Convergence Analysis and Network Properties of Wagner's Artificial Gene Regulatory Network Model

Original Research Article

Abstract

Wagner's gene regulatory network (GRN) model, which has mathematical roots origin from the Ising model and neural networks, is a powerful computation tool that helps integrate network thinking into biology, and motivated a new research theme focusing on the evolution of genetic networks. In this paper, I provide the convergence analysis for Wagner's GRN model by using the Markov chain theory. I show mathematically that if we consider the evolution process as an optimisation process, then the probability of finding the optimal configuration (a certain target phenotype) converges to probability one. Next, I investigate network characteristics such as stability, robustness and path length in initial populations. I find that generally small networks with a sparse connectivity have a higher initial stability. The robustness is also observed to be higher in initial stable networks with a low network connectivity. These results are partly explained by the pattern, as shown in this paper, that small networks with a sparse connectivity generally have a shorter path length and, therefore, they are not only able to quickly reach equilibrium phenotypic states but also more likely to resist perturbations.

Keywords: Gene Regulatory Networks; Robustness; Systems Biology; Optimization; Evolutionary Dynamics

1 Introduction

Understanding the history of life is to uncover mechanisms underlying the evolution of innovation at different life scales, ranging from molecular to cellular to tissue and organ levels [1]. One of the

most important forms of innovation can be attained through regulation, which refers to a process that controls a gene product at a particular time and place [2]. In particular, the transcriptional regulation, which is mediated by the binding of proteins to specific DNA sequences or cis-regulatory elements, is essential to life as it governs all levels of gene outcomes that enable cell survival and numerous cellular functions [3].

However, the evolution of transcriptional regulation is extremely difficult to study experimentally. The main reasons, as summarised in [1], are 1) DAN regions where they regulate transcription can span hundreds of kilobase pairs upstream and downstream of a regulated gene, 2) cis-regulatory elements can function regardless of their orientation and distance from a regulated gene, and 3) DNA regions surrounding them often evolve rapidly, not only through changes of individuals nucleotides but through insertions and deletions of large swaths of DNA. These reasons all make substantial challenges to characterise the transcriptional regulation through experiments due to limitations of currently available technologies.

In the past decades, researchers have made tremendous efforts in modelling regulatory networks using computational approaches. Kauffman introduced the basic boolean networks to study the behaviour of large, randomly constructed nets [4, 5]. Shmulevich et al. further developed the probabilistic boolean networks to include global dynamics and cope with uncertainty [6]. Petri nets initially proposed by [7] are used to study large metabolic networks. Friedman et al. introduced Bayesian networks as a probabilistic framework for discovering interactions between genes based on multiple expression measurements [8]. Differential equations are used to study network dynamics by explicitly modelling the concentration/activity changes of molecules over time [9].

In the mid-1990s, Andreas Wagner introduced a novel computational model where he explicitly modelled the developmental process in the system [10, 11]. Wagner's gene regulatory networks (GRN) model, which has mathematical roots origin from the Ising model [12] and neural networks [13] (see a nice review article of [14] on gene network family tree), has helped integrate network thinking into biology, and motivated a new research theme focusing on the evolution of genetic networks (see a recent review in [15]). In this paper, I first summarise the implementation details of Wagner's GRN model. Next, I provide the convergence analysis for Wagner's GRN model by using the Markov chain theory. Finally, I investigate network characteristics such as stability, robustness and path length in initial populations.

2 Wagner's Artificial GRN Model

In Wagner's GRN model,the genotype is presented as a network which contains interactions among transcriptional genes [11, 16, 17, 18, 19, 15]. Formally, for each individual in a finite population of size M, an $N \times N$ matrix W can be considerred as an artificial gene network that contains the regulatory interactions among N genes. Each element $w_{i,j}$ $(i, j = 1, 2, \cdots, N)$ represents the regulatory effect on the expression of gene i of the product of gene j. The connectivity parameter c determines the proportion of non-zero elements in the network W. Through gene interactions the regulatory effect acts on each gene's expression pattern. This can be denoted by a state vector $\mathbf{S}(t) = (s_1(t), s_2(t), \cdots, s_N(t))$ where $s_i(t)$ represents the expression pattern of gene i at time t. Each value of the expression state $s_i(t)$ can be varied continuously between -1 (complete repression) and $+1$ (complete activation). For a given gene regulatory network W, the dynamics of $\mathbf{S}$ for each gene i can be modelled by the following equation

$$s_i(t+1) = f\left(\sum_{j=1}^{N} w_{i,j} s_j(t)\right), \tag{2.1}$$

where $f(x)$ is a sigmoidal function, which is normally defined as in [16, 17]: $f(x) = 2/(1 + e^{-ax}) - 1$, where a is the activation constant determining the rate of change from complete repression to complete activation.

