REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
11-07-2018	Final Report	1-Aug-2014 - 31-Jul-2017

4. TITLE AND SUBTITLE

Final Report: Biomathematics - Canalization: a fundamental design principle of gene regulatory networks

5a. CONTRACT NUMBER
W911NF-14-1-0486

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
611102

6. AUTHORS

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES

University of Connecticut - Health Center
263 Farmington Ave.

Farmington, CT 06032 -1956

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
64888-MA.1

12. DISTRIBUTION AVAILIBILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Reinhard Laubenbacher
UU	UU	UU	UU		19b. TELEPHONE NUMBER 860-436-7516

RPPR Final Report
as of 01-Aug-2018

Agency Code:

Proposal Number: 64888MA **Agreement Number: W911NF-14-1-0486**
INVESTIGATOR(S):

Name: PhD Christopher D Heinen
Email: cheinen@uchc.edu
Phone Number: 8606798859
Principal: N

Name: Ph.D. Reinhard C Laubenbacher
Email: laubenbacher@uchc.edu
Phone Number: 8604367516
Principal: Y

Organization: **University of Connecticut - Health Center**
Address: 263 Farmington Ave., Farmington, CT 060321956
Country: USA
DUNS Number: 022254226 EIN: 521625543
Report Date: 31-Oct-2017 Date Received: 11-Jul-2018
Final Report for Period Beginning 01-Aug-2014 and Ending 31-Jul-2017
Title: Biomathematics - Canalization: a fundamental design principle of gene regulatory networks
Begin Performance Period: 01-Aug-2014 **End Performance Period:** 31-Jul-2017
Report Term: 0-Other
Submitted By: pam vachon Email: vachon@uchc.edu
Phone: (860) 679-8708
Distribution Statement: 1-Approved for public release; distribution is unlimited.

STEM Degrees: **STEM Participants:**

Major Goals: See attached document

Accomplishments: See attached document

Training Opportunities: See attached document

Results Dissemination: See attached document

Honors and Awards: See attached document

Protocol Activity Status:

Technology Transfer: Nothing to Report

PARTICIPANTS:

Participant Type: PD/PI
Participant: Reinhard Laubenbacher PhD
Person Months Worked: 1.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Co-Investigator

RPPR Final Report
as of 01-Aug-2018

Participant: Christopher Heinen PhD
Person Months Worked: 1.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Postdoctoral (scholar, fellow or other postdoctoral position)
Participant: Russell Posner
Person Months Worked: 12.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

Participant Type: Other Professional
Participant: Ulysses Andrews
Person Months Worked: 6.00 **Funding Support:**
Project Contribution:
International Collaboration:
International Travel:
National Academy Member: N
Other Collaborators:

FINAL PROJECT REPORT
GRANT NR. W911NF-14-1-0486

Biomathematics-Canalization: A fundamental design principle of gene regulatory networks

PI: Reinhard Laubenbacher, UConn Health

1. INTRODUCTION

The mathematical objective of the project is to derive mathematical principles governing network dynamics of gene regulatory networks, focusing on the role of so-called microRNAs, that is, regulatory elements in the genome that act through certain types of feedforward loops. They are hypothesized to confer robustness on gene regulatory networks, stabilizing their dynamics in the face of extrinsic and intrinsic noise and other stochastic features. Thus, their action represents a potential "law of biology" for the stabilization of organismal phenotype in the face of uncertainty.

While the role of many individual microRNAs in gene regulatory networks has been studied experimentally, it is very difficult or impossible to study their collective, genome-scale role, especially from the point of view of deriving general features of the dynamics of the genome-scale transcriptional networks in mammals. (Note that there are estimated to be thousands of microRNAs in the genome, each of which can regulate potentially hundreds of transcription factors.) This leaves mathematical modeling as one of a few options, combined with genome-scale information on the human genome, e.g., features of transcriptional networks and microRNA features.

2. KEYWORDS

Dynamical system, feedforward control, stability, gene regulatory network, microRNA.

3. ACCOMPLISHMENTS

Our initial hypothesis was that miRNAs provide some kind of buffer against stochastic perturbations in trajectories reaching particular steady states. But ultimately, we were guided in our mathematical analysis by an experimental result in the literature we discovered, that shows that, when ALL miRNAs are disabled in a zebrafish embryo, then it still develops all the components of a mature organism, but in malformed and nonfunctional fashion. This led us to hypothesize that the most important role of microRNAs might not be to buffer embryonal development against environmental perturbations. Instead, their role might be to delay the differentiation of stem cells until a strong enough gene expression signal is present within the embryo to ensure faithful pattern formation protocols. And this is precisely what our model analysis has shown. Thus, combination of experimental results and model analysis has led us to formulate the fundamental hypothesis about the biological law we were investigating.

We constructed a stochastic genome-scale model of a gene regulatory network consisting of transcription factors (see attached manuscript for details). The model and the network were designed to have features known to exist In mammalian transcriptional regulatory networks. We then carried out a number of computational experiments on simulated networks, across a range of sizes, from tens to thousands of nodes:

- We introduced progressively more miRNA feed-forward loops in different distributions across the network.
- We varied the distribution of the miRNAs across the network.

- We investigated the effect of miRNAs on protein production.
- We carried out a variety of other computational and theoretical experiments.

We have obtained the following main results.

i. miRNAs delay cell differentiation. Undifferentiated stem cells show gene expression patterns with uniform low expression levels for all transcription factors. As the extracellular environment changes and a strong signal is produced, the cell differentiates into high levels of transcription for a subset of genes, establishing a particular phenotype. miRNAs delay the time to differentiation in a way that is directly proportional to the fraction of the genome composed of miRNAs (Fig. 1B), and is independent of network size (Fig 1A).

ii. miRNAs reduce the production of proteins while the stem cell remains in an undifferentiated state. This can be viewed as an efficiency measure, helping the cell conserve energy. The savings observed in our model is in remarkable agreement with experimentally measured values.

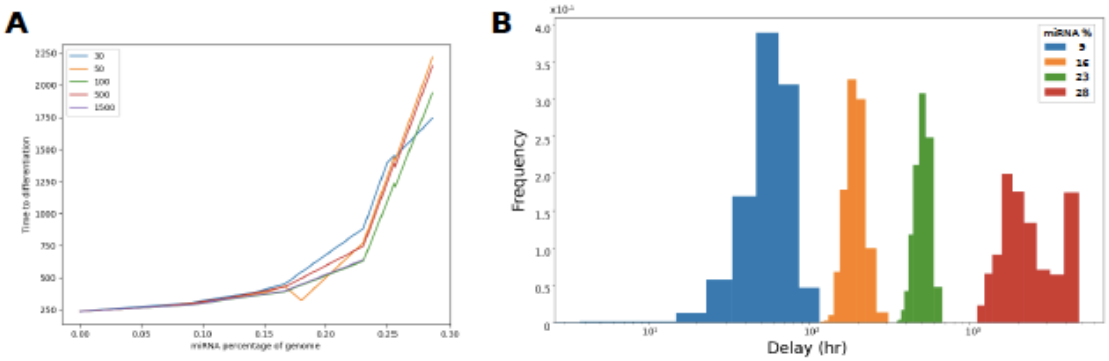


Fig 1. The graph in panel A shows that the delay mechanism is independent of genome size. Panel B shows the dependence of the size of the delay on the proportion of miRNAs in the genome.

These unique regulatory characteristics can help explain the significance of miRNA-mediated gene regulation in the developing embryo. As a stem cell differentiates, it focuses on the expression of a smaller set of transcription factors characteristic of the differentiated phenotype. One important role of the miRNA action is to “clean up” the transcripts of the other genes by limiting their translation, thereby preventing the cell from engaging in premature undesirable differentiation toward another phenotype. **In essence, miRNAs are the mechanism that maintains the unstable cell type of a stem cell until the conditions are created for a robust transition to another phenotype.**

We have not had an opportunity to extensively pursue the mathematical underpinnings of this phenomenon in the context of stochastic dynamical systems. One hallmark of robust gene regulatory networks is the existence of feedback controls. While transcription factors can form such feedback controls without other regulatory entities, miRNAs function as feedforward controllers, and cannot participate in feedback loops without the assistance of transcription factors. Classically, feedforward controls can compensate for disturbances more

quickly than feedback control, and are particularly useful for mitigating high-frequency disturbances. These features generally can decrease energy and resource consumption and prevent hysteresis. On the other hand, such controls require previous knowledge of system behavior, and do not generate new long-term stability in previously unstable systems. Our findings show that miRNA-mediated control provides these benefits to gene regulatory networks. We are continuing the investigation of a rigorous mathematical formulation of this feedback/feedforward control combination, with support from internal funds.

4. IMPACT

This project investigated the role played by miRNAs, a component of the genome, in development. The gist of our findings can be summarized in terms of Figure 2, an analogy made by the evolutionary biologist Conrad Waddington in the 1940s, who coined the term of so-called canalization. As cells develop and differentiate, characterized by their gene expression patterns, they travel along a landscape of mountains and valleys and eventually settle in one of the valleys, corresponding to a particular phenotype. This landscape is shaped by the action of the genes, through “pulleys” that determine the topography (Figure 3). The result of our project can be described by saying that miRNAs help to hold the cell at the top of the landscape in Figure 2 until the landscape is fully shaped by the action of the entire genome (Figure 3) to assure that the cell ends up in the “right” valley.

