

# **Acquiring randomized sequence for Repositioned Substitution Test and calculating divergence and nucleosome affinity for two Homologous DNA sequences**

**Priyank Jasani**

*Graduate Student, Computational Biosciences Program*

*Arizona State University*

## **Abstract**

Divergence between two sequences is normally smaller and the alignment between two sequences is better in coding regions than in non-coding regions. However in non coding regions due to high divergence and poor alignment between two sequences, it is very difficult to calculate the divergence just by calculating no of substitutions. We therefore need a better method in terms of calculating divergence which is introduced by Shannon Jensen equation. In this project I have used Shannon Jensen equation to calculate divergence between two sequences. This project also aims to develop tools for analyzing nucleosome affinity, which depends upon the pattern of AA and TT dinucleotide present in the sequence. The nucleosome affinity pattern for TATA-less and TATA-containing genes is different. I have worked on a C code which is used to calculate nucleosome affinity and divergence for any two sequences from yeast genome. I have also worked on a C code to acquire randomized sequences from 2 sequences for performing Reposition Substitution Test which is used to detect preserved, neutral or positive selection on the sequence-dependent nucleosome positioning.