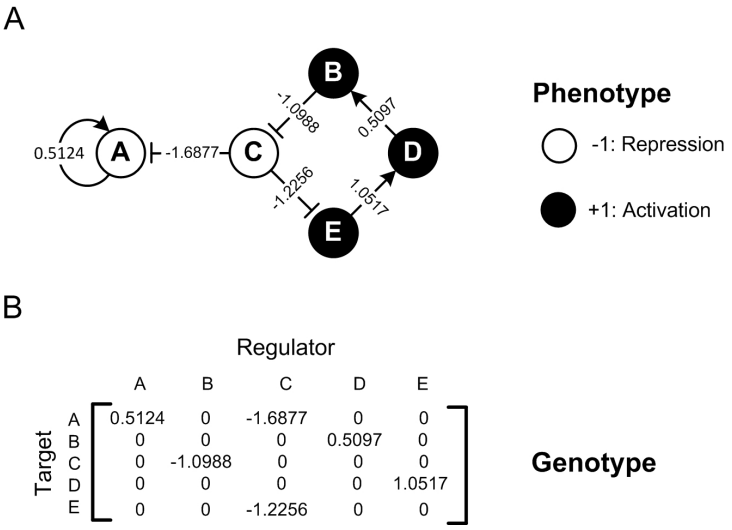


Figure 1: **An example of Wagner's gene regulatory network.** This example is re-used from [15]. (A) Network representation of regulatory interactions among five genes. Open and filled circles represent genes that are completely activation $(+1)$ or repression (-1), respectively. The initial gene expression pattern is $\mathrm{s}(0) = (-1, +1, -1, +1, +1)$. This example network is stable as it can reach an equilibrium pattern, which is $\mathbf{S}_{\mathrm{EQ}} = (-1, -1, +1, +1, +1)$ by iterating Equation (2.1) using the sigmoidal mapping function with $a = 100$. (B) Interaction matrix (W) represents the network in (A). Each element in row i and column j, i.e., w_{ij} $(i, j = 1, 2, \cdots, 5)$, represents the regulatory effect on the expression of gene i of the product of gene j.

A novel feature of Wagner's GRN model is that it introduces the selection for phenotypic stability [15], which is defined as the progression from an arbitrary initial expression state to an equilibrium expression state (reaching a fixed pattern) by iterating Equation (2.1) a fixed number of times, $devT$. If a given network W can achieve stability over this developmental time period, it is termed *stable*, otherwise it is labelled *unstable*. An equilibrium expression state can be reached when the following equation is met:

$$\frac{1}{\tau} \sum_{\theta=t-\tau}^{t} D\left(\mathbf{S}(\theta), \overline{\mathbf{S}}(t)\right) \leq 10^{-4}, \tag{2.2}$$

where $D(\mathbf{S},\overline{\mathbf{S}}) = \sum_{i=1}^{N} \left(s_i - {s'}_i\right)^2 \Big/ 4N$ measures the difference between the gene expression pattern $\mathbf{S}$ and $\overline{\mathbf{S}}$, and $\overline{\mathbf{S}}$ is the average of the gene expression levels over the time interval $[t-\tau, t-\tau+1, \cdots, t]$, where τ is a time-constant characteristic for the developmental process under consideration. An example of Wagner's GRN can be found in Figure 1.

For networks that achieve developmental stability (reaching an equilibrium state, $\mathbf{S}_{\mathrm{EQ}}$), then the fitness can be calculated as [16]:

$$F(\mathbf{S}_{\mathrm{EQ}}) = \exp\left(-\frac{D(\mathbf{S}_{\mathrm{EQ}}, \mathbf{S}_{\mathrm{OPT}})}{\sigma}\right), \tag{2.3}$$

where σ is the selection pressure that we imposed on the population during evolution. $\mathbf{S}_{\mathrm{OPT}}$ is usually set to be the initial state, i.e., $\mathbf{S}(0)$. $D(\mathbf{S}_{\mathrm{EQ}}, \mathbf{S}_{\mathrm{OPT}})$ is the phenotypic distance between the equilibrium state and the optimal state.

In typical evolution of Wagner's GRN model, an individual is chosen at random to reproduce either by cloning itself, if asexually, or by recombining two parent networks, if sexually, and then subject to mutation. Next, the offspring network is subject to the stabilising selection — if the offspring network cannot achieve developmental stability, then it will be wiped out from the population immediately. For the stable offspring network, the survival probability is based on its fitness calculated by using Equation (2.3). This process is repeated until the same size of offspring population as in initial population is formed.

3 Convergence Analysis

In Wagner's GRN model, the evolution process has three operators: mutation, recombination and selection. Therefore, the evolution process to find a target phenotype can be regarded as an optimisation process where the goal is to minimise $D(\mathbf{s}_{\mathrm{EQ}}, \mathbf{s}_{\mathrm{OPT}})$ such that individuals' phenotypic state is close to the optimal phenotypic state. Suppose that the initial population have M individual networks and the search space is N dimensions.

The phenotypes of individual networks at the g^{th} generation can be represented as $\mathbf{S}(g) = [\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_j, \cdots, \mathbf{s}_M]$, where $\mathbf{s}_j = (s_1, s_2, \ldots, s_N)$ is a individual's phenotype at equilibrium in N dimensional solution space. Suppose $\mathbb{S} = \mathbb{R}^N$ be the solution space, $\mathbb{S}^M$ be the population space. Without loss of generality, suppose that the optimisation goal of the evolution process described in Wagner's GRN model is to find the target phenotype, formally defined as: Given f: $\mathbb{S} \to \mathbb{R}$ find $\mathbf{S}^* \in \mathbb{S}$ such that $f(\mathbf{S}^*) \leqslant f(\mathbf{S})$. Here, the objective function can be defined as $D(\mathbf{s}_{\mathrm{EQ}}, \mathbf{s}_{\mathrm{OPT}})$.

The basic operators in Wagner's GRN model are described as follows.

