

Invited Speaker
Changhui (Charles) Yan
Utah State University

Protein-DNA interactions and identification of HTH DNA-binding
Motifs

Abstract: Protein-DNA interactions play pivotal roles in gene regulation. Deciphering the mechanisms of protein-DNA interactions presents a challenging task for biologists. The ability to identify amino acids that contribute to the specificity and affinity of the interactions can significantly improve our understanding of the details of the interactions. In our research, we have developed a machine-learning method that predicts whether a given amino acid residue is a DNA-binding residue based on amino acid identities and solvent accessibility of its neighboring residues. Our results indicate the feasibility of identifying interface residues based on local sequence information. The method compares favorably with other more complicated methods. Proteins interact with DNA through several well-defined structure motifs. To reveal the details of protein-DNA interactions, we not only need to identify the residues that interact with DNA but also to specify the structure motif through which a protein interacts with the DNA. Helix-Turn-Helix (HTH) is one of the well-defined DNA-binding motifs. Proteins with low sequence similarities can bind to DNA using a similar HTH structure. Due to the big variations in these protein sequences, identifying HTH motifs from protein sequence is extremely difficult. We have developed a Hidden Markov model that models both amino acids and predicted secondary structure. Reduced-alphabets are used to encode protein sequences. Our results show that the method can identify HTH motifs with improved performance.

Dr Yan earned his PhD from Iowa State University in 2005 with majors in Computer Science, Bioinformatics and Computational Biology. His research focuses on the development of machine learning algorithms to solve specific bioinformatic problems, such as identification of protein functional sites, protein function/structure prediction, gene finding, and sequence classification. He is also interested in protein folding, protein docking, gene expression analysis, phylogenetic inference and protein interaction networks.