In this model, we endeavor to account for our current understanding of the modes of Yan regulation as shown in [Fig Network Schematic]. We explicitly model Yan dimerization and the regulatory influence of Yan dimer species. Specifically, we build on current work being done in the Rebay lab and use thermodynamic equilibrium binding site models to simulate the activity of the Mae, miR-7, and Yan promoters in response to competition between repression by Yan dimer and activation by PntP1 and phosphorylated PntP2. We model the formation of Mae-Yan and Mae-PntP2P complexes to incorporate the effects of inactivation by Mae interaction. Yan transcription and translation are modeled separately to include the translational repression of Yan by complex formation of miR-7 with Yan mRNA. The influence of external Notch and EGFR inputs to the developing cell are represented by changing levels of Su(H) and dpERK in time. In our model, Su(H) is an activator of Yan transcription, while phosphorylation of PntP2 and Yan by dpERK is approximated with Michaelis-Menten enzyme kinetics. In total, our model consists of 33 individual induction, degradation, or complex formation reactions.

|  |  |  |
| --- | --- | --- |
| **Parameter** | **Value** | **Reference** |
| Default protein lifetime | 5 hr |  |
| Default RNA lifetime | 10 min |  |
| Phosphorylation time per dpERK | 1 min |  |
| Default complex dissociation time | 5 min |  |
| Default protein copy number | 500 |  |
| Default RNA copy number | 20 |  |
|  |  |  |
| ETS Yan-DNA binding | 10 kcal/mol | [Szymczyna and Arrowsmith, J Biol Chem 2000] |
| ETS Pnt-DNA binding | 12.4 kcal/mol | [Szymczyna and Arrowsmith, J Biol Chem 2000 |
| Non-specific Yan-DNA binding | 5.8 kcal/mol | [Soumya De et. al, J Mol Biol 2014] |
| Yan-Yan SAM domain interaction | 7 kcal/mol | [Zhang et al. 2010] |
| Mae-Yan SAM domain interaction | 10.9 kcal/mol | [Zhang et al. 2010] |
| Phosphorylated Yan lifetime | 5 min |  |
| Mae-YanP dissociation time | 30 s |  |

**Parameter decisions**

The parameters used to specify our model derive from the combination of general timescale assumptions for certain biological processes and biochemical information specific to the Yan network (Table 1). We uniformly use default kinetic rates for reactions without specific prior biochemical knowledge to ensure that biological processes in our model are acting with appropriate timescales. We expect that subsequent tuning of these default parameters could alter the timescales of copy number fluctuations or the relative timings of species induction and downregulation, but would not affect the basic actions of regulatory mechanisms in the model. Biochemical investigations of ETS transcription factors from within the Rebay lab and other groups inform our choice of parameters describing the transcriptional regulation of Mae, miR-7, and Yan. Similar to how ratios of parameters tend to determine the behavior of biological models [Gutenkunst et al. 2007], measured biochemical quantities can determine other rate parameters in the context of specific biological mechanisms. For example, assuming the estimated promoter region binding energies, ~500 copies of Yan are required for efficient repression by Yan dimer. Given the basic rate for protein degradation, the induction rate for Yan is then determined to achieve the required copy number. Similarly, given the Mae-Yan SAM domain interaction energy and default protein and RNA copy numbers, the Michaelis constant describing Mae-Yan phosphorylation by dpERK and the dissociation constant describing the binding of miR-7 to Yan mRNA are both chosen to ensure effective Yan downregulation. In our model, the relative strength of these two mechanisms can be tuned by adjusting their respective parameters. Other parameters specific to the Yan network are also set with particular biological mechanisms in mind, including a short ~5min lifetime for phosphorylated Yan and a rapid ~30sec dissociation time for Mae-YanP. Continued experimental perturbation of specific network components ought to further inform the parameter choices in our model. Sensitivity analysis of the remaining incompletely specified parameters will elucidate the potential roles for different network components and determine the parameter ranges necessary for biological function.