Definition 1. ***Mutation Operator:*** *Mutation operator, $T_M : \mathbb{S}^M \to \mathbb{S}$, operates on non-zero entries in individual's genotype with mutation rate μ and can change individual's phenotype $\mathbf{s}_j$ into $\mathbf{s'}_j$, and can be given as*

$$\mathbf{x}_j = \begin{cases} \mathbf{s'}_j & \textit{if mutate,} \\ \mathbf{s}_j & \textit{otherwise.} \end{cases}$$

Then, its probability distribution is

$$P\{T_M(\mathbf{s}_j) = \mathbf{x}_j\} = \begin{cases} \mu & \mathbf{x}_j = \mathbf{s'}_j \\ 1-\mu & \mathbf{x}_j = \mathbf{s}_j \end{cases}$$

Definition 2. ***Recombination Operator:*** *Recombination operator, $T_R : \mathbb{S}^M \to \mathbb{S}$, operates on all individual networks by segregating rows in genotypes of two randomly selected parental networks and the resulting offspring network can have a new phenotype $\mathbf{s''}_j$, and can be given as*

$$\mathbf{y}_j = \mathbf{s''}_j$$

Then, its probability distribution is

$$P\{T_R(\mathbf{s}_j, \mathbf{x}_j) = \mathbf{y}_j\} = 1$$

Definition 3. ***Selection Operator:*** *Selection operator, $T_S : \mathbb{S}^M \to \mathbb{S}$, operates on the candidate solution by preserving the phenotype with better fitness value, and can be given as*

$$\mathbf{s}_j(g+1) = \begin{cases} \mathbf{y}_j & \textit{if } f(\mathbf{y}_j) \leqslant f(\mathbf{s}_j), \\ \mathbf{s}_j & \textit{otherwise.} \end{cases}$$

Here, I only consider the selection for the target phenotype, i.e., $\mathbf{s}_{\mathrm{OPT}}$, as the main selection operator.

Then, its probability distribution is

$$P\{T_S(\mathbf{s}_j, \mathbf{y}_j) = \mathbf{y}_j\} = \begin{cases} 1 & \text{if } f(\mathbf{y}_j) \leqslant f(\mathbf{s}_j), \\ 0 & \text{otherwise.} \end{cases} \tag{3.1}$$

Definition 4. ***Optimisation Process:*** *Based on the operators defined in Definitions 1–3, the optimisation process can be defined as*

$$\mathbf{S}(g+1) = \{\mathbf{s}_j(g+1) = T_S \circ T_R \circ T_M(\mathbf{S}(g)); j = 1, 2, \cdots, M\}.$$

Definition 5. ***Optimal Value and Optimal Population Set:*** *Define* $F(\mathbf{S}) = \min\{f(\mathbf{s}_j); j = 1, 2, \cdots, M\}$ *as the optimal value of the population* $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_M]$. *Define* $M^* = \{\mathbf{S}|F(\mathbf{S}) = \min\{f(\mathbf{s}); \mathbf{s} \in \mathbb{S}\}\}$ *as the optimal population set.*

Theorem 6. *In the optimisation process, evolutionary direction of the population is decreasing monotonically, that is,* $F(\mathbf{S}(g+1)) \leqslant F(\mathbf{S}(g))$.

Proof. According to Equation (3.1), the selection operator can be considered as a greedy strategy in which the phenotype with the best fitness can always be preserved in the next generation. In the optimisation process, the optimal value for the objective function therefore is decreasing monotonically. □

Lemma 7. *In the optimisation process, the population sequence* $\{\mathbf{S}(g); g \in \mathbb{N}^+\}$ *is a Markov chain.*

Proof. According to Definition 4, the optimisation process can be represented as the following iteration of population:

$$\mathbf{S}(g+1) = T(\mathbf{S}(g)) = T_S \circ T_R \circ T_M(\mathbf{S}(g)),$$

where T_S, T_R, and T_M do not depend upon which states the chain was in before the current state g, that is, $\mathbf{S}(g+1)$ only depends on $\mathbf{S}(g)$. Thus, $\{\mathbf{S}(g);\ g \in \mathbb{N}^+\}$ is proved as a Markov chain. □

Lemma 8. *In the optimisation process, the Markov chain,* $\{\mathbf{S}(g);\ g \in \mathbb{N}^+\}$*, is a time-homogeneous irreducible aperiodic Markov chain.*

Proof. According to Lemma 7, the transition probability of the population Markov chain is expressed as

$$\begin{aligned} P\left\{T(\mathbf{S}(g))_j = \mathbf{s}_j(g+1)\right\} \quad = \quad & P\{T_M(\mathbf{s}_j(g)) = \mathbf{x}_j\} \times \\ & P\{T_R(\mathbf{s}_j(g), \mathbf{x}_j) = \mathbf{y}_j\} \times P\{T_S(\mathbf{s}_j(g), \mathbf{y}_j) = \mathbf{y}_j\}. \end{aligned}$$

For $\forall\ \mathbf{X}(g) \in \mathbb{S}^M$, $\exists\ \mathbf{x}_j$ and $\mathbf{y}_j$, it holds that:

$$\begin{aligned} & P\{T_M(\mathbf{s}_j(g)) = \mathbf{x}_j\} > 0, \\ & P\{T_R(\mathbf{s}_j(g), \mathbf{x}_j) = \mathbf{y}_j\} = 1 > 0, \\ & P\{T_S(\mathbf{s}_j(g), \mathbf{y}_j) = \mathbf{y}_j\} > 0. \end{aligned}$$

Therefore, $P\{T(\mathbf{S}(g)) = \mathbf{x}_j(g+1)\} > 0$, and $\mathbf{S}(g+1)$ only depends on $\mathbf{S}(g)$. Thus, the transition probability is given as

$$P\{T(\mathbf{S}(g)) = \mathbf{x}(g+1)\} = \prod_{j=1}^{M} P\left\{T(\mathbf{S}(g)) = \mathbf{s}(g+1)_j\right\}.$$

Therefore, $P\{T(\mathbf{S}(g)) = (g+1)\} > 0$, and $\mathbf{S}(g+1)$ only depends on $\mathbf{S}(g)$. Hence, the Markov chain, $\{\mathbf{S}(g);\ g \in \mathbb{N}^+\}$, is a time-homogeneous irreducible aperiodic Markov chain. □

Here, the transition probability of the Markov chain is rewritten as

$$P\{\mathbf{S},\mathbf{R}\} = P\{\mathbf{S}(g+1) = \mathbf{R} | \mathbf{S}(g) = \mathbf{S}\}. \tag{3.2}$$

