

課題名(タイトル):

Prediction for functional divergence in wheat homoeologous gene pairs

利用者氏名: ○江副晃洋(1)

理研における所属研究室名: (1) 環境資源科学研究センター 植物ゲノム発現研究チーム

1. 本課題の研究の背景、目的、関係する課題との関係

シーケンス技術の進展に伴い、様々な生物の膨大なゲノム情報が取得されるようになってきている。これにより遺伝子機能予測やゲノムの進化過程の解明が、直接的に品種改良に資するようになり、生物学、育種学の研究においてますます重要性を増している。

本研究では、巨大なゲノムを持つコムギおよびその近縁種ゲノムを対象として、遺伝子機能の予測とゲノム進化の過程を行うことを目的としている。コムギは全ゲノム重複を何度も経験しているため、多様な遺伝子を持ち、数も多い。さらに、そのゲノムは巨大で複雑であるため、進化の過程を詳細に追跡することが非常に難しかった。そのため、遺伝子機能の予測とゲノム進化の過程を行うことで、作物の品種改良や農業生産に大きく貢献することが期待されている。

また、コムギに限らず、多様な植物種間におけるゲノム進化を比較解析することにより、進化の普遍的なメカニズムを明らかにし、植物進化における新たな知見を得ることができると期待されている。このような大規模な解析には、膨大な計算リソースを必要とするため、大規模計算環境が必要だと考えられた。

2. 具体的な利用内容、計算方法

種内、種間の多様性や発現の比較を、多数の種、アクセッションを用いて行う。また、進化過程を解析するためのプログラムを使用する際には、大量のデータ処理が必要となる上に、ファイル入出力が頻繁に発生し、ストレージ負荷の問題が存在した。このような問題に対処し、本課題では効率的に大規模ゲノム解析を行うことを目指している。

3. 利用がなかった場合の理由

本年度は下記の理由から利用が進まなかった。

理由1: 種内多様性の計算について

種内多様性の計算を、大規模データを用いて行う予定だったが、小規模サンプルでのテストを行った後、ジョブが分割可能であることが判明し、手元の計算リソース内で、並列処理が完了した。そのため、大規模な計算を行うに至らなかった。

理由2: 種間の多様性の計算について

機能多様性の種間比較のため、大規模な遺伝子ファミリーの推定を行う予定だった。特に、この解析では、100程度の種を用いて、詳細な遺伝子ファミリー解析を行うという大規模計算を行う予定だったが、種数を削減することで、むしろ系統関係の精度が向上する可能性が判明した。そのため、分割した系統関係を手元のリソースで推定するのみで、大規模計算を行うに至らなかった。