The Influence of Whole Genome Duplication and Subsequent Diversification on Environmental Robustness and Evolutionary Innovation in Gene Regulatory Networks

Qinxin Pan[1], Christian Darabos[1], Anna L. Tyler[1], Jason H. Moore[1], and Joshua L. Payne[1]

[1]Dartmouth College, Hanover, NH 03755
Jason.H.Moore@Dartmouth.edu

Abstract

In biological systems, whole genome duplication and subsequent diversification constitute powerful mechanisms for the discovery of new phenotypes and for the protection of these phenotypes against environmental perturbation. Here, we use Random Boolean Networks to investigate the influence of these genetic mechanisms on the relationship between evolutionary innovation and environmental robustness in gene regulatory networks. We find that whole genome duplication is highly deleterious in ancestral environments, but provides fitness advantages in novel environments, which come at the cost of reduced environmental robustness. We then show that the subsequent diversification of duplicated networks, via the loss of regulatory interactions, can partly negotiate this trade-off, improving evolutionary innovation and environmental robustness. We conclude by discussing the implications, limitations, and future directions of our research.

Introduction

Biological systems exhibit two crucial and seemingly antagonistic properties: robustness and evolvability (Wagner, 2005). Regardless of the level of biological organization, living organisms display remarkable resilience to changing conditions, and at the same time, they are able to respond to these changes by developing novel phenotypes. At first glance, these qualities seem paradoxical, yet both empirical (Bloom et al., 2006; Ferrada and Wagner, 2008; Isalan et al., 2008) and theoretical (Aldana et al., 2007; Wagner, 2008; Draghi et al., 2010) analyses suggest their compatibility.

The relationship between robustness and evolvability has been investigated in biological systems ranging in scale from the molecule (Schuster et al., 1994; Cowperthwaite et al., 2008) to the cell (Aldana et al., 2007; Ciliberti et al., 2007a,b). At the cellular level, gene expression patterns are robust to changing environmental conditions, such as alterations in growth medium or the concentration levels of transcription factors (Alon, 2007). This insensitivity to environmental perturbation is largely influenced by the structure of the underlying gene regulatory network (GRN) (Aldana and Cluzel, 2003). A GRN consists of a set of genes, represented as vertices, linked by directed edges if the gene-product (e.g., protein, mRNA, microRNA) of the source gene has a regulatory influence on the target gene. Recent analyses of model GRNs have revealed that robustness is often correlated with the capacity for evolutionary innovation (Ciliberti et al., 2007a; Aldana et al., 2007).

One major form of structural change in GRNs comes from whole genome duplication (WGD) events, wherein the entire gene repertoire of an organism, including regulatory interactions, is doubled (Sémon and Wolfe, 2007). WGD has long been recognized as a driver of evolutionary innovation (Ohno, 1970) and recent genetic analyses have demonstrated that several major evolutionary transitions resulted from ancient WGD events (Kellis et al., 2004; De Bodt et al., 2005; Taylor et al., 2003). For example, the origin of the budding yeast *Saccharomyces cerevisiae* (Kellis et al., 2004) and the radiation of the angiosperms into over 250,000 species (De Bodt et al., 2005) have both been attributed to WGD. The duplication of genetic material has implications for environmental robustness, as redundant genes diverge to compartmentalize the original function of the ancestral gene (subfunctionalization) (Sémon and Wolfe, 2007). In *S. cerevisiae*, for example, this occurs through the differential expression of redundant genes under various growth conditions (Kafri et al., 2005). WGD also has implications for evolutionary innovation, as duplicate genes diverge to acquire new functions (neofunctionalization) (Sémon and Wolfe, 2007). In *S. cerevisiae*, the ability to consume glucose and grow anaerobically have both been attributed to the genetic diversification that followed a WGD event (Piškur, 2001).

Despite the known importance of WGD events for evolutionary processes, their influence on environmental robustness and evolutionary innovation in GRNs is not thoroughly understood. Here, we use Random Boolean Networks (RBNs) (Kauffman, 1969) to model the dynamics of GRNs. We simulate WGD events in RBNs and quantify their effect on environmental robustness and evolutionary innovation.

This paper is structured as follows. In the subsequent section, we present the key concepts of this work. We then present our model and the details of our simulations, ana-

lyze and discuss our results, and conclude with an outline of future research directions.

Background

Random Boolean Networks

Random Boolean Networks (RBNs) are abstract dynamical models of gene regulatory networks (GRNs) (Kauffman, 1969). RBNs consist of N nodes, which represent genes, and directed edges, which represent regulatory interactions. Node states are binary, representing the expression (1) or repression (0) of gene products. Node states are also dynamic, such that the state of a node in time step $t + 1$ is dependent upon the states of its regulating nodes in time step t. To model this dependence, each vertex is associated with a Boolean update function, which is captured by a look-up table that explicitly maps the output expression state for all possible combinations of input states. These output expression states are drawn at random with probability p_{expr} and are held fixed throughout the system's dynamics.