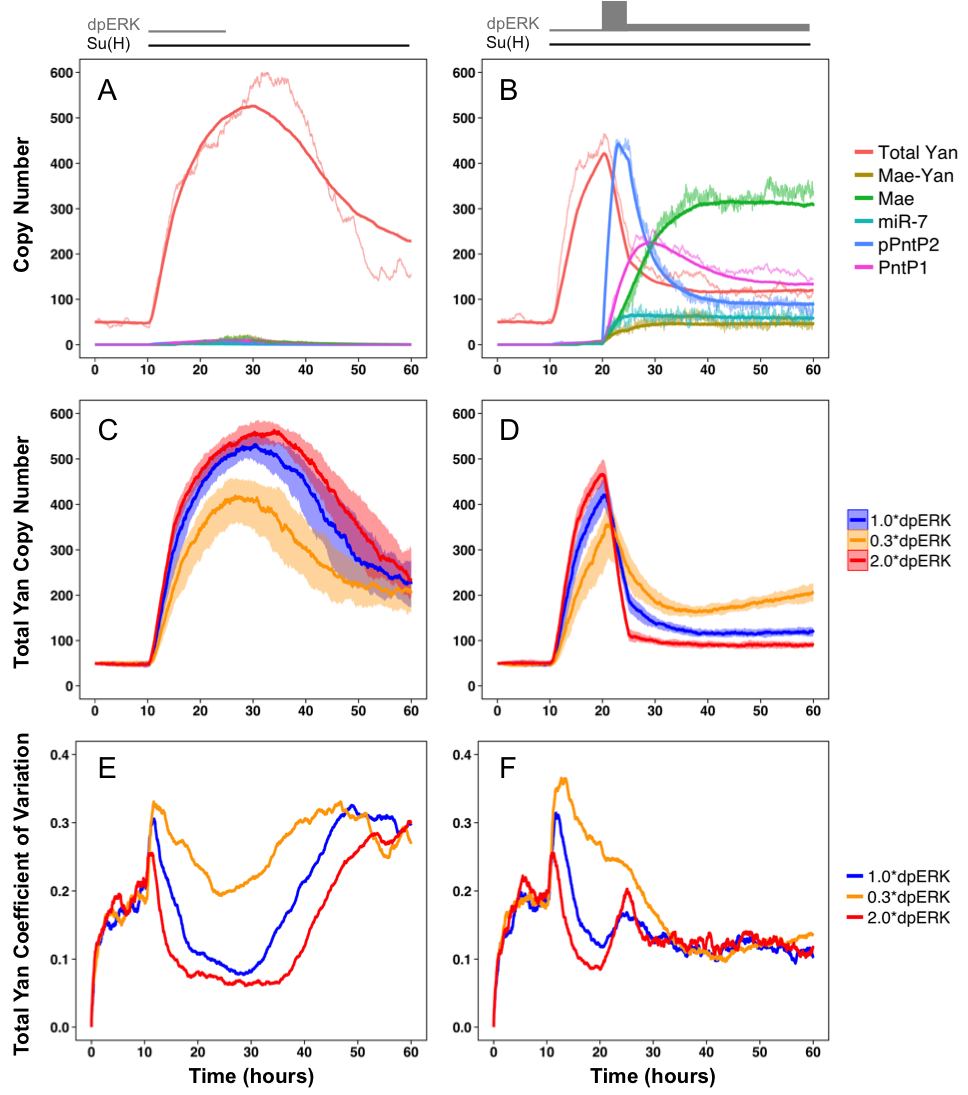
Results from our initial simulations largely reproduce experimentally observed and expected features of the Yan network [Fig Simulation]. Initially, Yan is produced and maintained at a low baseline level. At the 10hr mark, low levels of Su(H) and dpERK are introduced to the cells. The low level of dpERK phosphorylates a small fraction of PntP2, which then induces the expression of a low level of PntP1. Following experimental observations made from EGFR mutants, in our model PntP1 activates Yan transcription at low copy number. While some Mae and miR-7 are produced, their level of induction is not high enough to repress initial Yan induction. In simulated multipotent cells, this low level of dpERK is simply removed at hour 25, and Yan levels are slowly reduced as the remaining PntP1 degrades [Fig Simulation]. In contrast, simulated differentiating cells receive a strong pulse of dpERK from 20-25 hr, resulting in rapid phosphorylation of PntP2 and production of PntP1 [Fig Simulation]. This results in high levels of miR-7 and Mae that repress Yan mRNA and protein. Yan protein copy number is rapidly reduced by degradation of phosphorylated Yan produced via Mae. This low level of Yan is then maintained in differentiated cells by a persistent low level of dpERK that continues to repress Yan.

Modulating dpERK levels in simulation also qualitatively reproduce trends observed in EGFR mutant experiments. To emulate the EGFRts mutant, we scale the dpERK levels by 0.3 in simulation. In these simulated multipotent cells with deficient dpERK, Yan levels achieve a lower maximum and are reduced more rapidly. In simulated differentiating cells with deficient dpERK, Yan levels degrade more slowly upon differentiation and ultimately reach a higher copy number. When the level of dpERK levels is scaled by 2.0, emulating the RasV12 mutant, multipotent cells reach a higher maximum level of Yan and maintain higher levels of Yan at later times. In differentiating cells with increased dpERK, Yan levels degrade more rapidly upon differentiation and ultimately reach a lower copy number. The coefficients of variation of Yan copy number are quite similar in value to those observed in experiment. Also, the coefficient of variation tends to increase closely following peak Yan levels in simulations with deficient dpERK, a trend that appears in experiments with EGFRts mutant. Interestingly, in simulated multipotent cells, the coefficient of variation of Yan levels tends to increase in later times, while in corresponding experiments it tends to remain constant. This discrepancy indicates that our model mechanism for slow Yan decay, the eventual degradation of low levels of Pnt, is currently not sufficient to reproduce the noise characteristics found in the experimental data. These direct comparisons of noise properties between simulation and experiment will help identify instances in which noise is biologically regulated in unforeseen ways.

**Practical Simulation Info**

The simulation of a single 60 hour time course for differentiating cells record ~210,000 individual reaction events and require ~3sec to run on a standard desktop computer. The shortest reaction timescale in these simulations is ~30sec. For more detailed model representations or simulations of multicellular tissues, the compute time will scale inversely with the shortest reaction timescale and linearly with the number of cells simulated.

Is our current understanding of mechanisms within the network sufficient to describe the noise features observed, or are other biological mechanisms needed to generate them?



**Fig. XXX.** Stochastic simulation of the Yan network in response to Notch and EGFR input. **Left column:** Low level of transient EGFR simulates signaling interpreted by multipotent cells. **Right column:** Strong pulse of EGFR simulates signaling interpreted by differentiating photoreceptors. **(A-B)** Copy numbers of select regulatory species in time. Average copy numbers represented by thick lines and an example time course represented by thin lines. **(C-D)** Fluctuations in total Yan in response to varied dpERK levels. Median total Yan shown with overlaid interquartile range. **(E-F)** Coefficient of variation of total Yan in response to varied dpERK levels.