

課題名 (タイトル) :

エキソームシーケンス解析

利用者氏名 : 西部 弘純

所属 : 和光研究所 脳科学総合研究センター 疾患メカニズムコア 分子精神科学研究チーム

1. 本課題の研究の背景、目的、関係するプロジェクトとの関係

次世代シーケンス解析を使用したエキソームシーケンス解析は、バイオ解析の新しい技術として重要な位置を占める。

次世代シーケンス解析機から得られたデータのエキソームシーケンス解析は、大量のデータ量を処理する場合非常に時間がかかる。普通のデスクトップパソコンで処理を行った場合 10 日間、場合によってはそれ以上かかる解析ステップもある。そこで、RICC の高能率コンピューターを使用し解析を行うことにより時間と業務の効率化を図る。

Galaxy 解析ツールを利用したエキソームシーケンス解析は今後頻繁に行われるようになる。

私たちのラボでは、人の血液サンプルの解析などを行って精神疾患の病理を解明している。こうしたサンプルの解析結果を効率よく迅速に得るためには、RICC にあるような高能率コンピューターの利用は非常に有効である。

2. 具体的な利用内容、計算方法

RICC のコンピューター上に Galaxy 解析ツールをインストールしその解析ツールを利用しエキソーム解析をおこなう。

大まかな計算手順は、

- a. リードをゲノムにマッピング (BWA)
- b. 冗長性のあるリード・Duplicate を除去し (Samtools や Picard)
- c. キャプチャーした Exon 領域だけを取り出し (Bedtools)
- d. Samtools で SNV を抽出し
- e. SNV にアノテーションをつける
- f. non-synonymousSNV, missense, frame-shift など

3. 結果

Galaxy ツールに不具合が生じたためすべての解析が終わっていない。その問題を解決するまで解析を中断している。

4. まとめ

Galaxy ツールに不具合が生じたため解析を中断しているが、問題が解決次第解析を再度実行する。また、RICC のコンピューターを利用できることは時間と業務の効率化を図ることができ満足している。

5. 今後の計画・展望

Galaxy の不具合が解決次第エキソーム解析を引き続きおこなう。また、その解析で得られたデータは実験のデータ結果として論文に引用する。