Boolean Logic Promoter Architecture: Gene Expression Through Combinatorial Encoding in the Immune Response

Introduction

NFκB, AP1, and IRF are regulators in the immune response. NFκB is activated in response to inflammatory cytokines and bacterial products, driving expression of immune response-related genes. AP1 responds to growth factors and stress signals, while IRF signals are critical for antiviral responses. These three factors regulate immune genes together, and their combinatorial logic determines the cell's response. In this project, I will mathematically model two mixed logic gates that simulate the body's response to a combination of signals from these three transcription factors (TFs) that regulate downstream activation of gene transcription, and explore the ways that the two different gates respond to different combinations of TFs.

This question is motivated by the biological complexity of signal transduction. As demonstrated in lecture, gene expression is dependent on a variety of signaling pathways. In immune responses which have evolved to be highly specialized, combinatorial encoding enables immune cells to modify their gene expression to best respond to specific situations despite using the same cellular machinery. If any single signal were sufficient to trigger transcription, gene expression would become overly sensitive and occur too frequently, eliminating the level of specialization and control that cells require to carry out the correct function.

Results

To investigate how the different combinations of TFs regulate gene expression, we constructed two model genes with promoter regions using change equations and Boolean logic gates. The first gene, "NFκB or (AP1 AND IRF)", utilizes an AND gate inside an OR gate, meaning that expression will result from either NFκB activation, or both AP1 and IRF, the latter two having to both be on at the same time. The second gate, "(NFκB OR AP1) AND IRF", utilizes an OR gate inside an AND gate, meaning that either NFκB or AP1 must be on, and additionally, IRF must be on for expression to occur.

Both models, when active, influence the transcription of one gene. Promoter activation depends on concentration of these TFs, and additionally a Hill function. The Hill function describes the proportion of binding sites that are occupied by the TFs, and can be described as the function that approximates the probability that the gene is producing transcript. We created two Hill functions, one corresponding to each model. The first was constructed by creating a Hill function for the AND gate, and then combining this function with the proper format for a Hill function corresponding to an OR gate (Appendix entry 4). The second was similarly constructed, and consisted of an OR gate inside of an AND gate Hill function. The key difference between the two models is the way that the TFs combine, and how they interact to regulate transcription.

Model (a) expresses the gene whenever NFκB alone is active, or when AP1 and IRF are both active. The Hill function was constructed such that the gene is expressed when NFκB alone is present in the cell, and can independently produce transcription downstream. If NFκB were not present in this model, both IRF and AP1 must both be present at high enough levels to promote transcription. Model (b) requires IRF to be present in all cases, and an additional transcription factor, which could either be AP1 or NFκB. In this model, transcription requires IRF

to be active in order to begin transcription, and the addition of AP1 or IRF increases mRNA transcription, as seen in Figure 3.

To examine dynamic responses, we simulated both models under transcription factor inputs that turned each factor on and off in partially overlapping windows (Figure 2). This allowed us to compare and contrast how each promoter architecture interprets the combinatorial values of TFs. We plotted the resulting mRNA trajectories alongside the activation timelines of NFκB, AP1, and IRF. Results show that Model (a) produced higher mRNA output throughout the time period, particularly in intervals where NFκB or AP1 were active in the absence of IRF. In contrast, Model (b) showed little expression whenever IRF was absent, consistent with its requirement for IRF to be present in order to begin transcription.

To examine steady state responses in our models, we analyzed gene expression through modeling the amount of mRNA present compared with the maximum amount of mRNA possible (Figure 3). The maximum amount of mRNA produced was retrieved from a model in which all TFs were turned on for a prolonged period of time (20 hours), and mRNA concentration was allowed to reach a steady state level of approximately 15.9. Each heatmap in Figure 3 represents the percentage of mRNA formed during steady-state divided by the absolute maximum mRNA value, revealing the level of expression that the combination of TFs predicts.

