3K27me3).

Note that the last two states of the PRE/TRE (Active and Repressed) can have the EF either bound or unbound. The localization effect described earlier arises only for the bound states 2a and 3a. We do not explicitly include a bivalent state, as it can be effectively modeled with a gene whose repressed and active states are both present with nonnegligible proportions.

***Interaction between TFs and PREs/TREs.*** Our model also allows for interaction between the PRE/TRE state and TF-binding state, as shown in Fig. 1 *B* and *C*. This is motivated by the observation that PcG proteins cannot act on genomic loci under active transcription genes, as we have reviewed earlier. The interaction between PRE/TRE and TF-binding states can be modeled by disallowing a PcG protein from silencing a gene while an activating TF is bound to it. In addition, a TF cannot bind to a gene that has been silenced by PcG. Mathematical details concerning the implementation of this effect are provided in *SI Appendix*, section 2. Note that such an interaction can create regulatory feedback from the local GRN compartment to the EF compartment—active transcription at the target gene repels repressing EFs. Hence, the aggregate effect of the transcriptional activity of many target genes can alter the global level of an EF.

**The Genome-Wide Competition Context Reverses Expected Knockout Results.** In order to illustrate the model behavior, we consider first a toy example of a single gene that is only regulated by activating and repressing EFs. If the repressive EFs are dominant, then the gene is strongly repressed, while it is strongly active when the activating EFs are dominant. When none of the EFs are present at the gene, we assume that the gene is weakly active. In this scenario, knocking out an activating EF of a gene is expected to reduce expression, while knocking out a repressive EF is expected to increase expression. However, under a competition scenario, opposite effects might occur. We review several cases below.

***Two EFs.*** We first consider one activating EF and one repressive EF, as illustrated in Fig. 2 *A*, *B*, *D*, and *E*. When the competition effects are minimal, knockout of a repressor will induce strong activation of the gene, as shown in Fig. 2*A*. Similarly, the knockout of an activator will result in strong repression, as shown in Fig. 2*D*. However, when the two EFs compete for targets across the genome, then knocking out one of them can create many new targets for the competing EF. Depending on the EF's binding affinity to the gene under consideration, as compared to newly available targets, the competing EF can be diluted genome-wide and redistributed to new targets. This is illustrated in Fig. 2*B*, where knockout of the repressor does not produce strong activation due to the dilution of the activator. Similarly, knockout of the activator does not produce strong repression due to the dilution of the repressor, as shown in Fig. 2*E*.

The case with only two EFs is not sufficient to recapitulate all possible outcomes. For instance, it cannot capture a gene that is strongly repressed after the knockout of its only repressor. Such behavior can be recapitulated by a model with three EFs, as we show next.

***Three EFs.*** We consider cases with two activators and one repressor and with two repressors and one activator. As reviewed before, there are multiple EFs that have overlapping functions. For instance, H3K4 can be trimethylated by multiple factors. Therefore, the function of a knocked-out EF can be "rescued" by an alternative activator. For example, Hanna et al. (65) showed that H3K4me3 levels are elevated at many genomic locations after the knockout of SETD1B (which is an H3K4me3 methyltransferase) due to compensation by MLL2 (KMT2B),

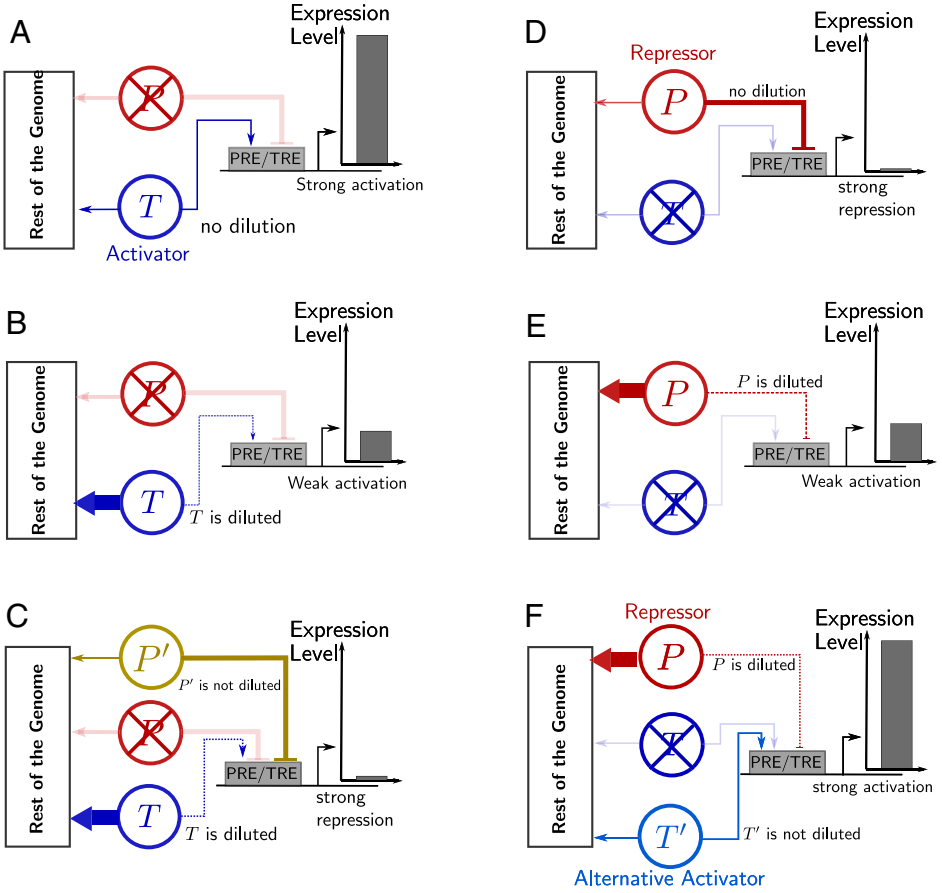



**Fig. 2.** All outcomes are logically possible after a knockout experiment due to the competition context. *A–C* schematically depict the impact of knocking out the repressor *P*, while *D–F* depict the impact of knocking out the activator *T*. (*A*) With minimal competition effects, the EF *T* can strongly activate its target after its competitor *P* has been knocked out. (*B*) The activator *T* gets diluted after the competitor has been knocked out. Hence, the PRE/TRE becomes mostly unmodified, rendering the gene weakly active. (*C*) The activator gets diluted, giving an alternative repressor the opportunity to repress the target gene. (*D*) The EF *P* represses its target after the competitor *T* has been knocked out. (*E*) Knockout of *T* and dilution of *P* renders the gene weakly active. (*F*) The alternative activator *T*′ activates the target gene in the absence of competition. A very thick arrow denotes sequestration of the corresponding EF by other genes, a dotted arrow depicts dilution of the corresponding EF, and a lightly shaded arrow denotes an absent regulatory link due to knockout.

an alternative methyltransferase. A similar pattern exists for repressors. For instance, it has been observed that the loss of DNA methylation and H3K9 trimethylation is rescued by silencing via H3K27 trimethylation (66).

In order to illustrate our modeling of the aforementioned behavior, Fig. 2*F* shows two activators, $T$, $T'$, and one repressor, $P$. Knockout of $T$ and the dilution of $P$ are not sufficient to explain the strong activation of the gene. Instead, the presence of an alternate activator, $T'$, is required for strong activation. To keep $T'$ undiluted, despite the knockout of $T$, the model requires asymmetry between $T'$ and $P$ in terms of their affinity to target sites across the genome. Finally, for completeness, we depict in Fig. 2*F* the case of two repressors and one activator, wherein the alternate repressor $P'$ rescues the respression of the gene, despite the knockout of the repressing EF $P$. This case corresponds to the experiment reported in ref. 66.

**The Model Explains Paradoxical Knockout Results.** Our modeling framework can explain multiple counterintuitive results from EF knockout experiments. To illustrate this, we consider the results of knocking out EED (a PRC2 component) and KMT2D (a component of the KMT2D–COMPASS complex) in the HMLER cell line (6). A total of 413 genes were identified as targets of PRC2 that had significant expression in the control or PRC2-KO cells. When examining the changes in the expression levels of these genes upon PRC2 and KMT2D knockout, we found multiple genes with paradoxical changes in expression levels. Below, we will use our modeling framework to interpret the observed behaviors.

***The balance between the global and local affinity parameters determines the model's behavior.*** In order to model the experiments, we study a single constitutively expressed gene that is subject to the effect of three EFs: one repressing ($P$; e.g., PRC2) and two activating ($T$, $T'$; e.g., KMT2D and an alternative activator). Consider the three experimental setups: control, $P$ knockout, and $T$ knockout. Under such an experiment, we are interested in six possible behavioral phenotypes. Each phenotype is characterized by an unambiguous ordering of the expression levels between the three cases. Fig. 3 shows that all phenotypes are possible if the global context is considered. In particular, it is shown that the interplay between global and local affinities of the EFs determines the observed phenotype.

We illustrate our framework by studying specific examples from the results reported in ref. 6. The majority of the PRC2 target genes (67.71%) exhibited their highest expression when PRC2 was knocked out, which is expected if the local context is dominant. One such example is CDH2 (N-cadherin), shown in Fig. 4*A*. Nevertheless, we still see, unexpectedly, that CDH2 becomes partially activated (compared to the control case) when KMT2D is knocked out. This can be interpreted either as PRC2 not being fully diluted upon KMT2D knockout, due to the weak affinity of the alternate activators (e.g., KMT2C) to CDH2, or a combination of both effects. Next, we study more striking examples.

***Activator knockout results in strong activation.*** In 61 of the PRC2 targets, a paradoxical behavior was reported. The highest expression level was observed when the activating EF KMT2D was knocked out. This includes multiple EMT-related genes, such as TWIST1, ZEB1, ZEB2, and PRRX1. Fig. 4*B* shows the case of TWIST1 as an example. In the control case, TWIST1 is strongly repressed. When PRC2 (the repressing EF) is knocked out, TWIST1's expression is increased, but is not strongly activated (when compared to the third case), which is counterintuitive.

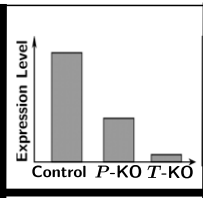
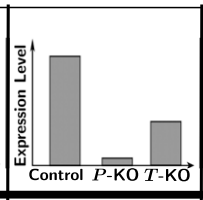
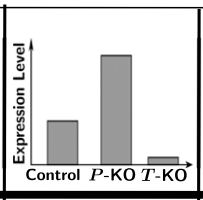
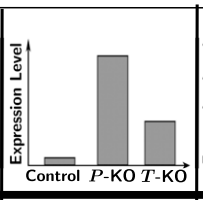
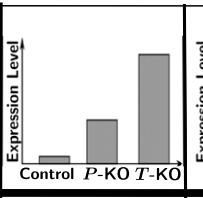
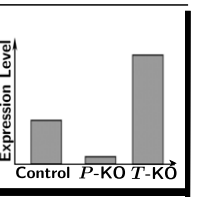
| Phenotype / Parameters | |  |  |  |  |  |  |
|---|---|---|---|---|---|---|---|
| | | Expression Level; Control $P$-KO $T$-KO | Expression Level; Control $P$-KO $T$-KO | Expression Level; Control $P$-KO $T$-KO | Expression Level; Control $P$-KO $T$-KO | Expression Level; Control $P$-KO $T$-KO | Expression Level; Control $P$-KO $T$-KO |
| Repressor $P$ | Global Association Ratio | Medium-High | Low-Medium | High | Medium-High | High | High |
| | Local Association Ratio | Low-Medium | Medium-High | Low | High | Medium | Low-Medium |
| Activator $T$ | Global Association Ratio | High | Medium | High | High | High | High |
| | Local Association Ratio | Medium | Medium | Low-Medium | Low-Medium | Low | Low |
| Alternative Activator $T'$ | Global Association Ratio | High | Medium-High | Medium-High | Low | Medium | Medium-High |
| | Local Association Ratio | Low | Low-Medium | Low | Low-Medium | Low-Medium | Low-Medium |

**Fig. 3.** The balance between local and global affinities determines the knockout phenotype. The behavior of each EF across the genome is characterized by a mix of local and global parameters. In this table, we consider a single constitutively expressed gene. We consider two parameters for each EF: the local association ratio to the gene of interest and the global association ratio to the rest of the genome. We keep the remaining parameters constant. Each entry in the table indicates the average of the corresponding parameter conditioned on the phenotype under consideration. The three experimental scenarios are: Control (both PRC2 and KMT2D are present), *P*-KO (PRC2 is knocked out), and *T*-KO (KMT2D is knocked out). A more detailed version of this table is provided in Fig. 8.

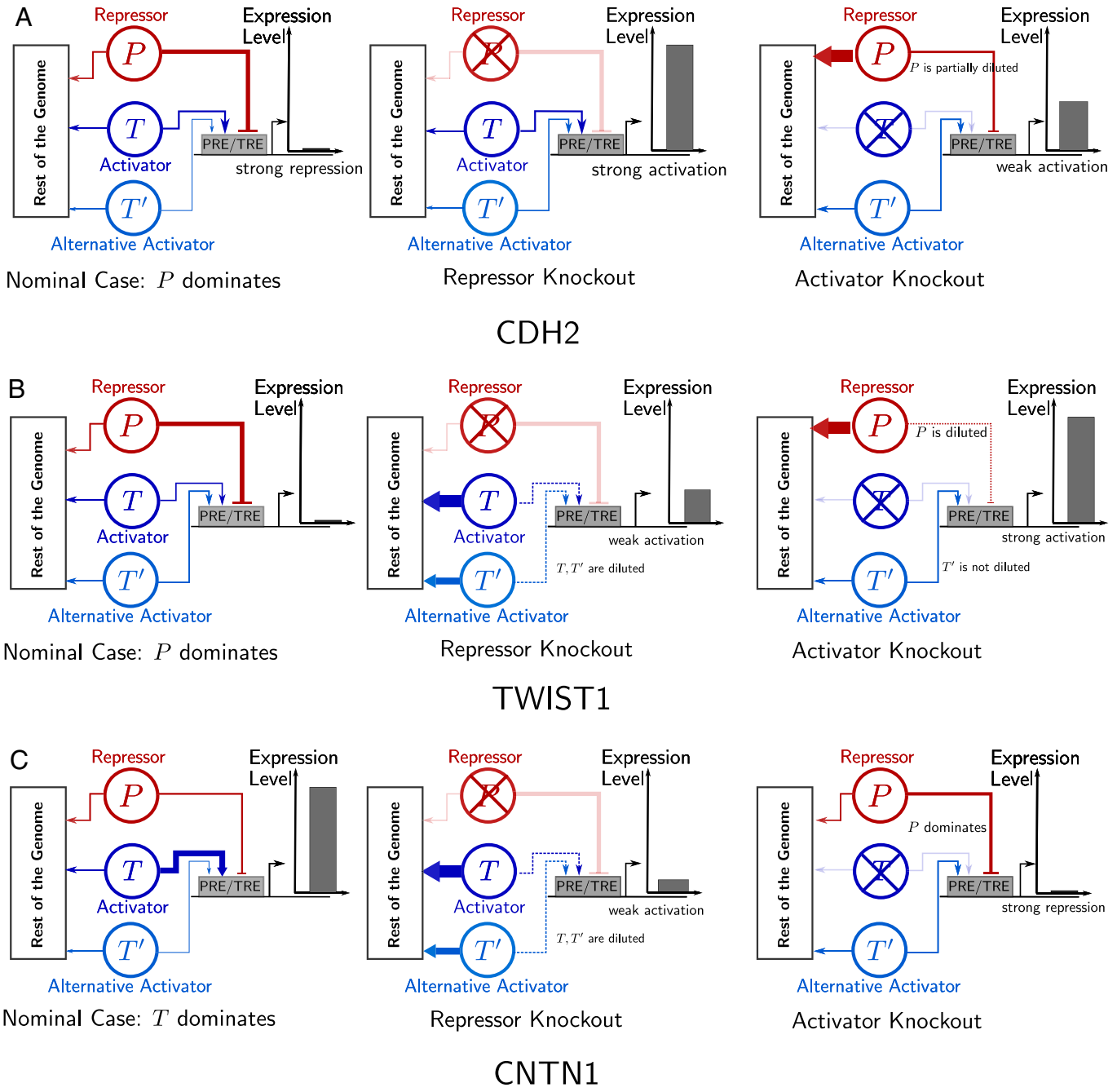



**Fig. 4.** Our model explains the differing behavior of PRC2 targets under two knockout experiments, as presented in ref. 6. Under PRC2 and KMT2D knockouts, the panels show our model's behavior for specific parameter sets that can explain the response of CDH2 (corresponds to the third column in Fig. 3) (*A*), TWIST1 (corresponds to the fifth column in Fig. 3) (*B*), and CNTN1 (corresponds to the first column in Fig. 3) (*C*). A very thick arrow denotes sequestration of the corresponding EF by other genes, a dotted arrow depicts dilution of the corresponding EF, and a lightly shaded arrow denotes an absent regulatory link due to knockout.

The small magnitude of the increase in TWIST1 expression upon PRC2 knockout can be explained by dilution of the activating EFs upon PRC2 knockout. Hence, our model's interpretation is that TWIST1 is operating at its nominal level without the presence of EF regulators.

The third case is even more surprising, where the knockout of KMT2D (an activating EF) causes TWIST1 to be strongly activated to an expression level that is multiple times greater than its expression level when PRC2 is knocked out. As explained in the previous subsection, PRC2 dilution cannot, by itself, explain this paradoxical disparity. This is since PRC2's dilution cannot be worse than a full PRC2 knockout. Mathematically, this implies the existence of an alternative activator, $T'$ (e.g., KMT2C), that rescues the expression of TWIST1 when PRC2 is diluted upon KMT2D knockout. This raises the following question: Why is the activating effect of $T'$ only observed when KMT2D is knocked out? One possibility within our modeling framework is that $T'$ binds weakly to targets across the genome compared to $P$. Therefore, when $T$ is knocked out, $P$ outcompetes $T'$ across the genome, and it gets diluted. This leaves $T'$ free to activate TWIST1.

***Repressor knockout results in repression.*** Another paradoxical behavior can be noticed when examining PRC2 targets that have the highest expression level in the control case. Such genes number 46 out of the 413 PRC2 targets. This set includes CNTN1 (Contactin 1) (Fig. 4*C*), which is highly expressed in the control

case, despite being a target of PRC2. Using our competition paradigm, this can mean that the activating EFs are dominant. Surprisingly, when the repressor PRC2 is knocked out, the expression of CNTN1 is significantly decreased. We interpret this as the outcome of dilution of the activators of CNTN1, caused by the knockout of PRC2. The result is less surprising when KMT2D is knocked out. As seen in Fig. 4*B*, loss of PRC2's competitor at CNTN1 allows for stronger repression. Another possible interpretation of the behavior under our modeling framework is that PRC2 is dominant locally, but its affinity to sites across the genome is much higher compared to the activator. Consequently, PRC2 cannot maintain repression of CNTN1 in the control case.

