EXECUTIVE SUMMARY [NON-CONFIDENTIAL, NON-TECHNICAL ABSTRACT FOR PUBLIC INFORMATION OR PROGRAM PROMOTION]: Early genomic sequencing efforts promised to revolutionize biology and medicine. Large-scale genome sequencing has provided the "blueprints" of many living organisms. However, this overwhelming abundance of genomic information is creating nearly as many problems as it provides answers. The information avalanche can only be dealt with through development of highly accurate bioinformatic methods for storing, annotating, and comparing the data. One key set of tools required are accurate methods to analyze and catalog "*important*" features of each protein sequence encoded within a genome. With new bioinformatic tools that identify such features in hand, the promised biological and biomedical insights will be more forthcoming. Sequence regions that are responsible for protein structural stability and proper enzyme function are frequently viewed as "*important*". It is generally accepted that protein motifs (conserved areas within sequence or structure) describe these critical regions. Previously, we have demonstrated that Phylogenetic Motifs (PMs) represent very good protein functional site predictions from sequence. In this proposal, we introduce an improved PMs detection algorithm, which resolves a long-standing problem inherent to our original PM detection algorithm. Implementation of the newly proposed Phylogenetic Similarity Maximization (PSM) algorithm should greatly improve the statistical significance of the identified PMs.