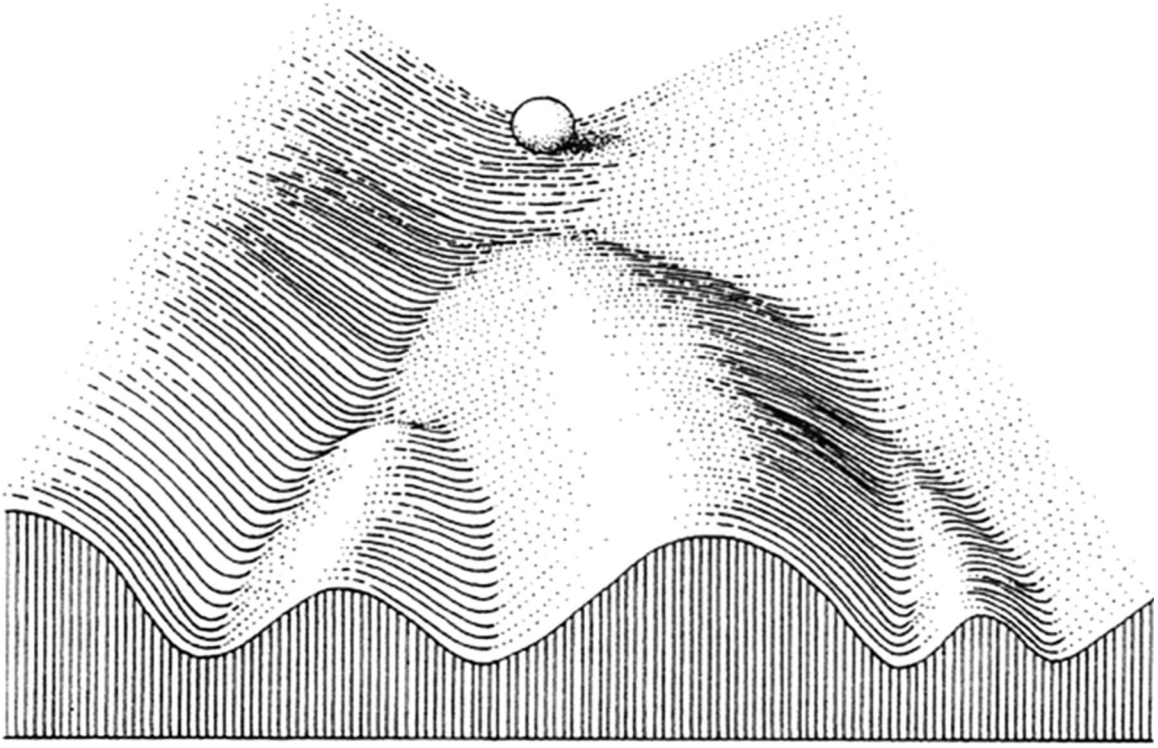


Figure 2. Waddington’s epigenetic landscape. (From Waddington, *The Strategy of the Genes,* 1958)

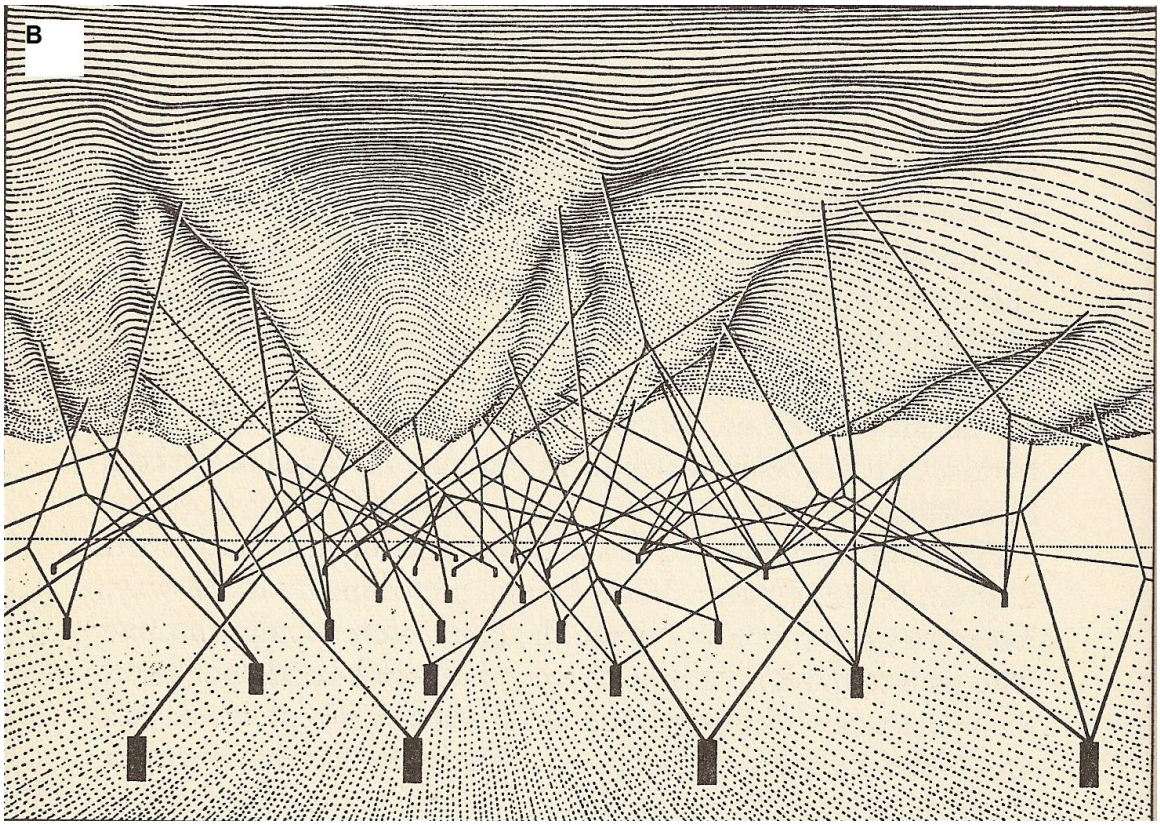


Figure 3. The action of the genes (blocks) through pulleys, shaping the epigenetic landscape. (From Waddington, *The Strategy of the Genes,* 1958)

The impact of this project is two-fold. Firstly, its impact on biology is that it proposes a mathematical paradigm of mixed feedforward/feedbackward control to guide evolutionary processes on a genome scale. Given the importance of this process, the mechanisms we identify can be considered a “fundamental law of biology.” Secondly, the impact on mathematics is that it presents a novel type of dynamical system with control that operates in an uncertain and changing environment. This provides new challeges for a mathematical analysis and new opportunities for potential applications to other fields, such as engineering.

5. CHANGES AND PROBLEMS

Our initial hypothesis was that miRNAs had a role in shaping the epigenetic landscape, that is, were part of the system of anchors and pulleys in Fig. 3. We spent a significant amount of time and effort trying to elucidate exactly how this would happen and what the mathematical basis for it would be, e.g., what the right modeling framework is. Through an extensive program of computational experimentation with the model and some fortuituous discoveries in the literature we finally settled on what turned out the actual mode of action, the delay mechanism in Figure 2.

One underlying reason for this problem is that there is relatively little information available at what happens at the genome scale, with most investigations focused on the effect of individual miRNAs. For instance, eliminating some specific miRNAs does in fact reshape the landscape.

6. PRODUCTS

We have a close to final draft of a manuscript (attached) we are planning to submit to the Proceedings of the National Academy of Sciences. We are preparing a second manuscript to be submitted to the Journal of Theoretical Biology.

7. PARTICIPANTS AND OTHER COLLABORATING INSTITUTIONS

In addition to the PI, the grant supported three graduate students:

i. Claus Kadelka, who received his Ph.D. in mathematics from Virginia Tech in 2015, and is starting this fall as a tenure-track assistant professor in the Mathematics Department at Iowa State University.

ii. Ulysses Andrews, who received his Ph.D. in mathematics from the University of Connecticut in 2016, and is now working in private industry

iii. Russell Posner, who will receive his Ph.D. in biomedical sciences from the University of Connecticut in 2019.

We collaborated with Dr. Chris Heinen, Ph.D., from the Department of Molecular Medicine at UConn Health, who performed several experiments for the project to validate some early predictions. We also collaborated with Drs. Pedro Mendes and Paola Vera-Licona, computational biologists at the Center for Quantitative Medicine at UConn Health. Finally, we collaborated with Dr. Jeff Chuang, a computational biologist at the Jackson Laboratory for Genomic Medicine, a not-for-profit research institute on the campus of UConn Health.

8. SPECIAL REPORTING REQUIREMENTS

N/A

9. APPENDICES

Appendix 1. Manuscript describing our main findings.
Appendix 2. Poster presentation on the project.

A Novel Proposed Role for miRNAs at the Genome Scale

Russell Posner[1] and Reinhard Laubenbacher[1]

[1]Center for Quantitative Medicine, UConn Health, 195 Farmington Ave. Farmington, CT

1 Introduction

The evolution of tissue-based organisms remains one of the most enigmatic and fascinating questions in biology. Numerous inroads have been made into understanding the origins first of multicellularity and later tissue-based structural organization, and this transition may have occurred over a dozen times in history[1]. Perhaps unsurprisingly, it was at this juncture that the number of genes in the primitive animal genome exploded, with an emphasis on new transcription factors, capable of manipulating a single common genome into a variety of different cell types[2].

While many different evolutionary trajectories led to the development of a variety of very different tissues and structures, all tissue-based organisms fundamentally depend on an "embryonic," or stem cell phenotype. Such a phenotype requires a number of unique properties: it must be capable of division and self-renewal in order to generate a sufficient number of "pre-precursor" cells, yet it cannot have the same level of long-term stability found in terminally-differentiated tissue types. Furthermore, a large part of this functionality must be hard-coded into the genome; without a diverse set of tissues to provide complex intercellular signals, the embryonic stem cell must be capable of producing much of this metastability internally. Finally, the timing of this differentiation process must be well controlled, so as to facilitate the simultaneous formation of codependent tissues.

One commonality of nearly all tissue-based organisms is their dependence on microRNAs (miRNAs) for proper cell differentiation and development[3;4]. These short, 16-22nt single-stranded RNA molecules facilitate post-transcriptional gene silencing across the genome in both plants and animals. Studies have found that miRNAs may be able to reduce stochastic noise in gene expression[5], and further that they evolved under selective pressure to target genes involved in cell differentiation and specification while avoiding genes responsible for more basic cellular processes[6]. Experiments in mouse and zebrafish embryos have also shown that disabling the miRNA regulatory system leads early developmental failure[7–9]. As a result, one popular theory for the role of miRNAs is that they may serve as canalizing agents - that is, they function to stabilize developing tissues and reinforce cell phenotype[4;10].

However, a number of features of the miRNA regulatory system have confounded our ability to better understand the role of miRNAs at the genome scale. While the number of distinct miRNA-coding elements in the genome may number in the thousands[11], there is significant overlap and redundancy in these elements; very similar (or even identical) miRNA-coding genes can be scattered across the genome[12;13]. Adding to the complexity, a single miRNA may have hundreds or even thousands of targets[14]. Despite this deep entanglement in the human gene regulatory network (GRN), most miRNAs appear to be dispensable, and large swaths of the "miRNA-ome" can be eliminated without producing any noticeable change in phenotype[15;16].

Given these complexities, we used mathematical modeling to understand the relationship between miRNA-mediated regulation and cellular phenotype. In this work, we designed a genome-scale model of transcription and translation which incorporated both miRNAs and protein-based transcription factors (TFs). Using our model, we found that miRNAs fill a unique niche in the gene regulatory apparatus: miRNAs appear to stabilize undifferentiated, transient cell types. Because tissue-based organisms depend on these metastable cell types at nearly every stage of development, our work suggests that miRNAs may have evolved as a facilitator of the embryo-based developmental model. Critically, our model shows that this effect is a consequence of the dynamics of miRNA-mediated gene regulation, and appears to be independent of the particular structure or size of the underlying gene regulatory network (GRN).