In Model A, when NFκB was strong, the gene was highly expressed across all concentrations of AP1 and IRF, consistent with NFκB's ability to influence transcription alone. When NFκB amplitude was low, expression required both AP1 and IRF to be high to reach full activation. In Model B, expression was almost completely dependent on IRF. When IRF amplitude was lowered, we saw decreased expression in mRNA, and very little increase as AP1 and NFκB were increased. Additionally, when IRF amplitudes were zero, we saw no mRNA (0%) present in the simulation, corresponding with Model B's dependence on IRF concentrations for expression.

Both models share the same promoter kinetics, Hill dissociation constants, and mRNA synthesis and degradation rates. Thus, any differences in expression arise purely from the combinatorial logic of the Boolean operator and inputs. Some similar cases between the models arise, one being when both AP1 and IRF are simultaneously active. This satisfies both Boolean conditions in the models, and transcription results from both. The presence of NFκB and IRF also satisfies both models, and results in transcription from both. These are the only situations in which the models behave in similar ways, aside from trivial cases like all TFs being on, or all being off.

To directly compare the models, we computed heatmaps of the difference in steady-state mRNA between Model (a) and Model (b) (Figure 4). Model (a) consistently produced more mRNA in all cases compared to Model (b). The difference was especially notable when NFκB was strong and IRF and AP1 were weak, and when the absence of IRF was limiting Model (b) from transcription. Regions where the models produced similar values of mRNA were when all TFs were strong. This confirms that the differences in expression are solely due to the difference in the Boolean gate that determines activation for each model.

For Figure 2, the original HW6 framework was modified to allow NFκB and AP1 to activate twice during the simulation, creating two distinct windows where they were present. We added this feature to create combinations of TFs within Figure 2. For the heatmap analyses

(Figures 3 and 4), transcription factor amplitudes were systematically varied from 0.0 to 1.0 in increments of 0.1, and steady-state mRNA levels were computed from the final time point of each simulation. We chose to examine steady-state responses because they better represent the transcriptional output of each promoter. Using these rather than transient values isolates the differences between the gates, allowing us to investigate them independently of timing or concentration effects. mRNA values were normalized to the global maximum mRNA concentration calculated from a simulation where all TFs were fully active for 20 hours, until mRNA reached a steady state concentration. Biologically, the “amplitude” corresponds to the concentration or activity level of that factor, which can vary substantially across different cell states. By varying amplitude and keeping other constants uniform, we can observe how each model reacts to weak, moderate, or strong activation of its regulatory inputs. We selected amplitudes of 0.1, 0.5, and 1.0: 0.1 representing a low level, weak signaling input, 0.5 representing an intermediate activation state, and 1.0 representing full activation. Using these three values clarifies that not only presence, but also concentration of TFs influences how much mRNA is transcribed.

Discussion

The goal of this project was to understand how different biological gates shape gene expression. Our results demonstrate that cells achieve specialized gene expression not through dedicated signaling pathways, but through promoter combinatorial logic. By arranging the same three transcription factors in different Boolean logic pathways, cells can convert environmental signals into the correct transcriptional response. Model A's logic (NFκB override) suits genes requiring rapid, broad activation. Model B's restrictive logic (IRF dependency) suits genes requiring high specificity, ensuring the immune response activates only when pathogens appear to indicate infection. Evolutionary pressure to balance sensitivity with specificity likely resulted in the use of combinatorial logic in gene expression to make unique responses possible with identical cell hardware. These results directly address our initial question by demonstrating how different logical architectures enable cells to react differently under similar signaling environments.

Figure 1:

