DOI: http://dx.doi.org/10.7551/978-0-262-31709-2-ch051

Bipartite Networks Show the Genotype-to-Phenotype Relationship in Biological Systems Models: A Study of the Robustness, Evolvability, and Accessibility in Linear Cellular Automata

Christian Darabos†, Britney E. Graham, Ting Hu, and Jason H. Moore

The Geisel School of Medicine at Dartmouth College, Hanover, NH 03755, USA
† Christian.Darabos@dartmouth.edu

Abstract

In biological organisms, a single genotype may map to several phenotypes and vice-versa. This many-to-many relationship is believed to be a major drive of the phenotypic robustness and genotypic evolvability found in all life forms. Given the inherent complexity of the genotype-to-phenotype (G2P) mappings, we use cellular automata (CAs) as rudimentary proxies for biological organisms. CA models have the same many-to-many G2P mappings, and their sensitivity to initial conditions allows the same genotype to differentiate into different phenotypes. We use a bipartite network to study the G2P landscape, and its projections in either space. The degree distributions of the network and its projections are all *heavy-tailed*, denoting the presence of highly connected hubs, implying that increased robustness is supported by the network structure. We also show a strong correlation between the phenotype's complexity and its robustness. We analyze the relationships between the robustness and the evolvability both at the genotypic and phenotypic level. Although we use different computational models, our results agree with those of previous similar studies, and with observations in biological organisms.

Introduction

For the past two decades, geneticists have been studying the intricate genotype-to-phenotype (G2P) relationship in biological organisms. Genome-wide association studies (GWAS), and the recent advances in modern high throughput sequencing technologies, have made understanding how metabolic reactions, cell signaling, and developmental pathways translate the genome of a living organism into its phenotype an achievable goal (Nuzhdin et al., 2012). However, GWAS have also unveiled unprecedented degrees of complexity, making clinical progress much slower than anticipated. As geneticists learn more about G2P mappings, it becomes more apparent that there is a many-to-many relationship. Indeed, several different genotypes, usually resulting from small perturbations or *neutral mutations*, result in the exact same phenotype. This feature is responsible for the phenotypic robustness of biological organisms, and their relative insensitivity to small genetic perturbations. On the other hand, identical genotypes may develop into dramatically different phenotypes, depending on a set of internal and external signals and factors. The embryonic stem cell, which may potentially develop into any cell type, is a prime example of a single genotype yielding several phenotypes. The ability to adapt to internal and external factors is believed to be at major factor of the evolvability of all life forms. Given the inherent complexity of the G2P mappings, we recognize the need for smart, adaptive mathematical, statistical or computational models to study this relationship.

In this work, we use a cellular automata (CAs) model to exhaustively explore the G2P relationship in basic models of biological organisms. We are specifically interested in the nature of CAs, which in their simpler form mimmic the many-to-many G2P relationship, and their sensitivity to initial conditions other model systems fail to encapsulate. CAs have been thoroughly studied in the past. However, in the context of this project, we will focus on the exhaustive description of all possible genotypes, phenotypes, and initial conditions. The results we gain from this are structured in a bipartite network, which consists of two types of nodes, in our case: genotype and phenotype. We then look at the projections of our bipartite network onto the genotype space and the phenotype space. Additionally, we study the distribution of robustness (also called neutrality) (Banzhaf et al., 1994) in the genotypic landscape of our model, and its effect on the phenotypic landscape. Similarly, we look at the genotypic and phenotypic evolvability, and the correlations between robustness and evolvability. Indeed, the seemingly contradictory effect of robustness and evolvability has been studied and disproved in many systems, where they in fact facilitate each other (Bloom et al., 2006; Ferrada and Wagner, 2008; Wagner, 2008a, 2005).

