Role of multiple pericentromeric repeats on heterochromatin assembly

Puranjan Ghimire[1], Mo Motamedi[2*], Richard I Joh[1,3*]

[1]Department of Physics, Virginia Commonwealth University, Richmond VA 23220

[2] Massachusetts General Hospital Center for Cancer Research and Department of Medicine, Harvard Medical School, Charlestown MA 02129

[3]Massey Cancer Center, Virginia Commonwealth University, Richmond VA 23220

*Coorresponding authors: Richard I Joh and Mo Motamedi

Email: rich.i.joh@gmail.com

mo_motamedi@hms.harvard.edu

RJ (ORCID: 0000-0003-0583-8032)

Author Contributions: RJ and MM perceived experiments. PG, MM and RJ designed models. PJ and RJ performed the analysis. PG performed a simulation. PG and RJ wrote the paper.

Competing Interest Statement: Authors declare no competing financial interests.

Keywords: Pericentromeric repeats, copy number, heterochromatin, H3K9me, mathematical modeling, siRNA

Abstract

Although the length and constituting sequences for pericentromeric repeats are highly variable across eukaryotes, the presence of multiple pericentromeric repeats is one of the conserved features of the eukaryotic chromosomes. Pericentromeric heterochromatin is often misregulated in human diseases, with the expansion of pericentromeric repeats in human solid cancers. In this article, we have developed a mathematical model of the RNAi-dependent methylation of H3K9 in the pericentromeric region of fission yeast. Our model, which takes copy number as an explicit parameter, predicts that the pericentromere is silenced only if there are many copies of repeats. It becomes bistable or desilenced if the copy number of repeats is reduced. This suggests that the copy number of pericentromeric repeats alone can determine the fate of heterochromatin silencing in fission yeast. Through sensitivity analysis, we identified parameters that favor bistability and desilencing. Stochastic simulation shows that faster cell division favors the desilenced state whereas cellular noise favors the silenced state. These results show the unexpected role of pericentromeric repeat copy number in gene silencing and elucidate how the copy number of silenced genomic regions may impact genome stability.

Significance Statement

Pericentromeric repeats vary in length and sequences, but their presence is a conserved feature of eukaryotes. This suggests that the repetitive nature of pericentromeric sequences is an evolutionarily conserved feature of centromeres, which is under selective pressure. Here we developed a quantitative model for gene silencing at the fission yeast pericentromeric repeats. Our model is one of the first models which incorporate the copy number of pericentromeric repeats and predicts that the number of repeats can solely govern the dynamics of pericentromeric gene silencing. Our results suggest that the repeat copy number is a dynamic parameter for gene

silencing, and copy-number-dependent silencing is an effective machinery to repress the repetitive part of the genome.

Main Text

Introduction

The eukaryotic centromere is comprised of a core region that is flanked by multiple repetitive DNA sequences, called pericentromeric repeats. Pericentromeric repeats in eukaryotic cells are organized in a dense chromatin structure known as heterochromatin which is vital for the establishment and maintenance of chromosomal stability. For example, pericentromeric heterochromatin is essential for accurate chromatin segregation and genome stability (1, 2). Heterochromatin assembly is also required for gene silencing and repression of recombination among repeats (3). The de-repression of pericentromeric repeats in human cancers (e.g., BRCA1-mutated breast cancer (4, 5)), and tumor formation caused by the forced transcription of pericentromeric satellite RNAs in mice (6) suggest that the transcriptional silencing of pericentromeric repeats is also crucial for preventing tumorigenesis. Heterochromatin is often characterized by DNA methylation, repressive histone modifications, and hypo-acetylation of histones. One of the hallmarks of heterochromatin formation in eukaryotic cells is the methylation of lysine 9 (K9) of histone 3 (H3) (7–9). This chromatin mark and the associated enzymes are conserved from a unicellular eukaryote, like the fission yeast, to a complex mammal, like a human (10).

In the fission yeast, there is one histone methyltransferase for H3K9, Clr4, which facilitates H3K9 mono- (me1), di- (me2), and tri- (me3) methylation. The nucleation of pericentromeric heterochromatin and its spreading by various chromatin-dependent mechanisms is extensively studied in fission yeast where the RNA interference (RNAi) pathway is predominantly responsible for pericentromeric heterochromatin assembly (11–14). RNAi recruits Clr4 via the RNA-induced transcriptional silencing (RITS) complex physically interacting with Clr4 (15–18). RITS is a small interfering RNA (siRNA)-containing effector complex, consisting of three proteins, the sole Argonaute homolog (Ago1), the GW domain protein, Tas3, and a chromodomain-containing protein, Chp1. RITS interacts with H3K9me via chromodomain protein Chp1 and nascent RNA via Ago1 (11–14). In addition to the interaction with Clr4, RITS can also interact with RNA-dependent RNA polymerase complex (RDRC) (16, 19), which together with Dicer (Dcr1) amplifies and processes nascent long non-coding RNAs (lncRNAs) into siRNAs creating a positive feedforward loop, where siRNAs increase H3K9me and H3K9me stimulate the subsequent synthesis of siRNAs (12).

Although the silencing of pericentromeric repeats is a conserved feature of eukaryotes, there is a considerable variation in those repeats in terms of their lengths, constituting sequences, and organization (20). However, the presence of multiple repeats, which flank the central core is a conserved feature (21). This suggests that the repetitive nature of pericentromeres may have a functional consequence. Here, we hypothesize that the copy number of the repeats is critical for pericentromeric gene silencing. Experimental manipulation and quantification of copy number-dependent gene silencing are difficult because of the presence of identical repeats in the pericentromeric region. In this work, we developed a mathematical model for copy-number-dependent gene silencing in pericentromeric repeats, and our simple model supports that multiple copies are required for gene silencing. We have tested the robustness of model parameters through sensitivity analysis and stochastic simulation using the Gillespie algorithm. Our results suggest that the copy number of pericentromeric repeats alone significantly impacts gene silencing, which may affect genomic stability and human diseases.

Results

Mathematical modeling of gene silencing by pericentromeric repeats: To investigate how the copy number of pericentromeric repeats affects gene silencing, we developed a mathematical model incorporating lncRNA, siRNA, and H3K9 methylation as key parameters involved in the establishment of heterochromatin. A schematic representation depicting a simplified molecular pathway used for mathematical modeling is shown in Fig. 1. Briefly, the model tracks the level of RNA (pericentromeric lncRNA), siRNA, and H3K9 methylation which provide the specificity needed for the recruitment of downstream complexes. These molecules also create H3K9me and siRNA feedforward loops, critical for the establishment of heterochromatin at pericentromeres (see Methods).

RNA. According to the nascent transcript model, long lncRNAs provide a platform for the assembly of RNAi and heterochromatin complexes (16–18). Their transcription and tethering to heterochromatin nucleate heterochromatin at centromeres.

siRNAs. In *S. pombe*, siRNA-bound RITS complex can recognize the nascent transcript by RNA-siRNA base-pairing interaction. Physical interaction between RITS and RDRC targets RDRC to single-stranded RNAs, thus converting them into dsRNAs, which can be processed by dicer (Dcr1) (16, 19) generating more siRNAs.