Overall, the power of our model stems from its versatility and its ability to account for local and global effects simultaneously. In the following sections, we will use our modeling framework to analyze the effect of PRC2 and KMT2D knockout on EMT in HMLER cells (6).

**Global EFs Modulate the Behavior of the Local GRN.** Zhang et al. (6) characterized the changes in the expression levels of multiple EMT markers in HMLER cells in response to the knockout of two histone methyltransferases, PRC2 and KMT2D. As shown in Fig. 5*A*, they reported an increase in the mesenchymal regulators ZEB1 and PRRX1 expression levels upon the knockout of both PRC2 and KMT2D, with a much higher fold change upon KMT2D knockout. Thus, both PRC2 and KMT2D knockout resulted in phenotypic change away from the epigenetic state, although to different extents. In contrast to changes in ZEB1 expression, SNAI1 expression, as compared to the control case, decreased upon KMT2D knockout and was the maximum under PRC2 knockout. This behavior is in disagreement with the traditional transcriptional picture, wherein SNAI1 is believed to activate ZEB1 expression, resulting in a positive correlation between their expression levels (67).

The change in ZEB1 and PRRX1 levels in response to PRC2 and KMT2D knockouts can be easily explained using the framework described in the previous section since both of these TFs are targets of PRC2, as reported by ref. 6. SNAI1, on the other hand, was not identified as a PRC2 target. To explain the unexpected changes in SNAI1 expression under epigenetic perturbations, we must integrate the transcriptional circuit involving ZEB1, PRRX1, and SNAI1 with our model of EF competition, as described below.

***Integration of EFs competition and local transcription regulation.*** We consider a system of coupled toggle switches, one involving ZEB1 and miR-200 (67) and another involving SNAI1 and PRRX1 (68). As shown in Fig. 5*E*, the two switches are coupled via the activation of ZEB1 and repression of miR-200 by SNAI1.

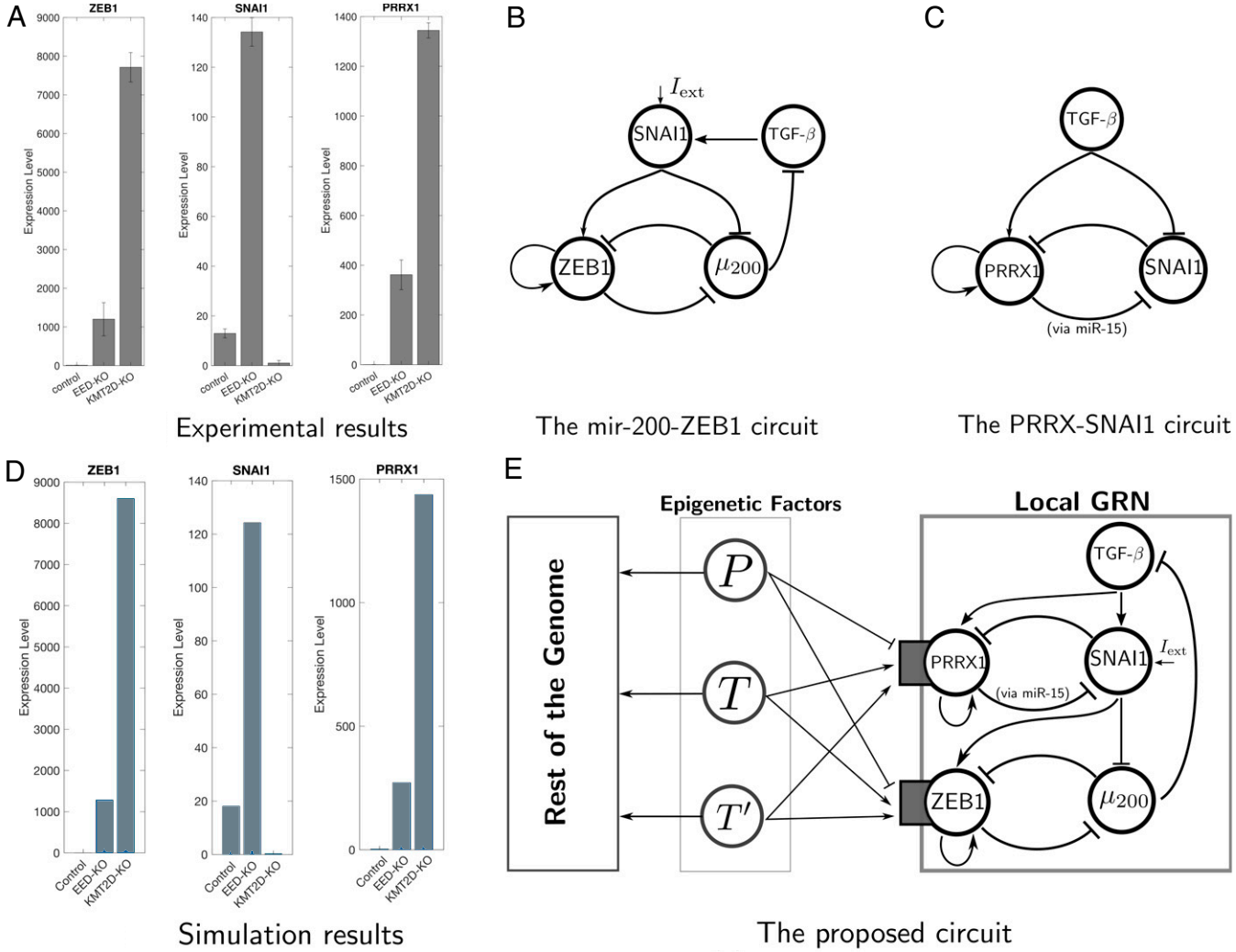



**Fig. 5.** The proposed modeling framework explains the EMT response under EF knockouts. (*A*) Experimental RNA-seq results from ref. 6. (*B*) The miR-200–ZEB1 circuit (67). (*C*) The PRRX1–SNAIL circuit (68). (*D*) Model simulations recapitulate the experimental results in *A*. (*E*) Our proposed model combining ternary epigenetic competition with the mir-200–ZEB1 circuit (67, 69) and the PRRX1–SNAIL circuit (70). *P* denotes PRC2, *T* denotes KMT2D, and *T*′ denotes an alternative epigenetic activator, such as KMT2C.

When PRC2 is knocked out, its targets (ZEB1 and PRRX1) are no longer strongly repressed, and they get modestly up-regulated. However, they do not get strongly activated due to the dilution of their activators (caused by PRC2 knockout). Consequently, the PRRX1–SNAI1 toggle switch will exhibit a high SNAI1 state. However, SNAI1 will be unable to activate ZEB1 in this scenario due to the dilution of ZEB1's epigenetic activator upon PRC2 knockout. When KMT2D is knocked out, PRC2 is diluted at ZEB1 and PRRX1, and, hence, the alternative activator $T'$ (which can be KMT2C) can fully activate both PRRX1 and ZEB1. The activation of PRRX1 results in the PRRX1–SNAI1 circuit switching to a high-PRRX1, low-SNAI1 state, in agreement with the experimentally reported behavior.

**The Local Transcriptional Context Determines the Effect of the EF Activity.** In the previous subsection, we have emphasized the effect of EFs on the local GRN. On the other hand, the transcriptional response in a GRN can also be influenced by epigenetic perturbations, mediated by the antagonistic interactions between PcG proteins and active transcription, as shown in Fig. 1*B* (see *SI Appendix* for detailed models). To demonstrate this effect, we consider the case of a single self-activating gene, here, ZEB1, as described below.

***Effect of single EF knockouts.*** ZEB1 is known to activate its own promoter (67). Therefore, we study the interaction between the self-activating feedback loop and the EF competition circuit. To that end, we consider change in the ZEB1 promoter activation level as a function of ZEB1 concentration under the knockout of individual EFs. The results are pictorially illustrated in Fig. 6*A*. In the control case (when both PRC2 and KMT2D are present), the activation level of the ZEB1 promoter increases very slowly with the ZEB1 protein concentration due to the inhibitory effect of PRC2. However, if PRC2 is knocked out, we see a sharp, Hill-function-like increase in ZEB1 promoter activation as a function of ZEB1 protein concentration. The activation level is low for low ZEB1 levels due to the dilution of KMT2D and other activators away from the ZEB1 promoter upon PRC2 knockout. However, when ZEB1 is abundantly available at its own promoter, the activation level increases due to the self-activatory loop. In the third case, if KMT2D is knocked out instead, the ZEB1 promoter remains highly active, even at low concentrations of ZEB1, and there is no substantial change in the activation level of the promoter with ZEB1 concentration. This is because KMT2D knockout is accompanied both by the dilution of repressive PRC2 away from the ZEB1 promoter and strong activity of the alternative activator (e.g., KMT2C) at ZEB1. Finally, the alternative activator knockout makes little difference on ZEB1 promoter activation, as compared to the control case. Thus, overall, Fig. 6*A* shows that epigenetic perturbations do not simply up-regulate or down-regulate their target genes: Such interventions can also change the response function of GRN, as shown here for the case of a GRN involving a single self-activating gene.

***Effect of double EF knockouts.*** To further illustrate the complexity of the transcription–epigenetic interplay, we next consider the effect of knocking out two EFs in different orders. Fig. 6*C* shows that starting from a GRN state with low ZEB1 expression (phenotypically corresponding to an epithelial state), the GRN will switch to a state with only modestly higher gene-expression level (corresponding to a quasi-mesenchymal phenotype). If this is followed by KMT2D knockout, the ZEB1 expression will decrease only slightly. PRC2 knockout, followed by KMT2D knockout, will thus result in epithelial cells switching to a quasi-mesenchymal state. In contrast, if KMT2D is knocked out in epithelial cells, Fig. 6 shows that the cells will switch to a mesenchymal state, one with very high ZEB1 expression. Thereafter, PRC2 knockout will lower the ZEB1 expression only slightly. Thus, double PRC2–KMT2D knockout in epithelial cells will have distinct phenotypic consequences: While PRC2 knockout followed by KMT2D knockout will cause these cells to switch to a quasi-mesenchymal state, KMT2D knockout followed by PRC2 knockout will result in the cells switching to a highly mesenchymal state.

**The Modeling Framework Provides Verifiable Predictions.** In the previous sections, we have developed a model that integrates the local GRN and global EFs. Using our framework, we can provide several predictions regarding the system studied in ref. 6, as discussed below.

***H3K27me3 redistributes and PRC2 is absent at the promoter of ZEB1 when KMT2D is knocked out.*** According to our model, an essential mechanism for the activation of ZEB1 and PRRX1 is PRC2's dilution when KMT2D is knocked out. The model predicts a redistribution of PRC2 across the genome, resulting in the redistribution of H3K27me3 marks. In addition, our model predicts PRC2's absence at the promoter of ZEB1 and PRRX upon KMT2D knockout. Indeed, it has been observed (6) that PRC2 is absent at the promoter of ZEB1 when KMT2D is knocked out and that H3K27me3 is redistributed to other genomic loci.

***EMT does not occur upon the knockout of the alternative epigenetic activator (e.g., KMT2C).*** Our proposed model (Fig. 5*E*) requires the existence of an alternative activator $T'$. Biologically, $T'$ can correspond to an alternate lysine methyltransferase, KMT2C. Using the same parameters used in the simulations depicted in Fig. 5*D*, we performed an in silico experiment by knocking out the

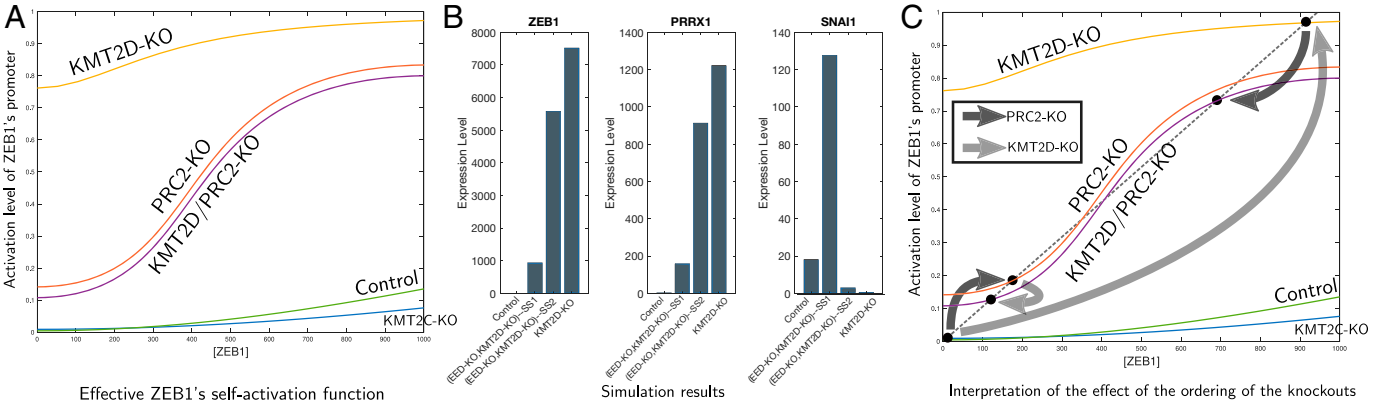



**Fig. 6.** The interplay between the epigenetic and transcriptional contexts. (*A*) The self-activation function of ZEB1 for various knockout scenarios, (*B*) Simulation results of the proposed network in Fig. 5*E* showing that the result of PRC2-KO is very different from KMT2D-KO followed by PRC2-KO. (*C*) Illustration of the serial knockout experiments using the concept of multistability.

alternative activator $T'$. The resulting behavior is indistinguishable from the control case. In other words, knocking out $T'$ fails to activate ZEB1 or PRRX1. This is indeed consistent with the screening performed in ref. 6, where knocking out KMT2C, for instance, did not result in EMT in HMLER cells.

***Simultaneous knockout of KMT2D and PRC2 will not result in strong activation of EMT genes.*** In silico experiments using the same parameters used for generating Fig. 5*D* show that the result of the simultaneous PRC2–KMT2D double knockout resembles the case of PRC2 knockout far more than the case of KMT2D knockout; i.e., ZEB1 and other EMT genes are not strongly activated under simultaneous PRC2–KMT2D double knockout.

***Knocking out PRC2 followed by KMT2D will not convert epithelial cells into a highly mesenchymal state.*** The activation curves shown in Fig. 6*A* provide us with a predication regarding the cellular response to the PRC2–KMT2D serial knockout experiment: PRC2 knockout will convert epithelial cells to a quasi-mesenchymal state. Thereafter, if PRC2 knockout is followed by KMT2D knockout, Fig. 6*C* shows our model prediction that the cells will stay in a quasi-mesenchymal state and will not switch to a highly mesenchymal state.

***Gradual knockouts of PRC2 and KMT2D have different signatures.*** Instead of the all-or-none knockout experiments analyzed before, we next consider the case of gradual EF knockouts. Indeed, inhibitors of EZH2 (a subcomponent of PRC2) are being investigated as therapy options in cancer (71–73). Fig. 7*A* shows that ZEB1 is more active when PRC2 is partially knocked out, compared to when it is fully knocked out. More precisely, it can be seen that ZEB1 becomes rapidly activated when PRC2's presence fraction goes from 0.6 to 0.5. This rapid activation would indicate the GRN behavior from a regime dominated by PcG to one dominated by TrX. However, when PRC2 is fully depleted, the levels of TrX get gradually diluted away from the ZEB1 promoter, leaving the PRE/TRE of ZEB1 in an unmodified state and weakly active.

Our model would suggest that this effect can be more pronounced: Simulations with different model parameters, as shown in Fig. 7*B*, would indicate the possibility of a situation where a small dip in the level of PRC2 can cause a rapid collapse in its activity at the ZEB1 promoter. This can be contrasted with the mechanism of activation in the case of a gradual KMT2D knockout. In that scenario, the activity of the ZEB1 PRE/TRE builds up slowly as PRC2 gets increasingly diluted and as the alternative epigenetic activator gets the full chance to activate ZEB1, as shown in Fig. 7*C*. From these results, we can make the counterintuitive prediction that that the partial knockout of an epigenetic repressor—here, PRC2—can have a stronger repressive effect on gene activity, as compared to a complete PRC2 knockout.

## Discussion

In this work, we have described a modeling framework that combines local transcriptional regulation with global epigenetic control and showed that complex interplay between transcriptional and epigenetic control can lead to rich gene-expression dynamics. We have used our modeling framework to understand the effects of various epigenetic perturbations on EMT, a crucial cellular process involved in both health and disease. We note that interplay between epigenetic competition, the EMT transcriptional network, and the baseline transcriptional context can result in counterintuitive experimental observations and generate unique paths for cells to transition between epithelial and mesenchymal states. Furthermore, our results indicate that the rest of the genome exerts an indirect effect on the behavior of the local GRN by competing for the same EFs. So when one of the EFs is knocked out, the new empty sites across the rest of the genome try to sequester the other EFs present at the local GRN. As a result, the local landscape is reshaped, leading to new steady states. A key property of the model is the asymmetry between the behaviors of different EFs, as shown in Fig. 3. In other words, the number of global sites that are made available to a specific EF depends on the identity of the eliminated and the competing EFs and the local and global affinities of the considered EFs.

**Further Evidence of Redistribution and Dilution of PcG Proteins.** Our modeling framework relies upon the competition between PcG and TrX proteins to modify histones at the same or nearby genomic sites and the redistribution of PcG that accompanies TrX knockout. Such an effect is not just restricted to TrX and PcG proteins. Another line of research has investigated the relationship between DNA methylation, H3K9me3, and H3K27me3. Despite the fact that all three are silencing marks, they do not usually mark the same genomic locations. DNA methylation and H3K9me3 mark constitutive heterochromatin, while H3K27me3 marks facultative heterochromatin (74, 75). Induced DNA hypomethylation can cause H3K27me3 to disappear from its normal locations and accumulate at new genomic sites in *Arabidopsis thaliana* (74), mouse somatic cells (76), and mouse ESCs (41). Hence, inhibition of DNA demethylation can open up new locations for PRC2 recruitment,

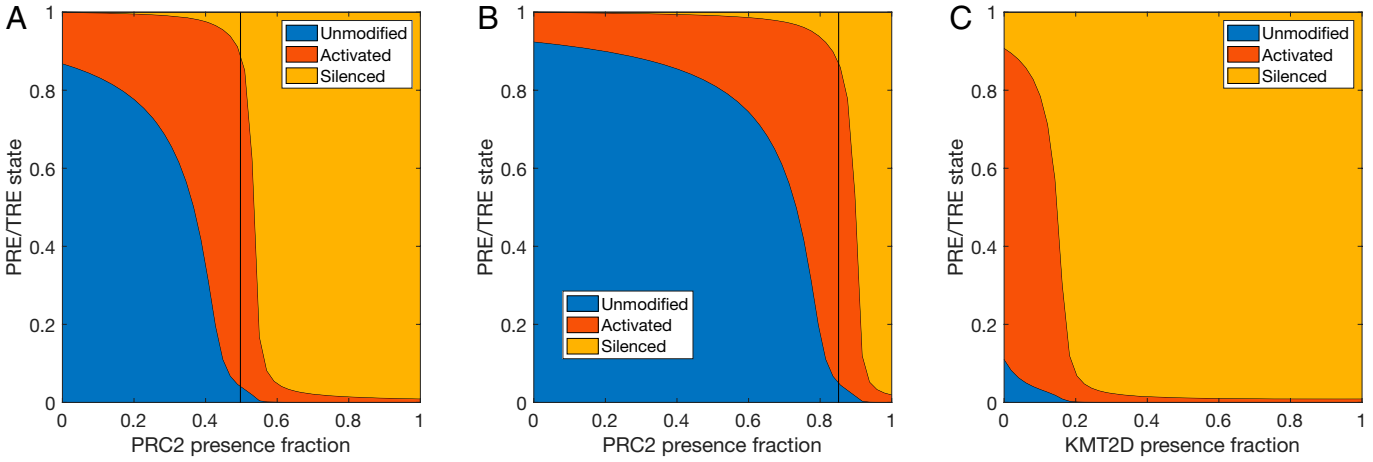



**Fig. 7.** Predictions of the effect of gradual knockout experiments on ZEB1. (*A*) Gradual PRC2 knockout. (*B*) Gradual PRC2 knockout: fast activation. (*C*) Gradual KMT2D knockout.

sequestering it from PcG-repressed genes (41). Similarly, elimination of H3K9me3 has been shown to cause H3K27me3 to disappear from its normal genomic locations and redistribute to new genomic loci in the fungus *Neurospora* (75, 77). It is worth noting that PRC2 sequestration and redistribution is not limited to competition scenarios. For instance, it has been observed that PRC2 redistributes across the genome upon modifying ATRX, a chromatin remodeler that assists PRC2's binding (78).