According to Lemma 8 and Equation (3.2), the population transition probability can be given as

$$P\{\mathbf{S}(g+1) = \mathbf{R} | \mathbf{S}(g) = \mathbf{S}\} = \begin{cases} \prod_{j=1}^{M} P\{T(\mathbf{S}(g)) = \mathbf{R}_j\} & \exists i_0 \in M^* \\ & \text{s.t. } \mathbf{R}_M = \mathbf{S}_{i_0}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

Theorem 9. *Let $\{\mathbf{S}(g);\ g \in \mathbb{N}^+\}$ be the population Markov chain of the optimisation process. Let M_0^* be a subset of the optimal population set, i.e., $M_0^* = \{\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_M);\ \mathbf{r}_j \in M^*\}$. Then, $\{\mathbf{S}(g);\ g \in \mathbb{N}^+\}$ converges to M_0^* with probability one, i.e.,*

$$\lim_{g\to\infty} P\{\mathbf{S}(g) = \mathbf{R} | \mathbf{S}(0) = \mathbf{S}_0\} = 1.$$

Proof. Suppose $\mathbf{r}^*$ is the unique optimum value for the given objective function. Then, according to Equation (3.2) and Equation (3.3), $P\{\mathbf{S},\mathbf{R}\}$ has the following properties:

1) If $\mathbf{S},\ \mathbf{R} \in M_0^*$, then $P\{\mathbf{S},\mathbf{R}\} > 0$ and $P\{\mathbf{R},\mathbf{S}\} > 0$. Therefore, the two states $\mathbf{S}$ and $\mathbf{R}$ are interconnected, i.e., $\mathbf{S} \leftrightarrow \mathbf{R}$.

2) If $\mathbf{S} \in M_0^*$ and $\mathbf{R} \notin M_0^*$, then $P\{\mathbf{S},\mathbf{R}\} = 0$. Therefore, the two states $\mathbf{S}$ and $\mathbf{R}$ are not interconnected, i.e., $\mathbf{S} \nleftrightarrow \mathbf{R}$.

Hence, M_0^* is a positive aperiodic irreducible closed set. For arbitrary initial state, we can obtain that

$$\lim_{g\to\infty} P\{\mathbf{S}(g) = \mathbf{R} | \mathbf{S}(0) = \mathbf{S}_0\} = \begin{cases} \pi(\mathbf{R}) & \mathbf{R} \in M_0^* \\ 0 & \mathbf{R} \notin M_0^* \end{cases}.$$

$\mathbf{S}(g)$ enters into M_0^* as $g \to \infty$, and satisfies one limiting probability distribution, $\pi(\mathbf{R})$. Therefore, $\lim_{g\to\infty} P\{\mathbf{S}(g) \in M_0^* | \mathbf{S}(0) = \mathbf{S}_0\} = 1$. □

4 Network Properties

To gain an impression of properties of initial gene regulatory networks, in this section, I have investigated the stability, robustness and path length.

4.1 Stability

I first tested the probability of stability in randomly-generated networks. As illustrated in Figure 2, smaller networks are more likely to be stable. Moreover, the relative frequency of stability in networks with low levels of connectivity is higher than that of networks with high levels of connectivity. This is in general accordance with previous work (typically done at connectivity $c = 0.75$, e.g. [17]) which indicates that larger networks with complex topology tend to be unstable. A similar pattern is also observed in networks with different values of activation constant a (see Appendix Figure A1). Generally, when a is small ($a = 1$), networks have a higher initial stability. Note that pattern is much more profound for networks with smaller sizes ($N = 5$, 10 and 15).

4.2 Robustness

Next, I explored the robustness of initial stable networks. That is, I investigated the probability that stable networks remain stable after a single round of mutation. Given that the initial stable networks were collected from the original randomly-generated ones, it would seem reasonable to predict that the small stable networks are more likely to break after one mutation round since they contain fewer

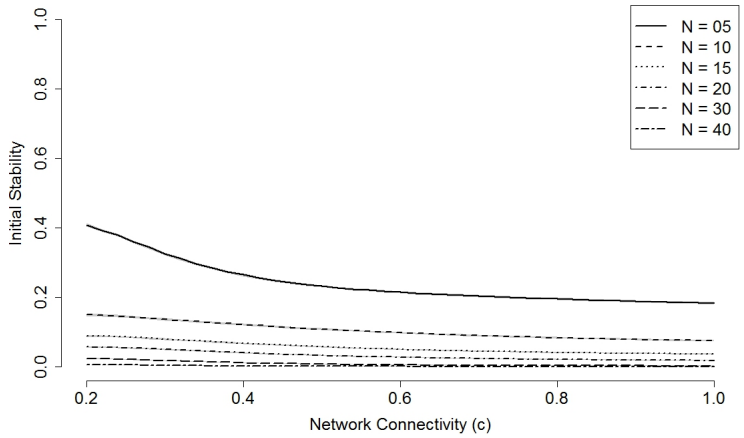


Figure 2: **Stability of randomly-generated networks.** For each network size ($N = 5,\ 10,\ 15,\ 20,\ 30$ and 40) with each connectivity given from a range of values in continuous interval ($[0.2, 1]$, step size 0.02), the initial stability (proportion of randomly-generated gene networks that are stable) was tested based on an initial $10,000$ randomly-generated gene regulatory networks. The system level parameters are set to be $a = 100$, $devT = 100$ and $\tau = 10$. Shaded areas represent 95% confidence intervals based on 100 independent runs.

pathways and a single mutation, therefore, has a greater proportional effect. However, the results in Figure 3 show the opposite — the stability of small networks is still high (cf. Figure 2). The mutation operation is effectively an alternative way of generating new networks, thus, the mutated networks have the same properties as the initial ones. A similar pattern is also observed in networks with different values of activation constant a (see Appendix Figure A2). Generally, when a is small, networks have a higher initial robustness.

