

Computational Regulatory Genomics

Martin Vingron

Max Planck Institute for Molecular Genetics

Abstract. Genome sequence encodes not only genes but also the regulatory relationships among genes. Thus, the time and spatial patterns of gene expression are also encrypted in the DNA sequence. In order to unravel this other genetic code, regulatory genomics attempts to integrate functional genomics data with sequence data. This talk will summarize several approaches developed in our group, starting with a biophysically motivated method for prediction of transcription factor binding sites. Main applications are the identification of tissue specific transcription factors and the prediction of regulatory changes due to SNPs. Further, the talk will describe some indications that the division of promoters into two classes with high and low CpG contents, respectively, is of functional importance and helps in understanding mammalian promoters. In fact, the two classes of promoters display different features when it comes to binding site usage and tissue specific regulation. The dichotomy is further supported by an analysis of histone modifications in the promoters. Taken together, we interpret this as indication that different regulatory mechanisms govern transcription in these two classes of promoters.