**Disorder in PcG/TrX Proteins Causes Disease and Is a Target for Therapeutics.** Given the global activity of PcG and TrX proteins, it is expected that their knockout will have far-reaching and detrimental effects on cells. Indeed, the inactivation or aberrant activation of such proteins has been shown to play a key role in the emergence of cancer (79). PRC2 has been studied extensively in that context (80, 81), and its catalytic component EZH2 has been tested as a therapeutic target in multiple clinical trials (71–73). Similarly, disorders in COMPASS proteins are very common in cancers (70, 82) and have been proposed as key regulators and potential therapeutic targets (83–87). Our results imply that such therapeutic interventions must proceed with the utmost caution by accounting for the global context.

**Similar EFs Can Play Different Roles.** Our results show that EFs with the same enzymatic activity, such as KMT2C/D, both of which deposit the same methylation mark on histone tails, can exhibit very different biological and functional behaviors. This has been reported in multiple experimental contexts. For example, Zhang et al. (6) found that while KMT2D knockout could induce transition to a mesenchymal state in HMLER cells, KMT2C was not identified to be among the key EMT inducers. Similarly, in MCF10A cells, TGF-$\beta$–induced EMT is accompanied by the up-regulation of the H3K27me3 demethylase KDM6B, while the enzymatically similar KDM6A is down-regulated during the process (88).

**Competition Effects in Molecular Biology.** Competition effects have been studied earlier in the context of synthetic biological circuits, where circuits compete for RNA polymerases and ribosomes. It has been shown that such competition can cause nontrivial coupling between isolated components and affect protein-expression performance (89, 90). It has also been investigated in the context in the design of Boolean genetic circuits via CRISPRi, which uses dCas9 as a shared resource (91). More detailed models of messenger RNA's competition for ribosomes during the transcription process have also been proposed (92, 93). Similarly, competition for the same micro-RNAs between transcripts that have the same or similar micro-RNA–binding sequence motifs has been shown to induce coupling between the expression levels of seemingly independent proteins (94, 95).

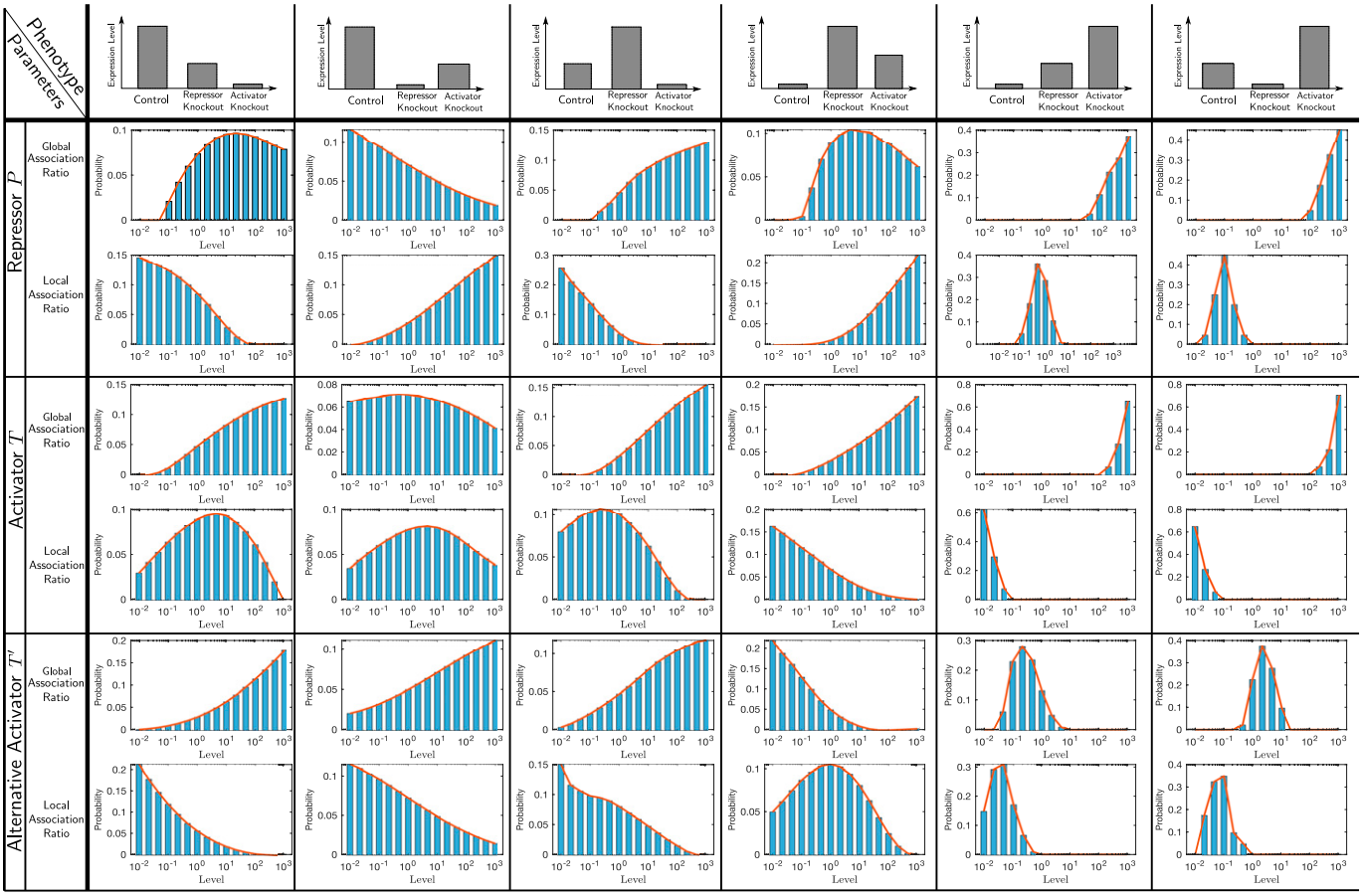



**Fig. 8.** All phenotypes are possible under the appropriate mix of global and local contexts. This is a more detailed version of Fig. 3. Each plot depicts the marginal probability distribution of the parameter under consideration conditioned on the phenotype under consideration. The marking ratios are fixed, while the association ratios are varied. The six association ratios are varied with 16 levels between $10^{-3}$ and $10^{3}$. This provides $6^{16} \approx 2.8211 \times 10^{12}$ sets of parameters, for each of which the steady states are calculated numerically for the control and the knockout cases. A set of parameters is said to give one of the six phenotypes if the highest expression level (among the three cases) is at least three times higher than the second highest and the latter is at least three times higher than the third highest. To generate this table, the total $P$, $T$, $T'$ levels are 1,000; 500; and 100, respectively. The local marking ratios are fixed to 1, while the global marking ratios are fixed to 0. The total copy number of the local gene is fixed to 1, while the total copy number of the rest of the genome is fixed to 1,000.

**Experiments Involving Epigenetic Perturbations Must Be Analyzed with Care.** One crucial takeaway from the analysis presented in this manuscript is the possibility of widespread cross-talk between the genomic targets of different EFs. Most experimental studies analyzing the effect of epigenetic perturbations follow a set procedure: Characterize the transcription profiles before and after knocking out an epigenetic modifier, identify the set of differently expressed genes (which often number in the hundreds), and carry out gene-set enrichment analysis (96) or Gene Ontology enrichment analysis (97) using the differently expressed genes. In light of the framework described here, it is unsurprising that the outcome of epigenetic perturbations is hundreds of differently expressed genes, a list that is then arbitrarily whittled down, depending on the biological interests of the researchers carrying out the analysis. This is usually followed by choosing a pathway or biological process of interest and analyzing how it is affected by the given epigenetic modifier. One would then conclude with identifying that epigenetic modifier as a key regulator of that biological process. Our analysis shows that any such conclusion could be highly unreliable outside the context of the specific experimental setup. For example, in the study by Zhang et al. (6), SNAI1 expression is up-regulated upon PRC2 knockout, even though it is not a direct target of PRC2. While a simple differential gene-expression analysis might lead one to identify PRC2 as a key regulator of SNAI1 expression, our analysis shows that the effect of PRC2 on SNAI1 expression can only be explained by the complex interplay between epigenetic control and the GRN involving SNAI1, ZEB1, and PRRX1. Thus, PRC2 knockout may have no effect on SNAI1 expression in cells wherein the SNAI1–ZEB1–PRRX1 GRN is inactive. Moreover, we show that the effect of epigenetic perturbations on gene expression will depend on the cell's transcriptional state at the time of the epigenetic perturbation. This would suggest that the same epigenetic perturbation could have very different effects on genotypically identical cells in different phenotypic states. For example, PRC2 knockout has been shown to trigger differentiation in primed mouse ESCs, but not in naive ESCs (98). Thus, careful analysis of the effects of epigenetic perturbations will also require that the effects be analyzed within the transcriptional context of the cells.

## Materials and Methods

**Datasets.** The datasets used in this study include the full RNA-sequencing (RNA-seq) data sheet and the list of PRC2 targets for the experiments reported in ref. 6. The data have been provided by Y.Z., who is the lead author in the aforementioned study.

**Mathematical Modeling.** The details of the mathematical models and numerical simulations are provided in *SI Appendix* and are very briefly summarized here.

***Constitutively expressed genes subject to EF competition.*** In *SI Appendix*, section 1, we describe the reaction network models for a general network of $N$ EFs and $n$ genes. We show that each interaction of an EF with a gene can be characterized by two parameters: the association ratio and the marking ratio. In particular, the results in Fig. 3 are generated by studying the effect of the parameters on the three experimental scenarios discussed in *The Model Explains Paradoxical Knockout Results*. A more detailed version of Fig. 3 is shown in Fig. 8.

***Self-activating genes subject to EF competition.*** In *SI Appendix*, section 2, we provide our mathematical model for the interaction between the transcriptional and epigenetic components of regulation for a group of $n$ self-activating genes and $N$ EFs. We illustrate this numerically by showing the effect of EF perturbations on the activation function of a single self-activating gene.

***A general local GRN subject to EF competition.*** In *SI Appendix*, section 3, we model a generic local GRN with TF and micro-RNA regulations subject to EF competition. We provide our concrete model for the network in Fig. 5*E*, along with the parameters utilized in the simulations.

***Parameter selection.*** The feasible parameters are not unique. The simulation parameters have been chosen to reproduce the qualitative behavior of the experimental results. For the epigenetic competition circuit, the parameters have been chosen based on a screen similar to those shown in Fig. 8 and *SI Appendix*, Tables 1–3. The parameters of the local GRN have been chosen by refining an initial parameter set generated by the software package RACIPE (9). The parameters are listed in *SI Appendix*.

**Software.** Numerical simulations have been performed via MATLAB R2020a on the discovery high-performance computing cluster at Northeastern University.

**Data, Materials, and Software Availability.** Previously published data were used for this work (GSE158115) (99). The code used to generate the figures is posted at https://github.com/malirdwi/EpigeneticFactorCompetition.

**ACKNOWLEDGMENTS.** M.A.A.-R., S.T., and H.L. were supported by NSF Grant PHY-2019745. M.A.A.-R. and E.D.S. were supported in part by NSF Grant NSF/DMS-2052455. E.D.S. was supported in part by Air Force Office of Scientific Research Grants FA9550-21-1-0289 and 22RT0159. Y.Z. was supported by Susan G. Komen Postdoctoral Fellowship PDF15301255. We thank Robert A. Weinberg for useful conversations and his support of the work of Y.Z. We also acknowledge fruitful discussions with Sendurai Mani.

Author affiliations: [a]Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115; [b]Center for Theoretical Biological Physics, Northeastern University, Boston, MA 02115; [c]PhD Program in Systems, Synthetic, and Physical Biology, Rice University, Houston, TX 77005; [d]Department of Physics, Northeastern University, Boston, MA 02115; [e]Whitehead Institute for Biomedical Research, Cambridge, MA 02142; [f]State Key Laboratory of Molecular Oncology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China; [g]Department of Bioengineering, Northeastern University, Boston, MA 02115; and [h]Laboratory of Systems Pharmacology, Harvard Medical School, Boston, MA 02115

1. L. Wolpert, C. Tickle, A. M. Arias, *Principles of Development* (Oxford University Press, Oxford, UK, ed. 6, 2019).
2. M. A. Nieto, R. Y.-J. Huang, R. A. Jackson, J. P. Thiery, EMT: 2016. *Cell* **166**, 21–45 (2016).
3. C.-Y. Wu, Y.-P. Tsai, M.-Z. Wu, S.-C. Teng, K.-J. Wu, Epigenetic reprogramming and post-transcriptional regulation during the epithelial–mesenchymal transition. *Trends Genet.* **28**, 454–463 (2012).
4. L. Eichelberger *et al.*, Maintenance of epithelial traits and resistance to mesenchymal reprogramming promote proliferation in metastatic breast cancer. *bioRxiv* [Preprint] (2020). https://doi.org/10.1101/2020.03.19.998823. Accessed 20 June 2022.
5. C. L. Chaffer *et al.*, Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. *Cell* **154**, 61–74 (2013).
6. Y. Zhang *et al.*, Genome-wide CRISPR screen identifies PRC2 and KMT2D-compass as regulators of distinct EMT trajectories that contribute differentially to metastasis. *Nat. Cell Biol.* **24**, 554–564 (2022).
7. G. Karlebach, R. Shamir, Modelling and analysis of gene regulatory networks. *Nat. Rev. Mol. Cell Biol.* **9**, 770–780 (2008).
8. S. Huang, G. Eichler, Y. Bar-Yam, D. E. Ingber, Cell fates as high-dimensional attractor states of a complex gene regulatory network. *Phys. Rev. Lett.* **94**, 128701 (2005).
9. B. Huang *et al.*, Interrogating the topological robustness of gene regulatory circuits by randomization. *PLoS Comput. Biol.* **13**, e1005456 (2017).
10. S. Kauffman, "Gene regulation networks: A theory for their global structure and behaviors" in *Current Topics in Developmental Biology*, A. A. Moscona, A. Monroy, Eds. (Elsevier, Amsterdam, 1971), vol. 6, pp. 145–182.
11. U. Alon, *An Introduction to Systems Biology* (CRC Press, Boca Raton, FL, 2006).
12. M. Lu *et al.*, Tristability in cancer-associated microRNA-TF chimera toggle switch. *J. Phys. Chem. B* **117**, 13164–13174 (2013).
13. V. Olariu, C. Lövkvist, K. Sneppen, Nanog, Oct4 and Tet1 interplay in establishing pluripotency. *Sci. Rep.* **6**, 25438 (2016).
14. T. Chen, M. A. Al-Radhawi, E. D. Sontag, A mathematical model exhibiting the effect of DNA methylation on the stability boundary in cell-fate networks. *Epigenetics* **16**, 436–457 (2021).
15. T. Thalheim, M. Herberg, M. Loeffler, J. Galle, The regulatory capacity of bivalent genes–A theoretical approach. *Int. J. Mol. Sci.* **18**, 1069 (2017).
16. Y. Zhang, N. Liu, W. Lin, C. Li, Quantifying the interplay between genetic and epigenetic regulations in stem cell development. *New J. Phys.* **21**, 103042 (2019).
17. T. Alarcón, J. Sardanyés, A. Guillamon, J. A. Menendez, Bivalent chromatin as a therapeutic target in cancer: An in silico predictive approach for combining epigenetic drugs. *PLoS Comput. Biol.* **17**, e1008408 (2021).
18. S. Bruno, R. J. Williams, D. Del Vecchio, Epigenetic cell memory: The gene's inner chromatin modification circuit. *PLOS Comput. Biol.* **18**, e1009961 (2022).
19. K. Sneppen, L. Ringrose, Theoretical analysis of Polycomb-trithorax systems predicts that poised chromatin is bistable and not bivalent. *Nat. Commun.* **10**, 2133 (2019).
20. W. Zhao, L. Qiao, S. Yan, Q. Nie, L. Zhang, Mathematical modeling of histone modifications reveals the formation mechanism and function of bivalent chromatin. *iScience* **24**, 102732 (2021).
21. J. Reinig, F. Ruge, M. Howard, L. Ringrose, A theoretical model of Polycomb/Trithorax action unites stable epigenetic memory and dynamic regulation. *Nat. Commun.* **11**, 4782 (2020).