4.3 Path Length

In the third sets of experiments, I measured the path length of initial stable networks. Here the path length refers to the amount of time steps, as used in Equation (2.2), that the network takes from an initial state $\mathbf{s}(0)$ to get to the equilibrium state $\mathbf{s}_{\mathrm{EQ}}$. From Figure 4, we can clear see that larger networks need more time to reach the equilibrium state. Moreover, networks with low levels of connectivity are able to stabilise faster than that of networks with high levels of connectivity, especially for networks with sizes of $N = 15,\ 20$ and 30. A similar pattern is also observed in networks with different values of activation constant a (see Appendix Figure A3). Generally, when a is small ($a = 1$), networks need much more time to get to the equilibrium state, especially for networks with size of $N = 5$ in comparison with the results when a is large (cf. Figure 4). However, the path length slightly decreases for networks with size of $N = 10,\ 15$ and 20 when $a = 1$.

Here, the time is defined as the minimum steps required for networks reaching the equilibrium state.

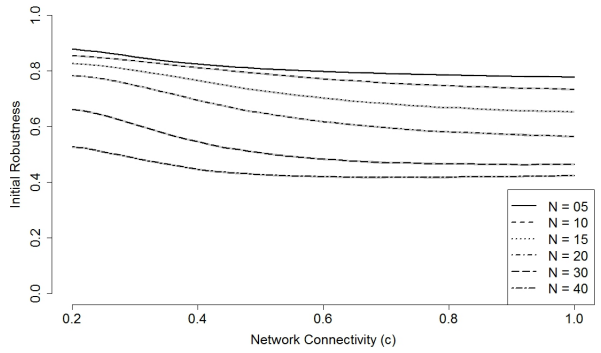


Figure 3: **Robustness of initial stable networks.** For each network size ($N = 5$, 10, 15, 20, 30 and 40) with each connectivity given from a range of values in continuous interval ($[0.2, 1]$, step size 0.02), the robustness (proportion of stable networks after exposure to a single round of mutation) was tested based on an initial $10,000$ randomly-generated stable gene regulatory networks. The system level parameters are set to be $a = 100$, $devT = 100$ and $\tau = 10$. Shaded areas represent 95% confidence intervals based on 100 independent runs.

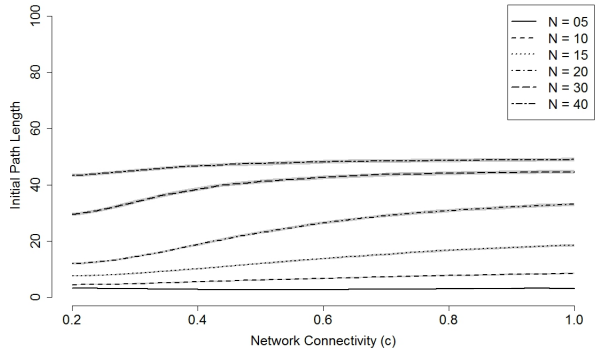


Figure 4: **Path length of initial stable networks.** For each network size ($N = 5$, 10, 15, 20, 30 and 40) with each connectivity given from a range of values in continuous interval ($[0.2, 1]$, step size 0.02), the path length (minimum time steps for reaching the equilibrium state) was tested based on an initial $10,000$ randomly-generated stable gene regulatory networks. The system level parameters are set to be $a = 100$, $devT = 100$ and $\tau = 10$. Shaded areas represent 95% confidence intervals based on 100 independent runs.

5 Discussion & Conclusion

Networks of regulatory transcription factors are essential to form developmental patterns in practically all organisms [20, 21, 22]. The process of development that reduces effects of genetic or environmental perturbations is due to the nonlinearity of genotype-phenotype mapping enhances the robustness of the system, whilst constraining phenotypic diversity, and consequently inhibits certain evolutionary pathways [23, 24]. Although many previous studies have shown that the process of development is critical for the study of evolution, the underlying mechanism, in particular, how the developmental process affects evolutionary dynamics that can drive evolutionary innovations, is still poorly understood.

Mutations in Wagner's GRN model or other similar models in natural systems are shown to be an important source of innovation. Previous studies have forced on separating two sources of mutations: genetic and non-genetic [25, 26, 27, 28, 29, 30, 24]. On the one hand, the genetic mutations refer to perturbations occur at the genotypes, usually have a weaker effect in altering gene's phenotypic state or causing instability of the network since the complex interactions among genes can buffer against mutations occurred at the genotype level. The non-genetic mutations, on the other hand, refer to perturbations caused by internal noise or environmental factors, and may sometimes have a strong effect in causing oscillatory dynamics to the developmental stability, especially changes occurring in gene initial expression patterns. Although previous studies have investigated many different types of mutations, how those mutations systemically affect phenotypic stability remains obscure.

In spite of mutation, recombination is also believed to be critical to affecting the underlying evolutionary dynamics in the context of genetic networks. The recombination modelled in Wagner's GRN model in a manner of free recombination of swapping rows between two parental genotypes. This operation follows the biological assumption that recombination happens more often between genes and tight linkage among regulatory elements within a promoter [10, 11]. Previous work has focused on benefits and low costs of recombination to reconcile the traditional antagonistic view that recombination is more likely to damage well-adapted lineages due to massively shifting patterns of gene regulation [17, 31, 32, 33, 34, 35, 19]. Although MacCarthy and Bergman and Lohaus et al. previously introduced a modifier of recombination into the model, different recombination modes have not yet been studied thoroughly [31, 33], given that the variety of mating systems and strategies in nature [36, 37].

