

Random Walks over Neutral Networks

P. Ferngani, E. Ferrada, Y. Gandica, C. Lamanna, S. Orozco, R. Vargas, J. Velázquez.

Complex Systems Summer School,

Bariloche, Argentina

Dec. 1-15, 2008

Paula, Evandro, Yérali, Christine, Sirio, Rodrigo, Jorge.

Abstract

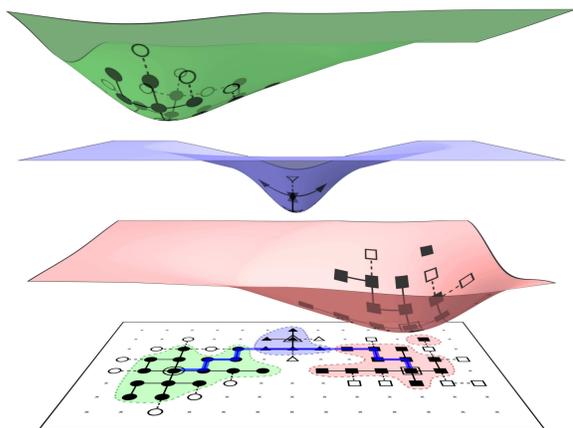
The evolutionary dynamics of RNA sequences is explored using random walks over a neutral network. The energy barrier between neighbor phenotypes is proposed to be proportional to the difference between the Free Energies of their phenotypes. The evolution of a genotype is expressed as the Probability $P_N(g_i)$ of reach the genotype g_i after N mutations. Point mutations as well as recombination can be simulated. The model shows that in a sufficient large number of mutation the probability of reaching a genotype with lower Free Energy associated, is bigger than the probability of mutate in other genotypes. Preliminary results show that a more precise estimation of the transition probabilities among phenotypes can be obtained.

1. Introduction

The genotype-phenotype relationship is a long standing problem in evolutionary biology. The interaction of phenotype and environment as well as the genotype-phenotype dynamics determine the available regions of the genotype space. Here, we use secondary structures of RNA embedded into a neutral network and random walks, to address the role of thermodynamic constraints on the probability of a given genotype in exploring the phenotype space.

2. Neutral Networks

The evolution of a biological system for which genotype and phenotype are well defined, can be represented as a network. In this network nodes represent genotypes and edges n-mutant neighbors. The mapping of phenotypes into the network reveals two interesting facts: phenotypes span over a small fraction of the possible space and the regions covered by these phenotypes differ considerably. Since this large sections of the network are able to explore a large number of different genotypes while keeping the same phenotype, they have been called 'neutral'.



3. Random Walks: The Model

To explore the possible evolutionary trajectories on an RNA neutral network, we start from an initial genotype, g_0 . A mutation carries g_0 from one site to another in the network with some transition probability $p(g_i|g_0)$, where g_i refers to a site in the network. After N steps, the Probability Distribution Function $P_N(g_i)$ of being in the site g_i is given by:

$$P_N(g_i) = \sum_{j=0}^{No.sites} p(g_i|g_j)P_{N-1}(g_j). \quad (1)$$

For a walker in a continuous space, the sum in 1 is replaced by an integral. For example, take a walker in one dimension that has the probability distribution of going from l to l' given by:

$$p(l|l') = \beta e^{-\beta(l'-l)}, \quad \text{for } (l-l') \geq 0 \quad (2)$$

$$p(l|l') = 0 \quad \text{for } (l-l') < 0. \quad (3)$$

$1/\beta$ is the mean step length and l is the distance to the origin. Notice that this walker can only step forward. For a walker starting at the origin, the Probability Distribution of finding this walker at a distance l after N steps is found by taking N times the convolution

$$P_N(l) = \int_0^\infty p(l|l')P_{N-1}(l')dl'. \quad (4)$$

With the initial condition $P_0(l) = \delta(0-l)$, where $\delta(0-l)$ is a Dirac delta function, a Poisson distribution is found:

$$P_N(l) = \frac{e^{-\beta l} (\beta l)^N}{N!}. \quad (5)$$

A mutation in the network is modeled as a step of a random walker in the network. The Energy landscape is taken into account by introducing an exponential weight of the energy differences in the transition probability, i.e.

$$p(g_i|g_j) = \frac{e^{-\beta[E(g_i)-E(g_j)]}}{\sum_{neighbors} e^{-\beta[E(g_j)-E(g_i)]}}. \quad (6)$$

4. Results

We fully characterized an RNA neutral network. Genotypes are represented by sequences of 15 nucleotides. Phenotypes correspond to their secondary structure estimated using ViennaRNA package. Figure 1 shows the observed number of genotypes per phenotype. The 4^{15} different genotypes fold into only 424 phenotypes.