2 Results

In order to better understand the interplay between TF- and miRNA-mediated gene regulation, we developed a mathematical model of stochastic transcription and translation (see 6). In particular, we chose to focus on the process of spontaneous differentiation by examining how stochastic transcription and translation can induce feedback control at the TF level. Using protein copy numbers as an endpoint, we evaluated how cells with low (zero) copy numbers of transcription factors reached steady state. This formulation is designed to correspond to both experimental observations in ES cells[17;18].

miRNAs delay cell differentiation MiRNAs had a profound effect on time to differentiation (Figure 1A). Time to differentiation largely depended on the fraction of the genome composed of miRNAs (miRNA fraction $= \frac{\#\text{miRs}}{\#\text{miRs}+\#\text{TFs}}$). This effect was independent of genome size (Figure 2A). Interestingly, the dependence on genome fraction was remarkably robust: with respect to differentiation time, increasing the number of miRNAs targeting a particular gene could not compensate for a reduction in miRNA fraction (Figure 1B,C).

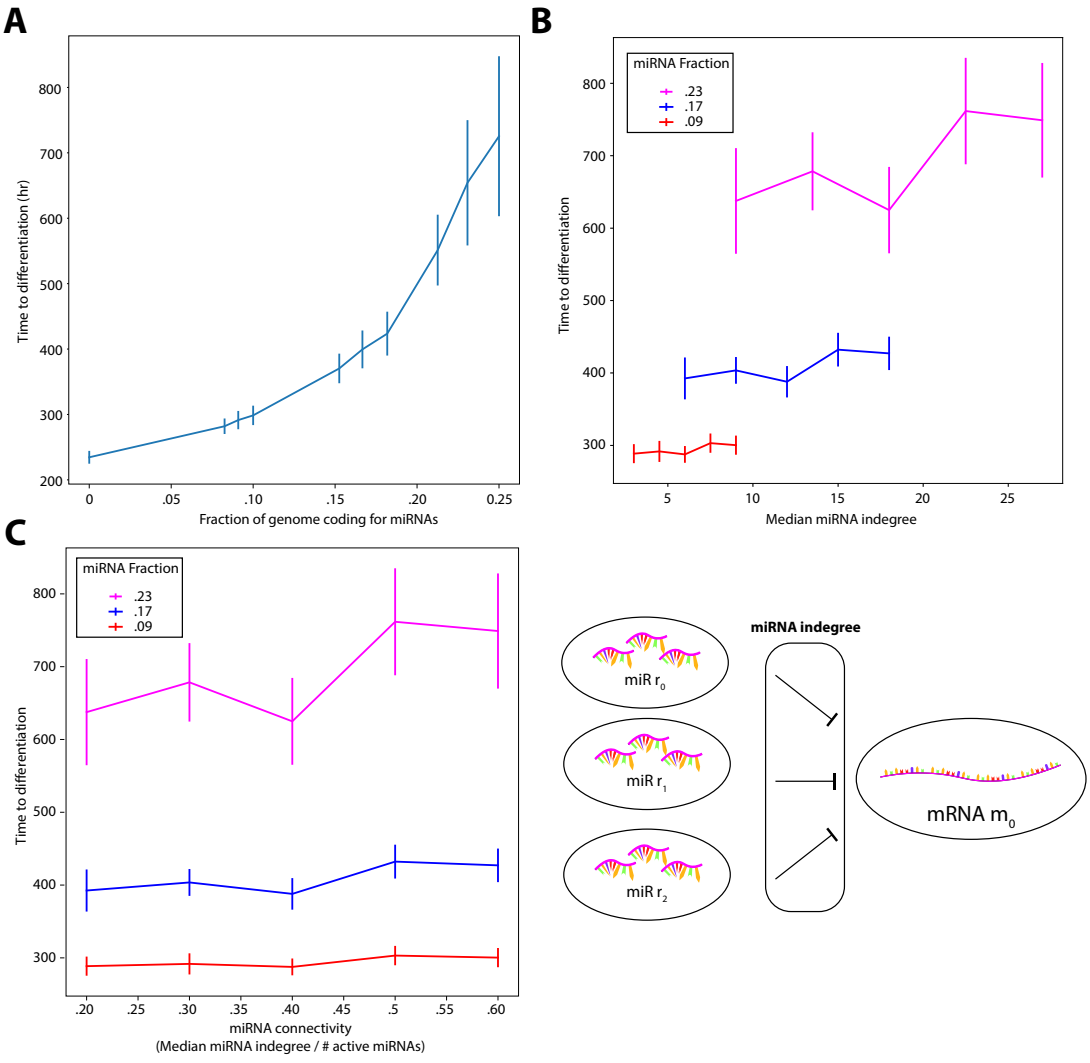


Figure 1: **A**, The model demonstrated a clear link between differentiation time and the fraction of the genome encoding miRNAs. This relationship was consistent even when miRNA indegree (that is, the number of miRNA genes which target a particular mRNA) was varied, both in the absolute sense (**B**) and relatively (**C**).

Delay magnitude varies with miRNA fraction of genome The previous findings suggested to us that the effect of miRNAs on differentiation time are a genome-scale property rather than being dependent on particular miRNAs in our networks. To further investigate this effect, we tested networks

of varying genome size with similar connectivity parameters (e.g. in- and outdegree distributions) to see what role genome size may play. Interestingly, genome size had little, if any, effect on time to differentiation (Figure 2A).

Despite a clear correlation between the fractional composition of miRNAs in the genome and time to differentiation, at sufficiently high levels - 25% or more - the predictability of differentiation time decreased substantially (Figure 2B), suggesting that at sufficiently high levels, control of differentiation time was lost. While the number of miRNAs in the genome is unknown, it is believed to be on the order of 10^3 [19;20], which would correspond to approximately 10% of expressed genes in a typical ES cell. In this range, we found that delay varied in a near-linear fashion with the fraction of miRNAs in the system and was on the order of $10^2 - 10^3$ hours. This finding was consistent with experimentally observed results, in which it has been shown that the presence of miRNAs appears to sustain the embryonic stem cell (ES cell) phenotype on a similar time scale [21;22].

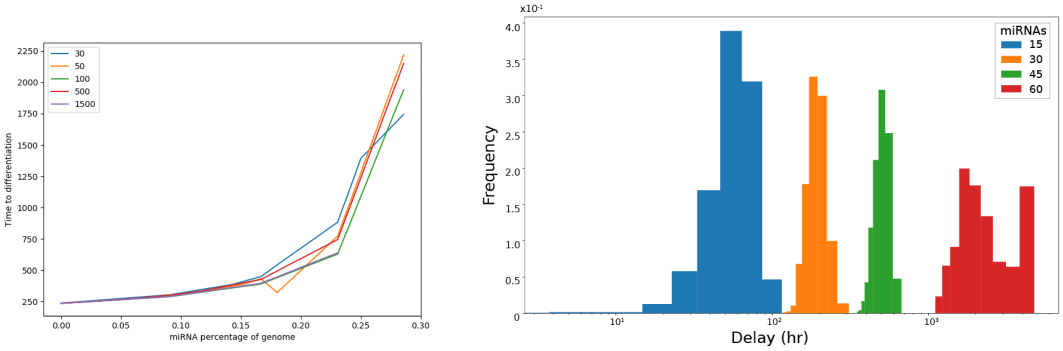


Figure 2: Time to differentiation increases with miRNA fraction of genome.

Effect on protein production As miRNAs function as repressors of translation, we observed that post-differentiation TF copy numbers decreased as miRNAs were added to the network as a result of a reduction in translation rates (Figure 3A,B)). As one might expect, miRNAs remove transcribable mRNAs from the RNA pool, thereby decreasing the total TF copy number. For a large range of miRNA levels, in the post-differentiated state, miRNA quantity had a near-linear effect on TF copy numbers.

Remarkably, this reduction amounted to a roughly 4.5% decrease in TF copy number per added miRNA, in good agreement with previous experimental estimates of a 3.5-5.6% reduction in protein copy number per miRNA [23]. At the highest levels, however, miRNAs appeared to decrease protein translation rates to a point of apparent breakdown, for which TF copy numbers were greatly suppressed. Interestingly, this trend was consistent with our findings for differentiation time: at the highest miRNA levels, differentiation time spanned a significantly wider range, suggesting a similar breakdown in control.

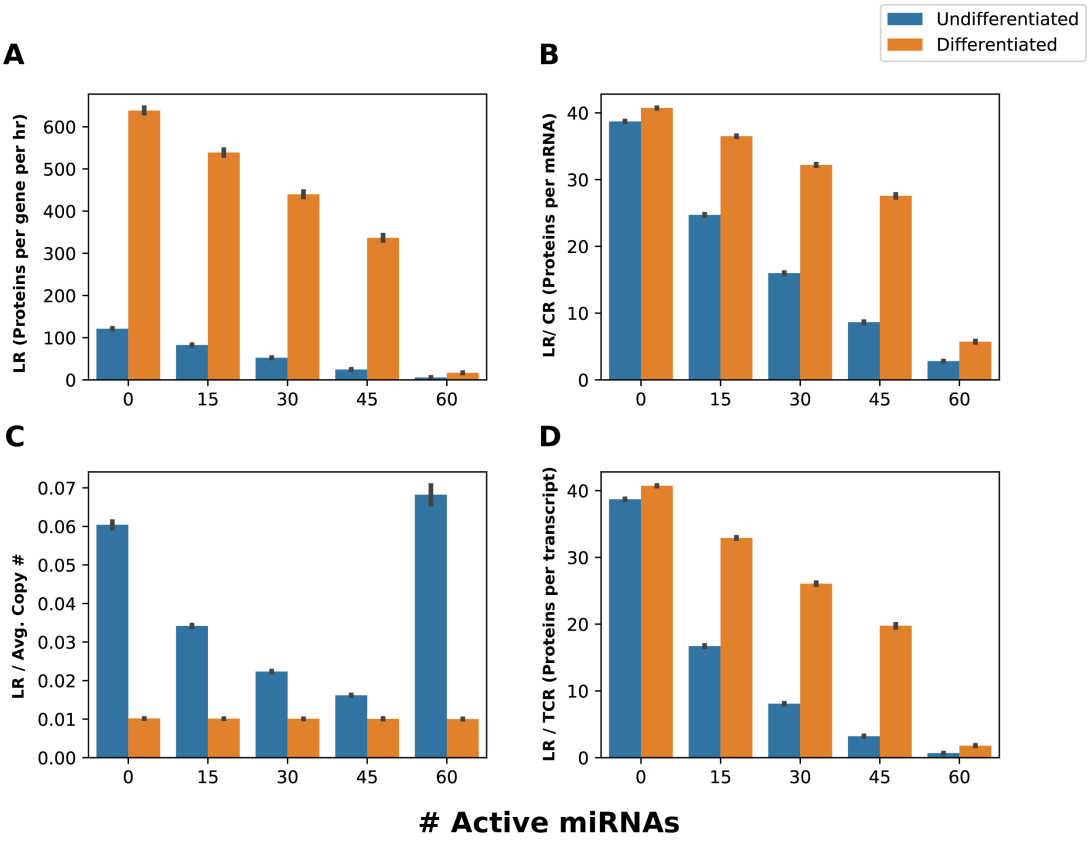


Figure 3: **A**, MiRNA quantity varied directly with average TF copy number in both the pre- and differentiated states. In the differentiated state in particular, the relationship was nearly linear with an approximately 4.5% reduction in average TF copy number per miRNA in the differentiated state. **B**, this reduction was largely caused by a reduction in protein yield per miRNA. This effect was much greater in the undifferentiated state, when both transcript and protein counts were lower. **C**, the ratio of the translation rate to the average TF copy number gives us information about the turnover rate of the cell; up to a limit, the addition of miRNAs decreases this rate, thereby potentially saving resources in the undifferentiated state. **CR**, transCription rate (mRNA only), **TCR**, total transcription rate (mRNA and miRNA), **LR**, transLation rate.