Background

Cellular Automata (CAs)

CAs (Codd, 1968) are dynamical, usually deterministic, discrete, abstract models used to simulate and study distributed computation. A standard CA consists of a finite number N of identical cells. Each cell can take one of a finite number of states s, here, the two Boolean states $s \in \{0, 1\}$. Each cell has a local knowledge of the state of a fixed num-

ber of n neighboring cells, including itself. The state of each cell is updated synchronously in discrete time steps, according to a local, identical update function or rule (these terms will be used interchangeably throughout this work). In all CAs, update functions are generally represented as a Boolean lookup table of all possible binary permutations of the cell's neighborhood. Cells are usually arranged on a d-dimensional grid, where usually $d \in \{1, 2, 3\}$. In the present case of a one-dimensional, or linear CAs, cells are arranged on a regular ring structure, connecting to a radius of r cells on each side. Thus, the neighborhood size is $n = 2r + 1$. At any given *discrete time step* t, the ensemble of all states s_i^t of all cells is called the *configuration* of the CA such that $c^t = (s_0^t, s_1^t, \ldots, s_{N-1}^t)$, thus CAs with N nodes have exactly 2^N possible configurations. Starting from an initial configuration (IC or c^0) at time $t = 0$, the CA will travel across transient configurations before reaching a previously visited state of the system. Because of its deterministic nature, the CA will, after at most 2^N time steps, start cycling deterministically through a subset of configurations, called an *attractor*. Figure 1 shows a small example of a 8 cell regular uniform CA, set to a random IC.

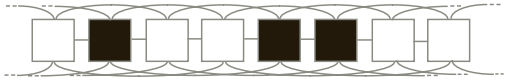


Figure 1: Regular uniform CA. Size $N = 8$, with a radius $r = 3$, and set to a random (initial) configuration. Black cells represent state 1 and white is 0. Dashed edges "wrap around" to connect the CA into a ring topology.

CA Genotype & Phenotype CAs have been used for years as a rudimentary proxy for biological organisms and phenomena. One prevalent example, using a generalized form of CAs, is Kauffmann's Random Boolean Network (RBN) model for genetic regulatory networks (Kauffman, 1969). RBNs use non-uniform, unstructured CAs where each cell has its own Boolean update function (BUF) and can arbitrarily be connected to any other cell. Other models, such as Genetic Programming, Bayesian systems, or differential equations have also been thoroughly studied. CAs possess, however, a modest advantage. Indeed, CAs have a genotype, a phenotype, and mimic the many-to-many G2P mappings. The update function of CAs is a direct equivalent of a genotype, which can be mutated at will, and is a set of rules followed by the system to achieve a steady state. The attractor reached by the CAs is the phenotype resulting from a genotype and an initial configuration. The same attractor can be reached by different BUFs, and a single BUF can result in different attractors depending on the IC. In this work, we explore the evolvability, robustness, and accessibility of pseudo-biological organisms modeled by a small CA. We exhaustively explore all G2P mappings of a CA by representing it in a bipartite network, and also projecting it onto the phenotype and genotype landscape respectively.

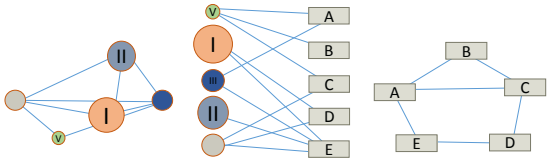
Alternative methods have been used to map the G2P relationship, such as genetic programming Hu et al. (2011), and random Boolean networks Kauffman (1969). Moreover, one can imagine using almost any systems where a genotype and a phenotype can be described, for example NK-landscapes Kauffman and Weinberger (1989); Ochoa et al. (2008), and many more. However, we determined that CAs were the best suited tools to simulate the many-to-many G2P mappings, while keeping the analytical simplicity.

Network Properties

A CA can be seen as a mathematical object known as a graph, where each cell resides on a vertex, and edges between vertices represent two neighboring cells. Therefore, formal definitions of graph theory do apply to CAs. For ease of reference, we summarize concepts used in subsequent sections particular to this work (see (Newman, 2010) for complete reference). In this work, a graph G, or network, consists of a set of v vertices V, and a set of e undirected, unweighted edges E. The degree k of a vertex is the number of edges connected to it. Thus the average degree $\bar{k}$ of G is the average of the degree over V. A path between vertices u and v is defined as the sequence of unique edges traversed when going from u to v. Its length is the number of edges in the sequence. The average path length (APL) of G is the average length of the *shortest* path between all pairs of vertices. The clustering coefficient C_j of a vertex j is defined as the ratio between the E_j edges that actually exist between the k_j neighbors of j and the number of possible edges between these nodes: $C_j = 2E_j/k_j(k_j - 1)$. The clustering coefficient (CC) of a graph is the average of C_j over V. C is thus independent of N for a regular lattice, and approaches $3/4$ as k increases. The degree distribution $P(k)$ of a graph G is a function that gives the probability that a randomly selected vertex has k edges incident to it. For a random graph $P(k)$ is a binomial peaked at $P(k)$. But most real networks do not show this kind of behavior. In particular, in scale-free graphs which are frequent in real-life, $P(k)$ follows a power-law distribution.