H3K9me. H3K9me is the main readout of eukaryotic heterochromatin. It suppresses transcription at heterochromatin by acting as a binding substrate for several complexes and recruits transcriptional- (TGS) and post-transcriptional gene silencing (PTGS) machinery (7–9). Clr4, the only histone methyltransferase capable of catalyzing mono-(me1) di-(me2) and tri-(me3) methylation in the fission yeast, is a member of the Clr4-Rik1-Cul4 complex (CLRC). Clr4 also possesses a chromodomain, capable of binding to H3K9me (22). This endows Clr4 with read/write capabilities: it can not only deposit H3K9me but also bind to it, thus catalyzing its spreading in *cis* by iterative cycles of H3K9 methylation and H3K9me binding (23). H3K9me can be removed by demethylation by Epe1 (24).

siRNA and H3K9me amplification loops. RITS contains Chp1, a chromodomain protein, capable of binding to H3K9me. This tethers the RITS-RDRC-Dcr1 siRNA biogenesis pathway (assembled on heterochromatin lncRNAs) to heterochromatin. Also, Clr4 physically interacts with RITS (15) and this interaction is critical for targeting Clr4 to heterochromatin. Once at heterochromatin, Clr4-mediated H3K9 methylation can spread in *cis*, and iterative cycles of siRNA generation, and RITS-Clr4-mediated H3K9 methylation amplify the H3K9me and siRNA signals at heterochromatin. H3K9me in turn represses the transcription (25), but even a small number of RNA can be efficiently turned into siRNA.

Repeat copy number. There are over 10 copies of pericentromeric repeats in fission yeast. Copy number is an explicit parameter in mathematical modeling. Each copy of the pericentromeric repeat can be transcribed independently, and without any feedback more copies promote more transcription linearly dependent on the copy number.

Bistability between silenced and desilenced states is possible at a small copy number: To understand the role of copy number on gene silencing, we first analyzed solutions of the ordinary differential equations (ODEs) (see eqs. (1-3)) while varying the copy numbers of pericentromeric repeats (CN). Our model assumes that repeats are transcribed and their copy number, an explicit parameter in our model, enhances transcription linearly for a given level of methylation. Fig. 2A-B depicts the dynamics of methylation from multiple initial conditions for different CNs, and Fig S1 shows the results of RNA, siRNA, and H3K9me. Specifically, our model revealed that there can be one (CN = 13) or two (CN = 5) stable steady states depending on the copy number. As described

above, H3K9me and lncRNA concentrations are coupled via the RNAi machinery. If the copy number is low, the amount of siRNA is low, and the system remains desilenced (Fig. S1 and Fig. 2A). If the copy number is high, enough RNA and siRNA can be generated, which in turn silences the repeats through methylation (Fig. S1 and Fig. 2B). If H3K9me is low, RNA degradation proceeds via conventional mechanisms, whereas if H3K9me is high, RNA is predominantly processed into siRNA.

To assess the effect of copy number change on gene silencing, we systematically varied the copy number and quantified the number of stable steady states. Fig. 2C depicts the steady states of methylation for various CNs, where the blue and red curves represent the higher and lower steady states of H3K9me, respectively. Fig 2C shows that for copy numbers up to 7, the system displays both silenced and desilenced steady states. However, for CN greater than 7, there is only one steady state, which is silenced. Fig S2 shows the dynamics from multiple initial conditions as CN increases, where the fraction of desilenced states decreases until the system becomes monostably silenced. The known copy number of lncRNA dg/dh in *S. pombe* wild-type (WT) cells is around 13, which is represented by the green star in Fig.2C, and the pericentromere of WT cells are silenced. Our results suggest that the heterochromatin-mediated silencing is nonlinearly dependent on the repeat copy number, which can change qualitatively from monostability to bistability. Overall, our model demonstrates that the repeat copy number is a critical dynamical parameter of RNAi-mediated silencing at the fission yeast centromeres.

A small change in copy number leads to a big change in steady states of H3K9me: Because RNA reaches steady-state much faster than siRNA, to test how CN modulates the steady states of the system, we used a quasi-steady-state approximation (QSSA), in which we assumed that RNA concentration reaches equilibrium at any given siRNA and H3K9me level. The time evolution of siRNA and H3K9me in the full model is similar to that in the QSSA (Fig. S3). QSSA reduces three variables in the full model into two variables, which allowed us to plot nullclines in 2-dimensions and thus find the steady-states. Fig. 2D shows the nullclines from eqs. (6-7), which represent curves where siRNA and H3K9me concentrations are not changing, respectively. The intersections of these two nullclines are the steady-states, which include both unstable and stable steady-states. The number of steady states can change between 1 and 3, with 1 or 2 stable steady states. The siRNA nullcline is CN-dependent as the nullcline moves up with high CN, and such changes control the number of steady states of the system. The H3K9me nullcline is CN-independent but the nonlinear shape of this nullcline allows it to have one to three intersections. The intersection point with low methylation (< 0.2) is a desilenced steady-state, and the intersection point with high methylation (~ 1) is a silenced steady-state. If there are three steady states, the middle steady state is unstable. For copy number 15, the siRNA nullcline intersects the H3K9me nullcline at the silenced state only, and this transition from bistability to monostability is due to saddle-node bifurcation (26).

Fig. 2E shows the bifurcation diagram from QSSA, revealing the stable and unstable steady states as a function of the copy number. As in the full model, the system is bistable for lower copy numbers and becomes silenced via saddle-node around copy number 7. Between copy numbers 7 and 8, the system drastically changes its steady-state dynamics between a single silenced state and the coexistence of both silenced and desilenced states. A big change in steady states of H3K9me with a small change in copy number suggests that the copy number is an important factor for the gene silencing of cells, especially if cells were initially desilenced.

Demethylation rate (δ_3) and methylation spreading rate (ϕ) drastically change the state of a system: To investigate the robustness of silenced and desilenced states in a cell, we have systematically varied parameters and tracked how the behavior of the system changes. Most of the parameters in our model are related to the proteins participating in the RNAi pathway of pericentromeric gene silencing. To identify the key parameters in copy number-dependent gene silencing, we tracked steady-state changes when a parameter was varied within 25% of its reference value while keeping all other parameters fixed at their reference values. In all cases,

such changes in the parameters can lead to 1) bistability between silencing and desilencing and 2) monostability of silencing. There are several factors in our model that can repress methylation like demethylation and RNA degradation, while other factors like methylation by siRNA, siRNA biogenesis, and spreading of methylation enhance methylation. Figs. 3A-D shows if the solution is silenced or bistable at the indicated copy number and four representative parameters (see Fig. S4 for other parameters). Enhancing anti-silencing processes can destabilize silencing and leads to switching from silencing to bistability. For example, a high RNA degradation (δ_1) and demethylation rate (δ_3) facilitates switching from bistability to silencing (Figs. 3A-B). On the other hand, enhancing H3K9me-associated processes can stabilize silencing. Figs. 3C-D shows the changes in the steady state behaviors where a high methylation rate (ε) and methylation spreading rate (ϕ) favor silencing.