Cellular Efficiency Examining transcription rates and translation rates more carefully (CR and LR, for "transCription" and "transLation," accordingly), we observed an interesting pattern in the ratio of the LR to the average TF copy number in the cell. This ratio gives an estimate of the "turnover" rate for maintenance of ambient TF levels. While miRNAs did not have a noticeable impact on this rate at the higher copy number levels of differentiated cells, they were able to decrease the cost of maintaining copy numbers in the undifferentiated state (Figure 3C). At the higher TF copy numbers observed in differentiated cells, the invariance of this rate with respect to miRNA content indicates that the protein decay rate was primarily responsible for the observed copy numbers, which is in line with experimental results[24]. At the same time, when fewer copies are present, translation rate plays a larger role in maintaining TF copy number. Nonetheless, when sufficient miRNAs were added, ambient copy numbers became sufficiently low so as to completely negate this potential gain in efficiency, again recapitulating the "breakdown" phenomenon observed earlier.

3 Discussion

MiRNAs as feedforward controllers One hallmark of robust gene regulatory networks is the existence of strong feedback controls[25]. While TFs can form such feedback controls without other regulatory entities, miRNAs function as feed-forward controllers, and cannot participate in feedback loops without the assistance of TFs. Classically, feedforward controls can compensate for disturbances more quickly than feedback control, and are particularly useful for mitigating high-frequency disturbances[26]. These

features generally can decrease energy and resource consumption and prevent hysteresis. On the other hand, such controls require previous knowledge of the system behavior, and do not generate new long-term stability in previously unstable systems[27]. Our findings show that miRNA-mediated control provides these benefits to transcription-translation GRNS. At biologically relevant levels, miRNAs can delay the onset of strong TF-mediated feedback controls, but not typically abort them, and exert stronger effects when systems are far away from their set points. In the case of miRNAs, which primarily function as negative feedforward controllers, these effects are strongest when the cellular system is below its set point, as is commonly observed in undifferentiated cell types[28].

Early pattern formation It has been found that a subset of transcription factors in the human genome have an unusually low number of miRNA target sites; these so called "anti-targets" appear to correspond to TFs associated with mor "basic" cellular functions, such as matrix formation, metabolism, and limited cell division[29;30]. At the same time, tissue-specific TFs appear to be more strongly regulated by miRNAs. Combined with our results, such a finding suggests that miRNAs can serve to delay *differentiation* without interfering with basic processes. In theory, this effect could suspend differentiation until the developing embryo has formed a sufficient cell mass and produced sufficiently strong chemical and mechanical gradients. Such an idea is appealing because it could provide an important context for the role of miRNAs. For tissue-based organisms, this function could help to better synchronize differentiation across the embryonic mass, permitting the coordinated formation of mutually dependent tissues. At the same time, one could imagine that making this early growth state too robust could lead to overgrowth of undifferentiated or improperly differentiated cell mass, which could be harmful to the developed organism. Our model shows that MiRNAs appear to be capable of extending the duration of this undifferentiated phenotype without conferring absolute robustness to it; because of their inability to exert feedback control, such a role may be uniquely suited for miRNAs as opposed to TFs to guarantee that the undifferentiated state ends in a timely manner. Furthermore, our results appear to be independent of precise genome size and construction, suggesting that this role may be applicable across a variety of organisms which all depend on miRNAs in one way or another for proper development. Finally, this delay effect appears to depend on having "enough" miRNAs as opposed to a specific, tailored subset, which may help to explain why many organisms often display normal phenotypes in the face of large variations in miRNA quantity and content.

Implications for embryonic fitness The number of miRNAs in the cell reduces the quantity of translatable mRNAs, which is bounded below by zero. As such, genes with few transcripts may appear to have none at all from the perspective of the translational apparatus. At the same time, the number of RNA-induced silencing complexes (RISCs) is limited and relatively constant across both stem and primary cell lines (Figure 7). This upper bound on the effective number of active miRNAs in a cell suggests that the attenuation of translatable elements saturates as mRNA copy numbers increase, thereby creating a one-way switch. On the single-gene level, such a role has been suggested[31;32], and our results support that finding on the genome scale. This structure stands in contrast to TF-DNA mediated gene regulation, which depends primarily on TF copy number (as the gene copy number is effectively constant).

These unique regulatory characteristics can help explain the significance of miRNA-mediated gene regulation in the developing embryo. In the embryonic stem cell, more genes are transcriptionally active, albeit at lower levels[33]. During differentiation, it appears that the embryo begins to restrict its transcriptional repertoire while focusing on a smaller principal set of genes to transcribe (see Figures 5 and 6). In this context, existing miRNAs could serve to "clean up" the translation of these incidental transcripts, which might otherwise lead the cell down an undesirable premature differentiation pathway. At the same time, suppressing these incidental transcripts would save the embryo from the expensive (both energy- and material-wise) process of translation, conserving its limited resources. In this way, miRNAs might help to explain the fundamental conflict of the "undifferentiated phenotype": despite being an observable, classifiable cell type, one of its fundamental properties is that it is *not* stable over time. To maintain this unstable state, more robust regulatory mechanisms that include TF-DNA interactions, DNA methylation, and chromatin formation may not be suitable. Instead, miRNAs could provide a "weak barrier" to protect the transitional phenotype.

4 Conclusions

In the context of the developing embryo, having a robust "differentiation clock" is critical. The formation of the morula and early blastula (week 1 post-fertilization) depend on cells maintaining an undifferentiated state without an external regulatory mechanism. During this time, cell divisions can occur and chemical gradients can be strongly established, setting the stage for the extremely delicate sequence of events as axes are established and organogenesis begins. Given the highly stochastic nature of gene expression, particularly in mRNA levels[34;35], a simple, intrinsic mechanism of "tamping down" accidental or incidental transcripts would be of particular use during this stage. As we have demonstrated, miRNAs appear to be excellent candidates for this role: they are capable of preventing the propagation of potentially harmful stochastic fluctuations while having a relatively modest effect on terminal phenotype. Although they may increase gene expression noise following differentiation, this expense may be tolerable when the organism is developed and self-sufficient.

In this work, we have used stochastic modeling to derive useful predictions regarding the role of miRNAs in tissue-based organisms. Our hypothesized role appears to be compatible with many of the molecular, cellular, and physiological experimental findings, and helps to explain why miRNAs are required specifically in tissue-based organisms.

5 Support

Funding provided by Army Research Office Grant W911NF-14-1-0486.

References

[1] Thibaut Brunet and Nicole King. The Origin of Animal Multicellularity and Cell Differentiation. Developmental Cell, 43(2):124–140, October 2017.

[2] Claire Larroux, Graham N. Luke, Peter Koopman, Daniel S. Rokhsar, Sebastian M. Shimeld, and Bernard M. Degnan. Genesis and expansion of metazoan transcription factor gene classes. Mol. Biol. Evol., 25(5):980–996, May 2008.

[3] Kathryn N. Ivey and Deepak Srivastava. MicroRNAs as Regulators of Differentiation and Cell Fate Decisions. Cell Stem Cell, 7(1):36–41, July 2010.

[4] Eran Hornstein and Noam Shomron. Canalization of development by microRNAs. Nat. Genet., 38 Suppl:S20–24, June 2006.

[5] Jörn M. Schmiedel, Sandy L. Klemm, Yannan Zheng, Apratim Sahay, Nils Blüthgen, Debora S. Marks, and Alexander van Oudenaarden. MicroRNA control of protein expression noise. Science, 348(6230):128–132, April 2015.

[6] Kyle Kai-How Farh, Andrew Grimson, Calvin Jan, Benjamin P. Lewis, Wendy K. Johnston, Lee P. Lim, Christopher B. Burge, and David P. Bartel. The Widespread Impact of Mammalian MicroRNAs on mRNA Repression and Evolution. Science, 310(5755):1817–1821, December 2005.

[7] Karin Lykke-Andersen, Michael J. Gilchrist, Joanna B. Grabarek, Partha Das, Eric Miska, and Magdalena Zernicka-Goetz. Maternal Argonaute 2 Is Essential for Early Mouse Development at the Maternal-Zygotic Transition. Mol. Biol. Cell, 19(10):4383–4392, October 2008.

[8] Erno Wienholds, Marco J. Koudijs, Freek J. M. van Eeden, Edwin Cuppen, and Ronald H. A. Plasterk. The microRNA-producing enzyme Dicer1 is essential for zebrafish development. Nat Genet, 35(3):217–218, November 2003.

[9] Antonio J. Giraldez, Ryan M. Cinalli, Margaret E. Glasner, Anton J. Enright, J. Michael Thomson, Scott Baskerville, Scott M. Hammond, David P. Bartel, and Alexander F. Schier. MicroRNAs Regulate Brain Morphogenesis in Zebrafish. Science, 308(5723):833–838, May 2005.

[10] Joana A. Vidigal and Andrea Ventura. The biological functions of miRNAs: lessons from in vivo studies. Trends in Cell Biology, 25(3):137–147, March 2015.

[11] Robin C. Friedman, Kyle Kai-How Farh, Christopher B. Burge, and David P. Bartel. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res., 19(1):92–105, January 2009.

[12] Paulina Paczynska, Adrian Grzemski, and Maciej Szydlowski. Distribution of miRNA genes in the pig genome. BMC Genet, 16(1), January 2015.

[13] Chunjiang He, Zejuan Li, Ping Chen, Hao Huang, Laurence D. Hurst, and Jianjun Chen. Young intragenic miRNAs are less coexpressed with host genes than old ones: implications of miRNA–host gene coevolution. Nucleic Acids Res, 40(9):4002–4012, May 2012.

[14] Lee P. Lim, Nelson C. Lau, Philip Garrett-Engele, Andrew Grimson, Janell M. Schelter, John Castle, David P. Bartel, Peter S. Linsley, and Jason M. Johnson. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature, 433(7027):769–773, February 2005.

[15] John L. Brenner, Kristen L. Jasiewicz, Alisha F. Fahley, Benedict J. Kemp, and Allison L. Abbott. Loss of individual microRNAs causes mutant phenotypes in sensitized genetic backgrounds in C. elegans. Curr. Biol., 20(14):1321–1325, July 2010.