22. M. K. Jolly, H. Levine, Computational systems biology of epithelial-hybrid-mesenchymal transitions. *Curr. Opin. Syst. Biol.* **3**, 1–6 (2017).
23. W. Jia, A. Deshmukh, S. A. Mani, M. K. Jolly, H. Levine. A possible role for epigenetic feedback regulation in the dynamics of the epithelial–mesenchymal transition (EMT). *Phys. Biol.* **16**, 066004 (2019).
24. W. Jia *et al.*, Epigenetic feedback and stochastic partitioning during cell division can drive resistance to EMT. *Oncotarget* **11**, 2611–2624 (2020).
25. L. G. Karacosta *et al.*, Mapping lung cancer epithelial-mesenchymal transition states and trajectories with single-cell resolution. *Nat. Commun.* **10**, 5587 (2019).
26. A. Piunti, A. Shilatifard, Epigenetic balance of gene expression by Polycomb and COMPASS families. *Science* **352**, aad9780 (2016).
27. B. Schuettengruber, H.-M. Bourbon, L. Di Croce, G. Cavalli, Genome regulation by polycomb and trithorax: 70 years and counting. *Cell* **171**, 34–57 (2017).
28. R. Margueron, D. Reinberg, The Polycomb complex PRC2 and its mark in life. *Nature* **469**, 343–349 (2011).
29. E. Conway, E. Healy, A. P. Bracken, PRC2 mediated H3K27 methylations in cellular identity and cancer. *Curr. Opin. Cell Biol.* **37**, 42–48 (2015).
30. A. P. Bracken, N. Dietrich, D. Pasini, K. H. Hansen, K. Helin, Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev.* **20**, 1123–1136 (2006).
31. F. Mohn *et al.*, Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol. Cell* **30**, 755–766 (2008).
32. A. Shilatifard, The COMPASS family of histone H3K4 methylases: Mechanisms of regulation in development and disease pathogenesis. *Annu. Rev. Biochem.* **81**, 65–95 (2012).
33. M. Wu *et al.*, Molecular regulation of H3K4 trimethylation by Wdr82, a component of human Set1/COMPASS. *Mol. Cell. Biol.* **28**, 7337–7344 (2008).
34. D. Hu *et al.*, The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. *Mol. Cell. Biol.* **33**, 4745–4754 (2013).
35. E. Froimchuk, Y. Jang, K. Ge, Histone H3 lysine 4 methyltransferase KMT2D. *Gene* **627**, 337–342 (2017).
36. D. Hu *et al.*, Not all H3K4 methylations are created equal: Mll2/COMPASS dependency in primordial germ cell specification. *Mol. Cell* **65**, 460–475.e6 (2017).
37. J. Zhang *et al.*, Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. *Nat. Med.* **21**, 1190–1198 (2015).
38. S. S. Dhar *et al.*, MLL4 is required to maintain broad H3K4me3 peaks and super-enhancers at tumor suppressor genes. *Mol. Cell* **70**, 825–841.e6 (2018).
39. S.-Y. Ang *et al.*, KMT2D regulates specific programs in heart development via histone H3 lysine 4 di-methylation. *Development* **143**, 810–821 (2016).
40. C. Wang *et al.*, Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 11871–11876 (2016).
41. D. Douillet *et al.*, Uncoupling histone H3K4 trimethylation from developmental gene expression via an equilibrium of COMPASS, Polycomb and DNA methylation. *Nat. Genet.* **52**, 615–625 (2020).
42. P. A. Steffen, L. Ringrose, What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory. *Nat. Rev. Mol. Cell Biol.* **15**, 340–356 (2014).
43. T. Klymenko, J. Müller, The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. *EMBO Rep.* **5**, 373–377 (2004).
44. F. Tie *et al.*, CBP-mediated acetylation of histone H3 lysine 27 antagonizes *Drosophila* Polycomb silencing. *Development* **136**, 3131–3141 (2009).
45. F. W. Schmitges *et al.*, Histone methylation by PRC2 is inhibited by active chromatin marks. *Mol. Cell* **42**, 330–341 (2011).
46. D. Pasini *et al.*, Characterization of an antagonistic switch between histone H3 lysine 27 methylation and acetylation in the transcriptional regulation of Polycomb group target genes. *Nucleic Acids Res.* **38**, 4958–4969 (2010).
47. F. Tie *et al.*, Polycomb inhibits histone acetylation by CBP by binding directly to its catalytic domain. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E744–E753 (2016).
48. E. Blanco, M. González-Ramírez, A. Alcaine-Colet, S. Aranda, L. Di Croce, The bivalent genome: Characterization, structure, and regulation. *Trends Genet.* **36**, 118–131 (2020).
49. E. Shema *et al.*, Single-molecule decoding of combinatorially modified nucleosomes. *Science* **352**, 717–721 (2016).
50. D. Pasini *et al.*, Coordinated regulation of transcriptional repression by the RBP2 H3K4 demethylase and Polycomb-Repressive Complex 2. *Genes Dev.* **22**, 1345–1355 (2008).
51. Y. Jin *et al.*, LSD1 collaborates with EZH2 to regulate expression of interferon-stimulated genes. *Biomed. Pharmacother.* **88**, 728–737 (2017).
52. Y. Guo, S. Zhao, G. G. Wang, Polycomb gene silencing mechanisms: PRC2 chromatin targeting, H3K27me3 "readout", and phase separation-based compaction. *Trends Genet.* **37**, 547–565 (2021).
53. D. Holoch, R. Margueron, Mechanisms regulating PRC2 recruitment and enzymatic activity. *Trends Biochem. Sci.* **42**, 531–542 (2017).
54. E. T. Wiles, E. U. Selker, H3K27 methylation: A promiscuous repressive chromatin mark. *Curr. Opin. Genet. Dev.* **43**, 31–37 (2017).
55. M. Uckelmann, C. Davidovich, Not just a writer: PRC2 as a chromatin reader. *Biochem. Soc. Trans.* **49**, 1159–1170 (2021).
56. M. Beltran *et al.*, The interaction of PRC2 with RNA or chromatin is mutually antagonistic. *Genome Res.* **26**, 896–907 (2016).
57. W. Yuan *et al.*, Dense chromatin activates polycomb repressive complex 2 to regulate H3 lysine 27 methylation. *Science* **337**, 971–975 (2012).
58. L. Pérez *et al.*, Enhancer-PRE communication contributes to the expansion of gene expression domains in proliferating primordia. *Development* **138**, 3125–3134 (2011).
59. L. J. Gaydos, A. Rechtsteiner, T. A. Egelhofer, C. R. Carroll, S. Strome, Antagonism between MES-4 and polycomb repressive complex 2 promotes appropriate gene expression in *C. elegans* germ cells. *Cell Rep.* **2**, 1169–1177 (2012).
60. C. Lu *et al.*, Histone H3K36 mutations promote sarcomagenesis through altered histone methylation landscape. *Science* **352**, 844–849 (2016).
61. A. Chaouch *et al.*, Histone H3.3 K27M and K36M mutations de-repress transposable elements through perturbation of antagonistic chromatin marks. *Mol. Cell* **81**, 4876–4890.e7 (2021).
62. C. M. Weber *et al.*, mSWI/SNF promotes Polycomb repression both directly and through genome-wide redistribution. *Nat. Struct. Mol. Biol.* **28**, 501–511 (2021).
63. T. Mortimer *et al.*, Redistribution of EZH2 promotes malignant phenotypes by rewiring developmental programmes. *EMBO Rep.* **20**, e48155 (2019).
64. S. Khazaei *et al.*, H3.3 G34W promotes growth and impedes differentiation of osteoblast-like mesenchymal progenitors in giant cell tumor of bone. *Cancer Discov.* **10**, 1968–1987 (2020).
65. C. W. Hanna *et al.*, Loss of histone methyltransferase SETD1B in oogenesis results in the redistribution of genomic histone 3 lysine 4 trimethylation. *Nucleic Acids Res.* **50**, 1993–2004 (2022).
66. M. Walter, A. Teissandier, R. Pérez-Palacios, D. Bourc'his, An epigenetic switch ensures transposon repression upon dynamic loss of DNA methylation in embryonic stem cells. *eLife* **5**, e11418 (2016).
67. M. Lu, M. K. Jolly, H. Levine, J. N. Onuchic, E. Ben-Jacob, MicroRNA-based regulation of epithelial-hybrid-mesenchymal fate determination. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 18144–18149 (2013).
68. H. Fazilaty *et al.*, A gene regulatory network to control EMT programs in development and disease. *Nat. Commun.* **10**, 1–16 (2019).
69. F. Bocci, H. Levine, J. N. Onuchic, M. K. Jolly, Deciphering the dynamics of epithelial-mesenchymal transition and cancer stem cells in tumor progression. *Curr. Stem Cell Rep.* **5**, 11–21 (2019).
70. R. J. Fagan, A. K. Dingwall, COMPASS ascending: Emerging clues regarding the roles of MLL3/KMT2C and MLL2/KMT2D proteins in cancer. *Cancer Lett.* **458**, 56–65 (2019).
71. K. H. Kim, C. W. Roberts, Targeting EZH2 in cancer. *Nat. Med.* **22**, 128–134 (2016).
72. R. Duan, W. Du, W. Guo, EZH2: A novel target for cancer treatment. *J. Hematol. Oncol.* **13**, 104 (2020).
73. M.-L. Eich, M. Athar, J. E. Ferguson III, S. Varambally, EZH2-targeted therapies in cancer: Hype or a reality. *Cancer Res.* **80**, 5449–5458 (2020).
74. A. Deleris *et al.*, Loss of the DNA methyltransferase MET1 induces H3K9 hypermethylation at PcG target genes and redistribution of H3K27 trimethylation to transposons in *Arabidopsis thaliana*. *PLOS Genet.* **8**, e1003062 (2012).
75. E. Y. Basenko *et al.*, Genome-wide redistribution of H3K27me3 is linked to genotoxic stress and defective growth. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E6339–E6348 (2015).
76. J. P. Reddington *et al.*, Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of Polycomb target genes. *Genome Biol.* **14**, R25 (2013).
77. K. Jamieson *et al.*, Loss of HP1 causes depletion of H3K27me3 from facultative heterochromatin and gain of H3K27me2 at constitutive heterochromatin. *Genome Res.* **26**, 97–107 (2016).
78. W. Ren *et al.*, Disruption of ATRX-RNA interactions uncovers roles in ATRX localization and PRC2 function. *Nat. Commun.* **11**, 2219 (2020).
79. M. A. Morgan, A. Shilatifard, Chromatin signatures of cancer. *Genes Dev.* **29**, 238–249 (2015).
80. A. Laugesen, J. W. Højfeldt, K. Helin, Role of the polycomb repressive complex 2 (PRC2) in transcriptional regulation and cancer. *Cold Spring Harb. Perspect. Med.* **6**, a026575 (2016).
81. M. Dockerill, C. Gregson, D. H. O' Donovan, Targeting PRC2 for the treatment of cancer: An updated patent review (2016–2020). *Expert Opin. Ther. Pat.* **31**, 119–135 (2021).
82. C. C. Sze, A. Shilatifard, MLL3/MLL4/COMPASS family on epigenetic regulation of enhancer function and cancer. *Cold Spring Harb. Perspect. Med.* **6**, a026427 (2016).
83. J. B. Dawkins *et al.*, Reduced expression of histone methyltransferases KMT2C and KMT2D correlates with improved outcome in pancreatic ductal adenocarcinoma. *Cancer Res.* **76**, 4861–4871 (2016).
84. K. Gala *et al.*, KMT2C mediates the estrogen dependence of breast cancer through regulation of ER$\alpha$ enhancer function. *Oncogene* **37**, 4692–4710 (2018).
85. W. Xiong *et al.*, MLL3 enhances the transcription of PD-L1 and regulates anti-tumor immunity. *Biochim. Biophys. Acta Mol. Basis Dis.* **1865**, 454–463 (2019).
86. Q. Yu *et al.*, Small molecule inhibitors of the prostate cancer target KMT2D. *Biochem. Biophys. Res. Commun.* **533**, 540–547 (2020).
87. S. S. Dhar, M. G. Lee, Cancer-epigenetic function of the histone methyltransferase KMT2D and therapeutic opportunities for the treatment of KMT2D-deficient tumors. *Oncotarget* **12**, 1296–1308 (2021).
88. J. H. Taube *et al.*, The H3K27me3-demethylase KDM6A is suppressed in breast cancer stem-like cells, and enables the resolution of bivalency during the mesenchymal-epithelial transition. *Oncotarget* **8**, 65548–65565 (2017).
89. Y. Qian, H.-H. Huang, J. I. Jiménez, D. Del Vecchio, Resource competition shapes the response of genetic circuits. *ACS Synth. Biol.* **6**, 1263–1272 (2017).
90. M. Ali Al-Radhawi, D. D. Vecchio, E. D. Sontag, "Identifying Competition Phenotypes in Synthetic Biochemical Circuits" in IEEE Control Systems Letters, vol. 7 (2023), pp. 211–216.
91. S. Zhang, C. A. Voigt, Engineered dCas9 with reduced toxicity in bacteria: Implications for genetic circuit design. *Nucleic Acids Res.* **46**, 11115–11125 (2018).
92. A. Raveh, M. Margaliot, E. D. Sontag, T. Tuller, A model for competition for ribosomes in the cell. *J. R. Soc. Interface* **13**, 20151062 (2016).
93. J. Miller, M. A. Al-Radhawi, E. D. Sontag, Mediating ribosomal competition by splitting pools. *IEEE Control Syst. Lett.* **5**, 1555–1560 (2021).
94. C. Bosia *et al.*, RNAs competing for microRNAs mutually influence their fluctuations in a highly non-linear microRNA-dependent manner in single cells. *Genome Biol.* **18**, 37 (2017).
95. L. Salmena, L. Poliseno, Y. Tay, L. Kats, P. P. Pandolfi. A ceRNA hypothesis: The Rosetta Stone of a hidden RNA language? *Cell* **146**, 353–358 (2011).
96. A. Subramanian *et al.*, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 15545–15550.
97. H. Mi, A. Muruganujan, D. Ebert, X. Huang, P. D. Thomas, PANTHER version 14: More genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. *Nucleic Acids Res.* **47**, D419–D426 (2018).
98. Y. Shan *et al.*, PRC2 specifies ectoderm lineages and maintains pluripotency in primed but not naïve ESCs. *Nat. Commun.* **8**, 672 (2017).
99. Y. Zhang, R. Weinberg, Single Cell RNA-seq data of C1-sgControl, C1-sgEED and C1-sgKMT2D cells. Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE201441. Deposited 30 April 2022.

# Supplement to "Epigenetic factor competition reshapes the EMT landscape"

M. Ali Al-Radhawi, Shubham Tripathi, Yun Zhang,
Eduardo Sontag and Herbert Levine

September 25, 2022

This supplement contains the details of the mathematical models utilized, additional numerical simulations, and the parameters used.

## Notation

We use the formalism of Biological Interaction Networks (BIN) (also known as, Chemical Reaction Networks) as expounded in [1–3]. We review basic notations below.

A reaction network consists of *species* $\mathscr{S} = \{\mathrm{X}_1, .., \mathrm{X}_n\}$ and *reactions* $\mathscr{R} = \{\mathrm{R}_1, .., \mathrm{R}_\nu\}$. Examples of species include binding sites, PRE/TREs, mRNAs, proteins, etc. Examples of reactions include binding, unbinding, dimerization, production, decay, etc. More formally, the reaction $\mathrm{R_j}$ is written in the following form: $\sum_{i=1}^{n} \alpha_{ij}\mathrm{X}_i \longrightarrow \sum_{i=1}^{n} \beta_{ij}\mathrm{X}_i$, where $\alpha_{ij}, \beta_{ij}$ are nonnegative integers. The stoichiometry matrix of the network is defined element-wise as $[\Gamma]_{ij} = \beta_{ij} - \alpha_{ij}$ and it describes the net gain or loss of the $i$th species at the $j$th reaction.

Each species $\mathrm{X}_i$ is a quantified by a concentration $X_i \geq 0$, while a reaction $\mathrm{R_j}$ is quantified by a reaction rate or velocity function $V_j$. We use the standard Mass-Action kinetics written as follows: $V_j(x) = \prod_{i=1}^{n} k_j x_i^{\alpha_{ij}}$, where $k_j's$ are the kinetic constants.

In order to describe the time-evolution of the network, the corresponding Ordinary Differential Equation (ODE) can be written as follows:

$$\dot{X} = \Gamma V(X), X(0) = X_\circ. \tag{S1}$$

BINs usually have conserved quantities. In our work, we assume that the genes and the Epigenetic Factors (EFs) are conserved. Mathematically, this means that there exists a nonnegative vector $d \in \mathbb{R}^n$ such that $d^T\Gamma = 0$. In addition $d^T X(t) = d^T X(0) =$ constant for all $t \geq 0$.

In this paper, we are mostly interested in steady-state analysis. Therefore, we will solve the following system of equations:

$$0 = \Gamma V(X), d_j^T X = c_{k,total}, k = 1, .., m, \tag{S2}$$

where $d_j, c_{k,total}, k = 1, .., m$ are the associated conservation laws and conserved quantities, respectively.

# S1 Modeling a single gene subject to epigenetic factor competition

## S1.1 Basic model

As shown in Figure 1, the network has $N$ EFs $\mathrm{F}_1, .., \mathrm{F}_N$. The local GRN has $n$ genes that are subject to the effect of the EFs. In addition, the EFs affect the rest of the genome which is modeled as a single "mega-gene" and we give it the index 0. Hence, the EFs have a total of $n+1$ targets.

**Competition network**. Let consider the $i$th gene, with $i \in \{0, 1, .., n\}$. Denote the corresponding PRE/TRE component by $\mathrm{G}^{(i)}$. Then, we assume it can have the following states:

1. *Unbound:* It is denoted by $\mathrm{G}_0^{(i)}$ which means that nothing is bound to $\mathrm{G}^{(i)}$ and there are no histone marks.

2. *Bound by the EF* $\mathrm{F}_j$*:* It is denoted by $\mathrm{G}_j^{(i)}$. The $j$th factor is bound to $\mathrm{G}_i$. For simplicity, we assume that the EF marks the corresponding histone immediately.

3. *Unbound and marked:* We denoted it either by $\mathrm{G}_+^{(i)}$ or $\mathrm{G}_-^{(i)}$ depending if the histone mark is activating or repressing, respectively.

Therefore, we can write the following of reactions for a repressing EF $\mathrm{F}_j$:

$$\mathrm{F}_j + \mathrm{G}_0^{(i)} \underset{a_{-ij}}{\overset{a_{+ij}}{\rightleftharpoons}} \mathrm{G}_j^{(i)} \xrightarrow{\gamma_{+ij}} \mathrm{G}_-^{(i)} + \mathrm{F}_j,\ \mathrm{G}_-^{(i)} \xrightarrow{\gamma_{-i}} \mathrm{G}_0^{(i)}. \tag{S3}$$

Similarly, we write the following for an activating EF $\mathrm{F}_j$:

$$\mathrm{F}_j + \mathrm{G}_0^{(i)} \underset{a_{-ij}}{\overset{a_{+ij}}{\rightleftharpoons}} \mathrm{G}_j^{(i)} \xrightarrow{\gamma_{+ij}} \mathrm{G}_+^{(i)} + \mathrm{F}_j,\ \mathrm{G}_+^{(i)} \xrightarrow{\gamma_{+i}} \mathrm{G}_0^{(i)}. \tag{S4}$$

In plain words, the reactions describe the binding/unbinding of $\mathrm{F}_j$ to the PRE/TRE $\mathrm{G}_0^{(i)}$, and the marking of the corresponding histones. The histone mark can be erased either constitutively, or via the activity of histone modifiers which are not explicitly modeled. The overall network has $N + (n+1)(N+3)$ species and $4N(n+1)$ reactions.