Our model revealed that demethylation (δ_3) and methylation spreading rates (ϕ) are sensitive parameters (Figs. 3B and 3D) as small changes in them can lead to qualitative differences between silencing and bistability. Next, we analyzed the model behavior for each possible combination of δ_3 and ϕ while keeping all other parameters fixed. Fig. 3E shows the silenced and bistable regime as a result of variations in both δ_3 and ϕ. The boundary of the bistable and monostable regime is almost linearly dependent on δ_3 and ϕ.

To test how each parameter affects system behavior, we also analyzed how the nullclines change from QSSA when one of the parameters is changed. Fig. 4 depicts the changes in nullclines due to variations of different parameters. Our model revealed that several parameters affect the H3K9me nullclines. Demethylation rate (δ_3, the activity of demethylases like Epe1), methylation spreading rate (ϕ, chromodomain activity of Clr4), and basal methylation rate (ζ, the RITS-independent activity of Clr4) affect the desilenced end of H3K9me nullcline (Fig. 4A), and these parameters are critical for bistability. This is in agreement with the experimental observation that the loss of Clr4 chromodomain, which is responsible for binding to H3K9me and equivalent to the reduction in ϕ, leads to variegated bistable heterochromatin establishment (22). Change in methylation rate (ϵ, catalytic activity of Clr4) shifts the H3K9me nullcline (Fig. 4B), and low ϵ can lead to a monostably desilenced system. On the other hand, siRNA biogenesis rate (γ, Dcr1 activity), transcription rate (α, RNAP activity), RNA degradation rate (δ_1, activities of ribonucleases like Dhp1), and siRNA degradation rate (δ_2, Eri1 activity) affect the siRNA nullclines, which primarily affect the desilenced state similar to changing CN (Fig. 4C). Mutation in exoribonuclease Dhp1 can lead to bistability which corresponds to the decrease in δ_1 (27). Hill coefficient (ρ) and half maximum methyl concentration rate (κ) can modulate the slope of the siRNA nullcline, which also controls the number of steady states (Fig. 4D), and this parameter determines the cooperativity of H3K9me-dependent transcriptional gene silencing. These results suggest that the bistability is not a unique behavior of the system, and the system can exhibit 1) bistability between desilenced and silenced states, 2) monostability of the desilenced state, and 3) monostability of the silenced state.

The model supports experimental results: To test if our model can recapitulate experimental observations, we changed our parameter values to match known genetic backgrounds. Figs. 5A-B shows the bifurcation diagram after setting various RNAi-associated parameters to zero. For example, spreading defect ($\phi = 0$) leads to a bistable state while RNA and siRNA degradation defect favors silencing. Our model predicts that at all copy numbers the system remains silenced if the demethylation rate is zero and desilenced if the methylation (ϵ) and siRNA biogenesis (γ) rates are zero. When the RNA degradation rate (δ_1) is zero, the system is monostable and silenced for a copy number greater than 1. If the siRNA degradation rate (δ_2) is zero, the silenced state is strongly favored, so that bistability can be observed only for small copy numbers less than 2 (Fig. 5B). On the other hand, the system seems to be in the bistable state for all copy numbers when methylation spreading rate is zero (Fig. 5B).

Figs. 5C-D shows the steady-state H3K9me and RNA concentrations, respectively, for copy number 13 that would correspond to various gene deletions. Our model recapitulates several

experimental results. Spreading defect ($\phi = 0$) corresponds to the loss of Clr4 chromodomain, which leads to variegated bistable heterochromatin establishment (22). Consistent with previous reports our model predicts that loss of H3K9me by, for example, *clr4* deletion (equivalent to $\varepsilon = 0$) leads to a loss in centromeric silencing (15, 17). Moreover, our model correctly predicts that loss of siRNA biogenesis pathway, e.g. by deletion of *ago1*, *chp1*, *tas3*, *dcr1*, etc. ($\gamma = 0$), leads to a loss of centromeric silencing (16–18, 28). siRNA degradation defect, $\delta_2 = 0$, is similar to Eri1 deletion, which stabilizes heterochromatin (29). RNA degradation (δ_1) and demethylation (δ_3) defect can be simulated computationally, but these are processes which are not specific to heterochromatin lncRNAs and H3K9me. For example, Epe1 is the primary known demethylase for H3K9me, but its activity affects not only H3K9me but also H3K36me.

Other studies have reported that the silencing in heterochromatin is variegated, suggesting that the system might be bistable. For example, two distinct outcomes of silenced and desilenced populations were observed with the deletion of histone deacetylase Sir2 and HDAC-associated Clr5 (11, 30, 31). In addition, transient overexpression of Swi6 can alter the epigenetic imprint and switch the desilenced state into the silenced state (32). These observations suggest that our quantitative model can recapitulate several experimental studies on heterochromatin gene silencing and bistability may occur in a wide range of genetic backgrounds.

Stochastic Simulation: ODE solutions represent the average behavior, so we simulated our system stochastically using the Gillespie algorithm (33, 34) (see Methods). Fig. 6 shows the typical dynamics of the system for copy number 7. To minimize the effect of the initial condition, data were sampled from 5000 to 10000 minutes, and the earlier part was excluded as a burn-in period. The time evolution of RNA, siRNA, and H3K9me within a cell is not static. Cell division can act as a 2-fold dilution and noise from external fluctuations can affect the abundance of RNA, siRNA, and H3K9me, which we incorporate with a Gaussian distribution. Assuming no cell division and no noise in the system, the system remains silenced (Fig. S5A). In a realistic scenario, we considered two additional variations to our model with external noise and cell division. Fig. 6A and Fig. S5E show the switching between the silenced and desilenced states where cell division time was incorporated from the lognormal distribution with an average of 60 minutes and Gaussian noise of standard deviation of 0.01.

At each cell division, RNA, siRNA, and H3K9me concentrations are reduced by the binomial distribution, which is illustrated by Fig. S6. At each cell division, H3K9me goes down by almost 50%, thus faster cell division can destabilize silencing. Fig. 6B shows the distribution of H3K9me over time at different cell division rates, and fast cell division can promote desilencing. With 60 min division rate, the system is almost always desilenced, whereas it is almost always silenced with 240 min division rate (Fig. 6B and Figs. S5B-C). We also incorporated external (to our model consideration) fluctuation to be a Gaussian noise. If the noise is large, the system becomes a diffusion model. Noise can facilitate switching from the desilenced state to the silenced state at different cell division rates (Fig. 6C and Figs. 5E-F), which suggests that the basin of attraction of the silenced state is much larger than that of the desilenced state. At a fixed division rate of 60 min, a large noise leads to a high fraction of silenced states (Fig. 6D). Our results suggest that the cell division rate plays a critical role in pericentromere silencing, which is a physiologically controllable parameter, and its interplay with cellular noise determines the time evolution of silencing in living cells.

