

平成 27 年 3 月 31 日

プログラム名：革新的研究開発推進プログラム

PM 名：鈴木 隆領

プロジェクト名：大規模ゲノム情報を活用した超高機能タンパク質の設計及び製造

委 託 研 究 開 発

実 施 状 況 報 告 書 (成 果)

平成 2 6 年度

研究開発課題名：

天然高機能タンパク質遺伝子の網羅的解析

研究開発機関名：

学校法人慶應義塾

研究開発責任者

富田勝

1. 当該年度の担当研究開発課題の目標と計画

天然の構造タンパク質素材遺伝子を網羅的(最終数値目標 1,000 サンプル)に取得し、これを理研・沼田班による天然構造タンパクの物性解析とあわせることにより、特定の物性につながる遺伝子塩基配列を明らかにし、任意の物性を持つ構造タンパク質を設計可能とすることが本研究開発における目標である。クモはゲノムサイズが数 Gbp と巨大なため、全ゲノム解析ではこの目標を達成することが現実的でないが、発現している遺伝子に限定して超並列ショートリードシーケンスを行い、リファレンスゲノム配列なしでバイオインフォマティクスにより de novo アセンブリーを行うことでこれを実現する。そこで、平成 26 年度は主にクモ類における de novo トランスクリプトームのサンプル調整及びバイオインフォマティクス解析をハイスループットに行うプロトコルを決定し、50 サンプルの解析を行うことを目指した。

2. 当該年度の担当研究開発課題の進捗状況と成果

2-1 進捗状況

まず、ジョロウグモを対象として、高純度に RNA 抽出・精製を行うためのサンプル保管方法や、対象となる組織、さらに破碎・抽出条件を検討した。続いて、Illumina MiSeq シーケンサーによって実際に de novo トランスクリプトーム解析を行い、クモ糸遺伝子を検出する際に適切なリード長及び量を決定した。続いて RNA 抽出段階(電気泳動による分解度合いの確認、微量分光光度計による純度の確認、蛍光色素による精密定量)、シーケンスライブラリ作製段階(電気泳動によるライブラリサイズの確認、蛍光色素による精密定量、リアルタイム PCR によるライブラリ構造の確認)、シーケンス後データの確認などの品質管理基準を決定し、破碎、RNA 抽出・精製、シーケンスライブラリ作製、シーケンシング、データ解析の全ステップを可能な限り並列処理するハイスループットプロトコルを確立した。

確立されたプロトコルをもとに、条件検討でシーケンスを行ったジョロウグモに加え、48 サンプルを NextSeq500 を用いたシーケンシング及びアセンブリーを行い、計 50 サンプルの解析を終了した。

2-2 成果

1週間で24サンプルのシーケンシングを行うことのできる、天然高機能タンパク質遺伝子のハイスループットな配列データ解析プロトコルを作成した。また、そのプロトコルを48の天然生体サンプルに実際に適用することで実用性を確認した。シーケンスリードは150bpのpaired-endで平均約40M read得られ、アセンブルの結果N50が約1500bpのコンティグが平均約6万個/サンプル得られた。公共データベースから取得した既知のクモ系関連配列を対象にBLAST検索を行ったところ、ほぼ全てのサンプルで複数のクモ系関連遺伝子を同定することができた。解析対象のクモには、ジョロウグモのように手足を除いた体の部分が数cmの比較的大きな個体から、ヒメグモなどの数mm程度の微量のRNAしか得られない個体まで多様な種が含まれており、天然生体サンプルからその遺伝子の配列データを網羅的に取得すること自体はプロトコルとして確立し、シーケンシングその後の物性との関連などの詳細なデータ解析にはさらに時間を要するものの、次年度以降の年間数百サンプルの解析は問題なく行えるものと考えられる。

2-3 新たな課題など

クモ系の遺伝子は通常の酵素などをコードする遺伝子に比べるとその配列が極めて長く、また、リピートに富むため、ショートリードを用いた現状の解析手法では断片は得られるものの完全長の配列を得る事が困難である。そこで、複数のK-merによるアセンブリーを統合することによるde novoトランスクリプトームアセンブリーアルゴリズムの改良と、リピート配列部分のカバレッジに基づく完全長の推定など、バイオインフォマティクスによる改良、及び、既に得られている5'側及び3'側断片からのプライマーウォーキングによる追加シーケンシングにより、次年度以降完全長の取得を目指す。既に参照配列として現在ジョロウグモのゲノムDNAの解析を進めており、この際にはPacBioシーケンサーの最新P6-C4試薬を試しているが、ここでは平均長10kbp程度の極めて長いリードを得る事に成功しているため、2については両端からのlong-range PCR産物をPacBioでシーケンスすることを含め検討する。

3. アウトリーチ活動報告

なし