Bipartite Network A bipartite network consists of two disjoint sets, or types, of nodes U and V. The nodes are connected in such a way that the nodes of one set will have no connections between them, but can only be connected to nodes of the other set. The use of a bipartite network is natural when dealing with two different types of data sets (see Figure 2b). Two nodes of the same type cannot be connected to each other, so one node can only be connected to a node of the other data type. We are interested in using a bipartite network to represent the relationship of our data.

From the bipartite network, one can project the data onto either the space U or V (Figure 2a,c). In either single dataset



(a) (b) (c)

Figure 2: Bipartite Network schematic. A bipartite network (b) made of two data sets the circles, U, and the rectangles, V. Projections in U (a) and in V (c).

space, the nodes are connected to one another "through" a vertex of the the other space. By ignoring the different types of data, all network properties described above remain valid on the bipartite network (as a single data set network) and on either projection. This type of network gives us three degree distributions, one for each projection, and one for the bipartite network. Each degree distribution shows how many links each node has. Nodes in a projection of a bipartite network are connected if they share at least one node in the other group. This gives us the ability to see the interactions within a group.

Methods

In order to fully explore the G2P relationship in our CA model, we exhaustively explore all possible genotype mappings for all possible ICs. Unfortunately, the (super)exponential nature of the genotype and phenotype spaces, we are limited to a small number of cells, $N = 5$, and a *radius* $r = 1$, where the radius defines the number of neighbors each cell arranged on a ring can reach on either side. Therefore, a radius $r = 1$ results in a neighborhood size of $n = 3$. In CAs, there are $2^{2^n} = 2^{2^3} = 256$ possible genotypes. This is the limit, as the next possible regular neighborhood size being $n = 5$, which results in $\approx 4.3 \times 10^9$ genotypes. This figure is computationally too expensive to search exhaustively. CAs have $2^N = 2^5 = 32$ possible ICs, and the same number of possible point (i.e. single configuration) attractors, and at most $2^{2^n} \times 2^N = 8192$ possible attractors of any length, as every combination of genotype and IC can potentially result in a different phenotype. We have successfully simulated CAs of size up to $N = 12$, the resulting networks are however not suited for representation.

Because we have two types of data, we can build a bipartite network. In our case, one set of nodes represents genotypes and the other, phenotypes. Our network is directed and shows the decent of a phenotype from a genotype. The degree distribution of the set of genotypes gives the number of phenotypes that are associated with each genotype and vice versa.

From our bipartite network, we can build the 2 other networks, one of each set of nodes. From this, we obtain a phenotype network and a genotype network. The nodes in each of these networks are connected if, in the bipartite network, they share at least one node of the other type. For example, in our phenotype network, two nodes are connected if they share at least on genotype. We carry out the same process for the genotype network.

Robustness, Evolvability, and Accessibility

Several measures of robustness and evolvability exist in the literature, at both the genotypic and phenotypic scales. Following (Wagner, 2008b), we define genotypic robustness as the fraction of the total number of possible point mutations to a given genotype that are neutral. Genotypic evolvability is defined as the fraction of the total number of possible phenotypes that are accessible through non-neutral point mutations to a single genotype. Phenotypic robustness is defined as the size of the phenotype's underlying genotype network. For phenotypic evolvability, the proportion of the total number of phenotypes that can be reached via non-neutral point mutations from a given phenotype.

In addition to measuring the propensity to mutate away from a phenotype, we also measure phenotypic accessibility (Cowperthwaite et al., 2008), denoted by A, which is formally defined as:

$$A_i = \sum_j f_{ij},$$

where

$$f_{ij} = \begin{cases} \frac{v_{ij}}{\sum_{k \neq i} v_{ik}}, & \text{if } i \neq j \\ 0, & \text{if } i = j \end{cases}$$

This metric takes on high values if phenotype i is relatively easy to access from other phenotypes, and low values otherwise.

Results

This section describes the analyses resulting from the exhaustive simulation of the G2P relationship under all possible initial conditions in a small CA. We begin by describing the G2P spaces in the form of its bipartite network and its projections, and studying the networks' statistical properties. The second part of this section focuses on the correlations between the genotypic and phenotypic robustness, evolvability, and accessibility.