Node states are updated synchronously and in discrete time. The dynamics of a RBN begin with a prespecified initial configuration of node states, which represents regulatory factors upstream of the GRN (Ciliberti et al., 2007a). After at most 2^N time steps, the system will encounter a configuration previously visited, thus entering a cycle of one or more configurations, which is referred to as an attractor.

An important aspect of RBNs is that their dynamical behavior falls into one of three regimes: ordered, critical, or chaotic. Systems in the ordered regime exhibit short attractors that are relatively insensitive to environmental perturbation. At the other end of the spectrum, systems in the chaotic regime possess longer attractors that are highly sensitive to environmental perturbation. The critical regime lies at the transition between the ordered and chaotic regimes, offering a balance between the ability to withstand environmental perturbation (robustness) and the ability to utilize these perturbations for evolutionary innovation (evolvability) (Aldana et al., 2007).

WGD and Subsequent Diversification

Immediately following whole genome duplication (WGD), organismal stability is generally reduced, leading to a decrease in fitness (van Hoek and Hogeweg, 2009). However, duplicate genes supply new genetic material, which can be shaped via mutation and selection to produce novel functions. These functions may allow for more rapid adaptation if a new environment is encountered, providing potential fitness benefits (van Hoek and Hogeweg, 2009). The genetic reorganization that accompanies such diversification may occur via gene loss, gene rearrangements, or alterations in the circuitry of genetic regulation (Sémon and Wolfe, 2007).

Methods

In this section, we separately present our implementations of RBN generation, duplication, and diversification. We then quantify environmental robustness and evolutionary innovation, outline the evolutionary processes used in our analyses, and provide the details of our simulations.

RBN Topology

The degree distribution of a RBN has an important influence on system dynamics (Aldana and Cluzel, 2003; Oikonomou and Cluzel, 2006; Aldana et al., 2007). Here, we consider RBNs with Poisson input degree distributions and power-law output degree distributions, as empirical evidence suggests that such topologies are representative of the GRNs of several organisms (Aldana and Cluzel, 2003; Albert, 2005). RBN topologies are generated as described by Darabos et al. (2009).

Duplication

WGD is simulated by first creating a mirror-image of the original RBN and then linking the duplicate and original components by drawing edges from the source nodes in one component to the targets in the other (Fig. 1a,b). Each node in the duplicated RBN has twice as many inputs as the corresponding node in the non-duplicated RBN. As a result, the number of entries in the look-up table is squared. To populate the entries of each table, we follow Aldana et al. (2007): when the duplicate regulatory inputs are not expressed, the Boolean rules remain identical to those prior to duplication. However, when the duplicate regulatory inputs are expressed the Boolean rules are assigned at random with probability p_{expr}.

Diversification

To simulate the genetic diversification that follows a WGD event, we take a conservative approach and assume that only regulatory interactions can be lost (akin to structural simplification algorithms for neural networks (Le Cun et al., 1990)). This represents a mutation to the promoter region of a gene that prohibits the binding of one of its regulating gene products. While this type of mutation represents only a small subset of all possible forms of genetic reorganization, it offers a useful and parsimonious starting point. Further, empirical data suggest that (i) interactions are lost at a rate that is three orders of magnitude larger than the rate at which they are gained (Wagner, 2001) and (ii) rates of alternative forms of reorganization, such as gene loss, are significantly reduced among transcription factors (De Bodt et al., 2005), which are the primary gene products modeled by RBNs.

In our simulations, diversification occurs through the removal of all non-functional regulatory edges (Fig. 1c,d). These edges link a source to a target, where the state of the source does not influence the expression of the target. Such edges are referred to as canalyzed (Kauffman et al., 2004),