In the seminal paper of [10], for the first time, the mathematical foundation of his GRN model was formally described. In addition to the seminal paper where Wagner showed that given an initial state $\mathbf{s}(0)$, the developmental process converges ultimately to a stable equilibrium state $\mathbf{s}_{\mathrm{EQ}}$, in this paper, I have further mathematically shown that the evolution process modelled in Wagner's GRN model can be regarded as an optimisation process that converges to the target configuration with probability one. Besides the convergence analysis, a few other studies have employed theories for calculating periodic orbit to study the systematic behaviour of developmental process [38, 24]. However, it is still not clear that how mutation and recombination operators modelled in the system change periodic orbit and ultimately affect the underlying evolutionary dynamics.

Previous work has shown that sparse networks are more stable than dense networks [38], here I also have observed a similar pattern varying network sizes and activation constants (see Figure 2). Furthermore, I have shown that randomly-generated stable sparse networks also have a higher robustness against mutations than dense networks (see Figure 3), though sparse networks may evolve to be more sensitive to mutations than networks that are more densely connected under stabilising selection [11, 16]. However, Leclerc showed that if costs of complexity have been considered, then robust networks are more likely to be sparsely connected [39] . This may help explain that sparse networks tend to be favoured by evolution in natural systems [40]. As Wagner and Siegal and Bergman suggested, the path length, time for reaching phenotypic equilibrium, may partially account

In Wagner's GRN model, genetic mutations are assumed to be epistatic mutations that alter the gene's regulation strength to other genes, but not mutations occur at the coding sequence at the lowest level.

for the underlying mechanism of stability and robustness [11, 16]. This is because if the phenotypic stabilising process takes more time, then the network is more likely to accumulate deleterious mutations or perturbed by internal noise or environmental factors. Here, I have provided some evidence to support this likelihood by showing that sparse networks tend to have a shorter path length to reach the equilibrium (see Figure 4). Note that different activation constant, a, which indicates the sensitivity of regulatory response to output phenotypes, can quantitatively affect initial stability, robustness and path length (see more supporting information in Appendix). Generally, networks with a small value of a has a much higher initial stability and robustness than networks with a larger value of a.

It should be emphasised that parameters used in Wagner's GRN model, such as population size, number of genes, network connectivity, and activation constant, will not typically change the qualitative results of general properties or patterns emerged from the evolved system [11, 16, 17]. In particular, previous studies have suggested that many biological networks have a scale-free topology, that is, its degree distribution of nodes follows a power law [41, 42]. However, previous research has shown that the degree distribution itself does not have a major effect on functional properties associated with nodes [11, 17, 22, 38]. Therefore, although in this paper networks were randomly generated and the parameter space had not been thoroughly explored, it is expected that the patterns or properties I have observed generally could be applied to scale-free networks and results presented in this paper are representative.

References

[1] Andreas Wagner. *The Origins of Evolutionary Innovations: A Theory of Transformative Change in Living Systems*. Oxford University Press, 2011.

[2] Andreas Wagner. *Arrival of the Fittest: Solving Evolution's Greatest Puzzle*. Penguin, 2014.

[3] Guy Karlebach and Ron Shamir. Modelling and analysis of gene regulatory networks. *Nature Reviews Molecular Cell Biology*, 9(10):770–780, Oct 2008.

[4] Stuart A Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology*, 22(3):437–467, 1969.

[5] Stuart A Kauffman. *The Origins of Order: Self Organization and Selection in Evolution*. Oxford University Press, 1993.

[6] Ilya Shmulevich, Edward R. Dougherty, Seungchan Kim, and Wei Zhang. Probabilistic boolean networks: A rule-based uncertainty model for gene regulatory networks. *Bioinformatics*, 18(2): 261–274, 2002.

[7] C. A. Petri. Kommunikation mit Automaten. Technical report, Universität Hamburg, 1962.

[8] Nir Friedman, Michal Linial, Iftach Nachman, and Dana Pe'er. Using bayesian networks to analyze expression data. *Journal of Computational Biology*, 7(3-4):601–620, August 2000.

[9] Edda Klipp, Ralf Herwig, Axel Kowald, Christoph Wierling, and Hans Lehrach. *Systems Biology in Practice: Concepts, Implementation and Application*. John Wiley & Sons, 2008.

[10] Andreas Wagner. Evolution of gene networks by gene duplications: A mathematical model and its implications on genome organization. *Proceedings of the National Academy of Sciences of the United States of America*, 91(10):4387–4391, 1994.

[11] A. Wagner. Does evolutionary plasticity evolve? *Evolution*, 50(3):1008–1023, 1996.

[12] Ernst Ising. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik*, 31(1):253–258, 1925.

[13] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366, January 1989.

[14] Janna L Fierst and Patrick C Phillips. Modeling the evolution of complex genetic systems: The gene network family tree. *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution*, 324(1):1–12, January 2015.

[15] Yifei Wang. Review of wagner's artificial gene regulatory networks model and its applications for understanding complex biological systems. *COJ Robotics & Artificial Intelligence*, 1(1): COJRA.000501.2019, Feb 2019.

[16] M. L. Siegal and A. Bergman. Waddington's canalization revisited: Developmental stability and evolution. *Proceedings of the National Academy of Sciences of the United States of America*, 99(16):10528–10532, 2002.

[17] Ricardo B. R. Azevedo, Rolf Lohaus, Suraj Srinivasan, Kristen K. Dang, and Christina L. Burch. Sexual reproduction selects for robustness and negative epistasis in artificial gene networks. *Nature*, 440(7080):87–90, 2006.

[18] Yifei Wang, Stephen G Matthews, and Joanna J Bryson. Evolving evolvability in the context of environmental change: A gene regulatory network (GRN) approach. In *Artificial Life 14: International Conference on the Synthesis and Simulation of Living Systems*, pages 47–53. The MIT Press, July 2014.