Discussion

Here we developed a quantitative model for gene silencing and H3K9me using the fission yeast centromere as a model. This region is one of the best-studied models for heterochromatin formation with many experimental studies testing the function of various associated chromatin factors. We used the copy number of repeats as an explicit model parameter and showed that the copy number itself can be a critical model parameter that can alter gene silencing. Depending on the copy

number, cells can be bistable, with silenced or desilenced pericentromere, which was shown in certain genetic backgrounds experimentally (11, 27, 31). Typically, bistability requires nonlinear interactions through various feedback (35), which we implement by methylation-dependent biogenesis of siRNA. Although we showed how the variation in each parameter affects the steady-state behaviors, the full model in the cell involves many other proteins whose functions are not well known. We assumed that each copy acts independently, but it may be possible that the chromatin state of one repeat preferentially affects the nearest neighbors. In addition, the boundary elements of heterochromatin can affect each repeat in a distant-dependent manner (36, 37).

Our results suggest that RNAi may operate with a threshold copy number for the silencing and target genomic regions with multiple repeats like the pericentromeric repeats, and this leaves the euchromatic loci unaffected by this mechanism. Although few known viruses can infect yeast species, one of the functions of RNAi is a defense against viruses or transposable elements (38, 39). RNAi-mediated copy-number-dependent silencing would be effective against viruses or transposable elements if they generate multiple copies. The threshold copy number and systemic bistability often require strong cooperativity, which in the case of fission yeast is incorporated by the H3K9me-mediated transcriptional gene silencing. There may exist a threshold copy number with weaker cooperativity like RNAi against multiple copies of dsRNA fragments in the absence of chromatin-based feedback.

The copy number variation within heterochromatin is found in many organisms. In human, the satellite DNA is organized in long arrays in pericentromeric heterochromatin, and the copy number increases with age (40). Additionally, human satellite II can massively expand in solid tumors, and this progressive elongation of pericentromeric regions is specific to tumors (4, 41). Duplication of pericentromeric repeats was also observed in the laboratory mouse, suggesting that such expansion is not restricted to primates (42). Fly genome also shows a considerable variation in satellite DNA abundance (43) where the change in abundance is correlated in various isolated populations. Therefore, pericentromeric repeats may be highly dynamic and functionally important for cellular homeostasis. Our study demonstrates that such copy number variation among transcriptionally repressed repeats can alter gene silencing and genomic stability/instability. A more complicated quantitative model would be required to assess the gene silencing on pericentromere in higher eukaryotes with additional regulatory machinery for gene silencing like DNA methylation.

Materials and Methods

Mathematical modeling: Our mathematical model tracks the concentration of lncRNA, siRNA, and H3K9me (Fig. 1). The model includes three first-order ODEs.

$$\dot{x}_1 = \frac{C\alpha}{1+x_3} - \delta_1 x_1 - \frac{\gamma x_1 \left(\frac{x_3}{\kappa}\right)^{\rho}}{1+\left(\frac{x_3}{\kappa}\right)^{\rho}} \quad (1)$$

$$\dot{x}_2 = \frac{\gamma x_1 \left(\frac{x_3}{\kappa}\right)^{\rho}}{1+\left(\frac{x_3}{\kappa}\right)^{\rho}} - \delta_2 x_2 \quad (2)$$

$$\dot{x}_3 = \epsilon x_2(1-x_3) - \delta_3 x_3 + \phi(1-x_3)x_3 + \zeta(1-x_3) \quad (3)$$

In eqs. (1-3), x_1, x_2, and x_3 are the concentrations of RNA transcripts, siRNA, and H3K9me, respectively, and x_3 is normalized with a maximum value of 1. C represents the copy number of pericentromeric repeats which is an explicit parameter of the model. Each copy can be transcribed independently, and the overall transcription is linearly proportional to C. The transcription of RNA, as given by the first term in eq. (1), is inversely related to the H3K9me concentration. RNA can be degraded at a rate δ_1, or turn into siRNAs in a methylation-dependent manner. siRNA biogenesis

rate is represented as a Hill function which is nonlinearly correlated with methylation. ρ is the Hill coefficient and κ denotes the half-maximum methylation. Eq. (2) describes the change in siRNA concentration which includes the siRNA biogenesis and degradation marked by biogenesis rate γ and degradation rate δ_2. The first term in eq. (3) represents the siRNA-mediated methylation with rate ϵ, and available unmethylated H3K9 is represented as $(1 - x_3)$. The spreading of methylation is shown by the third term in eq. (3), where ϕ is the spreading rate. Methylation other than from the RNAi mechanism is incorporated as basal methylation with rate ζ. δ_3 is the demethylation rate.

Solutions of ODE: The system of eqs. (1-3) was solved in MATLAB (44) using an ode45 solver for 24 different initial RNA, siRNA, and H3K9me concentrations. Values of the parameters used for wild-type cells are listed in table S2, and 24 initial conditions used in the model are listed in table S3.

Bifurcation Diagram: To obtain a bifurcation diagram, eqs. (1-3) were solved for different values of copy numbers ranging from 1 to 20, applying the initial conditions listed in table 3 for each copy number to 25,000 minutes. All 24 H3K9me steady states for each copy number were plotted to decide how many clusters they group for Fig. 2. The system becomes completely silenced for copy numbers 8 and above. Therefore, the k-means algorithm (45), where k = 2, was applied to divide all steady-states into two clusters and get the average of each cluster. If the absolute difference between the average of one cluster and that of the other is smaller than 0.085, they were monostable and bistable if greater than 0.085. Finally, the average steady state of each cluster was plotted against the corresponding copy number. Fig. S2 depicts the fraction of silenced and desilenced states obtained after solving eqs. (1-3) for 1000 random initial conditions up to copy number 20. Fractions of both silenced and desilenced states were obtained for copy numbers less than 8, where the system is bistable.

Quasi-steady state approximation (QSSA): Fig. S1 shows that RNA and H3K9 methylation reach steady-state much faster than siRNA, therefore we can approximate the full system where RNA is in equilibrium with respect to siRNA and H3K9me at any given time. This can be done by setting the left-hand side of eq. (1) to zero and solving for x_1, we obtain:

$$x_1 = \frac{\frac{C\alpha}{1+x_3}}{\delta_1 + \frac{\gamma\left(\frac{x_3}{\kappa}\right)^{\rho}}{1+\left(\frac{x_3}{\kappa}\right)^{\rho}}} \tag{4}$$

Substituting eq. (4) into eq. (2) leads to:

$$\dot{x}_2 = \frac{\gamma\frac{C\alpha}{1+x_3}\left(\frac{x_3}{\kappa}\right)^{\rho}}{\delta_1 + (\gamma+\delta_1)\left(\frac{x_3}{\kappa}\right)^{\rho}} - \delta_2 x_2 \tag{5}$$

Setting LHS of eq. (5) and eq. (3) equal to zero gives the equations of siRNA and H3K9me nullclines respectively.

$$x_2 = \frac{\gamma\frac{C\alpha}{1+x_3}\left(\frac{x_3}{\kappa}\right)^{\rho}}{\delta_2\left(\delta_1 + (\gamma+\delta_1)\left(\frac{x_3}{\kappa}\right)^{\rho}\right)} \tag{6}$$

$$x_2 = \frac{\delta_3 x_3 - \phi(1-x_3)x_3 - \zeta(1-x3)}{\epsilon(1-x_3)} \tag{7}$$

Solving eqs (6-7) leads to steady states, both stable and unstable.

