Chromatin accessibility and gene regulation

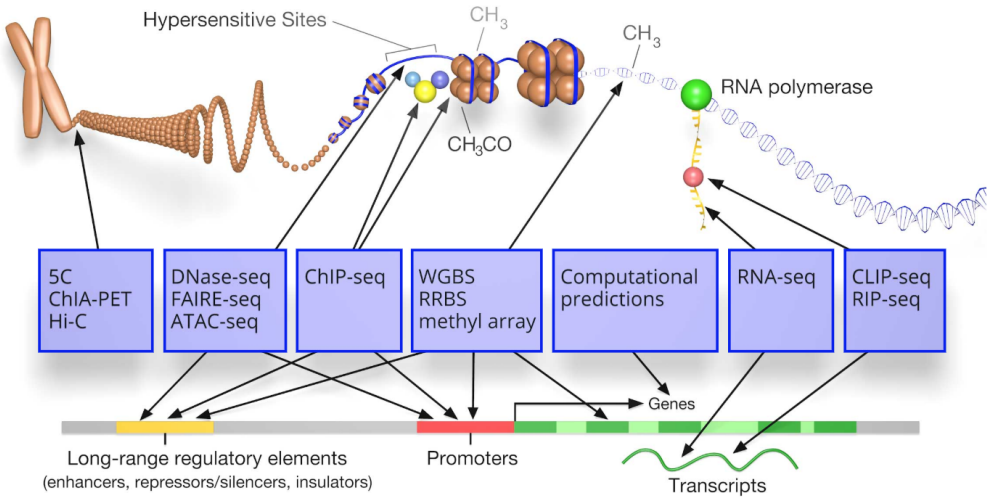
Sohyeon Park and Sam Morabito

March 2019

1 Introduction and motivation

The aim of this project is to investigate the effects of enhancer chromatin accessibility on gene expression using a mathematical model. Although there are many recent publications that characterize the relationship between chromatin accessibility and gene expression, there are very few that do so using a mathematical model. A number of genome-wide assays are commonly used to study the epigenome and its effects on gene expression, but these methods can only provide a snapshot of the system, missing out any time dynamics that may be at play. Even if a time series is constructed using genome-wide experiments, the resolution in time is coarse at best. Using differential equations to study gene regulation in the context of chromatin accessibility could lead to a more nuanced understanding of the system, especially with regards to understanding dynamics in time.

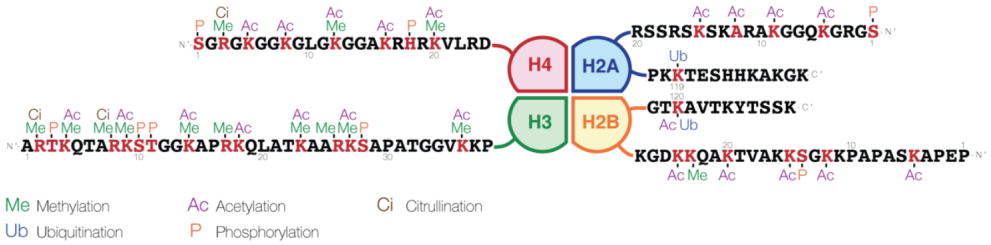
2 Biological background



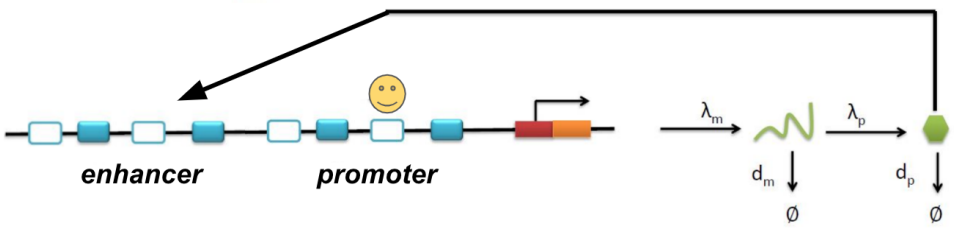
The above image (Encode) is a cartoon displaying how DNA is organized from a chromosome down to individual nucleotides. Roughly 150 nucleotides of DNA wrap around a protein complex called a nucleosome, which consist of a histone octamer. The distribution

of nucleosomes throughout a chromosome greatly affect gene expression. Highly condensed regions of nucleosomes are known as heterochromatin, while regions depleted of nucleosomes are called euchromatin. Intuitively, transcription machinery such as transcription factors cannot bind in a nucleosome dense region, because the binding site is physically blocked off.