Bipartite Network of G2P

We start by representing the G2P relationship in our CA model as a bipartite network. Figure 3 represents this bipartite network, the degree distributions, and both projections. Figure 3a shows the bipartite network, where genotypes (on top) only connect to phenotypes (on the bottom). To the best of our knowledge, this is a novel perspective on representing the G2P mappings. In all network Figures 3a-c-d, the size of the vertex is proportional to the number of members of

the opposite dataset associated. In other words, the genotypes vertices are proportional to the number of phenotypes mapped, and vice-versa. For readability reasons, we have filtered out vertices of a degree below 5. In the same figure, we also show the trends of degree distributions for the bipartite network, as well as both projections of the genotypes only and of the phenotypes only.

We identify the two most evolved phenotypes as the two single-point homogeneous phenotypes (00000 and 11111). The next most evolved phenotype is a two-state attractor made from the previous two. As expected, the identity update function (01010101) is the genotype that maps to the most phenotypes. In Figure 3b, all degree distributions are right skewed, with a heavy tail, denoting the presence of highly connected "hub" vertices and a "scale-free" like topology. In other words, the degree distribution of the network decays like a power-law or exponential function.

Beyond the degree distributions, we are also interested in showing a few more statistical characteristics of the networks, and how they differ between the bipartite network and the projections. Table 1 summerizes these measurements, described in the Methods Section.

network	bipartite	genotype	phenotype
#vertices	395	256	139
#edges	1398	1024	1268
$\bar{k}$	3.539	8.0	18.245
CC	0	0	0.816
APL	1	4.016	2.102

Table 1: Networks statistics for the bipartite network as a whole, and both the projection on the genotypes space and on the phenotypes space.

As we can see in Table 1, the bipartite network regroups all genotypes and phenotypes. Interestingly, we note that our networks are generally dense, with high $\bar{k}$ in the projections. In addition, we show that there is no clustering structure to speak of in the genotype network, and that the phenotype network is highly clustered. Finally, the phenotype network has a short APL. The high CC and the short APL of the phenotype network hints at interesting community properties in the phenotypes, which clusters phenotypes that are densely connected in subsets that share common biological or genetic information. This finding is in line with a previous, much more applied study on the Human Phenotype Network (HPN). In Darabos et al. (2013), we report that the HPN does have a strong community structure, and we are now reassured to note that the CA model phenotype network shares this characteristic.

Robustness, Evolvability, and Accessibility

In this section, we analyze the complex relationships between robustness and evolvability, both genotypic and phenotypic. We also report results of the influence of phenotypic accessibility. These links are, we believe, the most biologically relevant and could confer the most relevance to our model. In order to conduct this analysis, we study the statistical characteristics of the genotype and phenotype spaces, assigning robustness, evolvability, and accessibility "scores" to each genotype and phenotype, according to the description found in the Methods Section.

Firstly, we report in Figure 4 the strong, quasi linear positive correlation between the number of phenotypes mapping to a genotype, and the evolvability of those genotypes. This correlation is to be intuitively expected from biological organisms in which genotypes responsible for more phenotypes are also considered the most evolvable.

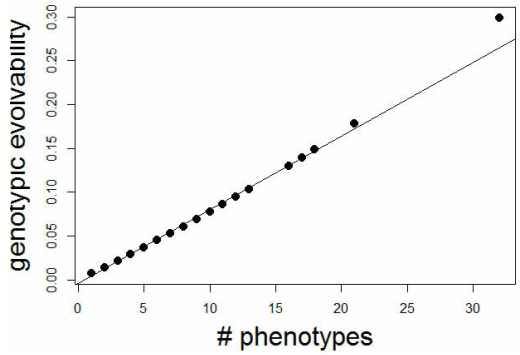


Figure 4: Strong correlation between the number of phenotypes mapped from a genotype and the genotypic evolvability.

One interesting aspect of robustness and evolvability in all systems, is their relationship to the perceived complexity of the phenotype. In our case, the complexity of the phenotype is measured by the size (or length) of the attractor. These relationships are depicted in Figure 5.