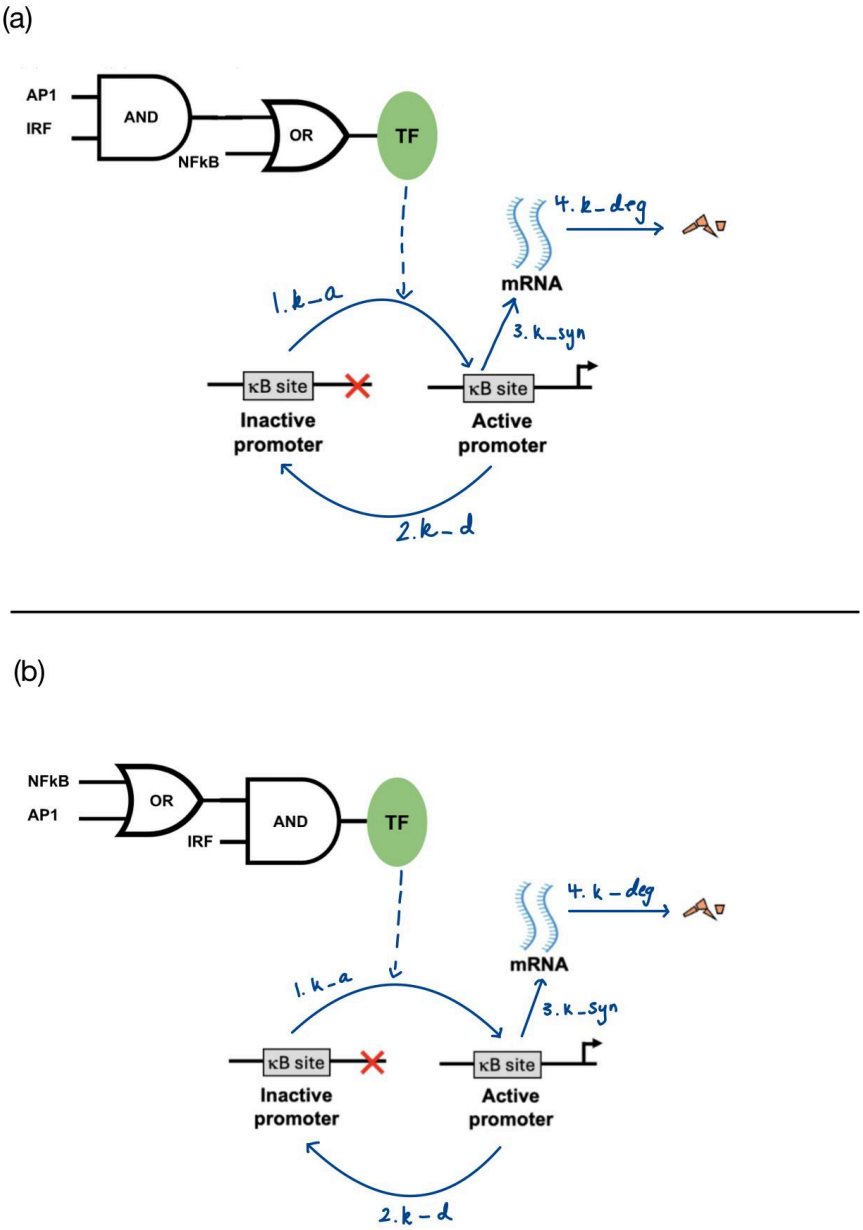


Figure 1: Network diagram that shows all reactions and rates in both models.

(a) A logic gate in which either NFKB must be present, or AP1 and IRF must both be present in order to activate the promoter at a rate of k_a. The active promoter then synthesizes mRNA at the rate k_syn. mRNA degrades at the rate k_deg, and the active promoter returns to inactive promoter at the rate k_d.

(b) A logic gate in which either NFKB or AP1 must be active, and additionally, IRF must be active for the promoter to become activated at the rate k_a. The active promoter then synthesizes mRNA at the rate k_syn, which degrades at the rate k_deg. The active promoter returns to an inactive state at the rate k_d.

Figure 2:

