

Math Seminar
Wed. Oct. 14, 2009, at 4pm in ELC 301

Dynamic linear model for the identification of miRNAs in next-Generation sequencing data

[Evan Johnson](#), Assistant Professor of the BYU Statistics Department

I'd first like to take 10-15 minutes and do some recruiting to BYU and then give a research talk. Also, I'm planning to provide pizza for any student who is interested in attending.

Abstract:

Next generation sequencing technologies are poised to revolutionize the field of biomedical research. The increased resolution of these data promise to provide a greater understanding of the molecular processes that control the morphology and behavior of a cell. However, the increased amounts of data require innovative statistical procedures that are powerful while still being computationally feasible. In this research, we present a method for identifying small RNA molecules, called miRNAs, which regulate genes by targeting their mRNAs for degradation or translational repression. In the first step of our modeling procedure, we apply an innovative dynamic linear model that identifies candidate miRNA genes in high-throughput sequencing data. The model is flexible and can accurately identify interesting biological features while accounting for both the read count, read spacing, and sequencing depth. Additionally, miRNA candidates are also processed using a modified Smith-Waterman sequence alignment that scores the regions for potential RNA hairpins, one of the defining features of miRNAs. We illustrate our method on simulated data sets as well as on a small RNA *Caenorhabditis elegans* data set from the Illumina sequencing platform. These examples show that our method is highly sensitive for identifying known and novel miRNA genes.