Single parameter sensitivity analysis: To test the sensitivity of a parameter (for example y) on copy number-dependent gene silencing, it was varied in the range $0.75p \leq y \leq 1.25p$, p being its default value, keeping all other parameters to their respective default values. Then we calculated the H3K9me steady-states as a function of y for a fixed copy number. This enabled us to visualize the range of values of a particular parameter for which the system was bistable and silenced, and thus the critical y at which the transition happened. This process was repeated for copy numbers 1 to 40 and all the parameters. The final result is depicted in the bar diagram (fig. 3 A-D and fig. S4 A-F). Each bar shows the range of values for which the system is bistable and silenced for a given copy number. The critical y at which the system changes from bistable to silenced or vice-versa is the border between these two ranges.

Two parameters sensitivity analysis: Two parameters ($y_1$and y_2) were varied in the range of $0.75p_1 \leq y_1 \leq 1.25p_1$, and $0.75p_2 \leq y_2 \leq 1.25p_2$, where p_1 and p_2 are the default values of parameters 1 and 2 respectively. Bifurcation diagrams for each combination of y_1 and y_2, keeping all other parameters to their respective default values, were obtained. (y_1, y_2) was grouped according to whether the corresponding bifurcation diagram is bistable or silenced. Finally, a grid where each point represents (y_1, y_2) was obtained with two distinct regions, bistable or silenced.

Stochastic simulation: We used the Gillespie algorithm for eqs. (1-3) for stochastic simulations (33, 34). The concentration of RNA, siRNA, and H3K9me was increased or decreased by 0.01, which is the step size of the random walk considering x_1, x_2, and x_3 do not represent the numbers of individual RNA, siRNA, and H3K9me molecules. Cell division time was chosen from a lognormal distribution. Consistent with known biology, at cell division, RNA, siRNA, and H3K9me will be divided into two daughter cells. To make the system truly stochastic, we chose the RNA, siRNA, and H3K9me concentrations that each daughter cell will get from the binomial distribution with a probability of $50\ \%$. Noise from Gaussian distribution with mean 0 and variable standard deviations depending on the simulation was added to $x_1, x_2,$ and x_3 after each reaction.

Acknowledgments

RJ was supported by Institutional Research Grant IRG-18-159-43 from the American Cancer Society. MM was supported by ACS Research Scholar grant (18-056-01-RMC) and R01 (GM125782). We are grateful to all the Joh lab members for the critical reading and helpful discussions. We would like to thank Dr. Turbasu Sengupta for his suggestions on figure editing.

References

1. A. H. Peters, *et al.*, Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323–337 (2001).

2. N. Dillon, Heterochromatin structure and function. *Biol. Cell* **96**, 631–637 (2004).

3. C. Ellermeier, *et al.*, RNAi and heterochromatin repress centromeric meiotic recombination. *Proc. Natl. Acad. Sci.* **107**, 8701–8705 (2010).

4. D. T. Ting, *et al.*, Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *science* **331**, 593–596 (2011).

5. Q. Zhu, *et al.*, BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature* **477**, 179–184 (2011).

6. Q. Zhu, *et al.*, Heterochromatin-Encoded Satellite RNAs Induce Breast Cancer. *Mol. Cell* **70**, 842-853.e7 (2018).

7. S. Rea, *et al.*, Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**, 593–599 (2000).

8. A. J. Bannister, *et al.*, Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120–124 (2001).

9. M. Lachner, D. O'Carroll, S. Rea, K. Mechtler, T. Jenuwein, Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116–120 (2001).

10. T. Kouzarides, Chromatin modifications and their function. *Cell* **128**, 693–705 (2007).

11. B. J. Alper, B. R. Lowe, J. F. Partridge, Centromeric heterochromatin assembly in fission yeast—balancing transcription, RNA interference and chromatin modification. *Chromosome Res.* **20**, 521–534 (2012).

12. D. Holoch, D. Moazed, RNA-mediated epigenetic regulation of gene expression. *Nat. Rev. Genet.* **16**, 71–84 (2015).

13. R. C. Allshire, K. Ekwall, Epigenetic regulation of chromatin states in Schizosaccharomyces pombe. *Cold Spring Harb. Perspect. Biol.* **7**, a018770 (2015).

14. R. Martienssen, D. Moazed, RNAi and heterochromatin assembly. *Cold Spring Harb. Perspect. Biol.* **7**, a019323 (2015).

15. E. L. Gerace, M. Halic, D. Moazed, The methyltransferase activity of Clr4Suv39h triggers RNAi independently of histone H3K9 methylation. *Mol. Cell* **39**, 360–372 (2010).

16. M. R. Motamedi, *et al.*, Two RNAi complexes, RITS and RDRC, physically interact and localize to noncoding centromeric RNAs. *Cell* **119**, 789–802 (2004).

17. K. Noma, *et al.*, RITS acts in cis to promote RNA interference–mediated transcriptional and post-transcriptional silencing. *Nat. Genet.* **36**, 1174–1180 (2004).

18. A. Verdel, *et al.*, RNAi-mediated targeting of heterochromatin by the RITS complex. *Science* **303**, 672–676 (2004).

19. T. Sugiyama, H. Cam, A. Verdel, D. Moazed, S. I. Grewal, RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. *Proc. Natl. Acad. Sci.* **102**, 152–157 (2005).

20. F. Brändle, B. Frühbauer, M. Jagannathan, Principles and functions of pericentromeric satellite DNA clustering into chromocenters in *Seminars in Cell & Developmental Biology*, (Elsevier, 2022).

21. J. Thakur, J. Packiaraj, S. Henikoff, Sequence, chromatin and evolution of satellite DNA. *Int. J. Mol. Sci.* **22**, 4309 (2021).

22. E. Akoury, *et al.*, Disordered region of H3K9 methyltransferase Clr4 binds the nucleosome and contributes to its activity. *Nucleic Acids Res.* **47**, 6726–6736 (2019).

23. A. R. Cutter DiPiazza, *et al.*, Spreading and epigenetic inheritance of heterochromatin require a critical density of histone H3 lysine 9 tri-methylation. *Proc. Natl. Acad. Sci.* **118**, e2100699118 (2021).

24. S. Isaac, *et al.*, Interaction of Epe1 With the Heterochromatin Assembly Pathway in Schizosaccharomyces pombe. *Genetics* **175**, 1549–1560 (2007).

25. M. R. Motamedi, *et al.*, HP1 proteins form distinct complexes and mediate heterochromatic gene silencing by nonoverlapping mechanisms. *Mol. Cell* **32**, 778–790 (2008).

26. S. H. Strogatz, *Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering* (CRC press, 2018).

27. J. F. Tucker, *et al.*, A novel epigenetic silencing pathway involving the highly conserved 5'-3'Exoribonuclease Dhp1/Rat1/Xrn2 in Schizosaccharomyces pombe. *PLoS Genet.* **12**, e1005873 (2016).

28. P. Provost, *et al.*, Dicer is required for chromosome segregation and gene silencing in fission yeast cells. *Proc. Natl. Acad. Sci.* **99**, 16648–16653 (2002).