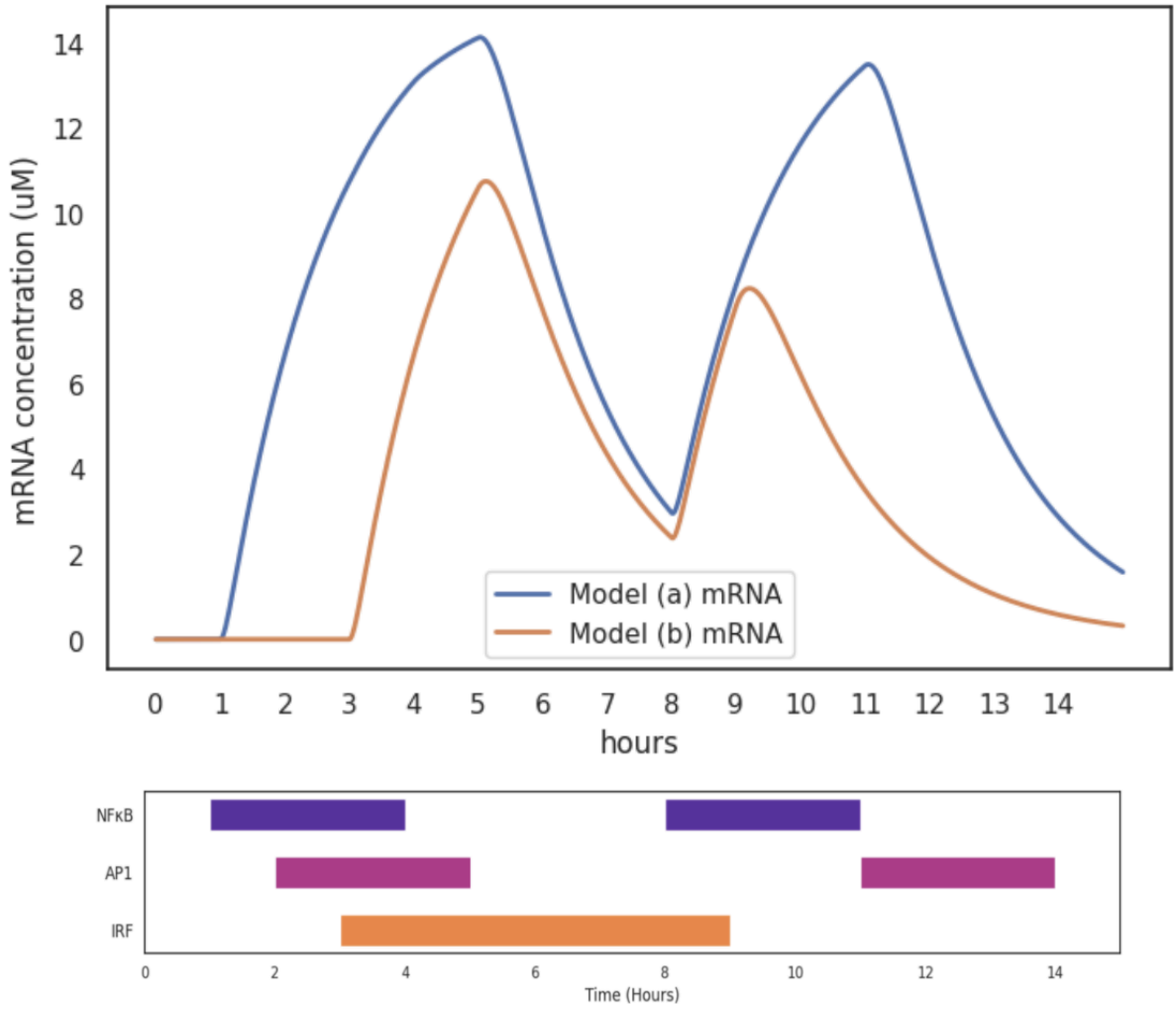
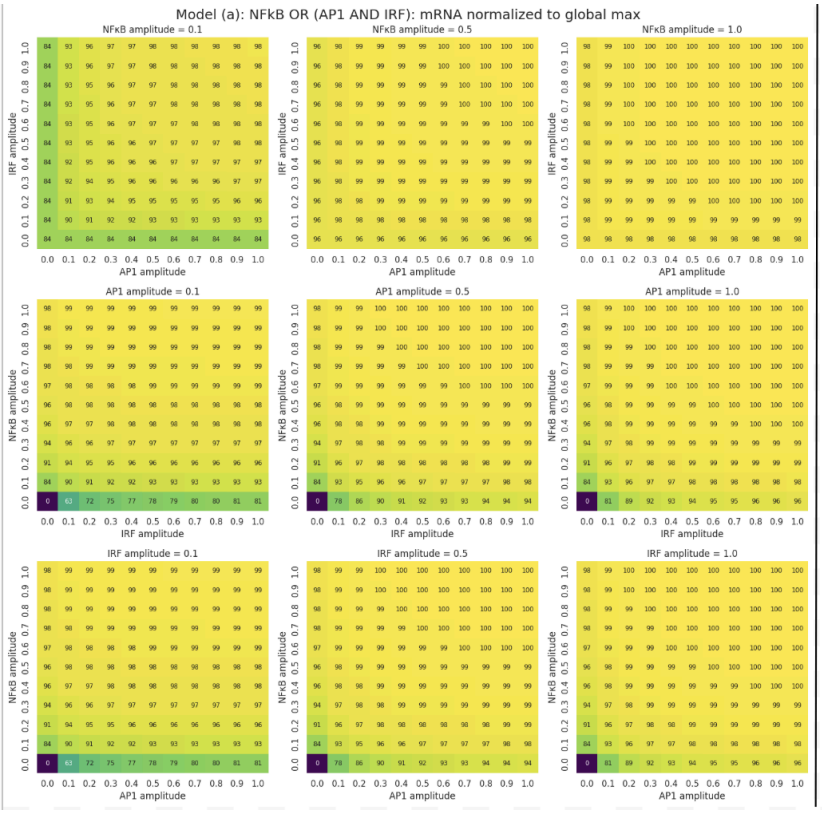