[16] Natalia Pinzón, Blaise Li, Laura Martinez, Anna Sergeeva, Jessy Presumey, Florence Apparailly, and Hervé Seitz. The number of biologically relevant microRNA targets has been largely overestimated. Genome Res., November 2016.

[17] Jovan Simicevic, Adrien W. Schmid, Paola A. Gilardoni, Benjamin Zoller, Sunil K. Raghav, Irina Krier, Carine Gubelmann, Frédérique Lisacek, Felix Naef, Marc Moniatte, and Bart Deplancke. Absolute quantification of transcription factors during cellular differentiation using multiplexed targeted proteomics. Nat Meth, 10(6):570–576, June 2013.

[18] Hendrik Marks, Tüzer Kalkan, Roberta Menafra, Sergey Denissov, Kenneth Jones, Helmut Hofemeister, Jennifer Nichols, Andrea Kranz, A. Francis Stewart, Austin Smith, and Hendrik G. Stunnenberg. The Transcriptional and Epigenomic Foundations of Ground State Pluripotency. Cell, 149(3):590–604, April 2012.

[19] Eric Londin, Phillipe Loher, Aristeidis G. Telonis, Kevin Quann, Peter Clark, Yi Jing, Eleftheria Hatzimichael, Yohei Kirino, Shozo Honda, Michelle Lally, Bharat Ramratnam, Clay E. S. Comstock, Karen E. Knudsen, Leonard Gomella, George L. Spaeth, Lisa Hark, L. Jay Katz, Agnieszka Witkiewicz, Abdolmohamad Rostami, Sergio A. Jimenez, Michael A. Hollingsworth, Jen Jen Yeh, Chad A. Shaw, Steven E. McKenzie, Paul Bray, Peter T. Nelson, Simona Zupo, Katrien Van Roosbroeck, Michael J. Keating, George A. Calin, Charles Yeo, Masaya Jimbo, Joseph Cozzitorto, Jonathan R. Brody, Kathleen Delgrosso, John S. Mattick, Paolo Fortina, and Isidore Rigoutsos. Analysis of 13 cell types reveals evidence for the expression of numerous novel primate- and tissue-specific microRNAs. Proc. Natl. Acad. Sci. U.S.A., 112(10):E1106–1115, March 2015.

[20] Ana Kozomara and Sam Griffiths-Jones. miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res, 42(D1):D68–D73, January 2014.

[21] Antonio J. Giraldez, Ryan M. Cinalli, Margaret E. Glasner, Anton J. Enright, J. Michael Thomson, Scott Baskerville, Scott M. Hammond, David P. Bartel, and Alexander F. Schier. MicroRNAs Regulate Brain Morphogenesis in Zebrafish. Science, 308(5723):833–838, May 2005.

[22] Collin Melton, Robert L. Judson, and Robert Blelloch. Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. Nature, 463(7281):621–626, February 2010.

[23] Daehyun Baek, Judit Villén, Chanseok Shin, Fernando D. Camargo, Steven P. Gygi, and David P. Bartel. The impact of microRNAs on protein output. Nature, 455(7209):64–71, September 2008.

[24] Björn Schwanhäusser, Dorothea Busse, Na Li, Gunnar Dittmar, Johannes Schuchhardt, Jana Wolf, Wei Chen, and Matthias Selbach. Global quantification of mammalian gene expression control. Nature, 473(7347):337–342, May 2011.

[25] Eduardo Sontag, Alan Veliz-Cuba, Reinhard Laubenbacher, and Abdul Salam Jarrah. The Effect of Negative Feedback Loops on the Dynamics of Boolean Networks. Biophysical Journal, 95(2):518–526, July 2008.

[26] S. Skogestad and I. Postlethwaite. Multivariable Feedback Control: Analysis and Design. Section 5.15: Controllability analysis with feedforward control. Wiley, 2005.

[27] Morten Hovd and Robert R. Bitmead. Feedforward for stabilization. IFAC Proceedings Volumes, 42(11):602–606, January 2009.

[28] M. Hu, D. Krause, M. Greaves, S. Sharkis, M. Dexter, C. Heyworth, and T. Enver. Multilineage gene expression precedes commitment in the hemopoietic system. Genes Dev., 11(6):774–785, March 1997.

[29] David P. Bartel. Micrornas: Target recognition and regulatory functions. Cell, 136.

[30] Alexander Stark, Julius Brennecke, Natascha Bushati, Robert B. Russell, and Stephen M. Cohen. Animal microrna confer robustness to gene expression and have a significant impact on 3 utr evolution. Cell, 123(6):1133 – 1146, 2005.

[31] Fei Shi, Wenbao Yu, and Xia Wang. Bistable switch in let-7 miRNA biogenesis pathway involving Lin28. Int J Mol Sci, 15(10):19119–19133, October 2014.

[32] Shankar Mukherji, Margaret S. Ebert, Grace X. Zheng, John S. Tsang, Phillip A. Sharp, and Alexander van Oudenaarden. MicroRNAs can generate thresholds in target gene expression. Nat Genet, 43(9):854–859, August 2011.

[33] Sol Efroni, Radharani Duttagupta, Jill Cheng, Hesam Dehghani, Daniel J. Hoeppner, Chandravanu Dash, David P. Bazett-Jones, Stuart Le Grice, Ronald D. G. McKay, Kenneth H. Buetow, Thomas R. Gingeras, Tom Misteli, and Eran Meshorer. Global transcription in pluripotent embryonic stem cells. Cell Stem Cell, 2(5):437–447, May 2008.

[34] Edward Yang, Erik van Nimwegen, Mihaela Zavolan, Nikolaus Rajewsky, Mark Schroeder, Marcelo Magnasco, and James E. Darnell. Decay Rates of Human mRNAs: Correlation With Functional Characteristics and Sequence Attributes. Genome Res., 13(8):1863–1872, August 2003.

[35] Björn Schwanhäusser, Dorothea Busse, Na Li, Gunnar Dittmar, Johannes Schuchhardt, Jana Wolf, Wei Chen, and Matthias Selbach. Global quantification of mammalian gene expression control. Nature, 473(7347):337–342, May 2011.

[36] Matthew Slattery, Tianyin Zhou, Lin Yang, Ana Carolina Dantas Machado, Raluca Gordân, and Remo Rohs. Absence of a simple code: how transcription factors read the genome. Trends Biochem Sci, 39(9):381–399, September 2014.

[37] José M. Franco-Zorrilla, Irene López-Vidriero, José L. Carrasco, Marta Godoy, Pablo Vera, and Roberto Solano. DNA-binding specificities of plant transcription factors and their potential to define target genes. PNAS, 111(6):2367–2372, February 2014.

[38] Stefan Ludwig Ameres, Javier Martinez, and Renée Schroeder. Molecular basis for target RNA recognition and cleavage by human RISC. Cell, 130(1):101–112, July 2007.

[39] Hervé Seitz. Issues in current microRNA target identification methods. RNA Biology, 14(7):831–834, July 2017.

[40] Smadar Ben-Tabou de Leon and Eric H. Davidson. Modeling the dynamics of transcriptional gene regulatory networks for animal development. Developmental Biology, 325(2):317–328, January 2009.

[41] Jasper Landman, Robert C. Brewster, Franz M. Weinert, Rob Phillips, and Willem K. Kegel. Self-consistent theory of transcriptional control in complex regulatory architectures. PLoS ONE, 12(7):e0179235, 2017.

[42] Yiming Zhou, John Ferguson, Joseph T. Chang, and Yuval Kluger. Inter- and intra-combinatorial regulation by transcription factors and microRNAs. BMC Genomics, 8:396, October 2007.

[43] Marco Antonio Valencia-Sanchez, Jidong Liu, Gregory J. Hannon, and Roy Parker. Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev., 20(5):515–524, March 2006.

[44] Anthony Mathelier, Xiaobei Zhao, Allen W. Zhang, François Parcy, Rebecca Worsley-Hunt, David J. Arenillas, Sorana Buchman, Chih-yu Chen, Alice Chou, Hans Ienasescu, Jonathan Lim, Casper Shyr, Ge Tan, Michelle Zhou, Boris Lenhard, Albin Sandelin, and Wyeth W. Wasserman. JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Res., 42(Database issue):D142–147, January 2014.

[45] V. Matys, O. V. Kel-Margoulis, E. Fricke, I. Liebich, S. Land, A. Barre-Dirrie, I. Reuter, D. Chekmenev, M. Krull, K. Hornischer, N. Voss, P. Stegmaier, B. Lewicki-Potapov, H. Saxel, A. E. Kel, and E. Wingender. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res., 34(Database issue):D108–110, January 2006.

[46] The ENCODE Project Consortium. A User's Guide to the Encyclopedia of DNA Elements (ENCODE). PLOS Biology, 9(4):e1001046, April 2011.

[47] Heonjong Han, Hongseok Shim, Donghyun Shin, Jung Eun Shim, Yunhee Ko, Junha Shin, Hanhae Kim, Ara Cho, Eiru Kim, Tak Lee, Hyojin Kim, Kyungsoo Kim, Sunmo Yang, Dasom Bae, Ayoung Yun, Sunphil Kim, Chan Yeong Kim, Hyeon Jin Cho, Byunghee Kang, Susie Shin, and Insuk Lee. TRRUST: a reference database of human transcriptional regulatory interactions. Sci Rep, 5:11432, June 2015.

[48] Shane Neph, Andrew B. Stergachis, Alex Reynolds, Richard Sandstrom, Elhanan Borenstein, and John A. Stamatoyannopoulos. Circuitry and dynamics of human transcription factor regulatory networks. Cell, 150(6):1274–1286, September 2012.

[49] Spyros Darmanis, Caroline Julie Gallant, Voichita Dana Marinescu, Mia Niklasson, Anna Segerman, Georgios Flamourakis, Simon Fredriksson, Erika Assarsson, Martin Lundberg, Sven Nelander, Bengt Westermark, and Ulf Landegren. Simultaneous Multiplexed Measurement of RNA and Proteins in Single Cells. Cell Rep, 14(2):380–389, December 2015.

[50] Fredrik Edfors, Frida Danielsson, Björn M Hallström, Lukas Käll, Emma Lundberg, Fredrik Pontén, Björn Forsström, and Mathias Uhlén. Gene-specific correlation of RNA and protein levels in human cells and tissues. Mol Syst Biol, 12(10), October 2016.

[51] Tobias Maier, Marc Güell, and Luis Serrano. Correlation of mRNA and protein in complex biological samples. FEBS Letters, 583(24):3966–3973, December 2009.

[52] Ute Bissels, Stefan Wild, Stefan Tomiuk, Angela Holste, Markus Hafner, Thomas Tuschl, and Andreas Bosio. Absolute quantification of microRNAs by using a universal reference. RNA, 15(12):2375–2384, December 2009.

[53] Benjamin Lewin. Genes VII. Oxford University Press, Oxford ; New York, 7th ed. edition edition, December 1999.