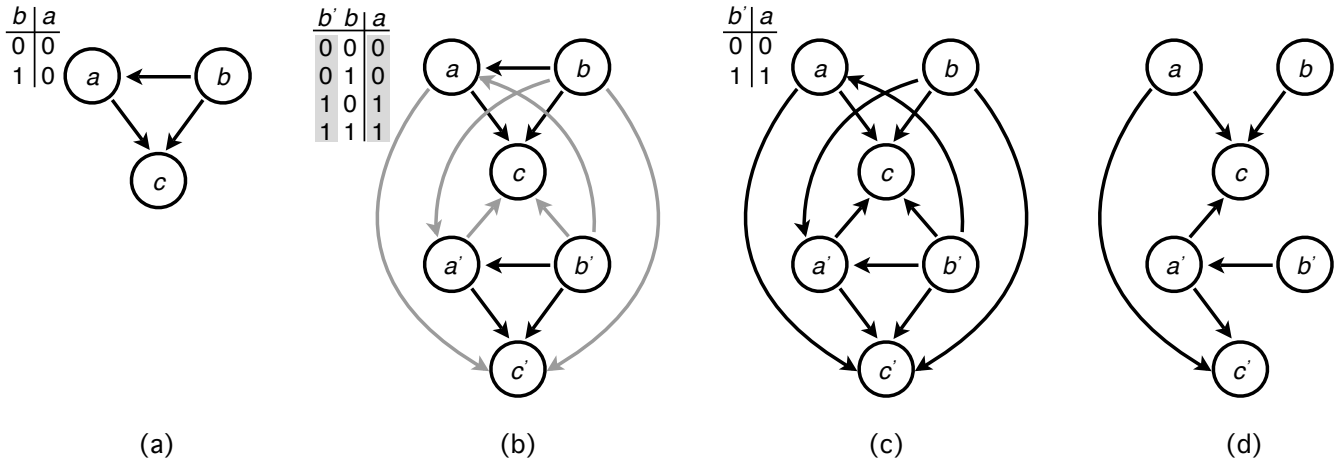

Figure 1: Schematic of whole genome duplication (WGD) and the subsequent diversification of a Random Boolean Network (RBN). The (a) non-duplicated RBN (b) undergoes WGD, wherein its entire gene repertoire is copied. Additional edges (gray) are drawn from the source nodes of one component to the target nodes of the other. The look-up tables of each node are expanded as described by Aldana et al. (2007). As an illustrative example, we depict one possible expansion of the look-up table of node a. (c) Diversification occurs via edge loss. E.g., $a \leftarrow b$. (d) All edges are removed where the state of the target node is independent of the state of the source node. The diversification process continues throughout the evolution of the population, resulting in RBN topologies that differ markedly from those that immediately followed WGD.

and their removal does not immediately affect the dynamics of the RBN.

Environmental Robustness

Environmental perturbations come in many forms, including alterations in temperature, growth medium, or biotic environment. A RBN is environmentally robust if its phenotype is insensitive to these non-genetic perturbations. We measure environmental robustness as the sensitivity of a RBN to the perturbation of a single, randomly chosen configuration of its attractor. Specifically, we systematically perturb the state of each node in the randomly chosen configuration, one at a time, and measure the proportion of perturbations in which the RBN returns to its original attractor.

Evolutionary Innovation

An evolutionary innovation can be thought of as a change in phenotype that confers a fitness advantage. To assess evolutionary innovation, we measure the fitness of a RBN as the ability of its attractor to match a randomly generated target attractor. This target attractor represents the gene expression pattern required for optimal adaptation to a given environment. Fitness thus provides a proxy for evolutionary innovation.

For each RBN, we randomly select a single output node and record the sequence of output states σ_{out} during its attractor. The fitness F of a RBN is then calculated as the Hamming distance between the output and target sequences (Oikonomou and Cluzel, 2006),

$$F = \max\left\{1 - \frac{1}{\operatorname{lcm}(L, L_{\text{c}})} \sum_{t=1}^{\operatorname{lcm}(L,L_{\text{c}})} |\sigma_{\text{out}}(t) - \sigma_{\text{target}}(t)|\right\}, \tag{1}$$

where L is the length of the output sequence, L_{c} is the length of the target sequence, and lcm denotes the least common multiple. To facilitate the comparison of sequences with $L \neq L_{\text{c}}$, both sequences are concatenated onto themselves until they are of length $\operatorname{lcm}(L, L_{\text{c}})$. To ensure that fitness is independent of the starting position of the output sequences, we take the maximum fitness over all cyclic permutations of σ_{out}.

Evolution

We simulate the evolution of randomly initialized populations of RBNs in discrete, non-overlapping generations. In every generation, the fitness of each RBN is assessed according to Eq. 1. RBNs are then selected with uniform probability, with replacement, to compete in binary tournaments. Within a tournament, the RBN with the highest fitness is selected to move on to the next generation, after undergoing mutation. Mutation only affects the RBN's look-up tables, such that the entries in the look-up tables associated with each vertex undergo bit-flip mutation with probability p_{mut}. This process of selection and mutation is repeated until the next generation is fully populated.

