July 16-21, 2017 at UC San Diego, CA, USA

Modular Growth and Structural Remodeling in Early RNA Evolution

M. Popović[1,3], C. Wei[2,3], A. Pohorille[2,3] and M. A. Ditzler[3],
[1]Blue Marble Space Institute of Science, [2]University of California San Francisco
[3]Space Science and Astrobiology Division, NASA Ames Research Center
* mark.a.ditzler@nasa.gov

Abstract: RNA molecules are widely believed to be early, if not the earliest, genetic and catalytic molecules. The earliest evolving RNAs are expected to be limited in length, with RNAs increasing in length over time. Combined phylogenetic and structural evidence suggests that complex modern RNA structures evolved from simple shared structures that evolved in the context of much shorter RNAs. Understanding the role of RNA in both the origin and evolution of life therefore requires an understanding of evolution as a function of polymer length. One way of conceptualizing this is with fitness landscapes. On these landscapes each genotype has a value of fitness, and the evolution of a phenotype to reach a fitness peak occurs through consecutive mutations. If a landscape consists of isolated peaks, then evolutionary optimization is possible only through recombination or alterations to the landscape. Alternatively, if the landscape contains large networks of near-neutral mutations, large volumes of genotypic space are crossed without marked effect on fitness, eventually resulting in more deterministic outcomes to the evolutionary process. The connectivity of fitness landscapes, and therefore the potential for optimizing fitness, is predicted to depend heavily on polymer length.

By combining exhaustive mapping of fitness landscapes for short RNAs with structure guided mapping for long RNAs, we investigated the RNA fitness landscapes as a function of polymer length. We evolved populations of ligase ribozymes of two lengths, one population with 20 fully random nucleotides (20N population) and one with 80 fully random nucleotides (80N population). We evaluated the evolved populations by way of combining high throughput sequencing data with comparative sequence and structure analysis. We examined the connectivity of fitness landscapes within both the 20N and 80N populations and found evidence for extensive neutral networks that include evolutionary paths connecting distinct secondary structures through a continuous series of active intermediates. We also observed that the optimal structures evolved within the short 20N ribozyme populations are present as modular components of the ribozymes evolved in the longer 80N populations. On the basis of the outcome of these evolution experiments, we assayed a range of structures of increasing length and complexity in which the longer RNAs preserve the functional structures of the shorter RNAs. Catalytic activity increases with these increasingly complex structures. This result demonstrates that the preservation of preexisting structures during the addition of new structural modules can be used to build upon the evolutionary success of shorter polymers. This is consistent with models according to which this evolutionary mechanism is responsible for modern ribosomal structures. Finally, we observed that the 80N ribozymes that utilize the structures present in the shorter ribozymes are less fit than other structurally unrelated 80N ribozymes. This result, along with evidence for evolutionary paths between distinct secondary structures, indicates that optimization of function for longer RNAs may lead to global structural rearrangements that erase evidence of the ancestral structure. The potential for such structural rearrangement complicates attempts to reconstruct ancestral RNAs based on extant structures. Understanding how polymer length impacts fitness landscapes provides both insight into early evolutionary processes in general and provides guidance to the interpretation of specific features of the molecular record preserved in modern RNA structures.