29. T. Iida, R. Kawaguchi, J. Nakayama, Conserved ribonuclease, Eri1, negatively regulates heterochromatin assembly in fission yeast. *Curr. Biol.* **16**, 1459–1464 (2006).

30. L. L. Freeman-Cook, *et al.*, Conserved Locus-Specific Silencing Functions of Schizosaccharomyces pombe sir2+. *Genetics* **169**, 1243–1260 (2005).

31. K. R. Hansen, *et al.*, H3K9me-Independent Gene Silencing in Fission Yeast Heterochromatin by Clr5 and Histone Deacetylases. *PLOS Genet.* **7**, e1001268 (2011).

32. J. Nakayama, A. J. Klar, S. I. Grewal, A chromodomain protein, Swi6, performs imprinting functions in fission yeast during mitosis and meiosis. *Cell* **101**, 307–317 (2000).

33. D. T. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* **22**, 403–434 (1976).

34. D. T. Gillespie, Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.* **81**, 2340–2361 (1977).

35. H. Qian, Cooperativity in cellular biochemical processes: noise-enhanced sensitivity, fluctuating enzyme, bistability with nonlinear feedback, and other mechanisms for sigmoidal responses. *Annu. Rev. Biophys.* **41**, 179–204 (2012).

36. J. E. Braun, E. Huntzinger, M. Fauser, E. Izaurralde, GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol. Cell* **44**, 120–133 (2011).

37. F. Lan, *et al.*, S. pombe LSD1 homologs regulate heterochromatin propagation and euchromatic gene transcription. *Mol. Cell* **26**, 89–101 (2007).

38. T. Muhammad, F. Zhang, Y. Zhang, Y. Liang, RNA Interference: A Natural Immune System of Plants to Counteract Biotic Stressors. *Cells* **8**, 38 (2019).

39. D. J. Obbard, K. H. Gordon, A. H. Buck, F. M. Jiggins, The evolution of RNAi as a defence against viruses and transposable elements. *Philos. Trans. R. Soc. B Biol. Sci.* **364**, 99–115 (2009).

40. E. S. Ershova, *et al.*, Copy number variation of human satellite III (1q12) with Aging. *Front. Genet.* **10**, 704 (2019).

41. F. Bersani, *et al.*, Pericentromeric satellite repeat expansions through RNA-derived DNA intermediates in cancer. *Proc. Natl. Acad. Sci.* **112**, 15148–15153 (2015).

42. J. W. Thomas, *et al.*, Pericentromeric duplications in the laboratory mouse. *Genome Res.* **13**, 55–63 (2003).

43. Y. Wei, *et al.*, Paternally induced transgenerational inheritance of susceptibility to diabetes in mammals. *Proc. Natl. Acad. Sci.* **111**, 1873–1878 (2014).

44. MathWorks - Makers of MATLAB and Simulink (September 20, 2022).

45. J. MacQueen, Classification and analysis of multivariate observations in *5th Berkeley Symp. Math. Statist. Probability*, (1967), pp. 281–297.

Figures

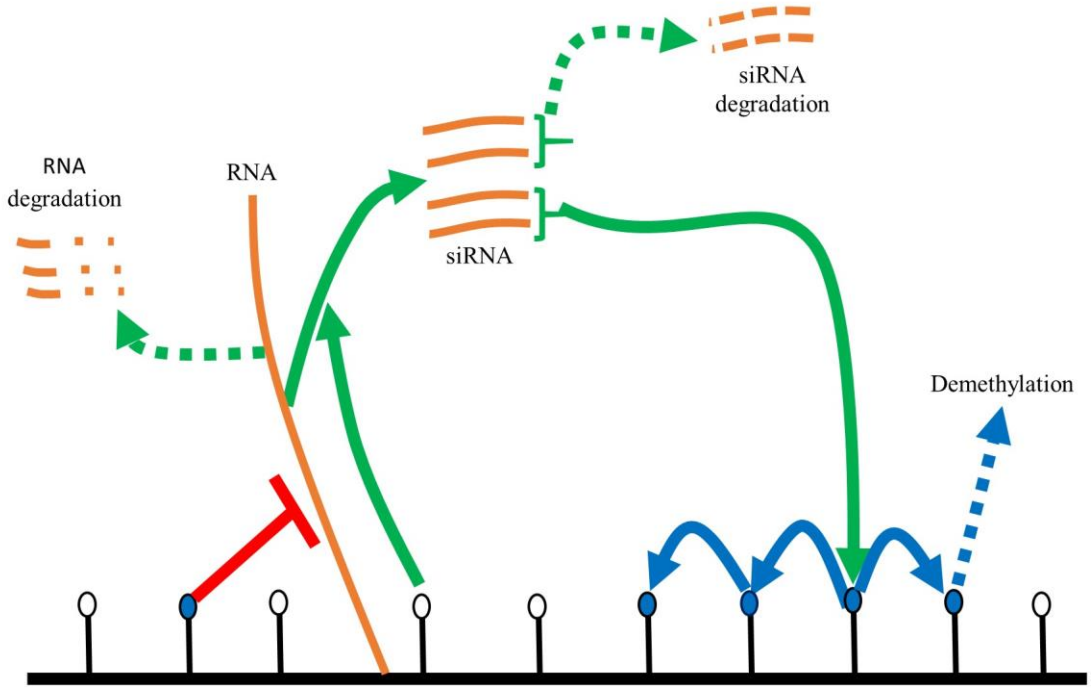

Figure 1. Schematic diagram for mathematical modeling. Circles represent histones, whereas blue ones are methylated. The spreading of H3K9me is shown by the blue arrow. RNA transcription can be repressed by methylation as shown. RNA is either degraded or turned into siRNA which promotes methylation.

