IJCSS International Journal of Complex Systems in Science

*Int. J. Complex Systems in Science*
vol. 1 (2011), pp. 42–46

# Determining evolutionary dynamics in complex networks using genetic algorithms

D. Aguilar[1,†], A. Córdoba[1] and M.C. Lemos[1]

[1] *Departamento de Física de la Materia Condensada, Universidad de Sevilla*

**Abstract.** Gene regulatory networks are one of the most important goals in the novel discipline of system biology. These regulatory networks, through the interaction of multiple genes, control and guide the proteic interactions and, in fact, the cellular behaviour. Understanding this regulation is, therefore, essential in the investigation in organogenesis, during the embryonic stages of the organisms, and in the research of genetic diseases. The whole set of simple genetic interactions inside the network converges in complex behaviours, so it is imperative to analyse the problem in the field of network theory and the evolutionary dynamic of complex systems. All this has led us to investigate the evolutionary dynamics in generic networks, this way, the results can be used in experimental researches in the field of system biology. This research aims to decode the evolutionary rules governing the transformation dynamics in a network. To do this, a genetic algorithm has been implemented, in which, starting from initial and ending network stages, it is possible to determine the transformation dynamics between these stages by using simple acting rules. The network description is the following: a) The network nodes states in the initial and ending stages can be active or inactive; b) The network links can act as activators or repressors; c) An evolutionary set of rules is established in order to transform the initial stage into the ending one; d) Due to the low connectivity in gene regulatory networks, each node will hold a maximum of three inputs with no restriction on outputs. The "chromosomes" of the genetic algorithm includes two parts, one related to the nodes links and another related to the evolutionary rules. The implemented evolutionary rules are based in certain genetic interactions behaviour. This set of rules includes the following: 1) Rule of Most: If the target node inputs consist on activators and represors, the node final state is determined by the action of most. 2) Absolute repressor rule: In this case, just one repressor is enough to inhibit the target and deactivate the node. 3) Logic gate AND with activators: The target node will be active only if the input consists on, at least, two activator links. 4) Logic gate AND with repressors: The target node will be inactive only if the input consists on, at least, two repressor links. These rules and their combinations are compound by logic conditions and set the bases to the network motifs formation, which are the building blocks of the network dynamics. The implemented algorithm results discovered satisfactorily appropriate dynamics in complex networks evolution between different stages for several cases.

*Keywords:* Systems Biology, Complex Networks, Evolutionary Dynamics, Genetic Algorithms
*MSC 2000:* 92C42, 92B05, 92B20, 90C35

## 1. Introduction

Gene regulatory networks and its evolution can be studied inside the complex network field. This problem is too complex to resolve it directly, so it should be divided in small modules or blocks which constitute the network construction units, or motifs[1]. These building blocks are based on simple logical rules which interactions in the complete motifs network leads to complex behaviours[2]. Once defined the acting rules it is required to develop a method to recognize these rules inside the complete network and identify its transformation dynamics.

## 2. The Model

A genetic algorithm (GA)[3] has been developed to reveal the transformation dynamics between both known initial ($A$) and final ($B$) stages through the acting of simple rules. The network nodes can be active or inactive in each stage. The links can be activators or repressors. The connectivity in gene regulatory networks is low so each node has up to 3 input links with no output restriction. The following network consists of 32 nodes and 75 links (Fig. 1.a).
***Genetic Algorithm.*** The GA optimizes the following parameters: nodes state in $B$, links among the nodes and applied rules on each node. The GA "chromosome" consists of the links and the rules. Each individual of the initial population begins from stage $A$, and is transformed through the evolutionary dynamics of the associated rules set leading to the final stage $B$. The evolution rules are the following: 1) Rule of Most: The node final state is decided by the action of most, remaining unchanged when equal contribution. 2) Absolute repressor: Just one repressor inhibits the target and deactivate the node. 3) Activator logic AND gate. 4) Repressor logic AND gate.
***Fitness.*** The GA fitness function is compound by two terms; the distance between the known stage $B$ and the GA's, $B^*$; the distance between the known links and the GA links.

$$\phi = \lambda \cdot dist[B, B^*] + (1 - \lambda) \cdot dist[links, links^*] \; ; \; Fitness = 1 - \frac{\phi}{\phi_{max}} \quad (1)$$

The new generation of individuals will keep just the best fitness solution from the previous generation. The rest are generated by crossover and mutation.
***Crossover.*** The individuals with better fitness are more likely to form new solutions (*roulette method*). A crossover point is needed for each "chromosome" related part, randomly in the links part, and defined in the rules part by the expression: $K = [l - N \cdot int(\frac{l}{N})]$, where $K$ is the rules crossover point, $l$ is the links crossover point and $N$ is the network nodes number. The new

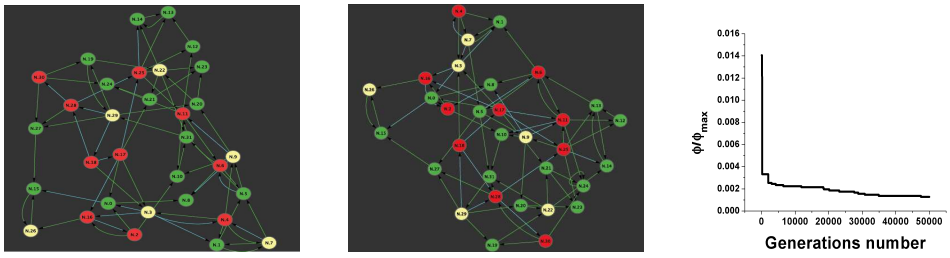


Figure 1: Left to right: (a) Initial Network: Nodes: $Red = A$, $Green = B$, $Yellow = Common$. Links: $Green = Activator$, $Blue = Repressor$; (b) GA Network; (c) Best solution fitness.

