

グラフ理論に基づく新規進化解析手法の開発

Development and application of the novel phylogenetic method based on graph theory

東京大学大学院 理学系研究科 生物科学専攻 岩崎研究室 松井 求

研究成果概要

「タンパク質スーパーファミリーレベルにおける系統樹の再構築」は進化学上の重要な課題の一つであるが、互いの配列距離が著しく遠い場合は **informative** な多重配列アライメントを作成することが難しく、この問題を一般的に解決できる手法は知られていなかった。そこで我々は新たな系統解析手法「**Graph Splitting 法 (GS 法)**」を考案した。GS 法は、まずペアワイズアライメントに基づいて配列類似性グラフを構築し、次にそのグラフを再帰的に分割することで系統樹を得る—つまり部分配列の類似関係を進化情報に変換する方法である。この操作は、多重配列アライメントに起因する従来からの諸問題を回避する。

本研究では、一連の従来手法(近隣結合法、最大節約法、最尤法、およびベイズ法)と GS 法の性能を比較する為に、京都大学化学研究所スーパーコンピュータシステムを利用し、大規模計算を行なっている。昨年度は進化シミュレーションに基づき、系統樹の予測精度と計算時間について検証し、GS 法がいずれについても従来手法を上回る性能を持つことを示した。本年度は、実際にタンパク質スーパーファミリーの進化過程を解き明かすことが GS 法によって可能になることを示すべく、様々な実データに GS 法を適用した。まず、最も大きく、かつ多様性に富むスーパーファミリーである G タンパク質共役受容体 (GPCR) に GS 法を適用したところ、これまでタンパク質の立体構造に基づいて予想されていた従来の系統樹と整合的な系統樹が構築され、さらに同一の機能が複数回獲得されたことが示唆されるなど、興味深い新規知見が得られた。また最大の転写因子ファミリーである HTH 転写因子ファミリーに GS 法を適用したところ高い内部枝支持率を持つ確からしい系統樹が得られ、平行進化を含む複雑な進化過程が示唆された。

発表論文(謝辞あり)

なし

発表論文(謝辞なし)

なし