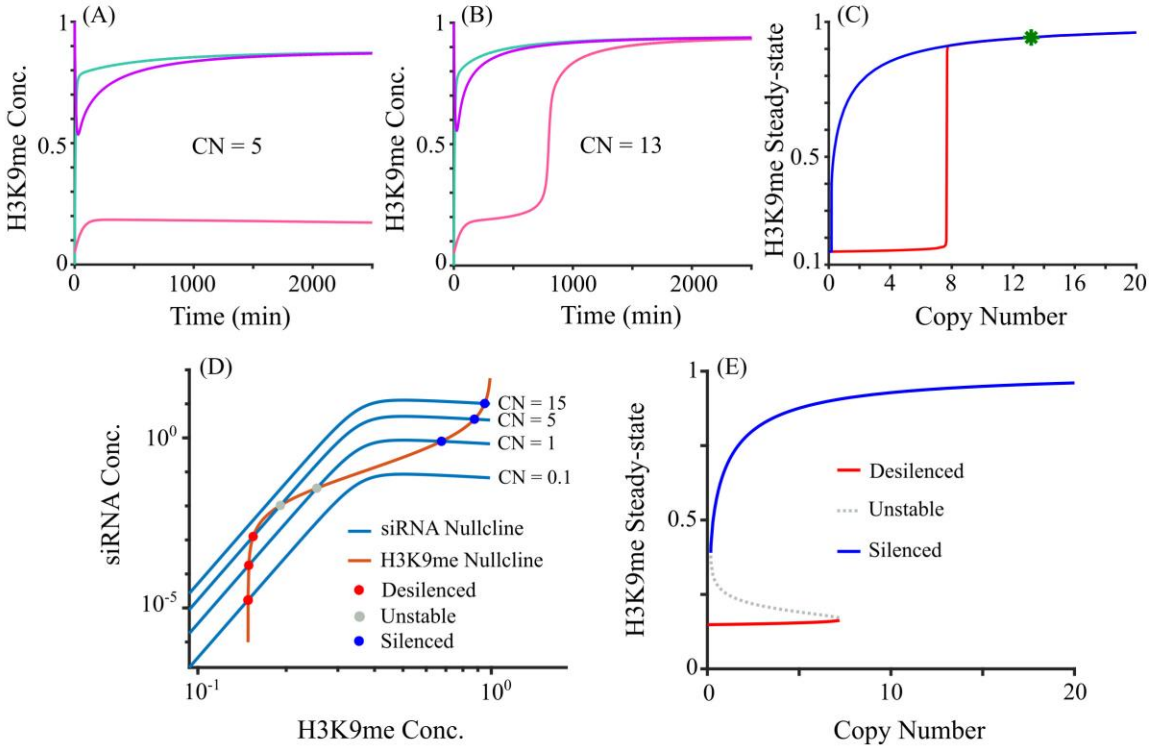

Figure 2. Change in the steady-state behavior as a function of repeat copy number. (A-B) Solutions of H3K9me for different CNs in the full model. (C) Steady-state H3K9me as a function of copy number. The system is bistable (silenced and desilenced) or monostable. Wild Type (WT) in fission yeast is shown at $CN = 13$ by a green star. (D-E) Quasi-steady state approximation. (D) Nullclines of siRNA and H3K9me using QSSA. Representative siRNA nullclines at indicated CNs are shown. (E) Bifurcation diagram showing the bistability with both silenced and desilenced states for smaller copy numbers and monostability at high copy numbers.

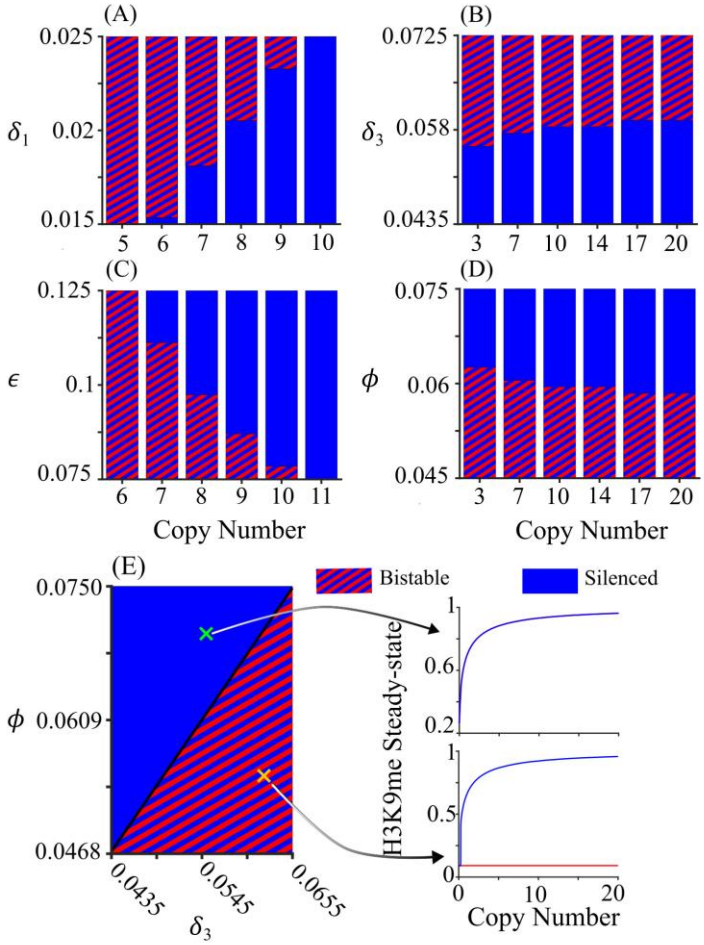

Figure 3: Parameter sensitivity analyses within 25% variation on default values of a specific parameter. (A) Colors indicate whether the system is silenced or bistable at the indicated parameter value and copy number. A large RNA degradation rate (δ_1) favors bistability for all copy numbers. All ranges of δ_1 favor bistable state for copy number 5 and below, favor silenced state for copy number 10 and above, and the range of δ_1 favoring silenced state increases with an increase in copy number from 6 to 9. (B) Like the RNA degradation rate, a higher value of demethylation rate (δ_3) favors bistability. Range of δ_3 favoring both silenced and bistable states exist for copy numbers at least up to 20. (C) A small methylation rate (ϵ) favors bistability. All ranges of ϵ favor the bistable state for copy number 6 and below, favor the silenced state for copy number 11 and above, and the range of ϵ favoring the silenced state decreases with the increase in copy number. (D) Like methylation rate, a lower value of methylation spreading rate (ϕ) favors bistability. Ranges of ϕ favoring both silenced and bistable states exist for copy numbers at least up to 20. (E) Silenced and bistable regions in the simultaneous variation of demethylation rate (δ_3) and methylation spreading rate (ϕ). Bifurcation diagrams corresponding to a point in the silenced region and bistable region are shown on the right side of the figure.