[19] Yifei Wang, Yinghong Lan, Daniel M. Weinreich, Nicholas K. Priest, and Joanna J. Bryson. Recombination is surprisingly constructive for artificial gene regulatory networks in the context of selection for developmental stability. In *Proceedings of the European Conference on Artificial Life 2015*, pages 530–537. The MIT Press, July 2015.

[20] Nabil Guelzim, Samuele Bottani, Paul Bourgine, and Francois Kepes. Topological and causal structure of the yeast transcriptional regulatory network. *Nature Genetics*, 31(1):60–63, April 2002.

[21] Eric H. Davidson, Jonathan P. Rast, Paola Oliveri, Andrew Ransick, Cristina Calestani, Chiou-Hwa Yuh, Takuya Minokawa, Gabriele Amore, Veronica Hinman, César Arenas-Mena, Ochan Otim, C. Titus Brown, Carolina B. Livi, Pei Yun Lee, Roger Revilla, Alistair G. Rust, Zheng jun Pan, Maria J. Schilstra, Peter J. C. Clarke, Maria I. Arnone, Lee Rowen, R. Andrew Cameron, David R. McClay, Leroy Hood, and Hamid Bolouri. A genomic regulatory network for development. *Science*, 295(5560):1669–1678, 2002.

[22] Mark L Siegal, Daniel E L Promislow, and Aviv Bergman. Functional and evolutionary inference in gene networks: Does topology matter? *Genetica*, 129(1):83–103, January 2007.

[23] Philipp Thomas, Nikola Popović, and Ramon Grima. Phenotypic switching in gene regulatory networks. *Proceedings of the National Academy of Sciences of the United States of America*, 111(19):6994–6999, May 2014.

[24] Ricardo Pinho, Victor Garcia, and Marcus W Feldman. Phenotype accessibility and noise in random threshold gene regulatory networks. *PLoS ONE*, 10(4):e0119972, April 2015.

[25] J Masel. Genetic assimilation can occur in the absence of selection for the assimilating phenotype, suggesting a role for the canalization heuristic. *Journal of Evolutionary Biology*, 17(5):1106–1110, September 2004.

[26] Volkan Sevim and Per Arne Rikvold. Chaotic gene regulatory networks can be robust against mutations and noise. *Journal of Theoretical Biology*, 253(2):323–332, July 2008.

[27] Tristan Kimbrell and Robert D Holt. Canalization breakdown and evolution in a source-sink system. *The American Naturalist*, 169(3):370–382, March 2007.

[28] Stefano Ciliberti, Olivier C Martin, and Andreas Wagner. Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLoS Computational Biology*, 3(2): e15, 2007.

[29] Olivier C Martin and Andreas Wagner. Multifunctionality and robustness trade-offs in model genetic circuits. *Biophysical Journal*, 94(8):2927–2937, 2008.

[30] Carlos Espinosa-Soto, Olivier C Martin, and Andreas Wagner. Phenotypic plasticity can facilitate adaptive evolution in gene regulatory circuits. *BMC Evolutionary Biology*, 11(1):5, 2011.

[31] Thomas MacCarthy and Aviv Bergman. Coevolution of robustness, epistasis, and recombination favors asexual reproduction. *Proceedings of the National Academy of Sciences of the United States of America*, 104(31):12801–12806, 2007.

[32] Olivier C Martin and Andreas Wagner. Effects of recombination on complex regulatory circuits. *Genetics*, 183(2):673–684, 2009.

[33] R Lohaus, C L Burch, and R B R Azevedo. Genetic architecture and the evolution of sex. *The Journal of Heredity*, 101(Supplement 1):S142–S157, April 2010.

[34] Andreas Wagner. The low cost of recombination in creating novel phenotypes. *BioEssays*, 33 (8):636–646, 2011.

[35] Yann Le Cunff and Khashayar Pakdaman. Reproduction cost reduces demographic stochasticity and enhances inter-individual compatibility. *Journal of Theoretical Biology*, 360:263–270, 2014.

[36] Stephen M. Shuster and Michael J. Wade. *Mating Systems and Strategies*. Monographs in Behavior and Ecology. Princeton University Press, May 2003.

[37] Stephen M Shuster. Sexual selection and mating systems. *Proceedings of the National Academy of Sciences of the United States of America*, 106(Supplement 1):10009–10016, June 2009.

[38] Ricardo Pinho, Elhanan Borenstein, and Marcus W Feldman. Most networks in Wagner's model are cycling. *PLoS ONE*, 7(4):e34285, April 2012.

[39] Robert D Leclerc. Survival of the sparsest: Robust gene networks are parsimonious. *Molecular Systems Biology*, 4:213, 2008.

[40] Nicholas M Luscombe, M Madan Babu, Haiyuan Yu, Michael Snyder, Sarah A Teichmann, and Mark Gerstein. Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature*, 431(7006):308–312, September 2004.

[41] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, October 1999.

[42] Mark Newman, Albert-Laszlo Barabasi, and Duncan J. Watts. *The Structure and Dynamics of Networks*. Princeton University Press, Princeton, NJ, USA, 2006. ISBN 0691113572.

