



The University of Chicago
Department of Statistics
Seminars for Fourth Year PhD Students

SU YEON KIM

Department of Statistics
The University of Chicago

**Testing for Natural Selection on Conserved
Non-genic Sequences in Mammals**

**MONDAY October 31, 2005 at 2:30 pm
110 Eckhart Hall, 5734 S. University Avenue**

ABSTRACT

The observation of high DNA sequence conservation across long periods of evolutionary time is thought to be a good signal of important regions. Otherwise, the similarity between sequences of species would have eroded by neutral mutation processes. This is also why, in general, higher conservation is expected in genic regions than non-genic regions. However, recently researchers have found highly conserved non-genic regions (CNGs), and have suggested that these regions contain functional elements that have roles in gene regulation. Given this new observation, we can ask biologically interesting questions such as: Are there CNGs that are positively selected on the human lineage but not on the other mammalian lineages? What are the interesting evolutionary patterns of CNGs? Do these evolutionary patterns vary with the levels of conservation?

In my proposal talk, I will present statistical methods that I have been developing to answer these questions. I will explain how I have constructed a database of CNGs and implemented an EM algorithm to estimate the parameters. I will also show preliminary results and talk about several statistical issues which need to be addressed in the future.