In Figure 5, we report a strong negative correlation between the length of the phenotype's attractor and its robustness. This result agrees with Kauffman's work on RBNs (Kauffman, 1969), who early on has reported that longer attractors are less stable in RBN systems. The same negative correlation appears with phenotypic evolvability. We speculate that more complex phenotypes have more difficulties accessing other phenotypes because of the basin of the attraction size. This is confirmed by the fact that more robust phenotypes are also less accessible (see below, Figure 7d).

We are especially interested in genetic robustness and evolvability and how they are distributed over the entire genotype space. These results are reported in Figure 6a-b. Moreover, we study the relationship between genetic robustness and evolvability in Figure 6c.

The "bell shaped" distributions of genotypic and phenotypic robustness (not shown here), or neutrality, are also aligned with results in similar studies (Hu et al., 2011). Most

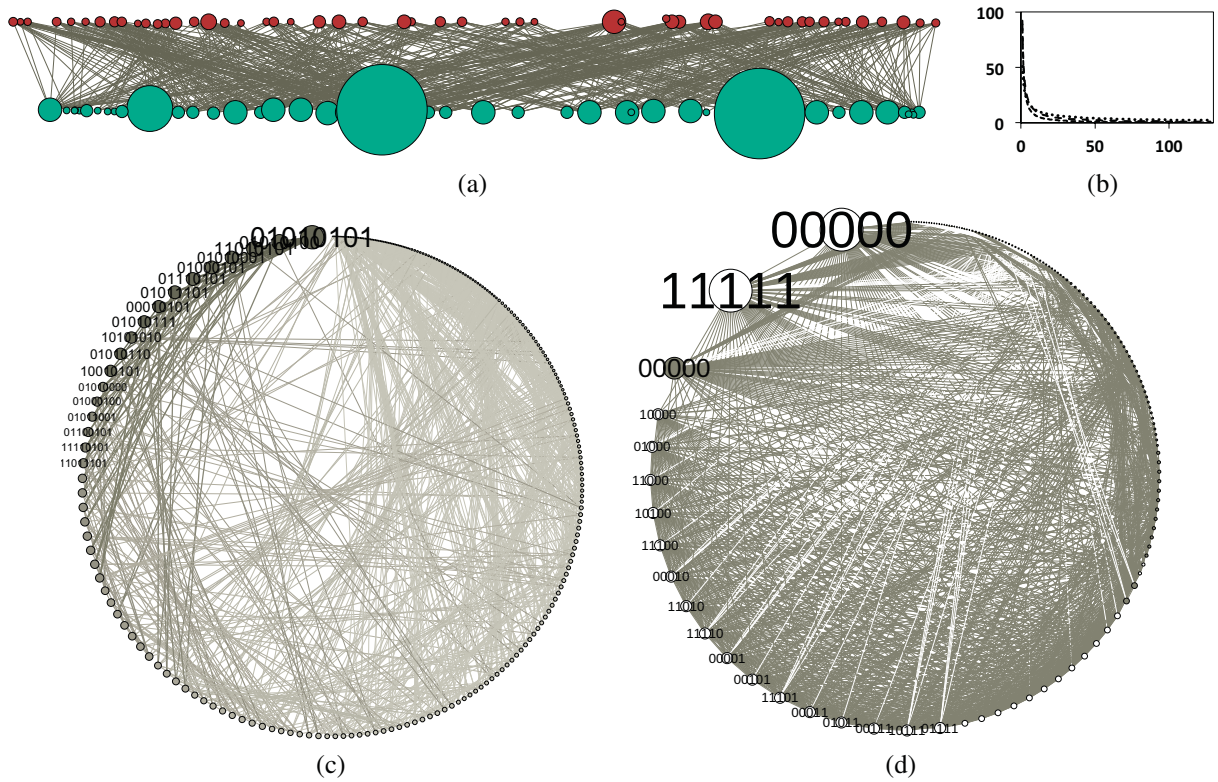


Figure 3: Filtered Bipartite G2P Network, Degree Distributions, and Projections. (a) The top row vertices represent the genotypes, and the bottom row vertices represent the mapped phenotypes. The vertex size is proportional to the degree (i.e. to the number of mapped phenotypes, or mapping genotypes respectively) Vertices of a degree below five are omitted for readability reasons. (b) the "heavy-tailed" degree distribution for the bipartite network and both phenotype and genotype projections. (c) the projection in the genotype space, the vertex sizes are proportional to the number of associated phenotypes, the darker the edge, the more phenotypes the genotypes have in common. (d) the projection in the phenotype space, with vertex sizes proportional to the number of mapping genotypes. The vertex is white if the phenotype is a point attractor, and grey if the attractor is longer (i.e. the phenotype is more complex). The darker the edge, the more genotypes the phenotypes have in common.

genotypes have a high robustness, and a rather low evolvability. However, the genetic evolvability distribution is right-skewed, with a heavy tail, therefore, a small number of genotypes have a very high evolvability. Interestingly, genotypes which have a middle range value robustness (≈ 0.5) tend to show a higher degree of evolvability. In other words, high or low genetic robustness yield lower evolvability than mid-range values.