Appendix

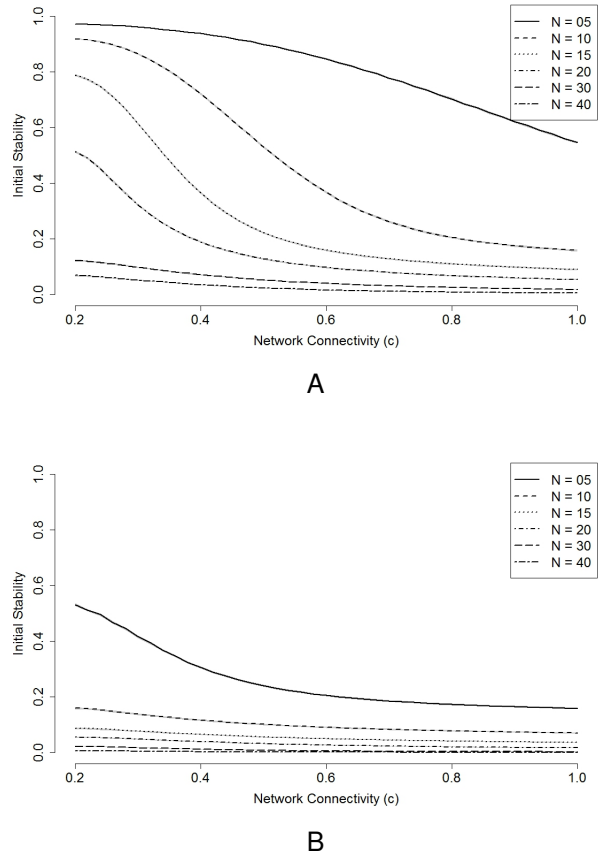
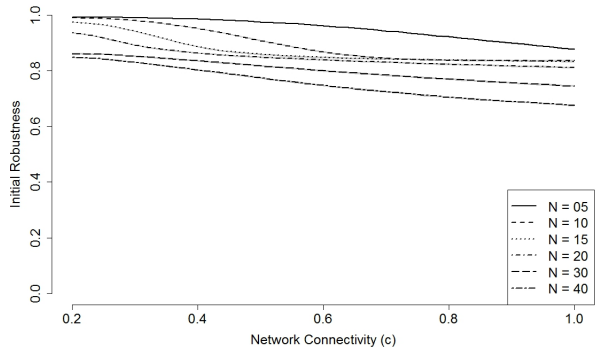
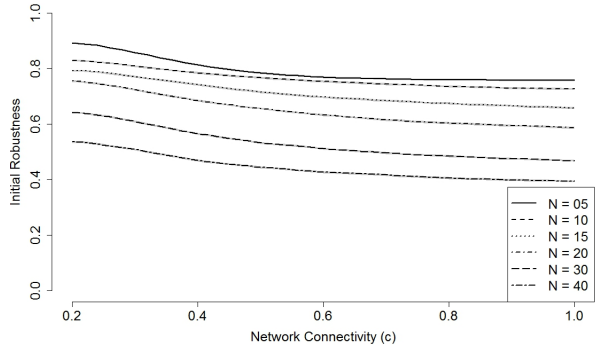


Figure A1: **Stability of randomly-generated networks varying activation constant.** For each network size ($N = 5,\ 10,\ 15,\ 20,\ 30$ and 40) with each connectivity given from a range of values in continuous interval ($[0.2, 1]$, step size 0.02), the initial stability (proportion of randomly-generated gene networks that are stable) was tested based on an initial $10,000$ randomly-generated gene regulatory networks. The system level parameters are set to be $a = 1$ in **(A)** or $a = 5$ in **(B)**, $devT = 100$ and $\tau = 10$. Note that Shaded areas represent 95% confidence intervals based on 100 independent runs.

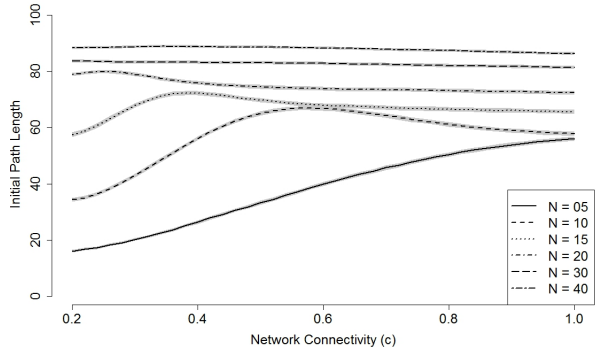


A

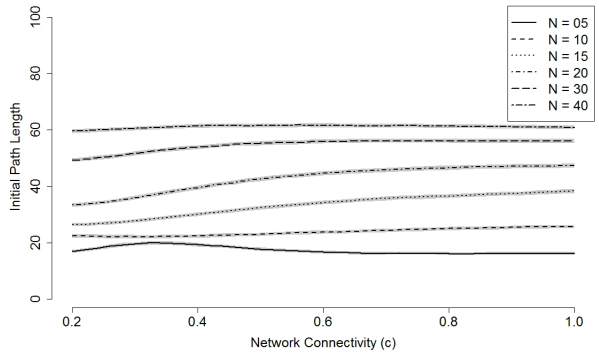


B

Figure A2: **Robustness of initial stable networks varying activation constant.** For each network size ($N = 5$, 10, 15, 20, 30 and 40) with each connectivity given from a range of values in continuous interval ($[0.2, 1]$, step size 0.02), the initial robustness (proportion of stable networks after exposure to a single round of mutation) was tested based on an initial $10,000$ randomly-generated stable gene regulatory networks. The system level parameters are set to be $a = 1$ in **(A)** or $a = 5$ in **(B)**, $devT = 100$ and $\tau = 10$. Note that Shaded areas represent 95% confidence intervals based on 100 independent runs.



A



B

Figure A3: **Robustness of initial stable networks varying activation constant.** For each network size ($N = 5$, 10, 15, 20, 30 and 40) with each connectivity given from a range of values in continuous interval ($[0.2, 1]$, step size 0.02), the path length (minimum time steps for reaching the equilibrium state) was tested based on an initial $10,000$ randomly-generated stable gene regulatory networks. The system level parameters are set to be $a = 1$ in **(A)** or $a = 5$ in **(B)**, $devT = 100$ and $\tau = 10$. Note that Shaded areas represent 95% confidence intervals based on 100 independent runs.