The propensity for a genomic region to be heterochromatic or euchromatic is highly dependent on post translational modifications on histones within that regions. Genome wide assays have been extensively used in order to characterize the effects of histone modifications on chromatin accessibility and gene regulation. Histone modifications alter the chemical properties of the nucleosome. A "histone code" has been compiled based on the known effects of individual histone modifications. The below figure displays a nucleosome with some examples of possible histone modifications on amino acid chains of the individual histones.



Our model is adapted from the 2018 paper "Cell-type switches induced by stochastic histone modification inheritance" (Huang 2018). Huang et al primarily focus on stochastic simulations, but they do present a deterministic ODE model that we are using for this project. The model encompasses transcription, translation, and autoregulation via protein binding to an enhancer, as shown in the cartoon below. We include a parameter for chromatin accessibility of the enhancer, which changes the amount that the enhancer contributes to an increase in transcription. This model assumes that a transcription factor is always bound to the promoter, so there will always be at least one stable steady state of mRNA and protein concentration determined by the production and degradation rates.



3 Model

$$\begin{aligned} \frac{dm}{dt} &= \lambda_{m1} + \lambda_{m2}\frac{p^n}{k^n + p^n} - d_m m \\ \frac{dp}{dt} &= \lambda_p m - d_p p \end{aligned} \tag{1}$$

Assume all the parameters given are positive real numbers. The biological meaning of each parameter is the following:

$m(t)$ = concentration of mRNA
$p(t)$ = concentration of protein
k = chromatin accessibility of enhancer
λ_{m1} = basal transcription rate
λ_{m2} = maximum increase in transcription due to enhancer
λ_p = translation rate
d_m = degradation rate of mRNA
d_p = degradation rate of protein
n = hill coefficient

4 Methods and Analysis

4.1 Solution of ODEs

We computed solutions for ODEs by using NDSolve function in Mathematica with arbitrary parameter values. The following plot shows how the solutions behave temporally. For this specific example, chosen parameter values are $\lambda_{m1} = 2, \lambda_{m2} = 18, k = 10, d_m = 0.8, \lambda_p = 1, d_p = 1, n = 4$

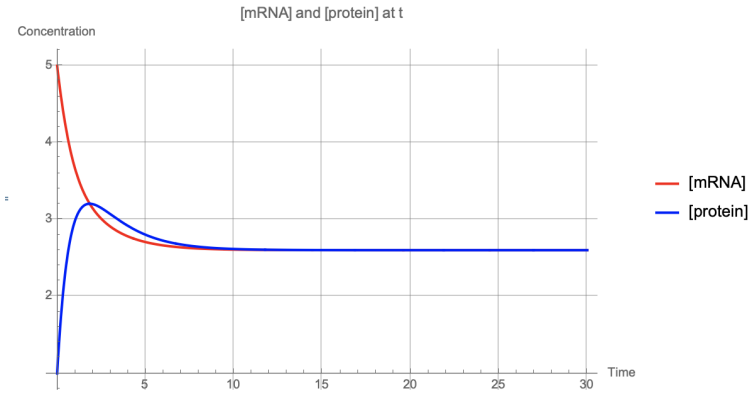


Figure 1: solution of ODE

4.2 Steady state equations

In order to gain insight about how this system behaves at steady state, we solve $m(t)$ and $p(t)$ using a steady state assumption. This enables us to plot the nullclines of the system and therefore see how changing parameter values affects stability within the system. Next we looked further into stability of the steady states by examining the Jacobian matrix.

$$\begin{aligned} m &= \frac{\lambda_{m1}}{dm} + \frac{\lambda_{m2}}{d_m}\frac{p^n}{k^n + p^n} \\ p &= \frac{\lambda_p}{d_p} m \end{aligned} \tag{2}$$

4.3 Jacobian and Stability Analysis

For convenience, let f be $\frac{dm}{dt}$ and g be $\frac{dp}{dt}$. Here we performed stability analysis by computing Jacobian matrix from the system above. Then we have

$$J = \begin{bmatrix} -d_m & \frac{\lambda_{m2} n p^{n-1}}{k^n + p^n} - \frac{\lambda_{m2} n p^{2n-1}}{(k^n + p^n)^2} \\ \lambda_p & -d_p \end{bmatrix}$$

