

グラフに基づく分子系統解析手法の開発と適用

Development and application of the graph-based phylogenetic method

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は、その長い進化史の中で配列と機能の多様性を獲得してきた。従って、その進化過程をまるごと解明することができれば、タンパク質の諸機能と進化を結ぶ統合的知見が得られるに違いない。しかしながら既存手法は、進化距離の大きな問題に対しては、多重配列アライメント (MSA) から抽出可能な情報量が著しく減少してしまうという問題を抱えており、いずれも適用困難であった。我々はこれまでに Graph Splitting (GS) 法という新たな系統解析手法を考案し、進化シミュレーションに基づいて GS 法がこの問題を解決することを示してきた。GS 法は MSA の代わりに All-to-All ペアワイズアライメントを行い、配列類似性グラフ (SSG) を構築し、その SSG を再帰的に分割することで MSA に由来する問題を回避しつつ系統樹を推定する方法である。GS 法をスーパーファミリーの網羅的解析に適用した結果、190 のスーパーファミリーについて信頼性の高い系統樹を得ることができた。特に T-fold, DNA/RNA ポリメラーゼ, トキシン膜輸送ドメインの各スーパーファミリーは系統樹の全体にわたって内部枝が強く支持されており、それぞれ代謝系、転写・複製、膜輸送系の初期進化を照らす新たな知見が示唆された。例えば DNA/RNA ポリメラーゼの系統解析結果からは RNA ワールドの時代から RNA-dependent RNA ポリメラーゼ → RNA-dependent DNA ポリメラーゼ → DNA-dependent DNA ポリメラーゼ → DNA-dependent RNA ポリメラーゼの順番で進化してきたことが示唆された。さらに GS 法を「初期進化」のような“古い”問題だけではなく、「病原菌・ウイルス・癌の短い時間スケールにおける高速な進化」といった“モダンな”問題に広く適用することも試みている。これまでにインフルエンザウイルスやバキュロウイルスなど多様なウイルスについて先行研究を大幅に超える規模と精度を持つ系統樹が得られており、それぞれのウイルスの感染の歴史の解明につなげて行きたいと考えている。

発表論文(謝辞あり)

Motomu Matsui, and Wataru Iwasaki, **Graph Splitting: A Graph-Based Approach for Superfamily-scale Phylogenetic Tree Reconstruction**, *Systematic Biology*, 2020, **69**(2):265–279.

発表論文(謝辞なし)

なし