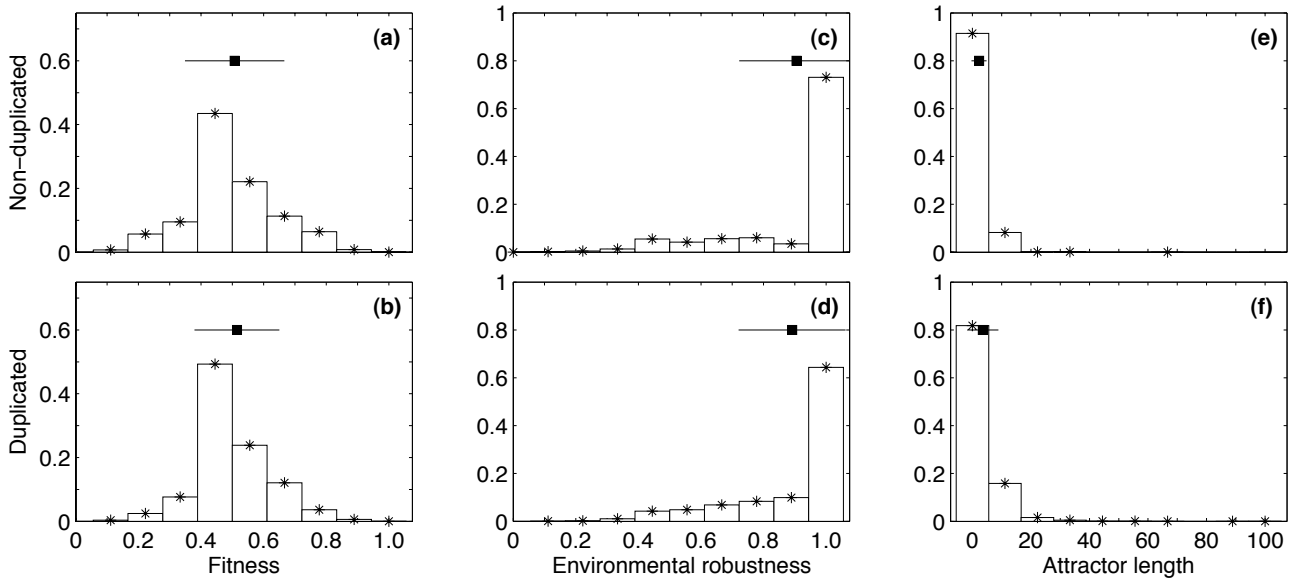

Figure 2: (a,b) Fitness, (c,d) environmental robustness, and (e,f) attractor length of non-duplicated (top row) and duplicated (bottom row) RBNs in novel environments. Each panel depicts the frequency distribution of data across 10,000 independent replications. The filled squares depict the mean of those data and the horizontal lines depict one standard deviation. The asterisk symbols are placed atop each non-empty bin as a visual aid. In (f), there was a single outlier with an attractor length $L = 254$, which is not shown.

Simulation Details

We consider RBNs with $N = 10$ nodes prior to duplication and $N = 20$ nodes after duplication. RBNs are initialized near the critical regime by setting the probability of gene expression to $p_{\mathrm{expr}} = 0.5$ and the scaling exponent of the output degree distribution to $\gamma = 1.894$, which yields criticality in RBNs with $N = 10$ (Aldana et al., 2007).

Evolutionary analyses are conducted with a population size of 500, wherein each RBN is paired with its own, randomly chosen initial state which does not change throughout the evolutionary trajectory of its lineage. In each experiment, we consider 100 independent replications that each consist of 5,000 generations. Mutation occurs with probability $p_{\mathrm{mut}} = 0.002$ per look-up table entry. In the experiments that include diversification, the deterministic edge-loss process only occurs every 10 generations, due to computational constraints.

Results

We present our results in four successive phases. First, we compare the immediate effects of WGD on the fitness of RBNs in their ancestral environments. Second, we compare the immediate effects of WGD on the fitness and robustness of RBNs in novel environments. Third, we consider the evolutionary dynamics of fitness and robustness for non-duplicated and duplicated RBNs. Fourth, we compare the evolutionary dynamics of these same quantities when duplicated RBNs are allowed to undergo diversification.

WGD in an Ancestral Environment

To simulate an ancestral environment, we simply assume that the expression profile of a randomly generated RBN is optimally adapted. To do this, we choose a random node from the RBN and define it as σ_{target}. We then simulate a WGD event, designate the expression profile of the same node in the duplicated RBN as σ_{out}, and compute its fitness (Eq. 1). To collect meaningful statistics, we repeat this process 10,000 times.

WGD is highly deleterious in an ancestral environment. Optimal fitness is maintained in only $\sim 42\%$ of WGD events. Of the remaining $\sim 58\%$ of duplicated RBNs, average fitness decreases to 0.37 ± 0.002.

WGD in a Novel Environment

To simulate a novel environment, we randomly generate σ_{target} of length $L_{\mathrm{c}} = 10$ (Oikonomou and Cluzel, 2006). We then generate a RBN, choose a random node, designate its expression profile as σ_{out}, and compute the RBN's fitness (Eq. 1). In addition, we measure the RBN's environmental robustness. We then collect these data for the same RBN after WGD. As in the previous analysis, this process is repeated 10,000 times.