Since we have $J \in \mathbb{R}^2$, without needs of computing eigenvalues, we can determine its stability by simply examining trace(τ) and determinant(det) of the Jacobian. Note that the system is stable when it has a Jacobian with negative trace, and a positive determinant.

$$\begin{aligned} \tau &= -d_m - d_p = -(d_m + d_p) \\ det &= d_m d_p - \lambda_p \left(\frac{\lambda_{m2} n p^{n-1}}{k^n + p^n} - \frac{\lambda_{m2} n p^{2n-1}}{(k^n + p^n)^2} \right) \\ &= d_m d_p + \lambda_p \frac{\lambda_{m2} n p^{n-1}}{k^n + p^n} \left(\frac{p^n}{k^n + p^n} - 1 \right) \end{aligned} \tag{3}$$

Note that we have $\tau < 0$ since all the parameters above are positive real numbers due to our assumption. Therefore a positive determinant determines a stable steady state in this system; and furthermore we know by choosing the parameter values carefully, we can enforce the system to have the stable or unstable steady states just from this equation.

5 Results

5.1 One steady state: Heterochromatin

We chose the following parameter values to enforce one steady state within the system: $\lambda_{m1} = 2, \lambda_{m2} = 5, k = 10, d_m = 0.8, \lambda_p = 1, d_p = 1, n = 2$. Below we show the nullclines and resulting vector field of the system using these parameter values. By looking at the vector field it appears that solutions are spiraling inwards at the fixed point, indicating a stable steady state. Thus, we can pictorially assume we have a stable steady state. Our calculation result matches with our observation since the determinant is positive with value of 0.681119. To reiterate, we demonstrated in section 4.3 that the trace is negative for all parameter values, so checking trace is not needed for stability analysis.

This mono-stable regime shows a low amount of protein and mRNA at steady state. This is aligns well with what we would expect if the enhancer were in a region of heterochromatin. If nucleosomes are densely clustered in the enhancer, the protein product of our gene cannot bind and therefore the positive feedback loop cannot contribute to the amount of mRNA being transcribed. In this case, mRNA concentration is only influenced by basal transcription and degradation.

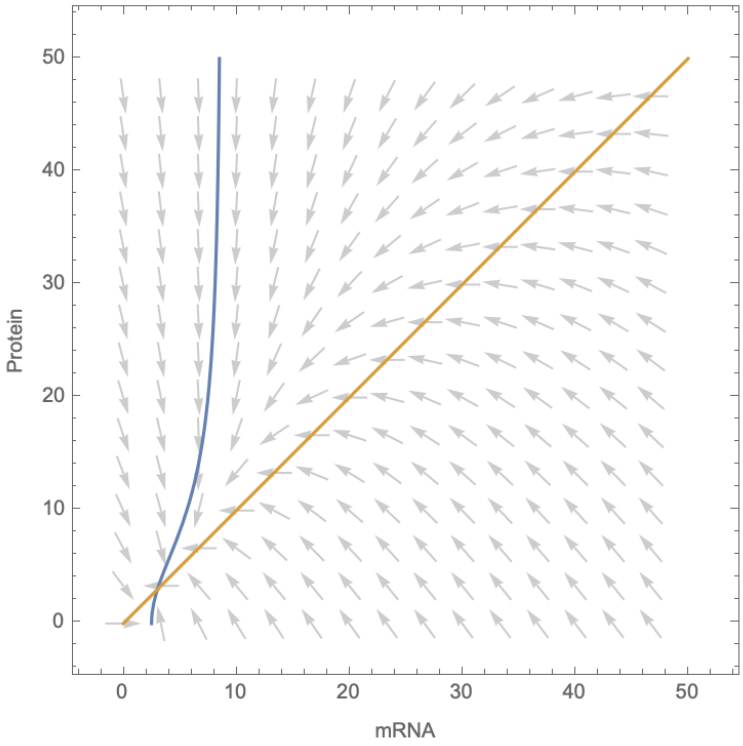


Figure 2: One stable steady state

5.2 Bistability: Heterochromatin and Euchromatin

Now that we have examinined a mono-stable parameter regime, we are interested in investigating the system when there are multiple steady states. There are many ways to induce multistability mathematically from our equations. However, we decided to narrow down our focus to hill number parameter n, and chromatin accessibility of enhancer parameter k, since those are the most biologically relevant and worthwhile to examine.