6 Methods

Model - molecular level We constructed a stochastic model of gene transcription and translation to help identify the interplay between TF/DNA-mediated and miRNA/mRNA-mediated gene regulation. In the model, genes, which have a fixed copy number, transcribe either messenger- or microRNAs as a simple Poisson process with a base rate α_i for gene i. Each gene has its own set of TF binding sites, which have an (amplifying or repressing) multiplier effect on the transcription rate when bound to a gene. At the same time, each mRNA has a corresponding base rate of translation, β_i, and has an exponentially distributed decay time with rate μ_i. Like genes, each mRNA has a certain set of miRNA-binding sites, which have a strong repressive multiplier effect on translation rate. Our model permits accelerated decay of mRNAs by miRNAs, although our results did not depend on this mode of action. A schematic overview of our model is given in Figure 1B, while reactions and parameters are given in Table 1.

Our model is based on the following assumptions (with justifications where applicable):

1. The genome is immutable over the time scale of our simulation (hours-weeks). Because both TFs and miRNAs typically target enhancers/promoters and 3'UTRs based on a preferred sequence[36–39], we consider each gene to have a predetermined number of binding sites for a subset of TFs at the DNA level, and miRNAs at the mRNA level. Binding site associations and dissociations (γ, δ, d, a in Figure 1B) represent the strength of these binding sites.

Table 1: Summary of reactions and their corresponding propensities in the stochastic model of transcription and translation.

DNA (Gene) G

Reaction	Rate
$G(\mathbf{o}) + P_j \rightarrow G(\mathbf{o} + e_j), j \in F$	$\gamma_j P_j$
$G(\mathbf{o}) \rightarrow G(\mathbf{o} - e_i) + P_i, i \in B$	δ_i
$G(\mathbf{o}) \rightarrow G(\mathbf{o}) + RNA$	$\beta \prod_{i \in B} \lambda_i$

mRNA R

Reaction	Rate
$R(\omega) + r_j \rightarrow R(\omega + e_j), j \in F$	$a_j R r_j$
$R(\omega) \rightarrow R(\omega - e_i) + r_i, i \in B$	$d_i \sum_{i=1}^{\lvert S \rvert} \omega_i$
$R(\omega) \rightarrow R(\omega) + P$	$\alpha \prod_{i \in B} \epsilon_i$
$R(\omega) \rightarrow \emptyset$	μR

Protein P & miRNA r

Reaction	Rate
$P \rightarrow \emptyset$	δ
$r \rightarrow \emptyset$	μ

2. TFs are capable of both activation and repression, while miRNAs are only capable of repression. At the extremes, TFs should be able to either completely abolish transcription of a given gene, or multiply it up to some maximal constant[17] - similarly, miRNAs should be capable of completely abolishing translation of an mRNA when bound to it.

3. miRNAs are capable of acting by either sequestration or accelerating degradation. There is still much controversy over the balance of these two effects, so our model permits both. The results presented assumed that when an mRNA is degraded, all bound miRNAs also degrade, though the general results did not appear to depend on this particular assumption.

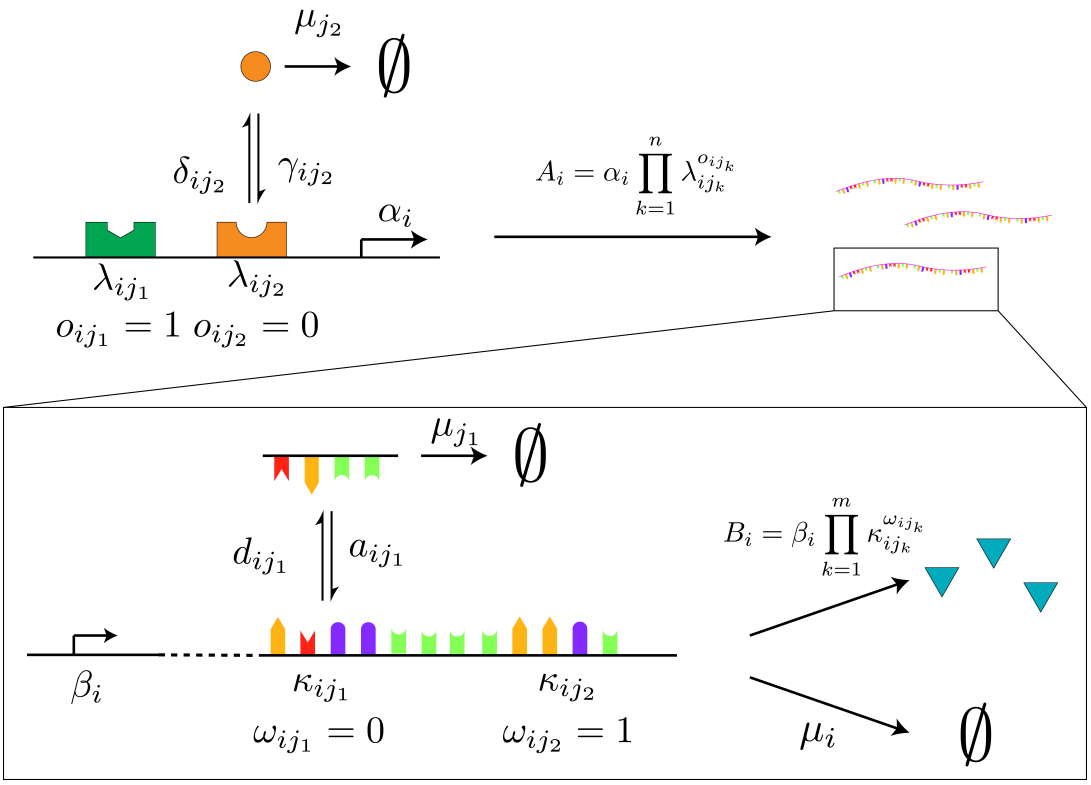
4. For our study, we wanted to focus on regulators which were capable of selectively modulating a subset of genes and not others; that is, we wanted to focus on the "steering" rather than the "throttle/brake" of gene regulation. On the time scale of cell differentiation (days-weeks), TFs and miRNAs serve as the primary *specific* regulators of transcription and translation, respectively. Studies have shown that TFs are the main specific regulators of transcription on this scale[40;41], while miRNAs serve as the dominant *specific* regulators of translation[42;43]. As such, our model does not include other (no less important) forms of gene regulation.

6.0.1 Implementation

Our model is a hybrid agent-based/population-based model, in which the genes (DNA) and mRNAs are treated as individual agents, while miRNAs and TFs are considered as a pool of identical molecules. It is implemented using a modified Gibson-Bruck algorithm and can simulate large numbers of events ($> 10^{10}$) using "genome-scale" GRNs (up to 10^5 protein-coding genes and miRNAs) on a Macbook Pro 2013 in under 1 hour. The system design allows for a very high level of specification and customization, and allows the import and use of "named" genes and miRNAs with individual kinetic parameters.

Human GRN Analysis and Fitting Naturally, the structure of the "true" human GRN is unknown. In order to derive sufficiently general results, we performed a careful bioinformatic analysis of available human genomic data in order to identify likely general characteristics of the human GRN. These include connectivity data for TFs and miRNAs[44–48] (see Supplementary Methods), both absolute and relative abundance data for protein and RNA abundances[35;49–52], and approximate synthesis and decay rates. Using these data, we identified suitable parameters to randomly generate "human genome-like" genomes; by examining phenomena on a wide range of such GRNs, we believe that it is reasonable that gross qualitative phenomena are likely applicable to the "true" human GRN.

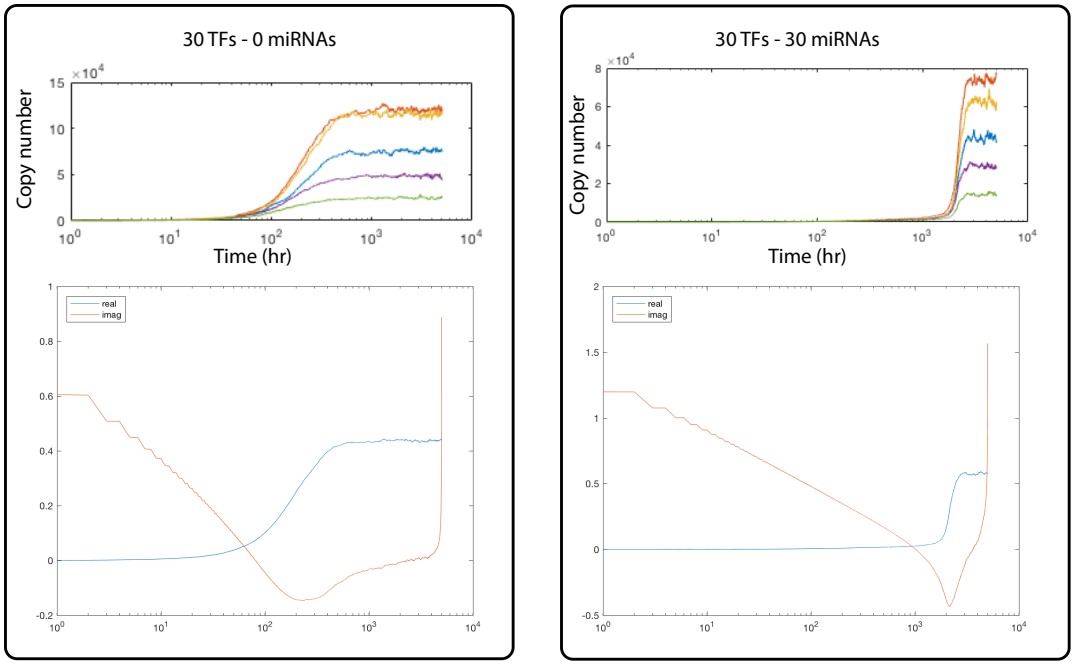
Initial conditions We defined differentiation abstractly as the time after which the state of the cell (determined by protein copy numbers) fundamentally "shifts" to a new equilibrium state from its initial



Supplementary Figure 1: Formulation of stochastic transcription-translation model. **A**, a small example network with 2 TFs and 1 miRNA with activation and inhibition (*blue, red*) at the transcriptional and translational levels. **B**, detailed model with associated rate laws. A_i, B_i correspond to transcription and translation rates for gene i, while the Boolean vectors $\boldsymbol{o}$, $\boldsymbol{\omega}$ correspond to the occupancies on the promoter and 3'UTR of the DNA and mRNA by TF and miRNA, respectively. Inset, dynamics for a single mRNA strand. For transcription, the rate multiplier λ can be greater than or less than 1, corresponding to activation or inhibition. All rate multipliers κ are strictly less than 1.

# TFs	# miRNAs (baseline)	# of networks tested
30	20	100
50	35	100
100	66	100
150	100	1000
500	350	1000
1500	1000	100

Table 2: Number of networks generated and simulated at each size. Baseline miRNAs represent the total pool of *potential* active miRNAs in system. For each simulation run, a random subset of this pool was activated and tested.

