Mathematical Methods Implemented in Promoter Recognition Systems

Vladimir B Bajic
BioDiscovery Group
Kent Ridge Digital Labs, Singapore
bajicv@krdl.org.sg

Abstract

Recognition of eukaryotic promoters is one of the most important fields in today's genomics. The problem, however, is extremely complex, since eukaryotic promoters do not have much in common, but are rather characterized as extremely diverse sequences. For this reason, promoter recognition is a challenging problem and it is not solved yet. We present here a number of the most important solutions implemented in promoter recognition and discuss the application of the underlying mathematical methods. We cover a number of publicly available systems such as: NNPP2.1, Promoter2.0, TSSW, TSSG, CpGpromoter, McPromoter, PromoterInspector, and Dragon Promoter Finder. The techniques used in these systems range from different classification methods, artificial intelligence, statistics, as well as nonlinear systems analysis and signal processing. The similarities and differences in designs are outlined, and general performances of these systems are also presented.