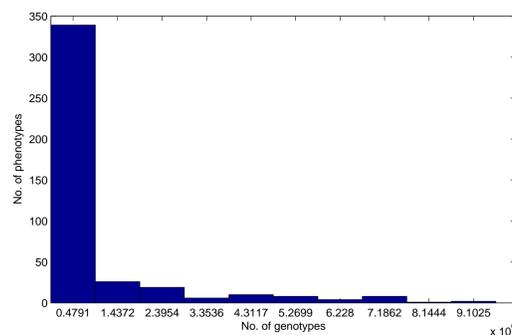


Figure 2 shows the resultant network. Nodes represent phenotypes and edges connect those found to be closer than 7 point mutations in average.

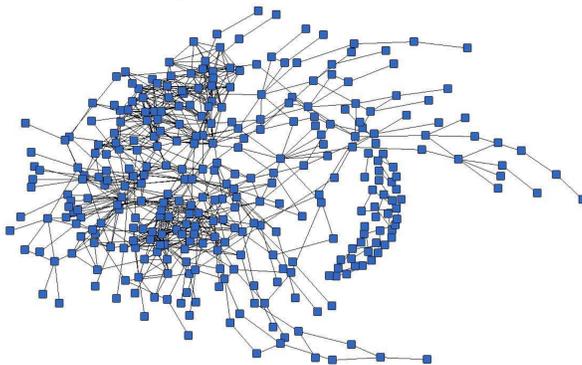
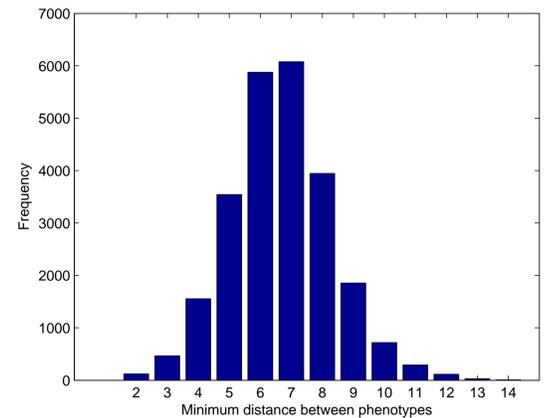
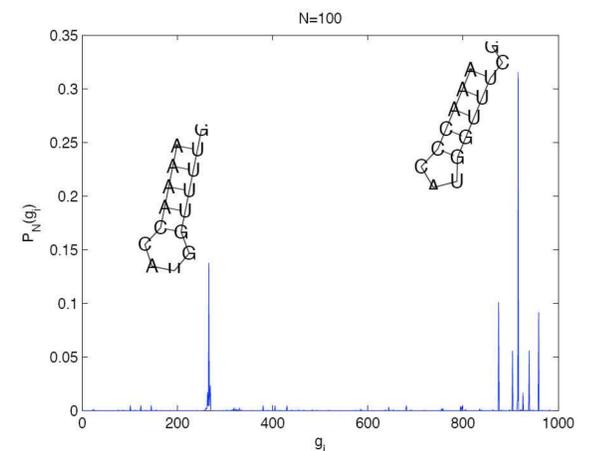


Figure 3 shows the minimum distance observed between genotypes belonging to different RNA structures.

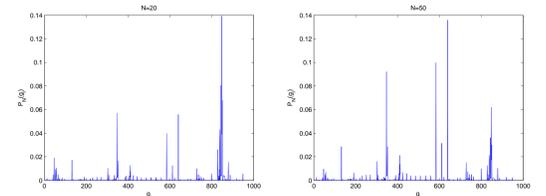


5. Examples

Figure 4 shows the stationary probabilities observed for a pair of two mutant neighbors belonging to two different phenotypes.



More complicated patterns can be observed for 3 and higher order mutant neighbors. Figure 5 shows the stationary probabilities for a 3 mutant neighbor exploration.



6. Conclusions

- A random walk methodology was implemented to explore the transition probabilities between phenotypes over a neutral network.
- Intrinsic properties of the nodes can be mapped into the network. When a free energy landscape is mapped we observed that the estimated available number of phenotypes is reduced.

References

- [1] B.D. Hughes, M.F. Schlesinger and E.W. Montroll, Proc. Natl. Sci. USA 78, 3287 (1981).