Figure 2: mRNA responses of the two gene models under staggered transcription factor levels.
mRNA production was simulated for Model (a) and Model (b) under an identical schedule of transcription factor activation, seen in the bottom panel. NFκB, AP1, and IRF were activated in overlapping time windows to reveal how both models react to these combinations of transcription factor presence. Model (a) produced higher mRNA levels compared to Model (b), especially during periods when NFκB or AP1 were active without IRF. Model (b) required IRF for any activation, resulting in lower overall expression. These results reveal how different combinatorial logic structures can generate distinct gene expression even under the same signaling inputs.

Figure 3:

3.a.

3.b.



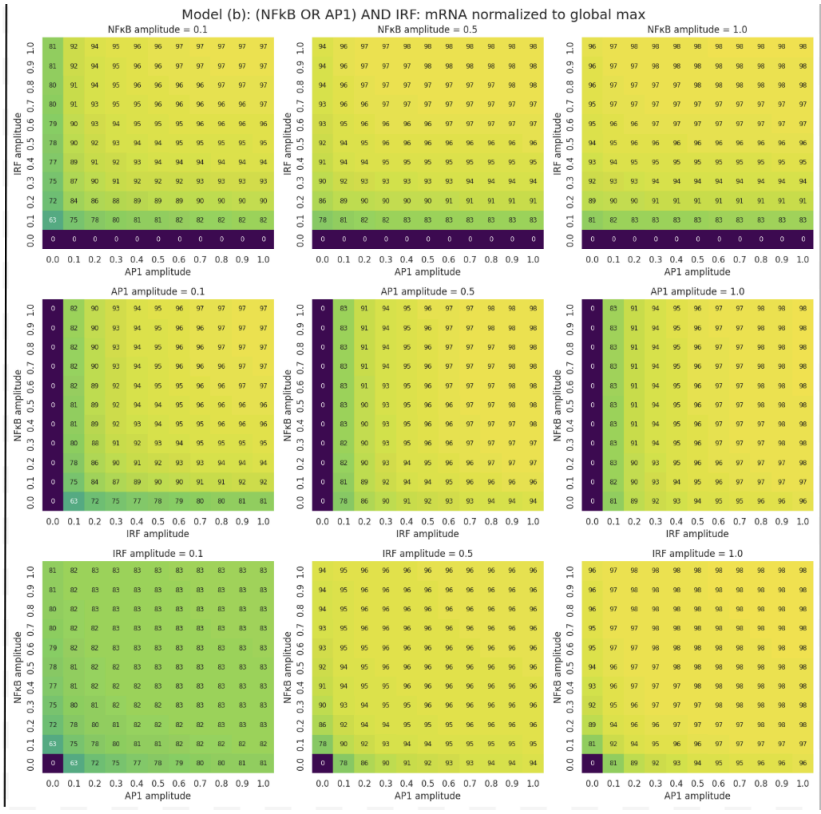


Figure 3: mRNA Levels measured as a function of TF amplitude.
mRNA concentration at steady state was measured as a function of transcription factor amplitude while they were all active simultaneously by taking the last value recorded from the simulation specific to these amplitudes. Data is reported as the percentage of maximum possible transcript concentration for each data point within each model gene. Maximum possible transcript concentration was observed by letting mRNA levels reach steady state while all TFs were active at their highest level (amp = 1.0), was found to be 15.948764030685542. 3.a.: mRNA transcription levels observed in Model (a). 3.b.: mRNA transcription levels observed in Model (b).