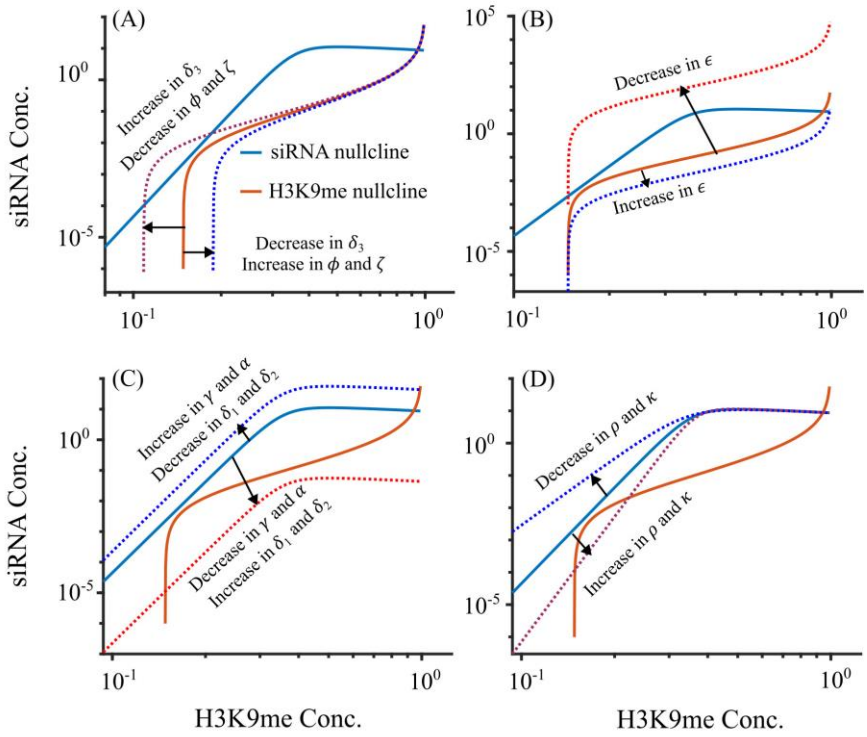

Figure 4: Changes in the nullclines as a result of changing parameters using QSSA. (A) Parameters $\delta_3, \phi,$ and ζ affect the stability of the desilenced state by modulating the tail of the H3K9me nullcline. (B) Parameter ϵ modulates the H3K9me nullcline. (C) Parameters $\gamma, \alpha, \delta_1,$ and δ_2 can shift the siRNA nullcline similar to changing CN. (D) Parameters ρ and κ modulate the tail of the siRNA nullcline, primarily affecting the desilenced state. δ_1, δ_2 and δ_3 are RNA degradation rate, siRNA degradation rate, and demethylation rate respectively. γ, α, ϵ, ρ, and κ are siRNA biogenesis rate, transcription rate, methylation rate, Hill coefficient, and half maximum methyl concentrations respectively.

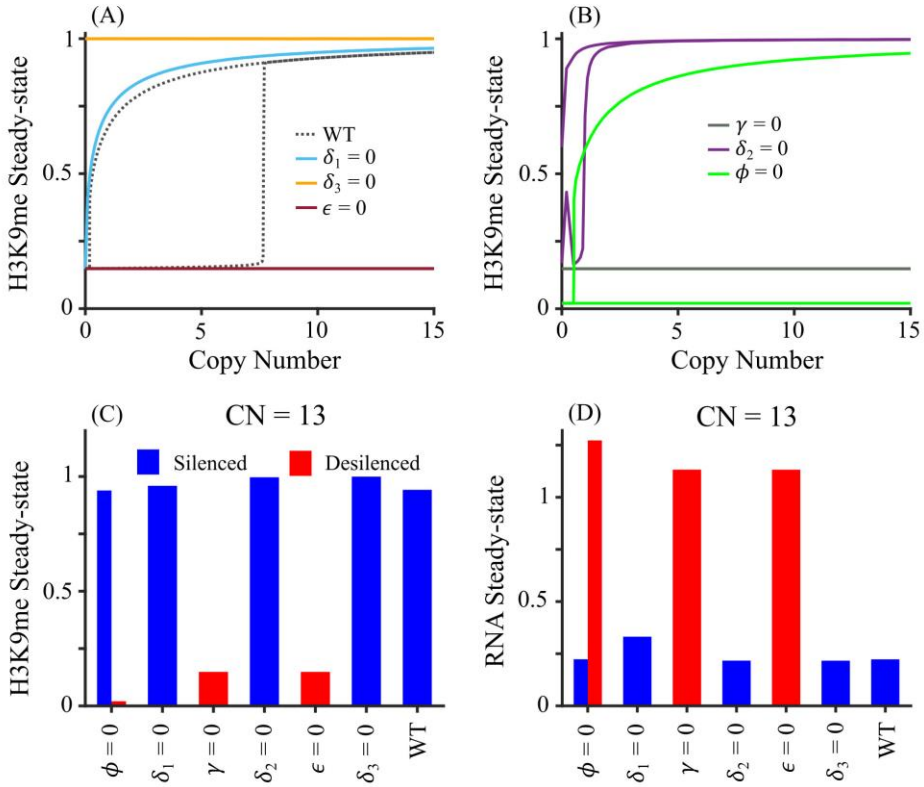

Figure 5. Comparison of model prediction for deletion experiments. (A) Bifurcation diagram corresponding to the deletion of the gene encoding RNAi protein, quantified by setting an associated parameter to zero, viz., RNA degradation rate (δ_1), demethylation rate (δ_3) and methylation rat (ϵ). (B) Bifurcation diagrams of siRNA biogenesis (γ), siRNA degradation (δ_2) and methylation spreading (ϕ) deletions. (C-D) The steady-state concentration of H3K9me (C) and RNA (D) for copy number 13 for deletion of genes encoding RNAi proteins. The blue bar represents the silenced state, the red bar represents the desilenced state.

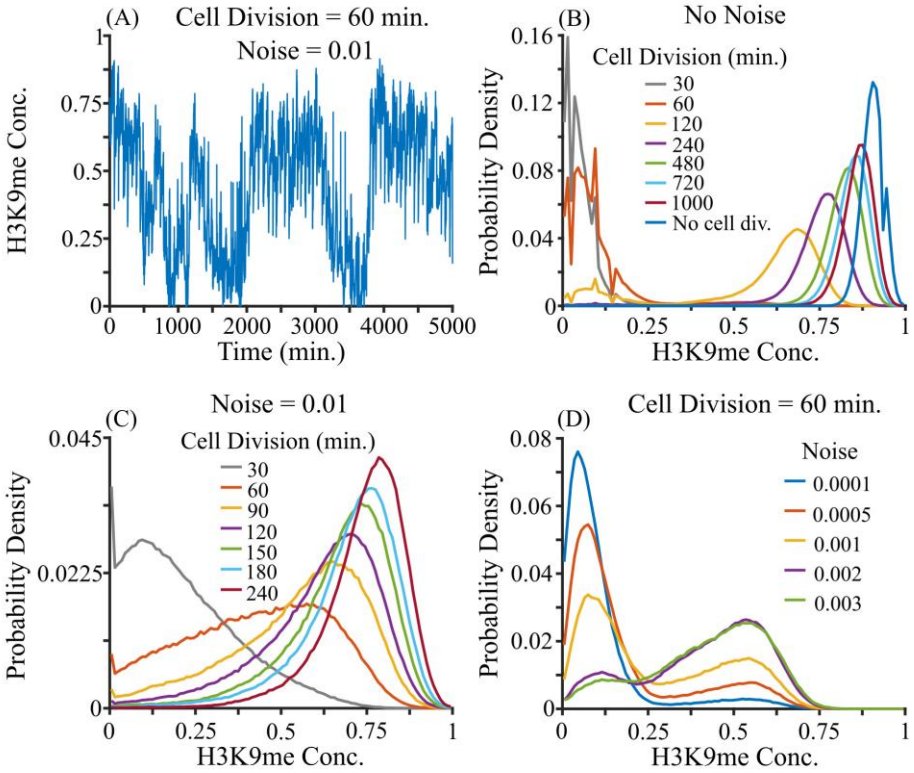

Figure 6. Stochastic simulation for copy number 7 using Gillespie algorithm. (A) A stochastic trajectory of a system of cells with an average cell division time of 60 min, and Gaussian noise of average zero and standard deviation 0.1. The system relatively spends more time in silenced states with occasional switching to desilenced states. The probability distribution of H3K9me concentration was simulated for 5000 minutes for different cell division times under (B) no noise and (C) Gaussian noise of average zero, and standard deviation 0.01. Faster cell division favors a desilenced state for a system with no noise (B) and with noise (C). (D) Probability distribution of H3K9me concentration was simulated for 5000 minutes at fixed cell division for different noises. Higher noise facilitates the silenced state.