Duplicated RBNs exhibit a marginal fitness advantage over their non-duplicated counterparts (Fig. 2a,b; Stu-

dent's t-test, $p = 5.43 \times 10^{-4}$). However, this advantage comes at the expense of a marginal decrease in environmental robustness (Fig. 2c,d; Kolmogorov-Smirnov test, $p = 3.46 \times 10^{-74}$). These subtle differences can be attributed to the increased attractor length of the duplicated RBNs (Fig. 2e,f; Student's t-test, $p = 8.9 \times 10^{-9}$), which have a higher probability of matching σ_{target} and exhibit a greater sensitivity to perturbation.

Evolutionary Dynamics of Duplicated RBNs

We now turn from a static analysis of environmental robustness and fitness to an evolutionary analysis of these quantities for populations of non-duplicated and duplicated RBNs. As in the previous section, we consider novel environments by randomly generating target sequences σ_{target} of length $L_c = 10$.

As observed in our previous analysis, the duplicated RBNs have an immediate, albeit slight, fitness advantage in a novel environment (Fig. 3a), but are marginally less robust (Fig. 3b). These differences in fitness and robustness become more pronounced throughout the evolutionary process. Duplicated RBNs reach a plateau of average fitness at 0.92 ± 0.006 (Fig. 3a, squares) while the non-duplicated RBNs stagnate at an average fitness of 0.89 ± 0.008 (Fig. 3a, triangles). Simultaneously, the duplicated RBNs drop to an average environmental robustness of 0.73 ± 0.042 (Fig. 3b, squares), while the non-duplicated RBNs retain a higher environmental robustness of 0.83 ± 0.018 (Fig. 3b, triangles). Thus, WGD in the absence of subsequent diversification leads to a trade-off between environmental robustness and evolutionary innovation.

Evolutionary Dynamics of Diversified RBNs

To investigate the effects of diversification after WGD, we conduct an evolutionary analysis of paired populations of duplicated RBNs, wherein diversification can only occur in one of the initially identical populations.

Diversification promotes evolutionary innovation, with populations reaching an average fitness of 0.95 ± 0.013 (Fig. 4a, open circles), a significant improvement over the fitness obtained with WGD alone (Fig. 4a, closed squares). Simultaneously, diversification increases environmental robustness (Fig. 4b, open circles), though not to the same levels observed prior to WGD (Fig. 3b, triangles). Thus, diversification allows for the partial negotiation of the trade-off between environmental robustness and evolutionary innovation that is induced by WGD.

The diversification process also leads to appreciable structural changes in RBN topologies, with average connectivity dropping rapidly (Fig. 4a, inset). RBN dynamics are also affected, with attractor lengths of diversified networks settling to an average of 9.48 ± 0.174, as compared to 9.92 ± 0.112 for non-diversified networks (Fig. 4b, inset). The probability of gene expression p_{expr} remains approximately constant

