# Research Experiences for Undergraduates (REU) 2013 Abstract

**ADAMS, R., OBALLE, C., RUDY, N., GILCHRIST, M. and T. PHAN. Predicting codon usage bias from a probabilistic model using codon position and gene expression. National Institute for Mathematical and Biological Synthesis, Knoxville, TN; Maryville College, Maryville, TN;  University of Texas, Austin, TX; Wofford College, Spartanburg, SC; University of Tennessee, Knoxville, TN.**

Codon usage bias (CUB) is the nonuniform usage of particular synonymous codons in genes. Current hypotheses suggest driving forces behind CUB include mutational biases, mRNA stability, and selection for translational efficiency. Under the translational efficiency hypothesis, there exist the ideas of maximizing ribosome efficiency and minimizing nonsense errors. In our research, we examine which of these two components contributes more greatly to codon usage patterns. Using equations for nonsense error costs and from population genetics, we developed a multinomial regression model predicting the probability of observing a particular codon amongst its synonyms given gene expression level and position. Our model provided a means for estimating mutation rates and NSE odds ratios for all codons. We were also able to compare the accuracy of our model to the accuracy of another multinomial regression model based on maximizing ribosome efficiency using AIC comparisons. Analysis of the S. cerevisiae genome showed that CUB is best explained by considering effects of both models. That is, nonsense error cost and efficient ribosome usage both contribute to codon usage patterns.