The network above models the four assumptions postulated in the main text. *Global-targeting* is captured by the PRE/TRE of the mega-gene denoted by $\mathrm{G}^{(0)}$. In our simulations, we set $G^{(0)}_{total} \gg G^{(1)}_{total}$ to model the fact that the total number of targets across the genomes is large. *Competition* is captured by the fact that two different EFs can not bind to a specific PRE/TRE $\mathrm{G}^{(i)}$ simultaneously. *Localization* is captured by the fact that the bound EFs $\{\mathrm{G}_j^{(i)}\}$ can not interact with other genes. Finally, *scarcity* is captured by the stoichiometric conservation of the EFs. In other words, there are no new EF molecules that are created or annihilated in the network above. More concretely, for each $j = 1, .., N$, we have

$$F_j + \sum_{i=0}^{n} G_j^{(i)} = F_{j,total}. \tag{S5}$$

The total copy number of each gene is fixed, hence, the PRE/TREs are also conserved. In particular, we have for each $i = 0, .., n$:

$$G_0^{(i)} + G_+^{(i)} + G_-^{(i)} + \sum_{j=1}^{N} G_j^{(i)} = G_{total}^{(i)}, \tag{S6}$$

where $G_{total}^{(i)}$ is equal to total concentration of the $i$th gene. Therefore, at each point of time, we can view the epigenetic state of the $i$th gene as a distribution of $N + 3$ states that span the unbound, positively and negatively marked, and the bound.

**Steady-state expressions**. By writing the ODEs for the competition model above, the steady-state values of the PRE/TREs states can be found. To that end, let us partition the EFs into the activating and repressing subsets: $J_+ = \{j \in \{1, .., N\} | \mathrm{F}_j \text{ is activating}\}$, $J_- = \{j \in \{1, .., N\} | \mathrm{F}_j \text{ is repressing}\}$. The interaction between a particular EF and gene pair can be characterized by two effective parameters:

1. The *EF association ratio* of $\mathrm{F}_j$ to $\mathrm{G}_j^{(i)}$ is defined as $a_{ij} := a_{+ij}/(a_{-ij} + \gamma_{+ij})$.
2. The *histone marking ratio*, and is defined as $\gamma_{ij} := \gamma_{+ij}/\gamma_{+i}^{(i)}$ if $j \in J_+$, and $\gamma_{ij} := \gamma_{+ij}/\gamma_{-i}$ if $j \in J_-$.

Therefore, using (S6), we can write:

$$\begin{aligned} G_j^{(i)} &= G_{tot}^{(i)} F_j \frac{a_{ij}}{1 + \sum_{j=1}^{N} c_{ij} F_j}, & G_{-i} &= G_{tot}^{(i)} \sum_{j \in J_-} F_j \frac{a_{ij}\gamma_{ij}}{1 + \sum_{j=1}^{N} c_{ij} F_j} \\ G_0^{(i)} &= G_{tot}^{(i)} \frac{1}{1 + \sum_{j=1}^{N} c_{ij} F_j}, & G_+^{(i)} &= G_{tot}^{(i)} \sum_{j \in J_+} F_j \frac{a_{ij}\gamma_{ij}}{1 + \sum_{j=1}^{N} c_{ij} F_j}. \end{aligned} \tag{S7}$$

where $c_{ij} := a_{ij}(1 + \gamma_{ij})$ is the *regulation ratio* which depends on the two aforementioned parameters.

By examining the expressions above, it can be seen that the different genes are only coupled via the the EFs $F_1, .., F_N$. Unlike the neat expressions above, determining the free EFs requires solving (S5). By substituting from (S7), we get

$$\left(1 + \sum_{i=1}^{n} \frac{G_{tot}^{(i)} a_{ij}}{1 + \sum_{j=1} c_{ij} F_j}\right) F_j = F_{j,total}, \tag{S8}$$

which yields the polynomial equation:

$$\prod_{i=1}^{n} \left(1 + \sum_{j=1} c_{ij} F_j\right) (F_j - F_{j,total}) + \sum_{i=1}^{n} G_{tot}^{(i)} a_{ij} F_j \prod_{\tilde{i} \neq i} \left(1 + \sum_{j=1} c_{\tilde{i}j} F_j\right) = 0. \tag{S9}$$

Hence, in order to find the free levels of the EFs we need to solve a system of coupled $(n + 2)$th-order polynomials. Even with a single local gene and two factors, this amounts to two coupled cubic equations which are infeasible to solve analytically. Luckily, it can be shown that the reaction network (S3)-(S4) is always injective [4], hence we can state the following result which can be proved by showing that the Jacobian is always $P_0$ [5]:

**Theorem 1.** *Let the EFs $F_1, .., F_N$, and genes $\mathrm{G}^{(1)}, ..., \mathrm{G}^{(n)}$ be given. For any fixed total EFs $F_{1,tot}, .., F_{N,tot}$ and total genes $G^{(1)}_{tot}, ..., G^{(n)}_{tot}$, the reaction network* (S3)-(S4) *can not admit more than a single positive steady.*

In our simulations, we recourse to numerical methods (Newton-Raphson or ODE solvers) to evaluate the unique solution of the system (S9).

**Gene expression network**. To model the manner in which the epigenetics affect transcription, we assume that the expression is most active when either an activating EF is bound or if the histone is positively marked. Furthermore, we assume a small residual expression when the gene is epigenetically unmodified, and zero expression when the gene is silenced. This can be written as the following network for the $i$th gene:

$$\begin{aligned} &\mathrm{G}^{(i)}_0 \xrightarrow{k_+} \mathrm{G}^{(i)}_j + \mathrm{X}^{(i)}, j \in J_+, \\ &\mathrm{G}^{(i)}_+ \xrightarrow{k_+} \mathrm{G}^{(i)}_+ + \mathrm{X}^{(i)}, \\ &\mathrm{G}^{(0)}_i \xrightarrow{\rho k_+} \mathrm{G}^{(0)}_i + \mathrm{X}^{(i)}, \\ &\mathrm{X}^{(i)} \xrightarrow{k_-} \emptyset. \end{aligned} \tag{S10}$$

where $0 < \rho < 1$ is the residual expression ratio of the unmodified state. Therefore, the steady state expression level is given as $X_i = \frac{k_+}{k_-}\Psi_i$, where $\Psi_i$ is the epigenetic activation function of the $i$th gene written as:

$$\Psi_i = \sum_{j \in J_+} G^{(i)}_j + G^{(i)}_+ + \rho G^{(0)}_i = G^{(i)}_{tot} \frac{\rho + \sum_{j \in J_+} c_{ij} F_j}{1 + \sum_{j=1} c_{ij} F_j}, \tag{S11}$$

and $F_1, .., F_N$ are the solutions of (S9).

## S1.2 Examples

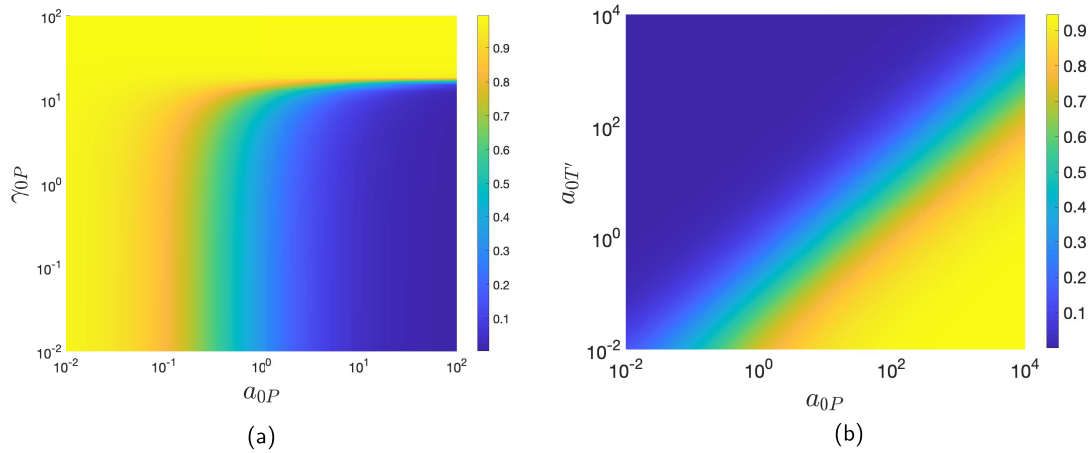
In order to find the activation functions, we will give few examples below.

**A single gene and two EFs**. As in the main text, we consider a toy example of a single gene $\mathrm{G}_1$ subject to the effect of two EFs: one activating (denoted by $T$) and one repressing (denoted by $P$). Recall that the rest of the genome is denoted by $\mathrm{G}^{(0)}$. Therefore, the BIN describing the system is given as follows:

$$\text{Target Gene} \begin{cases} \mathrm{T} + \mathrm{G}^{(1)}_0 & \rightleftharpoons \mathrm{G}^{(1)}_T \longrightarrow \mathrm{G}^{(1)}_+ + \mathrm{T}, \ \mathrm{G}^{(1)}_+ \longrightarrow \mathrm{G}^{(1)}_0 \\ \mathrm{P} + \mathrm{G}^{(1)}_0 & \rightleftharpoons \mathrm{G}^{(1)}_P \longrightarrow \mathrm{G}^{(1)}_- + \mathrm{P}, \ \mathrm{G}^{(1)}_- \longrightarrow \mathrm{G}^{(1)}_0 \end{cases} \tag{S12}$$

$$\text{Rest of the Genome} \begin{cases} \mathrm{T} + \mathrm{G}^{(0)}_0 & \rightleftharpoons \mathrm{G}^{(0)}_T \longrightarrow \mathrm{G}^{(0)}_+ + \mathrm{T}, \ \mathrm{G}^{(0)}_+ \longrightarrow \mathrm{G}^{(0)}_0 \\ \mathrm{P} + \mathrm{G}^{(0)}_0 & \rightleftharpoons \mathrm{G}^{(0)}_P \longrightarrow \mathrm{G}^{(0)}_- + \mathrm{P}, \ \mathrm{G}^{(0)}_- \longrightarrow \mathrm{G}^{(0)}_0 \end{cases} . \tag{S13}$$

Let $a_{1T}, a_{1P}, a_{0T}, a_{0P}, \gamma_{1T}, \gamma_{1P}, \gamma_{0T}, \gamma_{0P}$ be the EF association and histone marking ratios, respectively. As before, let $c_{1T} = a_{1T}(1 + \gamma_{1T}), c_{1P} = a_{1P}(1 + \gamma_{1P}), c_{0T} = a_{0T}(1 + \gamma_{0T}), c_{0P} = a_{0P}(1 + \gamma_{0P})$.




**Supplementary Figure 1: The Global context can reverse a perturbation**. (a) The fraction of the gene that is in a *repressed* state. It considers a scenario of two EFs with the activator knocked out. (b) The fraction of the gene that is in an *active* state. It considers a scenario of two activating EFs and one repressing EF with the first activator knocked out. In both examples, the local parameters are fixed.

By solving the associated equation (S2), the activation function takes the following form:

$$E = G^{(1)}_{tot}\frac{\rho + c_{1T}T}{1 + c_{1T}T + c_{1P}P}. \tag{S14}$$

In order to find $P, T$, we need to solve (S9) which can be written as a pair of cubic equations:

$$0 = (T - T_T)(1 + c_{1T}T + c_{1P}P)(1 + c_{0T}T + c_{0P}P) + G_{1,tot}a_{1T}T(1 + c_{0T}T + c_{0P}P) + G_{0,tot}a_{0T}T(1 + c_{1T}T + c_{1P}P)$$
$$0 = (P - P_T)(1 + c_{1T}T + c_{1P}P)(1 + c_{0T}T + c_{0P}P) + G_{1,tot}a_{1P}P(1 + c_{0T}T + c_{0P}P) + G_{0,tot}a_{0P}P(1 + c_{1T}T + c_{1P}P)$$

As in the main text, we assume that the gene is only regulated by the PRE/TRE. Let the X be the protein expressed. Hence, we write $\emptyset \underset{k_-}{\overset{k_+E}{\rightleftharpoons}} \mathrm{X}$. Therefore, we get $X = (k_+/k_-)E$.

**A single gene and three EFs**. By adding a second EF (say an activator), we modify the network (S12)-(S13) to become as follows:

$$\text{Target Gene} \begin{cases} \mathrm{T} + \mathrm{G}^{(1)}_0 \rightleftharpoons \mathrm{G}^{(1)}_T \longrightarrow \mathrm{G}^{(1)}_+ + \mathrm{T}, \ \mathrm{G}^{(1)}_+ \longrightarrow \mathrm{G}^{(1)}_{10} \\ \mathrm{T}' + \mathrm{G}^{(1)}_0 \rightleftharpoons \mathrm{G}^{(1)}_{T'} \longrightarrow \mathrm{G}^{(1)}_+ + \mathrm{T}' \\ \mathrm{P} + \mathrm{G}^{(1)}_0 \rightleftharpoons \mathrm{G}^{(1)}_P \longrightarrow \mathrm{G}^{(1)}_- + \mathrm{P}, \ \mathrm{G}^{(1)}_- \longrightarrow \mathrm{G}^{(1)}_0 \end{cases} \tag{S15}$$

$$\text{Rest of the Genome} \begin{cases} \mathrm{T} + \mathrm{G}^{(0)}_0 \rightleftharpoons \mathrm{G}^{(0)}_T \longrightarrow \mathrm{G}^{(0)}_+ + \mathrm{T}, \ \mathrm{G}^{(0)}_+ \longrightarrow \mathrm{G}^{(0)}_0 \\ \mathrm{T}' + \mathrm{G}^{(0)}_0 \rightleftharpoons \mathrm{G}^{(0)}_{T'} \longrightarrow \mathrm{G}^{(0)}_+ + \mathrm{T}' \\ \mathrm{P} + \mathrm{G}^{(0)}_0 \rightleftharpoons \mathrm{G}^{(0)}_P \longrightarrow \mathrm{G}^{(0)}_- + \mathrm{P}, \ \mathrm{G}^{(0)}_- \longrightarrow \mathrm{G}^{(0)}_0 \end{cases} . \tag{S16}$$

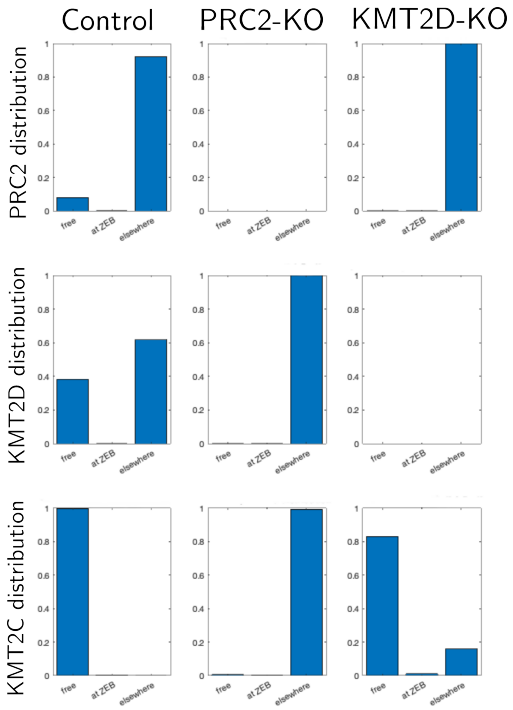
Similar to (S14), the activation function takes the following form:

$$\Psi = G^{(1)}_{tot}\frac{\rho c_0 + c_T T + c_{T'}T'}{c_0 + c_T T + c_{T'}T' + c_P P}, \tag{S17}$$

Supp. Fig. 1 shows how the global context can dictate the effect of a knockout. Supp. Fig. 1-a demonstrates that a low global marking ratio $\gamma_{0P}$ and a high global association ration $a_{0P}$ dilutes the repressing EF P in the absence of the activator $T$. Supp. Fig. 1-b also considers the case of activator knockout, and it shows that an alternative activator with a low global association ratio can rescue activation if the global association ratio of the repressor is high.

## S1.3 Modeling multiple knockout experiments

In the main text, we have shown pictorially how the results of different knockout experiments can be explained by the global context (Figures 2 and 4). To illustrate the underlying effect further, we show in Supplementary Figure 2 the distributions of PRC2, KMT2D and KMT2C for each of the experimental scenarios. In the case of PRC2-KO, the rest of the genome sequesters most of the free KMT2C/D which results in their dilution at the promoter of the gene and lackluster activation. In the KMT2D-KO case, most of the free PRC2 gets sequestered and it outcompetes KMT2C across the genome which leaves a sufficient number of unbound KMT2C complexes to activate the target gene.



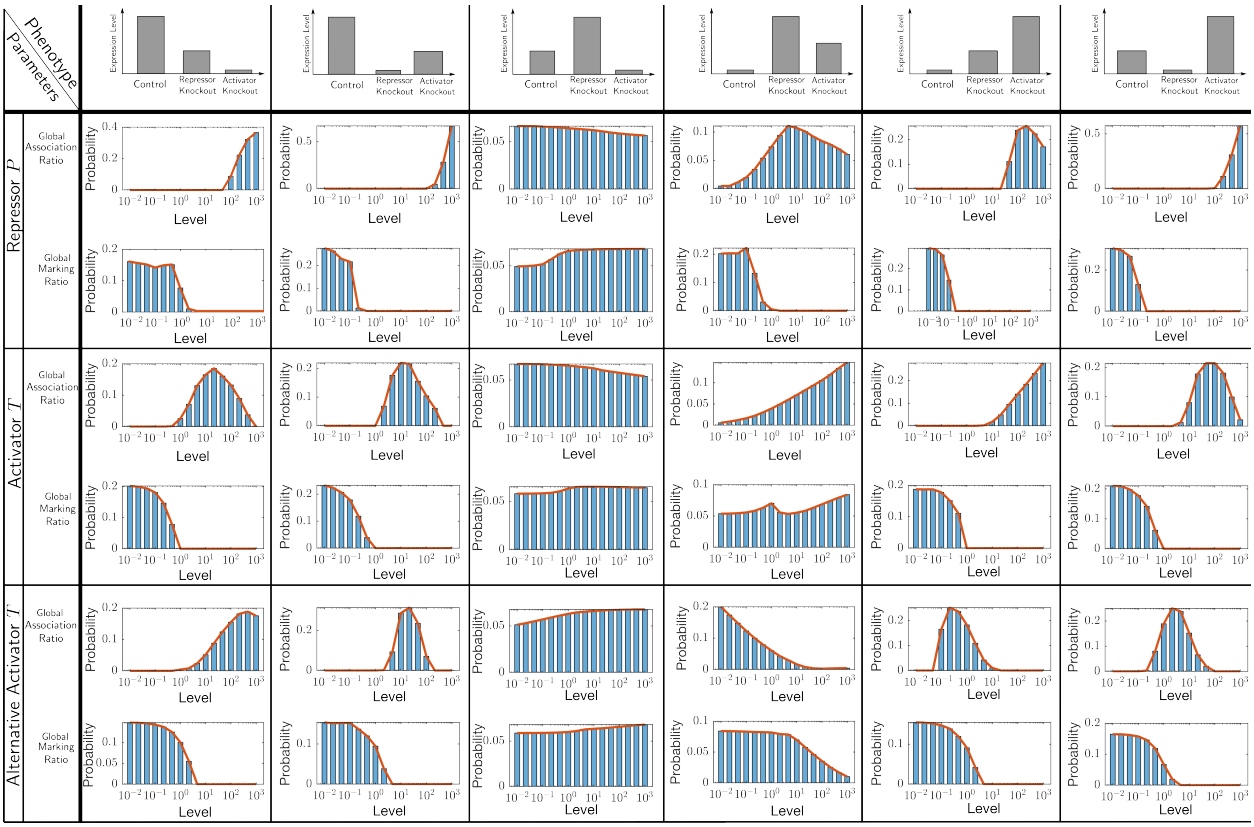

**Supplementary Figure 2:** The distribution of the three factors under the three experimental scenarios for a single gene model of ZEB1.