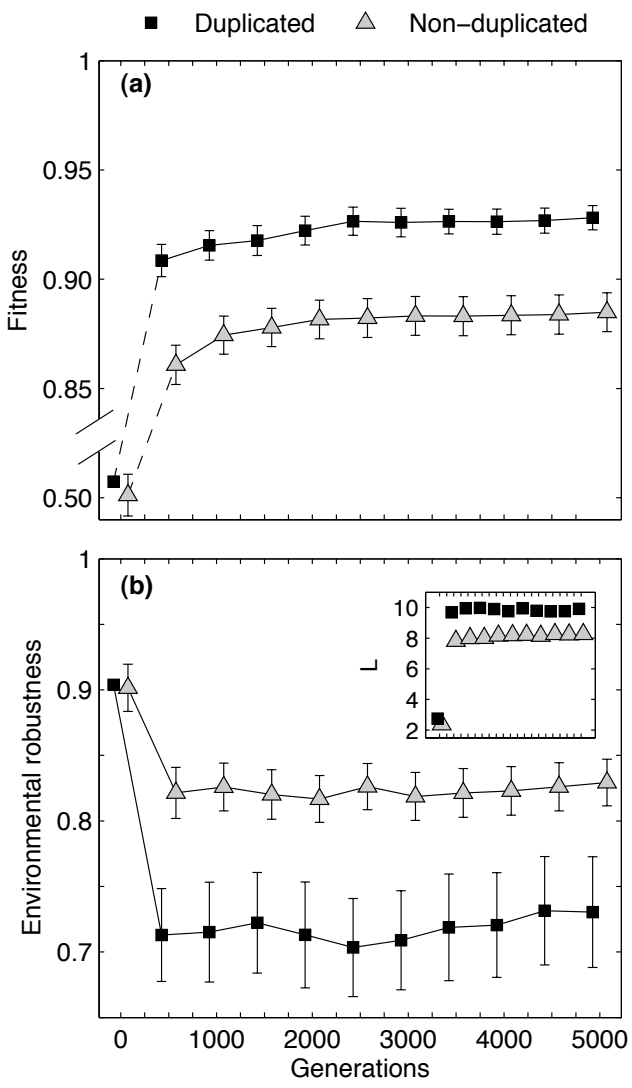

Figure 3: Evolutionary dynamics of (a) fitness and (b) environmental robustness for populations of duplicated and non-duplicated RBNs in novel environments. Data represent the mean of 100 independent replications and error-bars denote a single standard deviation. The inset in (b) depicts the average attractor length L. Data are deliberately offset in the horizontal dimension for visual clarity. Note the break in scale on the y-axis of (a). The scale of the x-axis is the same in all panels, including insets.

at 0.5 throughout the evolutionary process (data not shown).

Discussion

We have used Random Boolean Networks (RBNs) to investigate the influence of whole genome duplication (WGD) and subsequent diversification on evolutionary innovation and environmental robustness in gene regulatory networks (GRNs). There are some limitations to our approach that

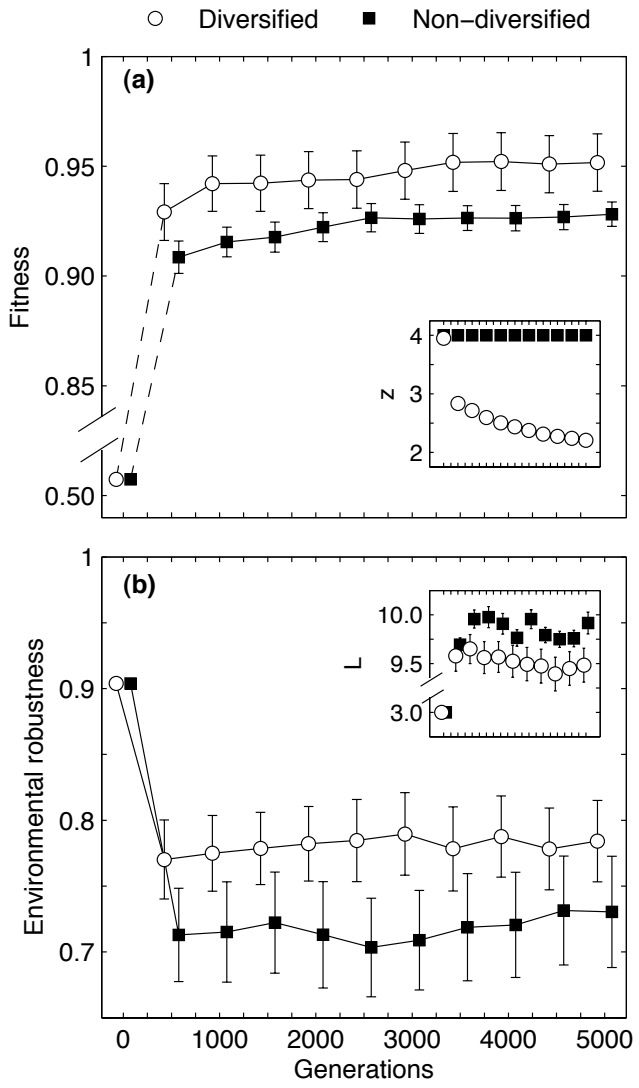

Figure 4: Evolutionary dynamics of (a) fitness and (b) environmental robustness for populations of RBNs following WGD in novel environments, with (open circles) and without (closed squares) subsequent diversification. Data represent the mean of 100 independent replications and error-bars denote a single standard deviation. The inset in (a) depicts the average network connectivity z and the inset in (b) depicts the average attractor length L. Data are deliberately offset in the horizontal dimension for visual clarity. Note that the data represented by the closed squares are the same as in Fig. 3. Also note the break in scale on the y-axis of (a) and the inset of (b). The scale of the x-axis is the same in all panels, including insets.

are worth highlighting. First, while the genotype produced by a WGD event is purely redundant, its resulting phenotype may differ immediately from that of the non-duplicated genotype. This occurs because both the original and duplicated nodes acquire new regulatory connections, necessitating the random initialization of entire segments of the expanded look-up tables. Further, the duplicated node cannot, under most circumstances, act as a "backup" because alterations to the original node may lead to phenotypic alterations that are not easily compensated for by the duplicate. Second, because many phenotypes yield identical fitness, and the phenotypic contribution of a single gene cannot be separated from its interaction partners, it is not possible to discern whether the observed changes in environmental robustness and evolutionary innovation are due to subfunctionalization, neofunctionalization, or a combination thereof (He and Zhang, 2005).

