

MINI-SEMINAR FOR SECOND-YEAR PH.D. STUDENTS

Department of Statistics

Discovery of Conserved Sequence Patterns

by

Baoguan Ke

Department of Statistics, University of Chicago

Tuesday, March 9, 2004, 4:30 pm in Eckhart 110
5734 S. University Avenue

ABSTRACT

This presentation is based on a paper by Mayetri Gupta and Jun Liu. The gene promoter regions that represent binding sites for certain proteins usually show some short repetitive patterns (motifs). To find these patterns out of sequences containing such regions, the paper builds up a Stochastic Dictionary Model which treats the conserved patterns and individual nucleotides as stochastic words. I will show how the models are built, and describe the effectiveness of the models using examples.