In addition, Figure 8 (in the main text) has demonstrated that all possible phenotypes are possible depending on the balance of local and global parameters. It is worth noting that the conditional marginal probability distributions presented in Figure 8 depend on the values of the fixed parameters. Here we recompute the conditional marginal probability distributions by choosing different groups of fixed parameters.

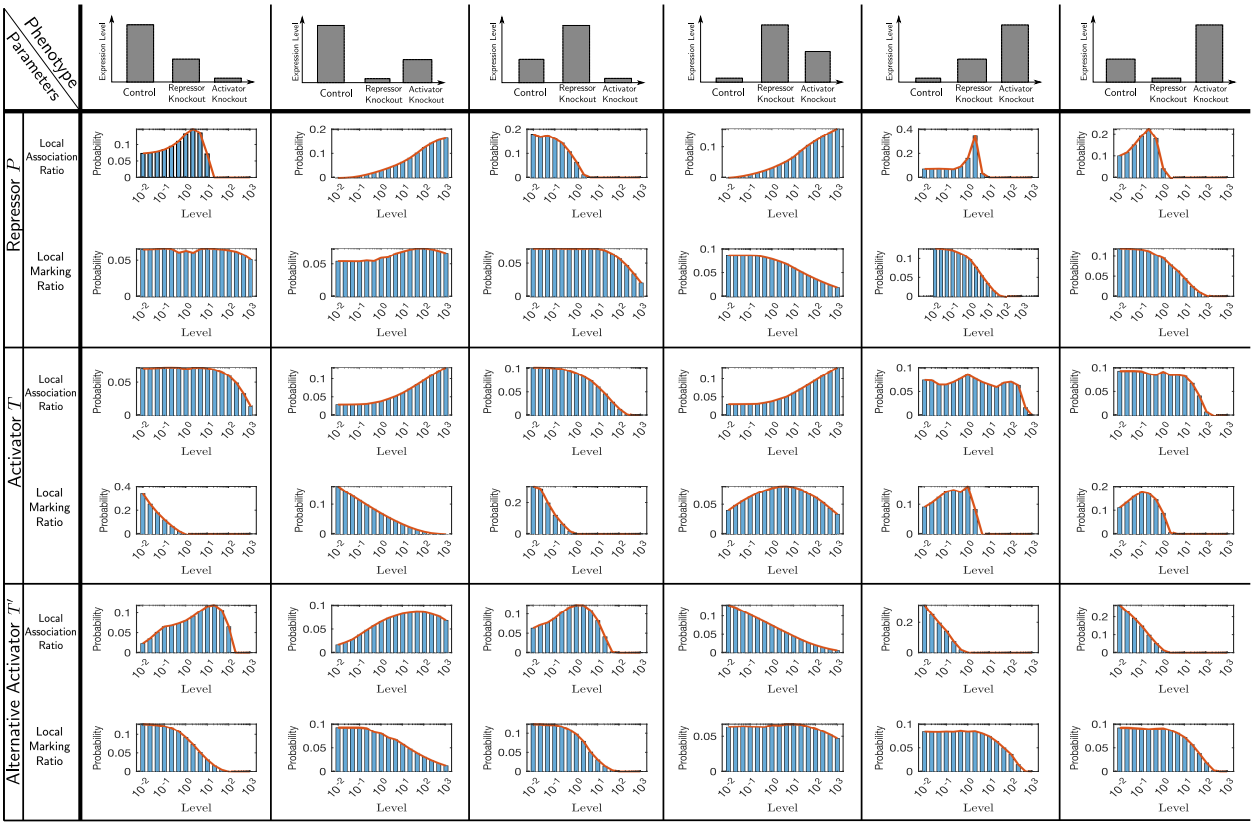
To be more concrete, recall that we have twelve parameters: The local association ratios $a_{1P}, a_{1T}, a_{1T'}$, the local marking ratios $\gamma_{1P}, \gamma_{1T}, \gamma_{1T'}$, the global association ratios $a_{0P}, a_{0T}, a_{0T'}$, and the global marking ratios $\gamma_{0P}, \gamma_{0T}, \gamma_{0T'}$. Figure 8, in the main text, had the marking ratios fixed, leaving the association ratio variable.

First, we fix the six local parameters with P being dominant, while leaving the six global parameters variable. Supp. Table 1 shows the marginal distribution of each global parameter conditioned realizing on a specified phenotypes. For example, consider the plot depicted in the fifth column and first row in the table. Among the parameter sets that give the corresponding phenotype, the range of values of the global association ratio $a_{0P}$ is higher than $10^2$ (compared to a local association ratio of 2). This means that the repressing EF P has to have a strong tendency to bind to targets across the genome. The figure shows, in addition, that the most "natural" phenotype (shown in the third column) does not require specific values of the parameters and it has wide distributions. While the more "paradoxical" phenotypes are only realized under the conjunction of multiple stringent parameter ranges.

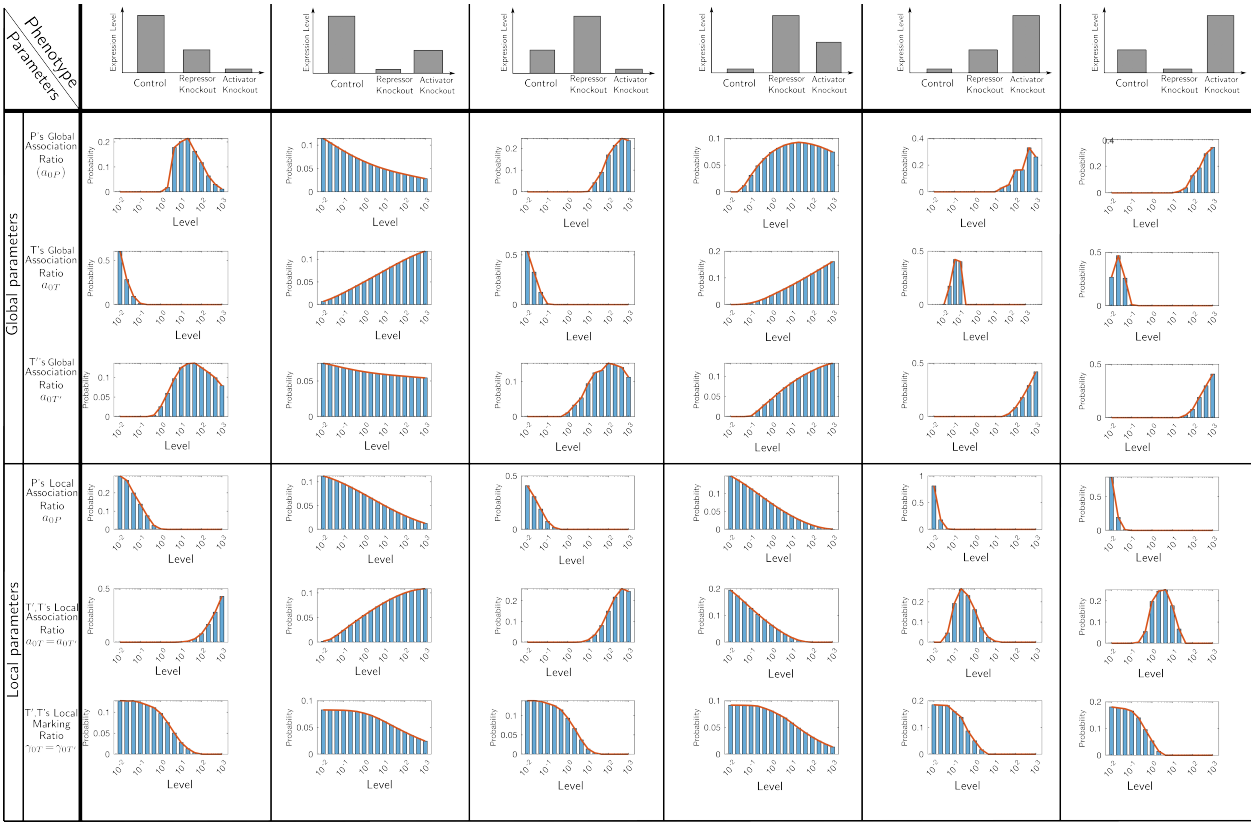
Two other Tables are provided in Supp. Table 2 and 3. For the first, we fix the global parameters while varying the local parameters. For the second, we make the two activators identical locally, and have the same total levels globally. In both tables, we see that all phenotypes are achievable.



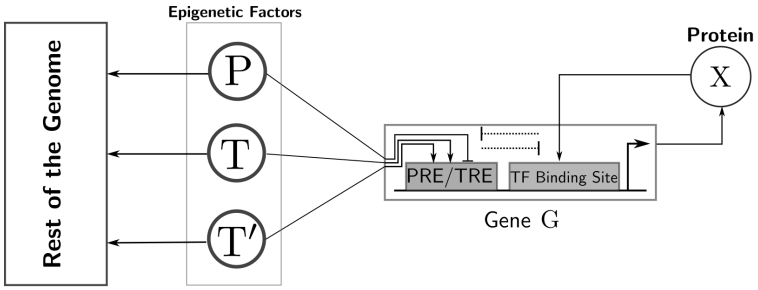

**Supplementary Table 1: All phenotypes are possible under the appropriate global context.** Each plot depicts the marginal probability distribution of the parameter under consideration conditioned on the phenotype under consideration. The local parameters are fixed so that the repressing EF P is dominant. In particular, we have $a_{1P} = 2, a_{1T} = 0.2, a_{1T'} = 0.2, \gamma_{1P} = \gamma_{1T} = \gamma_{1T'} = 1$. The six global parameters are varied with 16 levels between $10^{-3}$ and $10^{3}$. This provides $6^{16} \approx 2.8211 \times 10^{12}$ sets of parameters. For each of which, the steady states are calculated numerically for the control and the knockout cases. A set of parameters is said to give one of the six phenotypes if the highest expression level (amongst the three cases) is at least 50% higher than the second highest, and latter is at least 50% higher than the third highest.




**Supplementary Table 2: All phenotypes are possible for a fixed global context.** Each plot depicts the marginal probability distribution of the parameter under consideration conditioned on the phenotype under consideration. The global parameters are fixed as follows: $a_{0P} = a_{0T} = 100, a_{0T'} = 1, \gamma_{0P} = \gamma_{0T} = \gamma_{0T'} = 0$. The total levels are $P_{tot} = 1000, T_{tot} = 500, T'_{tot} = 100$. The six other parameters are varied with 16 levels between $10^{-3}$ and $10^{3}$. This provides $6^{16} \approx 2.8211 \times 10^{12}$ sets of parameters. For each of which (except for 1.732% of the space of parameters which gave rise to numerical problems), the steady states are calculated numerically for the control and the knockout cases. A set of parameters is said to give one of the six phenotypes if the highest expression level (amongst the three cases) is at least 50% higher than the second highest, and latter is at least 50% higher than the third highest.




**Supplementary Table 3: All phenotypes are possible even when the two activators act the same locally.** Each plot depicts the marginal probability distribution of the parameter under consideration conditioned on the phenotype under consideration. The following parameters are fixed: $\gamma_{0P} = \gamma_{0T} = \gamma_{0T'} = 0, \gamma_{1P} = 1$. Furthermore, the activators are assumed to act the same locally. Hence, we have $a_{1T} = a_{1T'}, \gamma_{1T} = \gamma_{1T'}$. The total levels are $P_{tot} = 1000, T_{tot} = T'_{tot} = 500$. The six other parameters are varied with 16 levels between $10^{-3}$ and $10^{3}$. This provides $6^{16} \approx 2.8211 \times 10^{12}$ sets of parameters. For each of which, the steady states are calculated numerically for the control and the knockout cases. A set of parameters is said to give one of the six phenotypes if the highest expression level (amongst the three cases) is at least 50% higher than the second highest, and latter is at least 50% higher than the third highest.




**Supplementary Figure 3:** A single self-activating gene subject to one repressing and two activating EFs.

# S2 Modeling a self-activating gene subject to epigenetic factor competition

## S2.1 Modeling a self-activating gene

We review here the standard framework for modeling self-activation. Consider a gene G expressing a protein X that activates its own gene. Then, X binds to G and influences its expression. This can be represented using the following reactions:

$$p\mathrm{X} + \mathrm{G}_0 \underset{a_-}{\overset{a_+/p}{\rightleftharpoons}} \mathrm{G}_1,\ \mathrm{G}_0 \xrightarrow{k_+} \mathrm{G}_0 + \mathrm{X},\ \mathrm{G}_1 \xrightarrow{\lambda k_+} \mathrm{G}_1 + \mathrm{X},\ \mathrm{X} \xrightarrow{k_-} \emptyset,$$

where $p$ denotes the cooperativity index, $\mathrm{G}_0, \mathrm{G}_1$ denotes the unbound and bound states, respectively. Expression of the gene G has a basal rate $k_+$ that is affected by the binding of the X by a factor of $\lambda$. We must have $\lambda > 1$ for X to be self-activating.

Solving (S2) for the promoter dynamics, and letting $X_0 := pa_-/a_+$, we get the standard equations [6, 7]:

$$\dot{X} = k_+ \mathcal{H}(X; \lambda, n, X_0) - k_- X, \tag{S18}$$

where

$$\mathcal{H}(Y; p, \lambda, Y_0) := \frac{1 + \lambda (Y/Y_0)^p}{1 + (Y/Y_0)^p}. \tag{S19}$$

The function $\mathcal{H}$ is the shifted Hill function.

## S2.2 The interaction between self-activating genes and epigenetic factor competition

To simplify the notation, we study the case of two activating EFs $\mathrm{T}, \mathrm{T}'$ and one repressing EF P acting on $n$ self-activating genes $\mathrm{G}^{(1)}, .., \mathrm{G}^{(n)}$ expressing protein $\mathrm{X}_1, .., \mathrm{X}_n$. The rest of the genome denoted by $\mathrm{G}^{(0)}$ as before. In order to combine the TF and EF effects, we refer to the pictorial representation in Supp. Figure 3. For a given gene G, we use the notation $\mathrm{G}_{ij}$. The first subscript denotes the histone state, while the second subscript denotes the occupancy of the TF binding site. Based on the evidence reviewed in the main text, the antagonism between active transcription and histone silencers (such as PRC2) is formalized

| TF site | PcG/TrX response element | Histone mark | Symbol |
|---|---|---|---|
| 0 | 0 | 0 | $\mathrm{G}_{00}$ |
| 0 | P | $-$ | $\mathrm{G}_{P0}$ |
| 0 | T | $+$ | $\mathrm{G}_{T0}$ |
| 0 | T$'$ | $+$ | $\mathrm{G}_{T'0}$ |
| 0 | 0 | $+$ | $\mathrm{G}_{+0}$ |
| 0 | 0 | $-$ | $\mathrm{G}_{-0}$ |
| $\mathrm{X}_i$ | 0 | | $\mathrm{G}_{0X_i}$ |
| $\mathrm{X}_i$ | T | $+$ | $\mathrm{G}_{TX_i}$ |
| $\mathrm{X}_i$ | T$'$ | $+$ | $\mathrm{G}_{T'X_i}$ |
| $\mathrm{X}_i$ | 0 | $+$ | $\mathrm{G}_{+X_i}$ |

**Supplementary Table 4:** All possible states of a self-activating gene subject to epigenetic competition between three EFs. P denotes the repressing EF, while $\mathrm{T}, \mathrm{T}'$ denote the activating EFs.

by assuming that the state in which the self-activating TF X and the repressing EF P are both bound is rare and hence it is neglected in the model. Similarly, we assume that a gene marked with repressing histone mark is inaccessible to the TF. The gene states are listed in Table 4. Therefore, we can write the following reactions for a repressing EF P:

$$\mathrm{P} + \mathrm{G}_{00}^{(i)} \rightleftharpoons \mathrm{G}_{P0}^{(i)} \longrightarrow \mathrm{G}_{-0}^{(i)} + \mathrm{P},\ \mathrm{G}_{-0}^{(i)} \longrightarrow \mathrm{G}_{00}^{(i)}. \tag{S20}$$

Similarly, we write the following for the activating EFs $\mathrm{T}, \mathrm{T}'$:

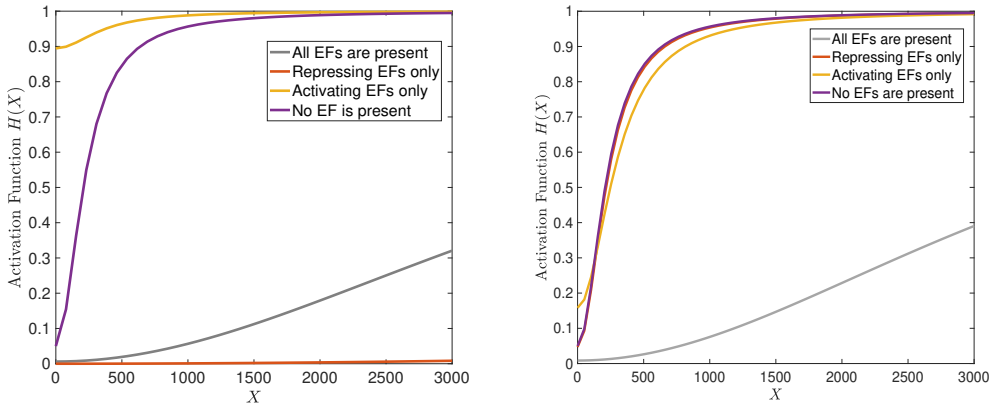
$$\begin{aligned}
&\text{First Activator} \begin{cases} \mathrm{T} + \mathrm{G}_{00}^{(i)} \rightleftharpoons \mathrm{G}_{T0}^{(i)} \longrightarrow \mathrm{G}_{+0}^{(i)} + \mathrm{T},\ \mathrm{G}_{+0}^{(i)} \longrightarrow \mathrm{G}_{00}^{(i)} \\ \mathrm{T} + \mathrm{G}_{0X_i}^{(i)} \rightleftharpoons \mathrm{G}_{TX_i}^{(i)} \longrightarrow \mathrm{G}_{+X_i}^{(i)} + \mathrm{T},\ \mathrm{G}_{+X_i}^{(i)} \longrightarrow \mathrm{G}_{0X_i}^{(i)} \end{cases} \\
&\text{Second Activator} \begin{cases} \mathrm{T}' + \mathrm{G}_{00}^{(i)} \rightleftharpoons \mathrm{G}_{T'0}^{(i)} \longrightarrow \mathrm{G}_{+0}^{(i)} + \mathrm{T}', \\ \mathrm{T}' + \mathrm{G}_{0X_i}^{(i)} \rightleftharpoons \mathrm{G}_{T'X_i}^{(i)} \longrightarrow \mathrm{G}_{+X_i}^{(i)} + \mathrm{T}', \end{cases} \\
&\text{TF Binding/ Unbinding} \begin{cases} p\mathrm{X}_i + \mathrm{G}_{00}^{(i)} \rightleftharpoons \mathrm{G}_{0X_i}^{(i)},\ p\mathrm{X}_i + \mathrm{G}_{T0}^{(i)} \rightleftharpoons \mathrm{G}_{TX_i}^{(i)}, \\ p\mathrm{X}_i + \mathrm{G}_{T'0}^{(i)} \rightleftharpoons \mathrm{G}_{T'X_i}^{(i)},\ p\mathrm{X}_i + \mathrm{G}_{+0}^{(i)} \rightleftharpoons \mathrm{G}_{+X_i}^{(i)}. \end{cases}
\end{aligned} \tag{S21}$$

In addition, we have (S13) which describes the competition of the EFs for sites on the rest of the genome.