Despite these limitations, our analyses have helped to clarify the influence of WGD and subsequent diversification on environmental robustness and evolutionary innovation in GRNs. While deleterious in ancestral environments, WGD provided marginal fitness benefits in novel environments, coming at the expense of reduced environmental robustness (Fig. 2). Over evolutionary time, these differences magnified, with duplicated RBNs achieving significantly higher fitness and significantly lower environmental robustness than their non-duplicated counterparts (Fig. 3). Genetic diversification, via the loss of non-functioning regulatory interactions, was able to partly negotiate this trade-off, leading to improvements in both fitness and environmental robustness (Fig. 4).

Environmental robustness and evolutionary innovation were therefore inversely related in this system. This occurred because fitness assignment was based solely on the ability of a RBN to match a target expression profile σ_{target} (Eq. 1). This induced selection pressure for longer attractors (insets in Figs. 3b and 4b), because increasing the duration of the expression profile of the output node increased the probability that some segment of that profile matched σ_{target}. In turn, environmental robustness decreased, because longer attractors were more sensitive to perturbation. Thus, while some aspects of robustness and evolvability are positively correlated in RBNs (Aldana et al., 2007), robustness to environmental perturbation and the ability to match a target phenotype are not amongst them.

Diversification increased environmental robustness (Fig. 4b) through a reduction in network connectivity (Fig. 4a, inset). This shifted the RBN dynamics closer to the critical regime and therefore reduced the average attractor length (Fig. 4b, inset), yielding more environmentally robust attractors. It is notable that this reduction in attractor length did not lead to a corresponding reduction in fitness (Fig. 4a). How the diversified RBNs were able to attain increased fitness using shorter attractors is not yet known. An analysis of the structural properties of evolved RBNs, such as network excitation (Draghi and Wagner, 2009) or degree distribution (Aldana et al., 2007), may provide more insight into the mechanisms by which diversification can simultaneously

increases fitness and environmental robustness.

In the absence of diversification, the selective advantage of WGD may depend heavily on the frequency with which environmental perturbations occur. Selection may favor phenotypes that consistently yield expression profiles of average fitness over those that inconsistently yield expression profiles of high fitness. By placing non-duplicated and duplicated RBNs in a head-to-head competition under varying levels of environmental perturbation, future work will seek to determine how selection moderates the trade-off between environmental robustness and evolutionary innovation, and to discover the conditions under which selection leads to the "survival of the flattest" (Wilke et al., 2001).

The environments considered in this study were static, meaning that the target gene expression profile did not change over time. Several studies have demonstrated the importance of dynamic environments in shaping a population's potential for evolutionary innovation (Kashtan et al., 2007; Draghi and Wagner, 2009). Future work will seek to understand how WGD and subsequent diversification influence evolutionary innovation and robustness in dynamic environments.

Future work will also seek to expand upon our usage of fitness as a proxy for evolutionary innovation. It may prove insightful to analyze not only the ability to move toward a specific fitness optimum, but also the ability to move toward arbitrary fitness optima. Such measurements of the diversity of accessible phenotypes are common in studies of evolvability (Ciliberti et al., 2007a; Cowperthwaite et al., 2008; Wagner, 2008), and could be incorporated into our analysis (Draghi and Wagner, 2009). In addition, we will also investigate alternative forms of genetic diversification, with a particular focus on gene loss, which will allow for a more direct comparison with alternative models of WGD and diversification in GRNs (Wagner, 1996). These extensions, among others, will lead to a more thorough understanding of how various genetic mechanisms influence the relationship between robustness and evolvability in gene regulatory networks.

Acknowledgements

This work was partially supported by NIH grants R01-LM009012, R01-LM010098, and R01-AI59694. C.D. was supported by the Swiss National Science Foundation, grant PBLAP2-129416/1. J.L.P. was supported by NIH grant K25-CA134286.

References

Albert, R. (2005). Scale-free networks in cell biology. *Journal of Cell Science*, 118:4947–4957.

Aldana, M., Balleza, E., Kauffman, S., and Resendiz, O. (2007). Robustness and evolvability in genetic regulatory networks. *Journal of Theoretical Biology*, 245:433–448.

Aldana, M. and Cluzel, P. (2003). A natural class of robust networks. *Proceedings of the National Academy of Sciences*, 100:8710–8714.

Alon, U. (2007). *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman and Hall.

Bloom, J. D., Labthavikul, S. T., Otey, C. R., and Arnold, F. H. (2006). Protein stability promotes evolvability. *Proceedings of the National Academy of Sciences*, 103:5869–5874.

