Model transcriptional networks with continuously varying expression levels

M. Carneiro, D. Hartl

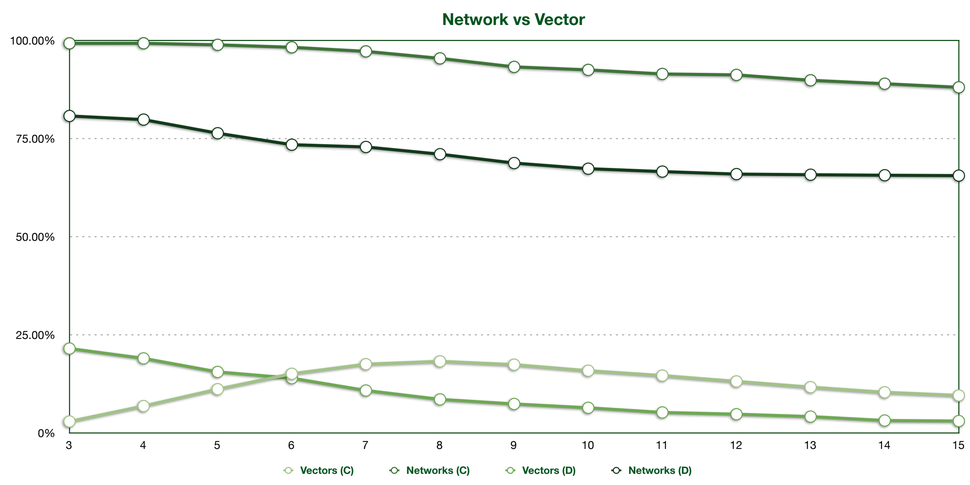
Department of Organismic and Evolutionary Biology, Harvard University

The Model

Imagine a gene regulatory network of N genes represented by the NxN matrix W, where W_{ij} measures how the abundance of the product of gene j affects the production of the product of gene i. Given a vector S_0 of initial expression levels for each of the N genes. An individual is tested for viability by a stepwise development process that quantifies the production of the transcription factors in the network over time.

$$S_{t+1} = f\left(\sum_{j=1}^{N} W_{ij} S_t\right) \qquad \varphi(S_t) = \frac{1}{\tau}\sum_{\theta=t-\tau}^{t} D(S_t, \overline{S_t}) \qquad D(X,Y) = \frac{\sum_{i=1}^{N}(X_i - Y_i)^2}{4N}$$

By multiplying W with the vector of gene products, we scale the effects of the direct and indirect interactions, adding together the weighted effect of every interaction between each gene and every other gene in the network.



Stability and Viability

A "viable individual" is an individual that develops a stable output vector S. The concept of stability is based on the development of the individual. Each developmental step is represented by a multiplication between the matrix W and the vector S_t , yielding a new vector S_{t+1} which is multiplied by the same matrix until the variation in S_{t+1} is less than a sensitivity constant σ when compared to previous values.

Haploid and Diploid Mating

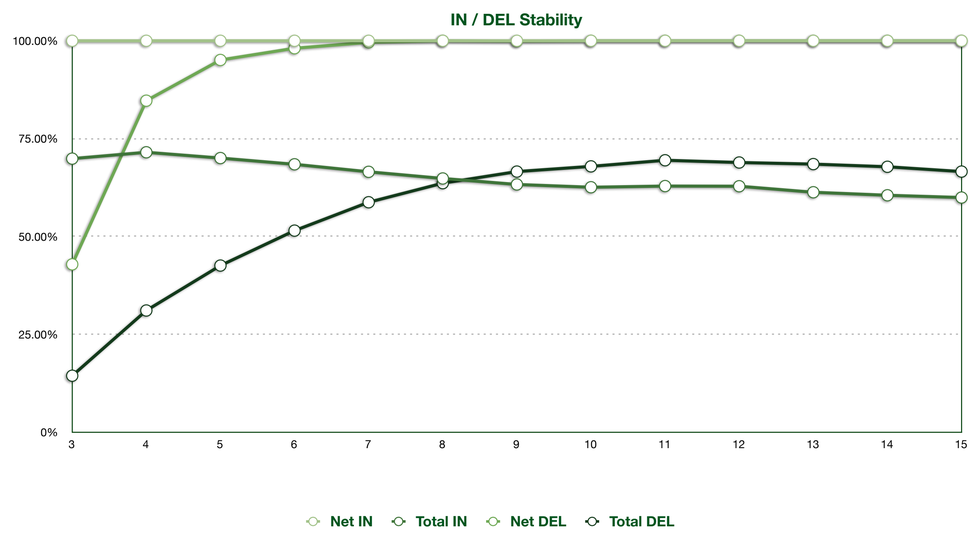
Haploid mating was the first type of mating tested to analyze how we could generate complex individuals by combining two viable individuals. The haploid model assumption is that, from any mating, a given gene is inherited from either the father or the mother with equal likelihood. Accordingly, in the haploid mating process we randomly select individual columns from within the paternal or maternal network and copy them to the offspring. This passes on one parent's gene without modification from one generation to the next. The column values represent the impact of the gene being copied on all other genes in the system. This way we keep the effects of the gene intact. Repeating the selection process N times yields a new offspring with a random set of both parents' genes.

Diploid individuals benefit from heterozygosity to curb the effects of damage or deleterious mutations as well as increased diversity through the recombination events between the parents chromosomes. The possibility to recombine genetic information before passing it to the offspring reduces the risk of carrying the negative impact of a gene on to next generations. In the diploid mating process, instead of randomly selecting columns from the parents, we calculate averages of the genes from the two parents to create the offspring network, averaging the values column by column. By doing this, we are representing the impact of each gene as an average of the impacts of this same gene in each parent. The goal is to mitigate the negative (and positive) effects of the parents when generating the offspring.

Insertions and Deletions

A gene insertion represents the phenomenon of a new gene being fully incorporated by the genome and now interacting with the other genes in the network. As it is a new gene, all interaction values are chosen at random from the uniform distribution [-1,1]. All other pre-existing genes also receive new randomly generated values for interaction with the newly inserted gene. The stable vector also receives a new randomly generated value representing the initial concentration of the product of the new gene. The end result is a new individual with an extra gene that may or may not be viable.

The likelihood of remaining a viable complex individual after removing a gene was tested by deleting one gene at random from a viable individual and developing the new individual to asses if he would remain viable.



Conclusion

The continuous model gives us a more complete view of the evolution of interacting genes. It allows the addition of more genes to the individuals and is more efficient in maintaining stability. The notion of a continuous output vector also creates a closer relationship with the reality of gene networks and gene products where it is not sufficient to determine whether or not a gene is on or off. Gene product concentrations play an important role in determining the viability of individuals and aids in the creation and maintenance of complexity.

The different mechanisms to generate complexity tested had a strong impact in the probability of finding a viable individual even in higher degrees, especially for diploid mating. The difference between the probability of drawing a viable individual with 15 genes at random and by mating two viable individuals is an indication that sexual reproduction is a key component in the evolution of complexity. The phenomena of insertion and deletion could also play an important role in the evolution of complexity given the high probability of a viable individual to remain viable after undergoing an insertion or a deletion.

This model gives an insight on the mechanisms that regulate the evolution of complexity with a more general model that could be used to represent the concentration of the products of gene networks. The inclusion of gene product concentration brings the model closer to biology and perhaps gives us a better notion to the rarity of complex viable individuals in the real world.