We are interested in characterizing the effective activation function of the gene $\Psi(X)$, i.e an analogous function to (S19). Therefore, we need to describe the gene expression network:

$$\begin{aligned}
&\mathrm{G}_{TX_i}^{(i)} \xrightarrow{k_+} \mathrm{G}_{TX_i}^{(i)} + \mathrm{X}_i,\ \mathrm{G}_{T'X_i}^{(i)} \xrightarrow{k_+} \mathrm{G}_{TX_i}^{(i)} + \mathrm{X}_i,\ \mathrm{G}_{+X_i}^{(i)} \xrightarrow{k_+} \mathrm{G}_{+X_i}^{(i)} + \mathrm{X}_i, \\
&\mathrm{G}_{T0}^{(i)} \xrightarrow{\rho_X k_+} \mathrm{G}_{T0}^{(i)} + \mathrm{X}_i,\ \mathrm{G}_{T'0}^{(i)} \xrightarrow{\rho_X k_+} \mathrm{G}_{T0}^{(i)} + \mathrm{X}_i,\ \mathrm{G}_{+0}^{(i)} \xrightarrow{\rho_X k_+} \mathrm{G}_{+0}^{(i)} + \mathrm{X}_i, \\
&\mathrm{G}_{0X_i}^{(i)} \xrightarrow{\rho_U k_+} \mathrm{G}_{0X_i}^{(i)} + \mathrm{X}_i,\ \mathrm{G}_{00}^{(i)} \xrightarrow{\rho_X \rho_U k_+} \mathrm{G}_{00}^{(i)} + \mathrm{X}_i,\ \mathrm{X}_i \xrightarrow{k_-} \emptyset,
\end{aligned} \tag{S22}$$

where $0 < \rho_X, \rho_U < 1$ are the ratios describing the reduction in expression rate due to the absence of the self-activating TF and the absence of the activating EF, respectively.




(a) There are no global targets (b) There are global targets

**Supplementary Figure 4: The global context dictates the local activation curve of the gene under different perturbations.** (a) The activation curves with no global context, i.e. $G^{(0)}_{tot} = 0$, under different perturbations. (b) The activation curves with a global context, i.e. $G^{(0)}_{tot} \gg G^{(1)}_{tot}$, under different perturbations.

Therefore, the $i$th activation function can be written as:

$$\Psi_i(X_1, .., X_n) = G^{(i)}_{TX_i} + G^{(i)}_{T'X_i} + G^{(i)}_{+X_i} + \rho_X(G^{(i)}_{T0} + G^{(i)}_{T'0} + G^{(i)}_{+0}) + \rho_U(G^{(i)}_{0X_i} + \rho_X G^{(i)}_{00}). \quad \text{(S23)}$$

Compared to (S19), the activation function depends on the local parameters, the EF levels, the other genes in the network, as well as the global parameters. The cross-talk between the different genes is indirect and is mediated via the EFs. As explained earlier, an analytical solution is not feasible. Hence, we solve the equations (S9) numerically for each given tuple $(X_1, .., X_n)$.

**Example: Multiple Knockouts**. We consider here the fifth phenotype depicted in Supp. Table 1 where the repressing P protein is more dominant locally. However, it also has strong affinity across the genome. The highest expression is achieved when the activator in knocked-out, followed by when the repressor is knocked-out. Here, we study how the EF competition model can interact with TF self-activation for a single gene. Supp. Figure 4 shows the effect of the global context by comparing two scenarios. The first is depicted in Supp. Figure 4-a where the global context is negligible. When there are no EFs present, the resulting curve is the TF Hill activation function (S19). When all EFs are present, the repressing EF is dominant as assumed, hence a much higher level of the protein X is required to activate the gene. However, when the activators are knocked-out, the repressing EF does not allow any expression for the simulated range of the protein. On the other hand, when there is no repressor, the gene is activated even with very little TF available. The picture changes drastically when the global context is considered as in Supp. Figure 4-b. It can be seen that knocking out the repressor will dilute the activators making the activation function resemble the a gene operating in the absence of EFs. We get a similar effect when both activators are knocked out, since the repressor gets diluted as a result.

# S3 Modeling a generic GRN subject to epigenetic factor competition

## S3.1 Review of GRN modeling

In order to combine our epigenetic regulatory model with standard GRN models, we review GRN models using BIN formalism. A self-activating gene has been reviewed in §2.1. In particular, we have defined the Hill function (S19) which is activating if $\lambda > 1$ and repressing if $\lambda < 1$.

### S3.1.1 Multiple TFs

Assume that a gene G is subject to the action of multiple TFs $\mathrm{Y}_1, .., \mathrm{Y}_n$ which can bind simultaneously. Hence, there are $2^n$ binding states based on the TFs that are bound. In order to simplify the analysis, we assume that the TFs bind to their respective sites and affect expression of the gene *independently* of the other TFs. We will make this precise below.

The binding state of a gene denoted as $\mathrm{G}_s$ where , $s \in \{0,1\}^n$. For instance, $\mathrm{G}_s$ with $s = [0,1,1]$ means that the first TF is unbound, while the second and third TFs are bound. We also write it as $s = 011$ for brevity. Let $\mathcal{G}_{+j}$ be the set of states where the $k$th TF is bound, i.e., $\mathcal{G}_{+j} := \{\mathrm{G}_s | s_j = 1\}$. Similarly, $\mathcal{G}_{-j} := \{\mathrm{G}_s | s_j = 0\}$. For example, in the case of three TFs, $\mathcal{G}_{+2} = \{\mathrm{G}_{010}, \mathrm{G}_{110}, \mathrm{G}_{011}, \mathrm{G}_{111}\}$.

Given the notation above, we can state our exact assumptions regarding the binding of the TFs to the gene:

1. *Independence:* The TFs bind independently of each other, meaning that the cooperativity index $p_j$, and the binding/unbinding rates $a_{+j}, a_{-j}, j = 1, .., n$, of each TF are independent of the presence of other TFs. More precisely, fix $j$. Then, for all $\mathrm{G}_{s^-} \in \mathcal{G}_{-j}$, the model contains the following binding/unbinding reaction:

$$p_j \mathrm{Y}_j + \mathrm{G}_{s^-} \underset{a_{-j}}{\overset{a_{+j}}{\rightleftharpoons}} \mathrm{G}_{s^+},$$

   where $s^+_j = 1$, i.e., $\mathrm{G}_{s^+} \in \mathcal{G}_{+j}$.

2. *Uniformity:* The effect of each TF on the expression of the gene is independent of the presence of other TFs. In other words, each TF $\mathrm{Y}_j$ has a corresponding factor $\lambda_j$. Furthermore, for any two $s^-, s^+$ which are identical expect for the $j$th site, the model contains the following two reactions

$$\mathrm{G}_{s^-} \xrightarrow{\tilde{k}} \mathrm{G}_{s^-} + \mathrm{X}, \ \mathrm{G}_{s^+} \xrightarrow{\lambda_j \tilde{k}} \mathrm{G}_{s^+} + \mathrm{X}.$$

   Hence, the ratio between the production rate of any state with $\mathrm{Y}_j$ unbound and the corresponding state with $\mathrm{Y}_j$ is bound is always equal to $\lambda_j$ regardless of the presence of other TFs

If the above assumptions are satisfied, then it can be shown that the standard expression holds [7]:

$$\dot{X} = k_+ \prod_{j=1}^{n} \mathcal{H}(Y_j; \lambda_j, p_j, Y_{0j}) - k_- X, \tag{S24}$$

where $\mathcal{H}$ is defined in (S19), and $Y_{0j} = p_j a_{-j}/a_{+j}$.

### S3.1.2 Multiple TFs and multiple genes

The above formalism can be extended to any GRN with $n$ genes. Hence, a general model can be written as follows:

$$\dot{X}_i = k_{+i} \prod_{j=1}^{n} \mathcal{H}(X_{ij}; \lambda_{ij}, p_{ij}, X_{0ij}) - k_{-i} X_i, i = 1, .., n. \tag{S25}$$

where $X_{0ji} = \infty$ if $X_j$ does not act on the gene expressing $X_i$. External factors can be added to (S25) if needed.

### S3.1.3 Micro-RNA regulation

Another form of regulation is via micro-RNAs (miRNAs). They inhibit target genes by binding to mRNAs and impeding translation. [6, 8, 9]. We review the modeling framework presented in [6, 8]. Consider a gene G that expresses a protein $X$ via an mRNA molecule which has has $m$ binding sites that miRNAs can bind to. Let $i \in 0, 1, .., m$, then $\mathrm{M}_i$ denotes the mRNA molecule with exactly $i$ miRNA molecules bound to it. Denote the miRNA molecule by µ. We assume that the rates of binding and unbinding of µ to a particular binding site on the mRNA molecule are given as $\nu_+, \nu_-$, respectively. Consider $\mathrm{M}_i$. Then, it has $m - i$ sites available for miRNAs. Hence, the total binding rate is $(m-i)r_+$. Similarly, $\mathrm{M}_{i+1}$ has $i+1$ bound miRNAs and the total unbinding rate is $(i+1)r_-$. Therefore, we write the following reactions:

$$\mathrm{G} \xrightarrow{k_+} \mathrm{G} + \mathrm{M}_0,\ \mathrm{M}_0 \xrightarrow{\beta_0} \emptyset,\ \emptyset \underset{k_{-\mu}}{\overset{k_{+\mu}}{\rightleftharpoons}} \mu,$$

$$\mu + \mathrm{M}_{i-1} \underset{(i+1)r_-}{\overset{(m-i)r_+}{\rightleftharpoons}} \mathrm{M}_i \xrightarrow{\beta_i^{(m)}} \mu,\ \mathrm{M}_i \xrightarrow{\beta_i^{(\mu)}} \mathrm{M}_{i-1}$$

$$\mathrm{M}_i \xrightarrow{\ell_i} \mathrm{M}_i + \mathrm{X},\ \mathrm{X} \xrightarrow{k_-} \emptyset, i = 1, .., m,$$

where the superscripts $(m), (\mu)$ are used to distinguish the decay rates of mRNA and microRNA in the corresponding mRNA-microRNA complex. It is generally assumed that $0 < \beta_0^{(m)} < .. < \beta_m^{(m)}$, and $\ell_0 > \ell_1 > .. > \ell_m > 0$.

As assumed in [6, 8], we assume that the binding and binding rates $r_+, r_-$ are very fast. Hence, we write $r_+ = \tilde{r}_+/\varepsilon, r_- = \tilde{r}_-/\varepsilon$ for some small $\varepsilon > 0$. Let $\mathbf{M} = [M_0, .., M_m]^T$ be the vector of all mRNA configurations. Therefore, we can write the following ODE to describe the time-evolution of $M(t)$:

$$\dot{\mathbf{M}}(t) = \frac{1}{\varepsilon} A\mathbf{M} - B\mathbf{M} + c,$$

where

$$A=\begin{bmatrix} -m\mu\tilde{r}_+ & \tilde{r}_- & 0 & \dots & & & & & 0 \\ \vdots & & \ddots & & & & & & \vdots \\ 0 & \dots & 0 & \mu\tilde{r}_+(m-i+1) & -i\tilde{r}_- - \mu\tilde{r}_+(m-i) & (i+1)\tilde{r}_- & 0 & \dots & 0 \\ \vdots & & & & & & & & \vdots \\ 0 & \dots & & & & & \mu\tilde{r}_+ & -m\tilde{r}_- \end{bmatrix}, B=\begin{bmatrix} \beta_0 & & \dots & 0 \\ 0 & \beta_1 & & \\ \vdots & & \ddots & \vdots \\ 0 & \dots & & \beta_n \end{bmatrix}, c=\begin{bmatrix} k_+G \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

We can write $\varepsilon\dot{\mathbf{M}} = A\mathbf{M} + \varepsilon(c - B\mathbf{M})$. By letting $\varepsilon \to 0$, we get $A\mathbf{M} = 0$. Since $\mathbf{1}^T A = 0$, we have the conservation law $\mathbf{1}^T\mathbf{M} = M$, where $M$ is *the total mRNA* which is constant in the fast-time scale. Hence, we need to solve the linear system $A\mathbf{M} = 0, \mathbf{1}^T\mathbf{M} = M$. By algebraic manipulations and defining $\mu_0 := r_-/r_+$, the result can be shown to be:

$$M_i = M\binom{n}{i}\frac{(\mu/\mu_0)^i}{(1+\mu/\mu_0)^m}, i = 0, .., M.$$

Let $\ell = [\ell_0, ..., \ell_n]^T$. Therefore, we can write the following system of equations:

$$\begin{aligned} \dot{M} =& \mathbf{1}^T(c - B\mathbf{M}) = k_+G - \frac{\sum_{i=0}^m \beta_i^{(m)}\binom{m}{i}(\mu/\mu_0)^i M}{(1+\mu/\mu_0)^m} - \beta M =: k_+G - Y^{(m)}(\mu)M, \\ \dot{X} =& \ell^T\mathbf{M} - k_-X = \frac{\sum_{i=0}^m \ell_i\binom{m}{i}(\mu/\mu_0)^i M}{(1+\mu/\mu_0)^m} - k_-x := L(\mu)M - k_-X, \\ \dot{\mu} =& k_{+\mu} - k_-\mu - \frac{\sum_{i=1}^m \beta_i^{(\mu)}\binom{m}{i}(\mu/\mu_0)^i M}{(1+\mu/\mu_0)^m} =: k_{+\mu} - k_-\mu - Y^{(\mu)}(\mu)M. \end{aligned}$$

Therefore, solving for $X$ at steady state, we get $X = (k_+/k_-)G\Lambda(\mu)$, where $\Lambda$ is given as:

$$\Lambda(\mu) := \frac{L(\mu)}{Y^{(m)}(\mu)} = \frac{\sum_{i=0}^m \tilde{\ell}_i(\mu/\mu_0)^i}{\sum_{i=0}^m \tilde{\beta}_i(\mu/\mu_0)^i}, \tag{S26}$$

where $\tilde{\ell}_i = \binom{m}{i}\ell_i, \tilde{\beta}_i = \binom{m}{i}\beta_i$. Compared to (S19) which has three degrees of freedom only, we notice that miRNA regulation gives rise to a rational function that is a ratio of two generic polynomials. For future reference, we define:

$$\mathcal{L}(\mu; \mu_0, \ell_0, ..., \ell_m) := \frac{\sum_{i=0}^m \ell_i(\mu/\mu_0)^i}{(1+\mu/\mu_0)^m},\ \mathcal{Y}(\mu; \mu_0, \beta_0, ..., \beta_m) := \frac{\sum_{i=0}^m \beta_i(\mu/\mu_0)^i}{(1+\mu/\mu_0)^m} \tag{S27}$$

### S3.1.4 Combining miRNA and TF regulations

A gene can be subject to both TF regulation and miRNA regulation. Let us consider genes $\mathrm{G}_1, .., \mathrm{G}_N$. Assume the first $n_p \leq n$ genes express proteins $\mathrm{X}_1, .., \mathrm{X}_{n_p}$, while the remaining genes express microRNAs $\mu_{n_p+1}, .., \mu_n$. The expressed proteins can act as TFs to activate or inhibit other genes including the genes expressing the microRNAs. For simplicity, we assume that a protein-expressing gene *cannot* be inhibited by more than one micro-RNA, but a single micro-RNA can target multiple genes.

For a protein $\mathrm{X}_i$ that is regulated by a microRNA $\mu_\ell$. The combined model can be written as follows,

$$\dot{M}_i = k_+ \prod_{j=1}^{n_p} \mathcal{H}_{ij}(X_j; \lambda_{ij}, p_{ij}, X_{0ij}) - Y_i^{(m)}(\mu_\ell) M_i,$$
$$\dot{X}_i = L_i(\mu_\ell) M_i - k_- X, i = 1, .., n,$$

while for the microRNA $\mu_\ell$, we write:

$$\dot{\mu}_\ell = k_{+\ell} \prod_{j=1}^{n_p} \mathcal{H}(X_j; \lambda_{\ell j}, n_{\ell j}, X_{0\ell j}) - k_{-\ell}\mu_\ell - \sum_{\{i:\mu_\ell \dashv X_i\}} Y_i^{(\mu)}(\mu_\ell) M_i, j = 1, .., q.$$

For a gene $\mathrm{X}_i$ that is only regulated by TFs, we write:

$$\dot{X}_i = k_{+i} \prod_{j=1}^{n_p} \mathcal{H}(X_{ij}; \lambda_{ij}, p_{ij}, X_{0ij}) - k_{-i} X_i, i = 1, .., N. \tag{S28}$$

where $X_{0ji} = \infty$ if $X_j$ does not act on the gene expressing $X_i$.

Using the equations above, the steady state equations describing the the protein $\mathrm{X}_i$ which is regulated by microRNA $\mu_\ell$ can be written as:

$$0 = k_{+i} \frac{L_i(\mu_\ell)}{Y_i^{(m)}(\mu_\ell)} \prod_{j=1}^{n_p} \mathcal{H}_{ij}(X_j; \lambda_{ij}, p_{ij}, X_{0ij}) - k_- X, i = 1, .., n, \tag{S29}$$

where $\mathcal{H}, \Lambda$ are defined in (S19),(S26), while the steady equation for a gene regulated by TFs only can be written as:

$$0 = k_{+i} \prod_{j=1}^{n_p} \mathcal{H}_{ij}(X_j; \lambda_{ij}, p_{ij}, X_{0ij}) - k_- X, i = 1, .., n, \tag{S30}$$

Finally, the steady state equation for $\mu_\ell$ can be written as:

$$0 = k_{+\ell} \prod_{j=1}^{n_p} \mathcal{H}(X_j; \lambda_{\ell j}, n_{\ell j}, X_{0\ell j}) - k_{-\ell}\mu_\ell - \sum_{\{i:\mu_\ell \dashv X_i\}} \frac{Y_i^{(\mu)}(\mu_\ell)}{Y_i^{(m)}(\mu_\ell)} \prod_{j=1}^{n_p} \mathcal{H}_{ij}(X_j; \lambda_{ij}, p_{ij}, X_{0ij}). \tag{S31}$$

## S3.2 Combining local GRN models with EF factor competition

A general model for a local GRN network that is subject to EF competition can be written using the model components discussed before.