Ciliberti, S., Martin, O. C., and Wagner, A. (2007a). Innovation and robustness in complex regulatory gene networks. *Proceedings of the National Academy of Sciences*, 104:13591–13596.

Ciliberti, S., Martin, O. C., and Wagner, A. (2007b). Robustness can evolve gradually in complex regulatory gene networks with varying topology. *PLoS Computational Biology*, 3:e15.

Cowperthwaite, M. C., Economo, E. P., Harcombe, W. R., Miller, E. L., and Meyers, L. A. (2008). The ascent of the abundant: how mutational networks constrain evolution. *PLoS Computational Biology*, 4:e10000110.

Darabos, C., Tomassini, M., and Giacobini, M. (2009). Dynamics of unperturbed and noisy generalized Boolean networks. *Journal of Theoretical Biology*, 260:531–544.

De Bodt, S., Maere, S., and Van de Peer, Y. (2005). Genome duplication and the origin of the angiosperms. *TRENDS in Ecology and Evolution*, 20:591–597.

Draghi, J. and Wagner, G. (2009). The evolutionary dynamics of evolvability in a gene network model. *Journal of Evolutionary Biology*, 222:599–611.

Draghi, J. A., Parsons, T. L., Wagner, G. P., and Plotkin, J. B. (2010). Mutational robustness can facilitate adaptation. *Nature*, 463:353–355.

Ferrada, E. and Wagner, A. (2008). Protein robustness promotes evolutionary innovations on large evolutionary time-scales. *Proceedings of the Royal Society London B*, 275:1595–1602.

He, X. and Zhang, J. (2005). Rapid subfunctionalization accompanied by prolonged and substantial neofunctionalization in duplicate gene evolution. *Genetics*, 169:1157–1164.

Isalan, M., Lemerle, C., Michalodimitrakis, K., Horn, C., Beltrao, P., Raineri, E., Garriga-Canut, M., and Serrano, L. (2008). Evolvability and hierarchy in rewired bacterial gene networks. *Nature*, 452:840–846.

Kafri, R., Bar-Even, A., and Pilpel, Y. (2005). Transcription control reprogramming in genetic backup circuits. *Nature Genetics*, 37:295–299.

Kashtan, N., Noor, E., and Alon, U. (2007). Varying environments can speed up evolution. *Proceedings of the National Academy of Sciences*, 104:13711–13716.

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

Kauffman, S. A., Peterson, C., Samuelsson, B., and Troein, C. (2004). Genetic networks with canalyzing Boolean rules are always stable. *Proceedings of the National Academy of Sciences*, 101:17102–17107.

Kellis, M., Birren, B. W., and Lander, E. S. (2004). Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature*, 428:617–624.

Le Cun, Y., Denker, J. S., and Solla, S. A. (1990). Optimal brain damage. *Advances in Neural Information Processing Systems*, 2:598–605.

Ohno, S. (1970). *Evolution by Gene Duplication*. George Allen and Unwin, London, UK.

Oikonomou, P. and Cluzel, P. (2006). Effects of topology on network evolution. *Nature Physics*, 2:532–536.

Piškur, J. (2001). Origin of the duplicated regions in the yeast genomes. *Trends in Genetics*, 17:302–303.

Schuster, P., Fontana, W., Stadler, P. F., and Hofacker, I. L. (1994). From sequences to shapes and back: a case study in RNA secondary structures. *Proceedings of the Royal Society London B*, 255:279–284.

Sémon, M. and Wolfe, K. H. (2007). Consequences of genome duplication. *Current Opinion in Genetics and Development*, 17:505–512.

Taylor, J. S., Braasch, I., Frickey, T., Meyer, A., and Van de Peer, Y. (2003). Genome duplication, a trait shared by 22,000 species of ray-finned fish. *Genome Research*, 13:382–390.

van Hoek, M. J. A. and Hogeweg, P. (2009). Metabolic adaptation after whole genome duplication. *Molecular Biology and Evolution*, 26:2441–2453.

Wagner, A. (1996). Genetic redundancy caused by gene duplications and its evolution in networks of transcriptional regulators. *Biological Cybernetics*, 74:557–567.

Wagner, A. (2001). The yeast protein interaction network evolves rapidly and contains few redundant duplicate genes. *Molecular Biology and Evolution*, 18:1283–1292.

Wagner, A. (2005). *Robustness and Evolvability in Living Systems*. Princeton University Press.

Wagner, A. (2008). Robustness and evolvability: a paradox resolved. *Proceedings of the Royal Society London B*, 275:91–100.

Wilke, C. O., Wang, J. L., Ofria, C., Lenski, R. E., and Adami, C. (2001). Evolution of digital organisms at high mutation rates leads to survival of the flattest. *Nature*, 412:331–333.