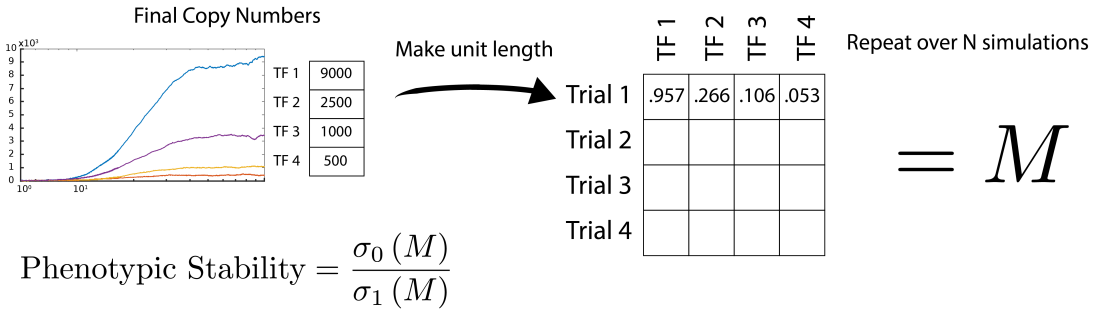


Supplementary Figure 2: Trace examples (upper, 5 randomly chosen TFs) shows that # of miRNAs can produce a visible delay in "time-to-differentiation" - however, one needs an objective, computational approach. Note that at some point - approximately 200 hr in the 0 miRNA case and 2000 hr in the 30 miRNA case, the traces, when shifted down, resemble odd curves. After scaling each trace by its 2-norm, summing the scaled values, and applying the Hilbert transform (lower), "oddness" corresponds to a minimum of the imaginary part (lower, red) of the discrete HT, which we used to identify differentiation time.

state. In general, we only required that our initial copy numbers be set at a level below their equilibrium values, and our results were robust to changes in initial conditions. However, to be able to more accurately determine differentiation time, we used zero initial conditions for our simulations.

Simulation Process The number of GRNs generated and tested at each network size is given in Table In order to balance computational time with the need to run a large number of simulations with a large number of networks, the main results presented were derived from simulations of randomly generated GRNs with 150 TF-coding and 100 miRNA-coding genes (not all necessarily active at once). For each such network, we ran 200 simulations each with either 0, 15, 30, 45, or 60 active miRNAs. With the exception of the miRNA-absent simulations, the set of active miRNAs was randomly selected from the pool of 100 pre-generated miRNAs: the remainder were inactivated. A schematic of the active miRNA selection process can be found in Figure S 4.

Identification of time-to-differentiation Although we used a relatively abstract definition of differentiation, sample simulation trajectories show that one can clearly see that these two phases exist Figure 2. To measure the time-to-differentiation, we use the Hilbert transform (HT) of the trajectories.



$$\text{Phenotypic Stability} = \frac{\sigma_0(M)}{\sigma_1(M)}$$

Supplementary Figure 3: Example calculation for phenotype stability with 4 TFs and 4 simulation trials. For each simulated network at each level of active miRNAs, final TF copy numbers were used to evaluate phenotype stability. After normalizing each copy number vector and constructing a final state matrix M. The stability was estimated by the ratio of the two largest singular values of M.

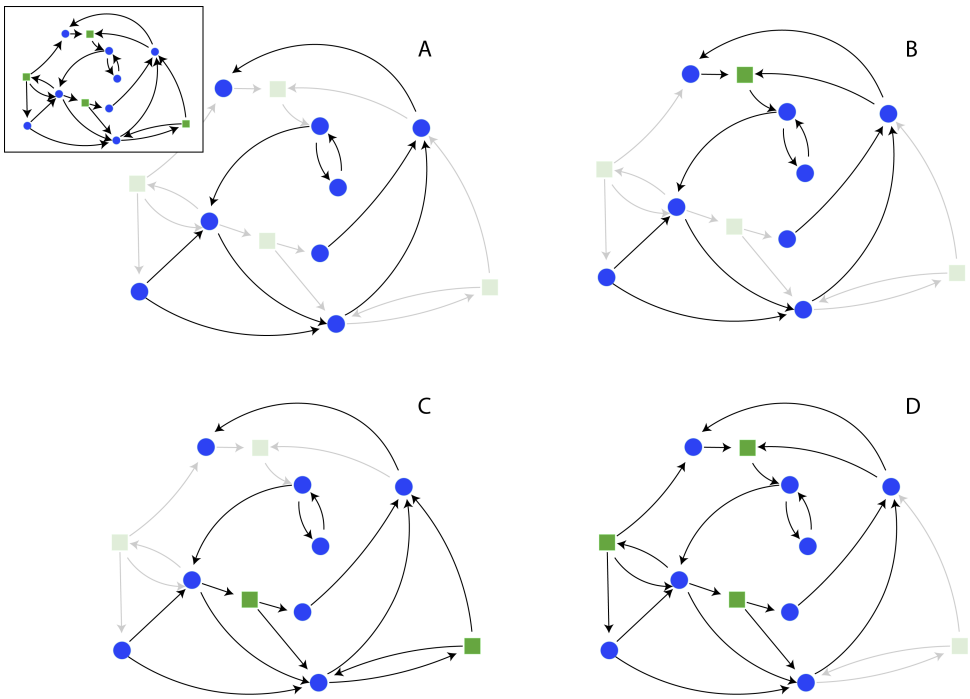
The HT is an invertible linear transformation on real-valued L^2 functions on the real line. For a function $f : \mathbb{R} \to \mathbb{R}$, $\mathcal{H}(f)(x) = g(x)$, where $F = f + ig$ is holomorphic; that is, the HT maps a function to its harmonic conjugate. The Cauchy-Riemann equations then show that an inflection point in the original function corresponds to a root of the derivative of the harmonic conjugate.

The closely-related discrete HT has both real and imaginary parts - for a real-valued signal, the real part generates an envelope for the original signal, while the imaginary part gives the "actual" HT. Use of this transformation is especially useful for a stochastic (real-valued) signal, as it reconstructs the signal by removing certain high-frequency oscillations. The calculation of time-to-differentiation is given by the following steps:

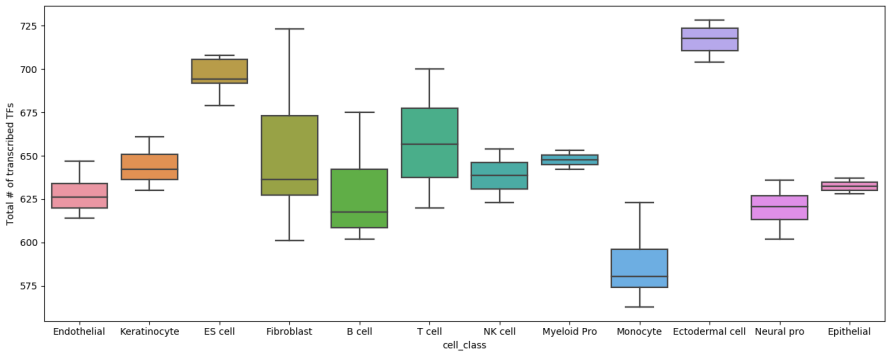
1. Scale each copy number trajectory by its ℓ^2 norm.
2. Sum the values of the scaled trajectories.
3. Apply the HT to the summed trajectory.
4. Find the time when the imaginary part of the HT is minimized.

Determination of phenotype stability In this study, we used post-differentiated TF copy numbers to determine phenotype stability. We considered a network, with its corresponding differentiated state, to be stable when identical initial conditions produced nearly identical post-differentiation TF copy numbers. For a series of simulations with a fixed number of active miRNAs, we took the vector of final TF copy numbers and made each unit length. Those vectors were stacked to form a final #trials $\times$ #TFs matrix M (Figure S 3).

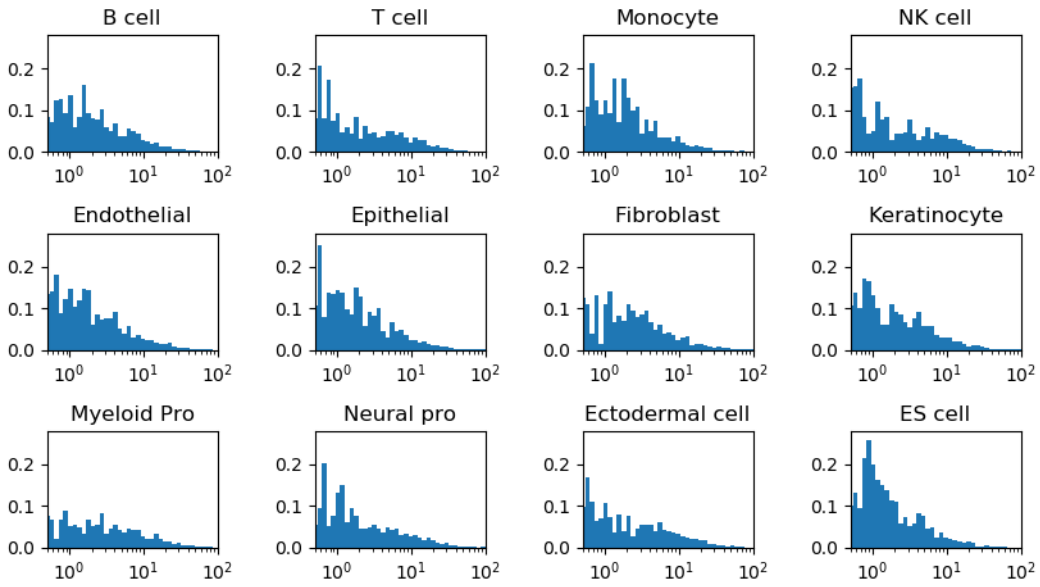
For a stable network, we would expect M to be of very low rank. The coefficient of phenotype stability was given by $\frac{\sigma_0(M)}{\sigma_1(M)}$, where σ_0, σ_1 correspond to the largest and second-largest singular values of M. For this analysis, the number of trials at each miRNA level was fixed at 200, and the number of TFs was fixed at 150. Fixing the size, as well as normalizing the copy number vectors, ensures that for a single fixed network structure, the stabilities for different numbers of active miRNAs are directly comparable.