Relating this research to biological organisms, we are particularly interested in the phenotypic aspect of the CA models and how they respond to external (i.e. environmental) and internal (i.e. genotypic) perturbations, and how these perturbations impact robustness. Therefore, we consider the phenotypic robustness of our systems, and study its relationships to genotypic and phenotypic evolvability, genotypic robustness, and the accessibility of the phenotypes. These results are reported in Figure 7.

The first, and natural correlation we are interested in is phenotypic robustness to evolvability. Figure 7a shows a strong correlation between phenotypic robustness and evolvability. This means that phenotypes that are more robust are also more evolvable, and vice-versa. Although it seems counterintuitive that more stable phenotypes are also more evolvable, our findings are in perfect agreement with Wagner's conjecture and work, in which stability in biological organisms actually supports evolvability (Wagner, 2008b). In Figure 7b, we see that more robust genotypes do not give rise to more robust phenotypes, and vice-versa. This is not actually surprising, although it is not in line with Hu's work (Hu et al., 2011). In our work, the robustness of the genotype does not guarantee that of the phenotype. Panel (c) shows that phenotypic robustness and genotypic evolvability are moderately negatively correlated, which indicates that more stable phenotypes are obtained by genotypes that can give rise to fewer phenotypes after a small perturbation. Finally, Figure 7d reveals that phenotypes that are the least accessible from other phenotypes, through small mutation of their genotypes, are also the most stable. This implies that ran-

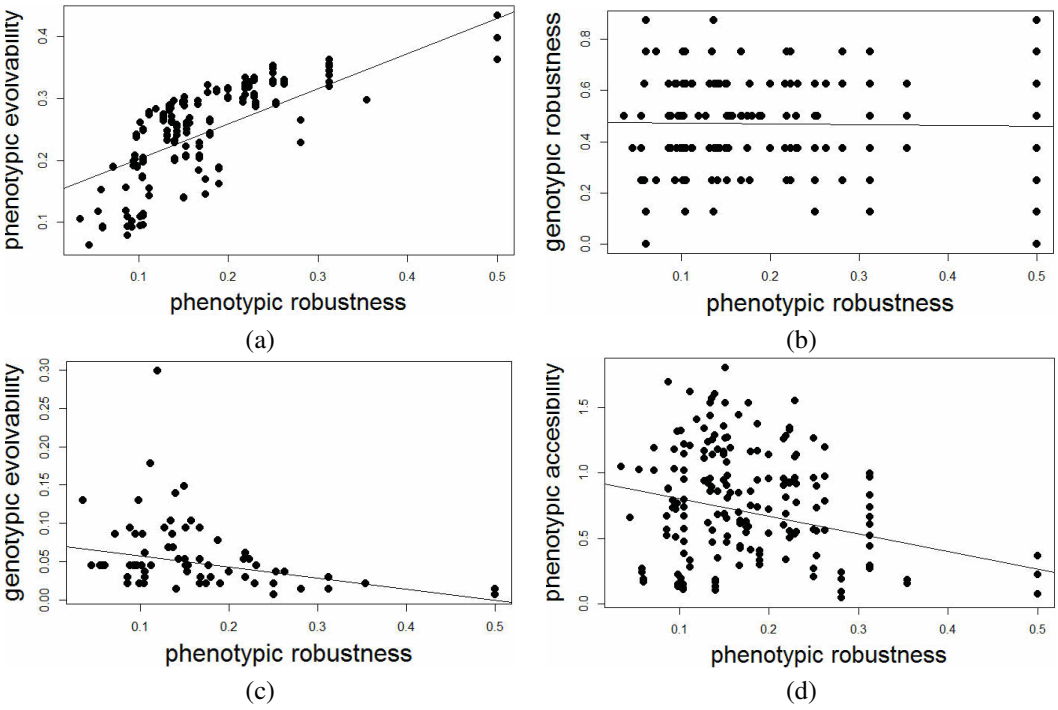


Figure 7: Phenotypic Robustness in relationship to (a) genotypic evolvability ($r = -0.46$), (b) genotypic robustness ($r = -0.02$), (c) phenotypic evolvability ($r = 0.79$), and (d) phenotypic accessibility ($r = -0.37$). We show trend lines on all panels and specify the correlation coefficients (r) in parenthesis.