5.2.1 Bistability induced by high hill number

Using the Manipulate function in Mathematica, we observed the nullcline and vector field plot as we vary the Hill coefficient n. Increasing the hill number enables the system to have three steady states. Just by observing the vector field, we can see two stable steady states and one unstable steady state. Stable points 1 and 3 on the below figure appear to sprial inwards while point 2 is a saddle. Once again this aligns with what we determine by examining the determinant of the Jacobian. Note that unlike our previous monostability case, we don't need to assume for low transcription rate anymore, we give relatively higher parameter value for λ_{m2}. Increasing the hill number allows for ultrasensitive heterochromatin spreading and this multiple stable steady states. Model parameters are the following:

$$\lambda_{m1} = 2, \lambda_{m2} = 18, k = 10, d_m = 0.8, \lambda_p = 1, d_p = 1, n = 4$$

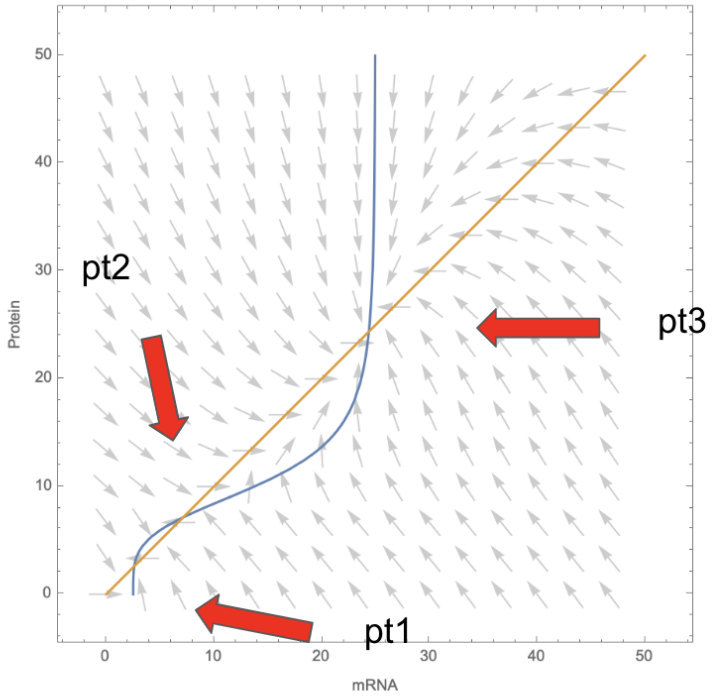


Figure 3: Bistability: Hill number 4

point	det value	sign
pt1	0.685301	$det > 0$
pt2	-1.04999	$det < 0$
pt3	0.726599	$det > 0$

5.2.2 Bistability induced by varying chromatin accessibility

Once again using the Manipulate function in Mathematica, we observed the nullclines and vector field plot aas we vary the chromatin accessibility parameter. In the plot below we show that multistability is achieved with a certain range of k values. Biologically, this implies that a certain amount of chromatin accessibility allows for a gene's transcription to be enhanced. However, chromatin accessibility does not necessarily mean that the enhancer will certainly increase transcription, which is why there is two In the plot below the chosen parameter values are same as the above case.