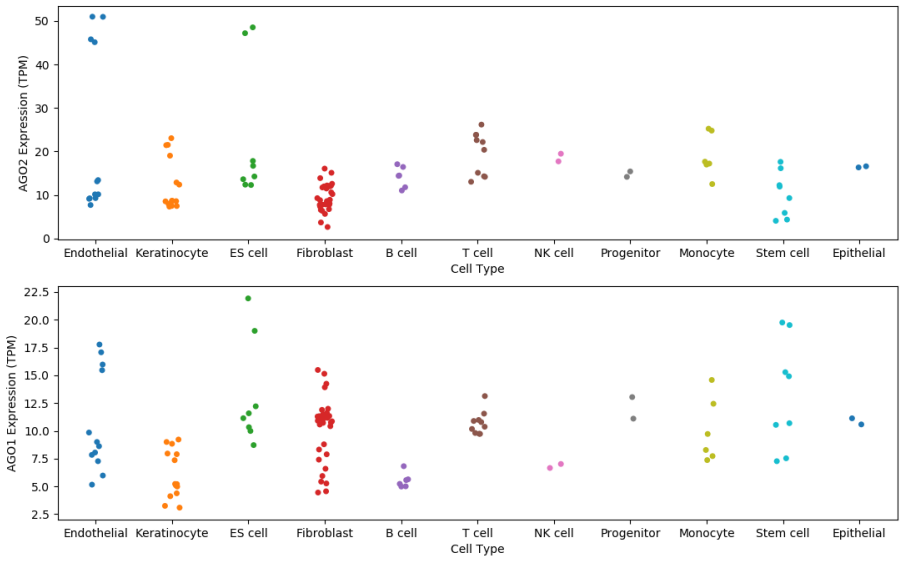
Supplementary Figure 4: Example of simulation process. **A**, baseline phenotypes were determined by deactivating all miRNAs (green squares) while keeping all TFs (blue circles) active, inactive elements shown greyed out. For each simulation for a given quantity of active miRNAs, the appropriate number of miRNAs were chosen randomly from the network and activated - in this example, we have **B**, 25%, **C**, 50%, **D**, 75% active when compared to original network (inset).



Supplementary Figure 5: Number of TF genes expressed at ≥ 1 transcript per million (TPM) across different cell types shows that stem cells express many more TFs than other, more differentiated cell types. Below, although the number of different genes transcribed is much higher, TPM counts by gene are much lower.



Supplementary Figure 6: Distribution of transcripts per gene (measured as transcripts per million, TPM) by cell type from ENCODE data. Despite the high number of TFs expressed by ES cells (lower right), one can see that more than 95% of TFs are expressed at fewer than 5 TPM.



Supplementary Figure 7: Expression of argonaute genes remains relatively consistent across cell types. The argonaute proteins, particularly AGO2, are critical for formation of the RNA-induced silencing complex (RISC). Their relative constancy across cell types suggests that the number of formed RISC complexes likely saturates at a number well below the number of ribosomes in a typical mammalian cell $(10^6 - 10^7)$ [53].

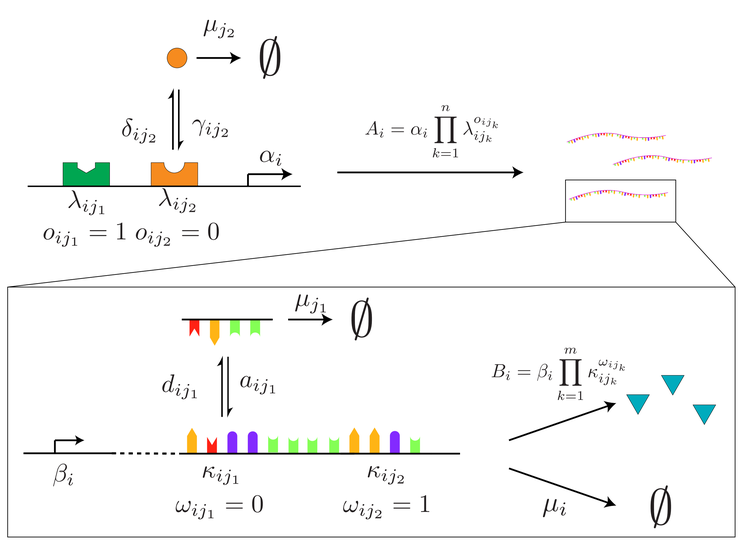
Control of differentiation time by microRNA-mediated regulation at the genome scale

Russell Posner, Reinhard Laubenbacher Ph.D.
Center for Quantitative Medicine, UConn Health, Farmington, CT

Motivation

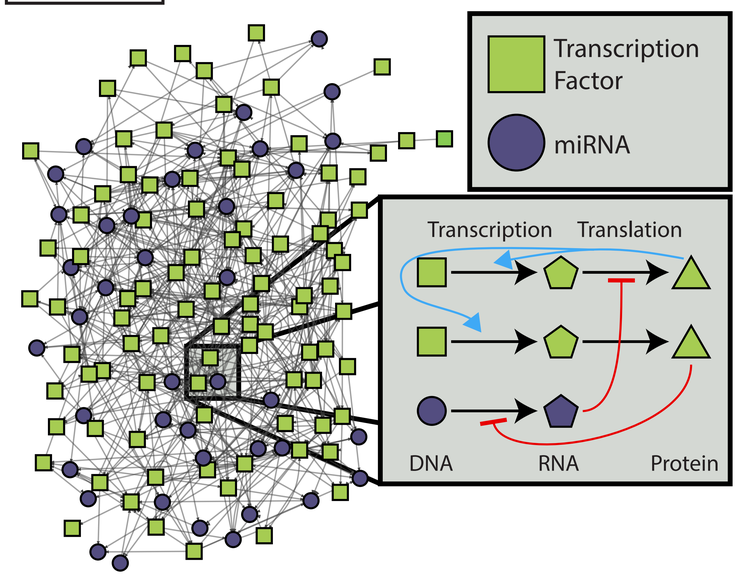
The evolution of tissue-based organisms remains one of the most enigmatic and fascinating questions in biology. One commonality of nearly all such organisms is their dependence on microRNAs (miRNAs), for proper cell differentiation and development[1]. While it is known that miRNAs function generally as repressors of protein translation, numerous experiments have shown that large perturbations of the miRNA regulatory apparatus produce no noticeable change in phenotype[2]. In this work, we use mathematical modeling and simulation to explore how miRNA-mediated regulation can influence cell differentiation.

Model Design

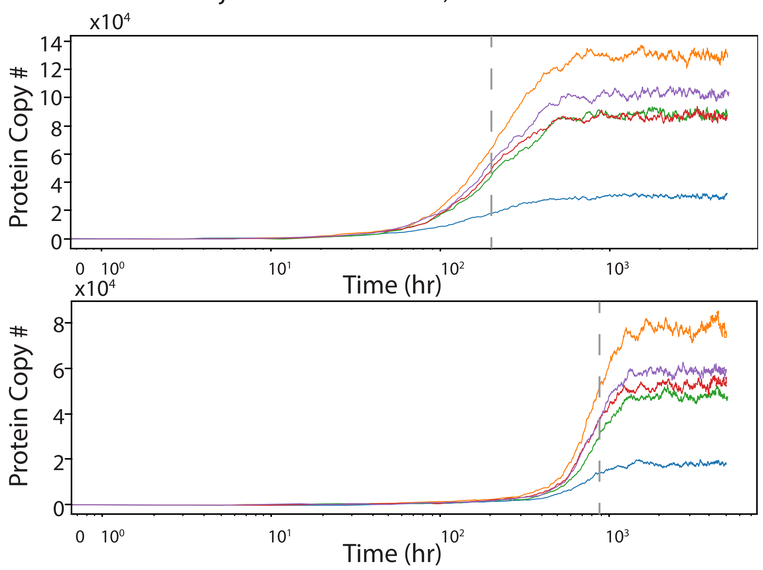


- Designed a stochastic model of transcription and translation
- Genes and mRNAs are set with prescribed rates of association/ dissociation with TF/miRNA, respectively.
- Effects of bound factors on synthesis rates are multiplicative
- All components except DNA undergo stochastic decay

Methods



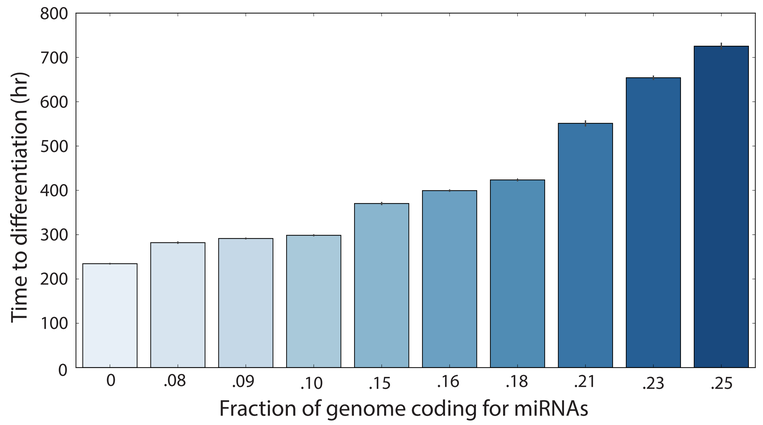
Build approximate representations of human gene regulatory networks from regulatory datasets (depicted, subnetwork derived from ENCODE Project and miRTarBase).



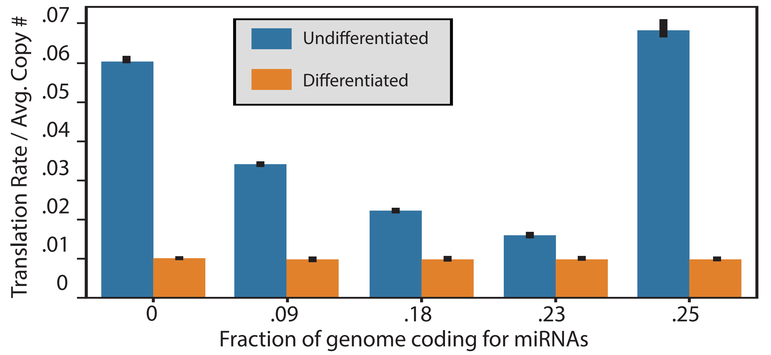
Example simulation traces of protein copy number for 5 TFs in a 150 TF network with miRNAs absent (top) and 15 miRNAs (bottom).

Results

- Addition of miRNAs increase time to differentiation independent of specific network structure and scale



- Magnitude of differentiation delay is proportional to miRNA density in genome



- MiRNAs can improve energy and material consumption by reducing translational burden in pre-differentiated period

Acknowledgments

Pedro Mendes Ph.D., Paola Vera-Licona Ph.D.
Army Research Office Grant No. W911NF-14-1- 0486 t

References

1. Wienholds, E., Koudijs, M. J., et al. (2003). The microRNA-producing enzyme Dicer1 is essential for zebrafish development. Nature Genetics, 35(3), 217–218.
2. Vidigal, J. A., & Ventura, A. (2015). The biological functions of miRNAs: lessons from in vivo studies. Trends in Cell Biology, 25(3), 137–147.
3. ENCODE Project Consortium. (2004). The ENCODE (ENCyclopedia Of DNA Elements) Project. Science (New York, N.Y.), 306(5696), 636–640.
4. Chou, C.-H., Chang, N.-W., et al. (2016). miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database. Nucleic Acids Research, 44(D1), D239–D247.