### S3.2.1 Uncoupled transcription and epigenetic competition

We first model the case in which the EFs bind to the PRE/TRE components of the gene independently of the TFs, and vice versa. In addition, we assume that the EFs exert their

effects uniformly regardless of the presence of TFs, and vice versa . This is similar to our earlier assumptions when modeling the action of multiple TFs on a single gene as discussed in §3.1.1. Therefore, for every protein, the different regulatory terms appear as a product multiplying the production rate. This simplifying assumption is relevant when considering the immediate aftermath of knockout experiments at turned-off genes like ZEB1 and PRRX1 as discussed in the main text.

In order to make the discussion more concrete and simplify the notation, we will illustrate our modeling framework by writing the equations for the proposed network in Figure 5-e that have been used to generate the simulation depicted in Figure 5-d.

**The EF competition circuit**. Let $\mathrm{P}, \mathrm{T}, \mathrm{T}'$ denote PRC2, KMT2D, and the third activator, respectively. Let $\mathrm{Z}, \mathrm{R}$ denote ZEB1 and PRRX1 proteins, respectively. Without loss of explanatory generality, we assume that the kinetic rates that describe the interactions of $\mathrm{P}, \mathrm{T}, \mathrm{T}'$ are identical for both PRRX1 and ZEB1. Hence, we can find $\Psi$ given in (S11) by solving (S9) for a network of two genes and three EFs. Hence, for constant kinetic rates, $\Psi$ is a function of the total levels $P_{tot}, T_{tot}, T'_{tot}$. *As a result, the value of* $\Psi$ *changes for each knockout.*

After finding $\Psi$, we can write $\Psi_X = \Psi G^{(Z)}_{tot}/(G^{(Z)}_{tot} + G^{(R)}_{tot})$, $\Psi_R = \Psi G^{(R)}_{tot}/(G^{(Z)}_{tot} + G^{(R)}_{tot})$, where $G^{(Z)}_{tot}, G^{(R)}_{tot}$ are the total copy numbers of the ZEB1, PRRX1 genes, respectively.

**Local GRN**. With reference to Figure 5-e, let $Z, R, S, T$ denote the levels of ZEB1, PRRX1, SNAI1, TGB-$\beta$. Let $\mu$ be the level of $\mu_{200}$. Using (S29),(S30),(S31), we can write the steady state equations as follows:

$$
\begin{aligned}
0 &= k_{+Z}\Psi_Z H_{Z\to Z}(Z)H_{S\to Z}(S)\frac{L^{(\mu)}_Z(\mu)}{Y^{(m)}_Z(\mu)} - k_{-Z}Z,\\
0 &= k_{+R}\Psi_R H_{R\to R}(R)H_{S\dashv R}(S)H_{T\to R}(T) - k_{-R}R,\\
0 &= I_{\text{ext}} + k_{+S}H_{R\dashv S}(R)H_{T\to S}(T)H_{S\to S}(S) - k_{-S}S,\\
0 &= k_{+T}\frac{L^{(\mu)}_T(\mu)}{Y^{(m)}_T(\mu)} - k_{-T}T,\\
0 &= k_{+\mu}H_{Z\dashv\mu}(Z)H_{S\dashv\mu}(S) - k_{-\mu}\mu - \frac{Y^{(\mu)}_Z(\mu)\Psi_Z H_{Z\to Z}(Z)H_{S\to Z}(S)}{Y^{(m)}_Z(\mu)} - \frac{Y^{(\mu)}_T(\mu)}{Y^{(m)}_T(\mu)}.
\end{aligned}
$$

**Parameters.**. We list the parameters used to produce the simulation in Figure 5-d. We start with the competition circuit:

$$
\begin{array}{c}
{\scriptstyle \mathrm{G}^{(\mathrm{ZEB1})} \atop + \atop \mathrm{G}^{(\mathrm{PRRX1})}} \\
\\
{\scriptstyle \text{Rest of} \atop \text{the Genome}}
\end{array}
\left\{
\begin{array}{l}
\mathrm{T} + \mathrm{G}_0^{(1)} \underset{425.8}{\overset{115.6}{\rightleftharpoons}} \mathrm{G}_T^{(1)} \xrightarrow{127.34} \mathrm{G}_+^{(1)} + \mathrm{T},\ \mathrm{G}_+^{(1)} \xrightarrow{815.94} \mathrm{G}_0^{(1)} \\
\mathrm{T}' + \mathrm{G}_0^{(1)} \underset{364.7}{\overset{939.1}{\rightleftharpoons}} \mathrm{G}_{T'}^{(1)} \xrightarrow{283.1} \mathrm{G}_+^{(1)} + \mathrm{T}' \\
\mathrm{P} + \mathrm{G}_0^{(1)} \underset{3.28}{\overset{9265.8}{\rightleftharpoons}} \mathrm{G}_P^{(1)} \xrightarrow{734.22} \mathrm{G}_-^{(1)} + \mathrm{P},\ \mathrm{G}_-^{(1)} \xrightarrow{737.56} \mathrm{G}_0^{(1)}
\end{array}
\right. \tag{S32}
$$

$$
\left\{
\begin{array}{l}
\mathrm{T} + \mathrm{G}_0^{(0)} \underset{0.1}{\overset{1000}{\rightleftharpoons}} \mathrm{G}_T^{(0)} \xrightarrow{0.1} \mathrm{G}_+^{(0)} + \mathrm{T},\ \mathrm{G}_+^{(0)} \xrightarrow{159.9} \mathrm{G}_0^{(0)} \\
\mathrm{T}' + \mathrm{G}_0^{(0)} \underset{1000}{\overset{5.753}{\rightleftharpoons}} \mathrm{G}_{T'}^{(0)} \xrightarrow{29} \mathrm{G}_+^{(0)} + \mathrm{T}' \\
\mathrm{P} + \mathrm{G}_0^{(0)} \underset{0.1336}{\overset{999.98}{\rightleftharpoons}} \mathrm{G}_P^{(0)} \xrightarrow{0.1} \mathrm{G}_-^{(0)} + \mathrm{P},\ \mathrm{G}_-^{(0)} \xrightarrow{556} \mathrm{G}_0^{(0)}
\end{array}
\right. . \tag{S33}
$$

The remaining parameters are: $P_{tot} = 894.14,\ T_{tot} = 494.46,\ T'_{tot} = 81.34, G^{(0)}_{tot} = 975, G^{(Z)}_{tot} = 0.5, G^{(R)}_{tot} = 0.5, \rho_U = 0.05$.
For the local GRN, the activation functions are given as:

$$
\begin{aligned}
H_{Z\to Z}(Z) &= \mathcal{H}(Z; Z_0, n_Z, 1.4), H_{S\to Z}(S) = \mathcal{H}(S; 500, 2, 2), Z_0 = 5097.411, n_Z = 4, \\
H_{S\dashv R}(S) &= \mathcal{H}(S; 250, 4, 0.5), H_{T\to R}(T) = \mathcal{H}(T; 10, 4, 4), H_{S\dashv S}(S) = \mathcal{H}(S; 250, 2, 0.75), \\
H_{R\to R}(R) &= \mathcal{H}(R; R_0, n_R, 1.6), R_0 = 1029.996, n_R = 2, \\
H_{R\dashv S}(R) &= \mathcal{H}(R; 485.27, 6, 0), H_{T\to S}(T) = \mathcal{H}(T; 6.2059, 4, 7.3225) \\
H_{Z\dashv\mu}(Z) &= \mathcal{H}(Z; 63.27, 1, 0), H_{S\dashv\mu}(S) = \mathcal{H}(S; 500, 2, 0.5), \\
Y_Z^{(m)}(\mu) &= \mathcal{Y}(\mu; 20, 1, 2, 2, 3, 3, 3, 3), Y_Z^{(\mu)}(\mu) = \mathcal{Y}(\mu; 20, 0, 0.6038, 1.2076, 1.8114, 7.2456, 10.868), \\
L_Z(\mu) &= \mathcal{L}(\mu; 20, 1, 0, 0, 0), Y_T^{(m)}(\mu) = \mathcal{Y}(\mu; 22.84, 1, 5.139, 5.139, 49.287, 100.66), \\
L_T(\mu) &= \mathcal{L}(\mu; 22.84, 1, 0, 0, 0), Y_T^{(\mu)}(\mu) = 0,
\end{aligned}
$$

while the remaining parameters are $k_{+Z} = 79.461, k_{-Z} = 0.01, k_{+\mu} = 65, k_{-\mu} = 2.5, k_{+R} = 6.3422, k_{-R} = 0.02, k_{+S} = 18.1, k_{-S} = 1, k_{+T} = 100, k_{-T} = 1, I_{\text{ext}} = 0.1$.

### S3.2.2 Coupling transcription and epigenetic competition

As reviewed in the main text, PRC2 interacts antagonistically with active transcription. Hence, the model presented in the previous subsection cannot accurately capture a local GRN in which ZEB1 and PRRX1 are highly expressed. This is especially relevant in the case of sequential knockout experiments.

Therefore, we modify the model presented in the previous subsection by the utilizing the model presented in §S2 that couples transcription and epigenetic competition. Here we show that such a model can be integrated into the local GRN model by assuming that the EF competition-self-regulation sub-circuit is independent of other regulators and exerts its effect uniformly (similar to the assumptions made in the previous subsection and in §3.1.1).

In order to make the discussion more concrete and simplify the notation, we will illustrate our modeling framework by writing the equations for the proposed network in Figure 5-e that have been used to generate the simulation depicted in Figure 6-b. Needless to say, the underlying principles are generalizable to arbitrary networks.

**Epigenetic competition/self-regulation subcircuit**. Similar to the previous subsection, for the purpose of simplifying the computations, we assume that PRRX1 and ZEB1 react with EFs in a similar way except for the manner in which each TF bind to its own promoter. To model the interaction between ZEB1, PRRX and the EFs, we use the model described by (S20),(S21), (S22) with two genes and three EFs. The first gene X refers to both PRRX1 and ZEB1. The rest of the genome is modeled by (S13). Using the above model, we are interested in computing the activation functions for ZEB1 and PRRX1 in an analogous way to (S23).

Therefore, we first solve the combined competition and self-activation function for one local gene and a global mega-gene to get a function $\Psi(X)$. Then, we write

$$\Psi_{_Z}(Z,R) = \frac{G^{(Z)}_{tot}}{G^{(Z)}_{tot}+G^{(R)}_{tot}}\Psi((Z/Z_0)^{n_Z}+(R/R_0)^{n_R}),$$
$$\Psi_{_R}(Z,R) = \frac{G^{(Z)}_{tot}}{G^{(Z)}_{tot}+G^{(R)}_{tot}}\Psi((Z/Z_0)^{n_Z}+(R/R_0)^{n_R}),$$

for some $R_0, Z_0 > 0$, and integers $n_Z, n_R \geq 1$.

**Local GRN**. In this case, the GRN is identical to the one presented in the previous subsection except for the new self-activation functions where $\Psi_Z H_{Z\to Z}(Z)$,$\Psi_R H_{R\to R}(R)$ are replaced by $\Psi_{_Z}(Z,R), \Psi_{_R}(Z,R)$. Using (S29),(S30),(S31), we can write the steady state equations as follows:

$$\begin{aligned}
0 &= k_{_{+Z}}\Psi_{_Z}(Z,R)H_{S\to Z}(S)\frac{L^{(\mu)}_Z(\mu)}{Y^{(m)}_Z(\mu)} - k_{_{-Z}}Z,\\
0 &= k_{_{+R}}\Psi_{_R}(Z,R)H_{S\dashv R}(S)H_{T\to R}(T) - k_{_{-R}}R,\\
0 &= I_{\text{ext}} + k_{_{+S}}H_{R\dashv S}(R)H_{T\to S}(T) - k_{_{-S}}S,\\
0 &= k_{_{+T}}\frac{L^{(\mu)}_T(\mu)}{Y^{(m)}_T(\mu)} - k_{_{-T}}T,\\
0 &= k_{_{+\mu}}H_{Z\dashv\mu}(Z)H_{S\dashv\mu}(S) - k_{_{-\mu}}\mu - \frac{Y^{(\mu)}_Z(\mu)\Psi_{_Z}H_{Z\to Z}(Z)H_{S\to Z}(S)}{Y^{(m)}_Z(\mu)} - \frac{Y^{(\mu)}_T(\mu)}{Y^{(m)}_T(\mu)}.
\end{aligned}$$

**Parameters**. Here we report the parameters used to generate the simulation in Figure 6-b. All the parameters are identical to the ones reported in the previous subsection. We only need to report the additional parameters that characterize the interaction between the epigenetic competition circuit and self-activation loop. Those parameters are listed below.

$$\begin{aligned}
&\mathrm{T} + \mathrm{G}^{(1)}_{0X} \underset{425.8}{\overset{115.6}{\rightleftharpoons}} \mathrm{G}^{(1)}_{TX} \xrightarrow{127.34} \mathrm{G}^{(1)}_{+X} + \mathrm{T},\ \mathrm{G}^{(1)}_{+X} \xrightarrow{407.97} \mathrm{G}^{(1)}_{0X}\\
&\mathrm{T}' + \mathrm{G}^{(1)}_{0X} \underset{93.91}{\overset{729.53}{\rightleftharpoons}} \mathrm{G}^{(1)}_{T'X} \xrightarrow{407.97} \mathrm{G}^{(1)}_{+X} + \mathrm{T}',
\end{aligned}$$

The reactions describing the binding/unbinding of X are listed below:

$$\begin{aligned}
&\mathrm{X} + \mathrm{G}^{(1)}_{00} \underset{1}{\overset{1}{\rightleftharpoons}} \mathrm{G}^{(1)}_{0X},\ \mathrm{X} + \mathrm{G}^{(1)}_{T0} \underset{1}{\overset{1}{\rightleftharpoons}} \mathrm{G}^{(1)}_{TX},\\
&\mathrm{X} + \mathrm{G}^{(1)}_{T'0} \underset{1}{\overset{1}{\rightleftharpoons}} \mathrm{G}^{(1)}_{T'X},\ \mathrm{X} + \mathrm{G}^{(1)}_{+0} \underset{1}{\overset{1}{\rightleftharpoons}} \mathrm{G}^{(1)}_{+X}.
\end{aligned}$$

Finally, we have $\rho_X = 1$.

**Local stability of the steady states**. Since the experimental results are reported at the steady state, the simulation results of our model are also reported at the steady state. Our model does not include assumptions on the relative time-scale separation between the state variables since it has no impact on the existence of steady states. However, for a general nonlinear system, it is theoretically possible for the asymptotic stability of a steady state to be lost when its subsystems evolve on different time-scales. To preclude this possibility, we have performed additional computations to verify that the reported steady states are asymptotically stable when the GRN states and the EFs evolve on different time-scales, and also with different time-scales for the EFs (which are responsible for writing/erasing of the histone marks). This is motivated by the observation that histone modification marks can have different half-lives depending on the type of modification [10].

More concretely, let us write the overall dynamical model as follows:

$$\begin{aligned} \dot{x} &= f(x, y), \\ \dot{y} &= \mathcal{E} g(x, y) \end{aligned} \tag{S34}$$

where $x = [Z, \mu, R, S, T]^T$, $y = [P, T, T']$, and $\mathcal{E}$ is defined as follows:

$$\mathcal{E} = \varepsilon_0 \begin{bmatrix} \varepsilon_P & 0 & 0 \\ 0 & \varepsilon_T & 0 \\ 0 & 0 & \varepsilon_{T'} \end{bmatrix},$$

where $\varepsilon_0, \varepsilon_P, \varepsilon_T, \varepsilon_{T'} > 0$. The parameter $\varepsilon_0$ controls the relative time-scale separation between the GRN state variables and the EFs. The other three parameters $\varepsilon_P, \varepsilon_T, \varepsilon_{T'}$ control the relative time-scale separation between the different EFs. For instance, a small $\varepsilon_P$ and large $\varepsilon_T, \varepsilon_{T'}$ mean that the kinetics of $P$ are slow compared to $T, T'$.

To test stability, we evaluate the Jacobian of (S34) at the steady state of interest, and check that it is Hurwitz, i.e., we check that all the eigenvalues have negative real parts.

We performed this calculation for the five knockout scenarios: control, PRC2-KO, KMT2D-KO, $T'$-KO, and (PRC2,KMT2D)-KO. Then, we have calibrated $\varepsilon_0$ so that the two subsystems in (S34) are in the same time-scale judged by having the corresponding eigenvalues in the same order of magnitude (for the control case). Afterwards, we test stability for eight different cases with $\varepsilon_P, \varepsilon_T, \varepsilon_{T'}$ being either high or low. More precisely, we let $(\varepsilon_P, \varepsilon_T, \varepsilon_{T'}) \in \{0.01, 1\}^3$. In all the tested cases and for all the steady states, the Jacobian has been verified to be Hurwitz. In other words, local stability of the steady-states is verified across widely different time scales.

# References


[1] M. Feinberg. Chemical reaction network structure and the stability of complex isothermal reactors–I. The deficiency zero and deficiency one theorems. *Chemical Engineering Science*, volume 42(10):pages 2229–2268, 1987.

[2] P. Érdi and J. Tóth. *Mathematical models of chemical reactions: theory and applications of deterministic and stochastic models*. Manchester University Press, 1989.

[3] M. Ali Al-Radhawi, D. Angeli, and E. D. Sontag. A computational framework for a Lyapunov-enabled analysis of biochemical reaction networks. *PLoS Computational Biology*, volume 16(2):page e1007681, 2020.

[4] G. Craciun and M. Feinberg. Multiple equilibria in complex chemical reaction networks: I. the injectivity property. *SIAM Journal on Applied Mathematics*, pages 1526–1546, 2005.

[5] M. Banaji, P. Donnell, and S. Baigent. P matrix properties, injectivity, and stability in chemical reaction systems. *SIAM Journal on Applied Mathematics*, volume 67(6):pages 1523–1547, 2007.

[6] M. Lu, M. K. Jolly, R. Gomoto, B. Huang, et al. Tristability in cancer-associated microRNA-TF chimera toggle switch. *The Journal of Physical Chemistry B*, volume 117(42):pages 13164–13174, 2013.

[7] B. Huang, M. Lu, D. Jia, E. Ben-Jacob, et al. Interrogating the topological robustness of gene regulatory circuits by randomization. *PLOS Computational Biology*, volume 13(3):page e1005456, 2017. doi: 10.1371/journal.pcbi.1005456.

[8] M. Lu, M. K. Jolly, H. Levine, J. N. Onuchic, et al. MicroRNA-based regulation of epithelial-hybrid-mesenchymal fate determination. *Proceedings of the National Academy of Sciences*, volume 110(45):page 18144–18149, 2013. doi:10.1073/pnas.1318192110.

[9] M. A. Al-Radhawi and E. D. Sontag. Analysis of a reduced model of epithelial–mesenchymal fate determination in cancer metastasis as a singularly-perturbed monotone system. In C. Beattie, P. Benner, M. Embree, S. Gugercin, et al., editors, *Realization and Model Reduction of Dynamical Systems*. Springer-Verlag, 2022.

[10] T. K. Barth and A. Imhof. Fast signals and slow marks: the dynamics of histone modifications. *Trends in biochemical sciences*, volume 35(11):pages 618–626, 2010.