

Selective Constraints in Evolution of Gene Families and TBP-Dependent Promoters

This talk is divided into two related parts. First, I explore the origins and impacts of constraints in evolution of gene families. Surprisingly, I find that selection is homogeneous for members of the same paralogous gene family. Thus, I partition gene families in several genomes into two classes: those that include at least one essential gene (E-families) and those without essential genes (N-families). I find that weaker purifying selection causes N-families to evolve in a more dynamic regime with higher rates both of duplicate fixation and pseudogenization. Because genes in E-families are subject to significantly stronger purifying selection than those in N-families, they survive longer and exhibit much greater sequence diversity. Longer average survival time also allows for divergence of upstream regulatory regions, resulting in larger changes of transcriptional control mechanisms among paralogs in E-families. Second, I explore the determinants of promoter architecture that differentiate promoters with a TATA-box that are regulated by the TATA Binding Protein (TBP) from promoters that have the TATA sequence but are not regulated by TBP. I show that some global properties of the extended promoter such as length and G+C content correlate with activity of TBP. I then discuss a physical model of TBP function in terms of efficiency in localization of the Pre-initiation complex. Finally, I find that there is differential selection on promoter sequences in light of the physical constraints imposed by the TBP. If time permits, I will discuss a putative relationship between selection in the evolution of gene families and TATA promoters.