Figure 4:

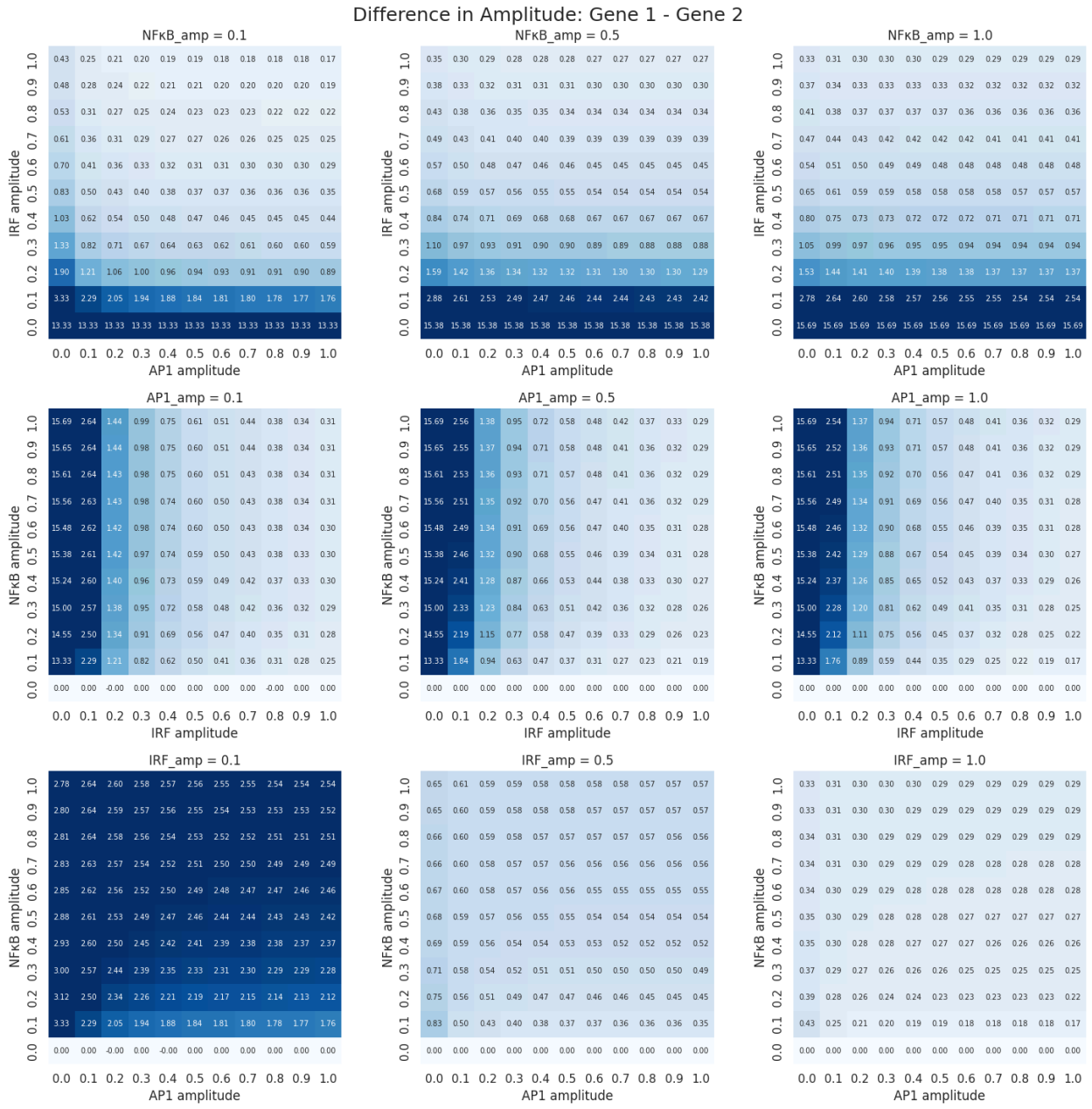


Figure 4: Difference in mRNA levels between Model (a) and Model (b).
Heatmaps show the difference in steady-state mRNA concentration. Figure 1 revealed that Model (a) consistently created a higher level of mRNA, so we computed Model (a) – Model (b) and plotted the difference. Higher values (darker blue) indicate a larger difference between Model (a) and Model (b), and lower values (light blue or white) indicate little to no difference. These results demonstrate that the combinatorial logic structures of the two models produce different levels of transcription based on which TFs are present, and at what level they appear.
Appendix

1. State Variables
pr(t): fraction of promoter in the inactive state
pr_a(t): fraction of promoter in the active state
tr(t): mRNA transcript concentration

Corresponding initial conditions:
pr0 = 1
pr_a0 = 0
tr0 = 0

2. Model Parameters
k_a = 0.2: Rate at which inactive promoter transitions into active state
k_d = 0.05: Rate at which active promoter returns to inactive state

K_d1 = 0.1: dissociation constant for NFKB binding
K_d2 = 0.1: dissociation constant for AP1 binding
K_d3 = 0.1: dissociation constant for IRF binding

k_syn = 0.2: mRNA synthesis rate
k_deg = 0.01: mRNA degradation rate

3. On/Off times used for steady-state observation in Figure 3 when amplitudes were varied

$$NF\kappa B(t) = \begin{cases} 1, & 0 < t < 1200 \\ 0, & \text{otherwise} \end{cases}$$

NFKB_on = 0 min
NKFB_off = 1200 min
NKFB_amp = (0, 1.0, 11)
(all values varied between 0 and 1 in intervals of 0.1)