| Node | N.0 | N.1 | N.2 | N.3 | N.4 | N.5 | N.6 | N.7 | N.8 | N.9 | N.10 | N.11 | N.12 | N.13 | N.14 | N.15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Rule | 1 | 1 | 2 | 3 | 2 | 3 | 2 | 4 | 1 | 2 | 1 | 2 | 1 | 1 | 1 | 1 |
| Node | N.16 | N.17 | N.18 | N.19 | N.20 | N.21 | N.22 | N.23 | N.24 | N.25 | N.26 | N.27 | N.28 | N.29 | N.30 | N.31 |
| Rule | 2 | 2 | 2 | 1 | 3 | 1 | 4 | 1 | 1 | 2 | 1 | 1 | 2 | 3 | 1 | 1 |

Table 1: Rules applied to each node to lead stage A to stage B

individuals should not exceed a critical links number in order to maintain the low connectivity.

***Mutation.*** Two different mutation probabilities are set. The links one $\mu_1 = 0.001$ and the rules one $\mu_2 = 0.5$.

***Results.*** A random network (Fig. 1.a) execution with a population of 100 individuals concludes after 50.000 generations in a new network (Fig. 1.b) with a fitness value of 99.87% (Fig. 1.c) and a transformation rules set to lead $A$ to $B$ stage (see Table 1). Eight "false positives" and five new links were found.

The resulting set of rules is part of the possible solutions family that allows the correct stage transformation.

## 3. Systems biology application

The model is applied to the evolution of a gene regulatory network between two different embryonic stages in the *mus musculus* mouse eye development, this is, 9 (E9) and 10.5 (E10.5) days gestation. The network is built from experimental retinome genetic expression data and a wide bibliographic review[4]. It consists of 92 interactions among 32 differentially expressed genes (see Fig. 2.a). An execution with a population of 100 individuals and 200.000 generations leads to a network (Fig. 2.b) with a fitness value of 99.92% (Fig. 2.c) and a transformation rules set to turn E9 ($A$) to E10.5 ($B$) stage (see Table 2). Four false positives and four new links have been determined.

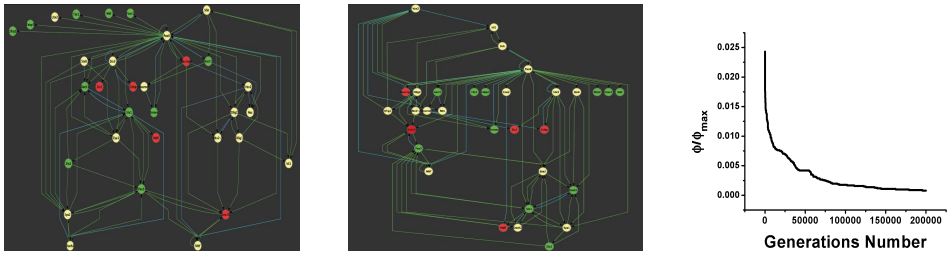


Figure 2: Left to right: (a) Experimental network; (b) GA Network; (c) Best solution fitness.

| Gene | Pax6 | Tbr1 | Sfrp2 | Neurog3 | Fabp7 | Sox10 | Foxc1 | Eomes | Mitf | Lhx2 | Ascl1 | Maf | Nes | Six3 | Eya2 | Zic2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Rule | 1 | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 3 | 1 | 3 | 1 | 4 | 4 | 2 | 2 |
| Gene | Cnbp | Six2 | Eya1 | Six1 | Six6 | Dach2 | Mdfi | Pax3 | Dach1 | Olig2 | Shh | Nkx2.2 | Isl1 | Neurod1 | Vsx2 | Olig1 |
| Rule | 2 | 1 | 2 | 1 | 4 | 2 | 2 | 1 | 1 | 4 | 4 | 1 | 4 | 1 | 3 | 2 |

Table 2: Rules applied to each gene to lead stage E9 to stage E10.5

## 4. Conclusions

The study of the evolutionary dynamics in complex networks through genetic algorithms reveals an appropriate and promising method for future research. The modelled genetic algorithm provides a way to determine a fitting dynamics between different stages. Its application to systems biology determines the evolution of gene regulatory networks and provides experimentally testable data to supply new unknown interactions between nodes finding false positives.

## Acknowledgements

To Dr. Fernando Casares and M$^a$ Ángeles Domínguez from CABD (CSIC) for their contributions and advices. This work has been partially financed by project P06-TIC-02025 of Junta de Andalucía and FIS-2008-04120 of the Spanish Government.

## References

[1] A.-L. Barabási, Z. N. Oltvai AND S. Wuchty, *Lect. Notes Phys.* **650**, 443-457 (2004).

[2] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, AND U. Alon, *Science* **298**, 824-827 (2002).

[3] M. Mitchell, *An introduction to Genetic Algorithms, The MIT Press*, Cambridge, MA (1999).

[4] W. Smith, J. Smith and F. Smith, *Gene Expression Patterns* **6**, 110-118 (2005).