dom mutations are less likely to lead to robust phenotypes than to non-robust phenotypes. Combining the observations in panels (a) and (d), we hypothesize that the most robust phenotypes are difficult to find (Figure 7d) and highly evolvable (Figure 7a).

Conclusions

In order to explore the complex genotypes-to-phenotypes (G2P) relationship in biological organisms, we propose to use small, simplistic, proxies in the form of a Cellular Automata (CAs) model. CAs share the many-to-many G2P mappings of living organisms. In this implementation, we keep the systems small enough so that we can exhaustively explore both the genotype and the phenotype space, and all initial conditions. To visually assess the G2P associations, we use a bipartite network and its projections onto either space. It shows a dense network, with some prominent genotypes that can access all phenotypes, and some highly popular phenotypes that can be reached by all genotypes.

Genotypes that are associated with the most phenotypes are also the most evolvable, which suggests that evolvable genotypes cluster together and tend to be genetically close. In agreement with Kauffman's work on RBNs, the phenotype's complexity (i.e. attractor length) is negatively correlated with both its evolvability and its robustness. In addition, we conclude that the more evolvable genotypes yield a robustness in the middle of the range of possible values. These results suggest that some robustness is necessary to promote evolvability in genotypes, but too much robustness will hinder genotypic evolvability.

Finally, we showed that phenotypic robustness and evolvability are closely related, which agrees with Wagner's results, and that phenotypic robustness supports phenotypic evolvability. Despite its counterintuitive nature, this finding is aligned with biological systems. We also see that stable phenotypes are not generated by stable or evolvable genotypes. Moreover, we note that popular (accessible) phenotypes tend to be less robust, which contradicts some evidence in natural systems.

Robustness, evolvability and accessibility correlations, or lack thereof, are highly system-dependent. In the case of CAs, we show that these relationships are, generally, in line with observations in biological organisms. However, the robustness-evolvability-accessibility relationship in biological organisms is also highly system-dependent. Nevertheless, models are unavoidable when studying interactions as complex as G2P, and CAs remain a viable option, considering that they also model the many-to-many G2P relationship.

In a future line a research, we plan on exploring the G2P relationships in much larger CAs in order to inch closer to simple biological organisms. Unfortunately, this will come

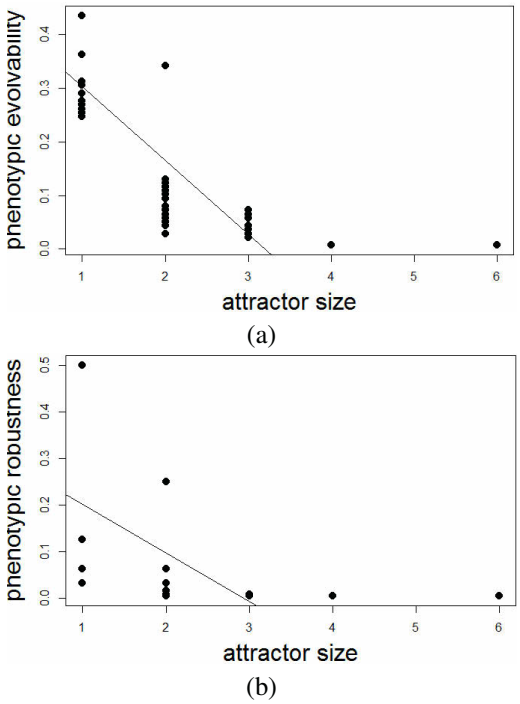


Figure 5: The effect of Attractor size on (a) phenotypic evolvability ($r = -0.74$) and (b) phenotypic robustness ($r = -0.39$). We show trend lines on all panels and specify correlation coefficients (r) in parenthesis.