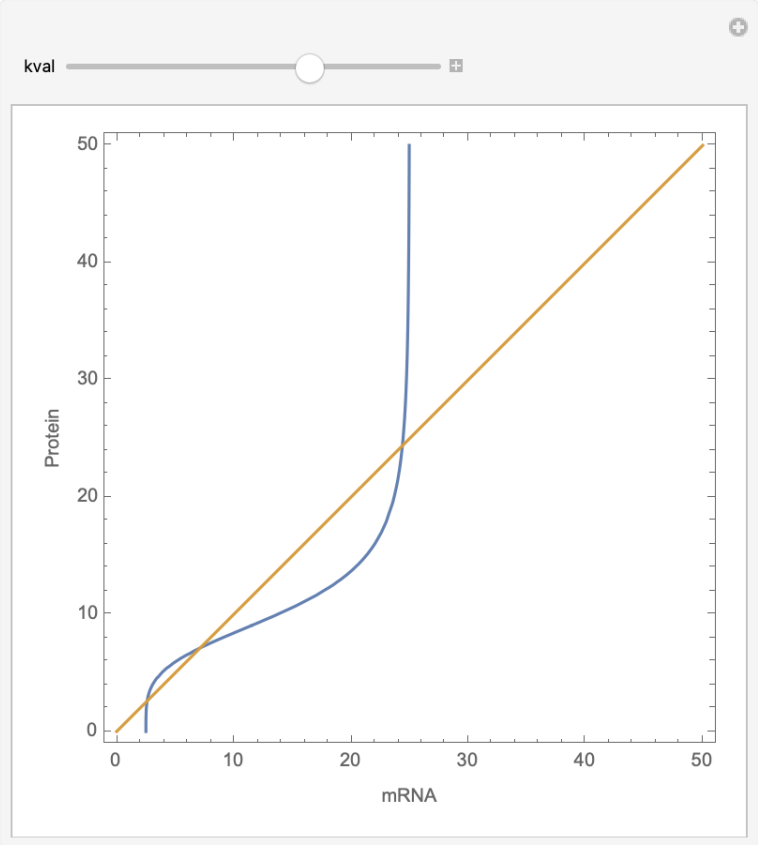


Figure 4: Bistability: k value

We wish to determine what range of k value gives bistability in more systematic way, so we solve our ODE for k in terms of p.

$$k = p\left(\frac{\lambda_{m2}}{d_m \frac{d_p}{\lambda_p} p - \lambda_{m1}} - 1\right)^{1/n}$$

By plotting protein concentration versus chromatin accessibility, we can see when k value is in range of middle section denoted as euchromatin, we have bistability. By computing critical points, and plug back that values in our k equation, we actually can figure out corresponding k values are $7.52506 < k < 14.7696$ for our choice of parameters. A monostable region with a low amount of protein is consistent with what we expect from heterochromatin, while bistability or monotstable with high protein makes sense with respect to euchromatin.

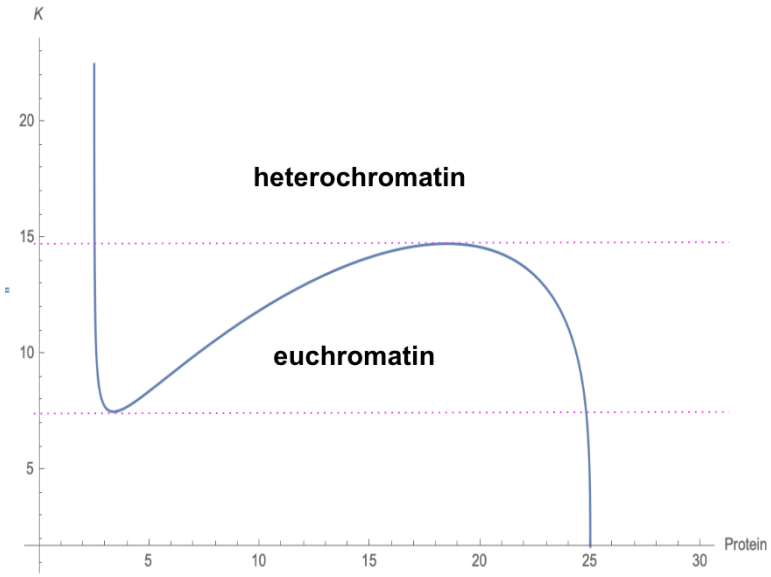


Figure 5: Range of k

6 Conclusion

By adapting a deterministic model from Huang et al, we built a model that captures some aspects of chromatin accessibility's affects on gene regulation. We solved the system and performed stability analysis using Mathematica. We demonstrated that for certain parameters the model shows monostability that aligns well with heterochromatin features, while for another set of parameters we have bistability that displays euchromatin features. This model could be extended to include multiple enhancers acting on a single gene, to more accurately capture what really goes on in biology.

7 Works cited

Huang, Rongsheng, and Jinzhi Lei. "Cell-Type Switches Induced by Stochastic Histone Modification Inheritance." Biorxiv, 2018, doi:10.1101/419481.