$$AP1(t) = \begin{cases} 1, & 0 < t < 1200 \\ 0, & \text{otherwise} \end{cases}$$

AP1_on = 0 min
AP1_off = 1200 min
AP1_amp = (0, 1.0, 11)
(all values varied between 0 and 1 in intervals of 0.1)

$$IRF(t) = \begin{cases} 1, & 0 < t < 1200 \\ 0, & \text{otherwise} \end{cases}$$

IRF_on = 0 min
IRF_off = 1200 min

IRF_amp = (0, 1.0, 11)

(all values varied between 0 and 1 in intervals of 0.1)

4. Hill Functions

a. Hill Function for model (a)

$$H = 1 - (1 - H_{NF\kappa B})(1 - H_{AP1,IRF})$$

$$H_{NF\kappa B} = \frac{NF\kappa B}{NF\kappa B + K_{d1}}$$

$$H_{AP1,IRF} = \left(\frac{AP1}{AP1 + K_{d2}}\right)\left(\frac{IRF}{IRF + K_{d3}}\right)$$

b. Hill function for model (b)

$$H = H_{NF\kappa B,AP1} \cdot H_{IRF}$$

$$H_{NF\kappa B,AP1} = 1 - \left(1 - \frac{NF\kappa B}{NF\kappa B + K_{d1}}\right)\left(1 - \frac{AP1}{AP1 + K_{d2}}\right)$$

$$H_{IRF} = \frac{IRF}{IRF + K_{d3}}$$

5. Change Equations

pr_prime = -k_a*H*pr +k_d*pr_a

pr_a_prime = +k_a*H*pr -k_d*pr_a

tr_prime = +k_syn*pr_a -k_deg*tr

6. Chemical Reactions

1) k_a*H

pr -> pr_a

2) k_d

pr_a -> pr

3) k_syn

pr_a -> pr_a + tr

4) k_deg

tr ->

7. Acknowledgements

All team members contributed to the foundational components of the project. Ryan, Kailyn, and Josephine developed the model structure together, constructed the network diagrams, and derived the change equations. After establishing the shared code, each team member independently performed their own simulations, designed their own test conditions, and generated their individual figures and analyses. OpenAI's ChatGPT assisted with code for visualizing results (coding Figures 2-4).