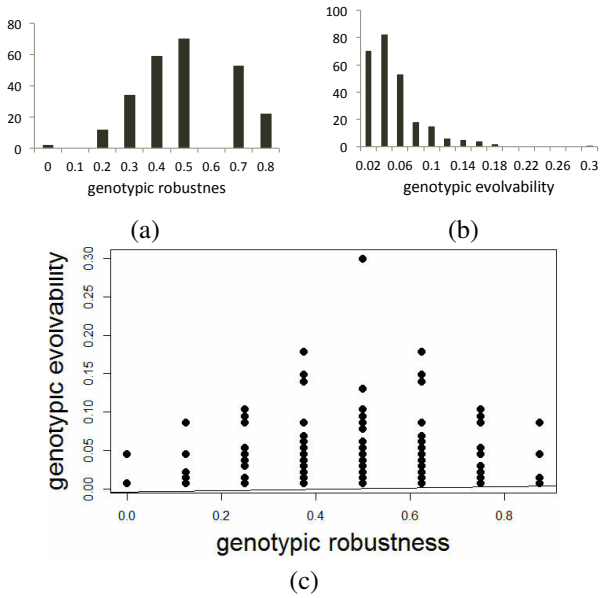


Figure 6: Genetic Robustness and Evolvability and their relationship to one another. Top row: distribution of (a) genetic robustness and (b) genetic evolvability over all possible genotypes. Bottom figure: the relationship, with a slight positive correlation between robustness and evolvability (correlation coefficient $r = 0.10$).

at the cost of the exhaustive search, and force us to select a limited number of sample genotypes and phenotypes mappings. Moreover, we are planning on addressing the phenotypic plasticity, or adaptability of the phenotype during its lifetime, and explore ways to integrate this aspect into our CA model without loosing its attractive simplicity. Finally, it would be interesting to compare our results with a real-life biological case, if the quality of the data permits.

Acknowledgements

This work was partially supported by NIH grants AI59694, LM010098 and LM009012.

References

Banzhaf, W., Davidor, Y., p. Schwefel, H., and (eds, R. M. (1994). Genotype-phenotype-mapping and neutral variation - a case study in genetic programming.

Bloom, J. D., Labthavikul, S. T., Otey, C. R., and Arnold, F. H. (2006). Protein stability promotes evolvability. *Proc Natl Acad Sci U S A*, 103(15):5869–5874.

Codd, E. F. (1968). *Cellular Automata*. Academic Press, New York.

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 Comput Biol*, 4(7):e1000110.

Darabos, C., Desai, K., Cowper-Sallari, R., Giacobini, M., Graham, B., Lupien, M., and Moore, J. (2013). Inferring human phenotype networks from genome-wide genetic associations. In Vanneschi, L., Bush, W., and Giacobini, M., editors, *Evolutionary Computation, Machine Learning and Data Mining in Bioinformatics*, volume 7833 of *Lecture Notes in Computer Science*, pages 23–34. Springer Berlin Heidelberg.

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.

Hu, T., Payne, J. L., Banzhaf, W., and Moore, J. H. (2011). Robustness, evolvability, and accessibility in linear genetic programming. In *Proceedings of the 14th European conference on Genetic programming*, EuroGP'11, pages 13–24, Berlin, Heidelberg. Springer-Verlag.

Kauffman, S. A. (1969). Metabolic stability and epigenesis in randomly constructed genetic nets. *J. Theor. Biol.*, 22:437–467.

Kauffman, S. A. and Weinberger, E. D. (1989). The NK model of rugged fitness landscapes and its application to maturation of the immune response. *Journal of Theoretical Biology*, 141:211–245.

Newman, M. (2010). *Networks: An Introduction*. Oxford University Press, Inc., New York, NY, USA.

Nuzhdin, S. V., Friesen, M. L., and McIntyre, L. M. (2012). Genotype-phenotype mapping in a post-gwas world. *Trends Genet*, 28(9):421–426.

Ochoa, G., Tomassini, M., Vérel, S., and Darabos, C. (2008). A study of nk landscapes' basins and local optima networks. *CoRR*, abs/0810.3484.

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

Wagner, A. (2008a). Neutralism and selectionism: a network-based reconciliation. *Nat Rev Genet*, 9(12):965–974.

Wagner, A. (2008b). Robustness and evolvability: a paradox resolved. *Proc Biol Sci*, 275(1630):91–100.