

課題名 (タイトル) :

次世代シーケンスデータの大規模解析

利用者氏名 : 沼田 興治

所属 : 筑波研究所 バイオリソースセンター バイオリソース関連研究開発プログラム
動物変異動態解析技術開発チーム

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

生物試料における遺伝情報を解析する DNA シーケンサー装置が輩出するデータのスループットはここ数年で飛躍的に増大した。一度の解析で数十〜数百ギガバイトのデータを生成することが可能なため、これらのデータを全ゲノムワイドで解析するためには膨大なストレージと計算コストが必要となる。本研究課題では、上記の技術をもちいて、主にマウスの生殖細胞におけるゲノムの化学修飾状態 (シトシンのメチル化や、ヒストンの修飾) や、遺伝子の発現状態の変化を明らかにするため、その解析やデータマイニングに RICC をもちいる。

2. 具体的な利用内容、計算方法

本課題が RICC においてもちいるデータは、マウス細胞における ChIP-seq (ゲノムの化学修飾状態) や RNA-seq (遺伝子発現状態) などのデータで、マウスゲノム上へのマッピングを完了しているものである。これらのデータは、マウスゲノムの領域毎の遺伝子発現量やゲノム修飾状態を反映した定量的な情報である。このなかから特定のゲノム領域 (数百万から数千万箇所の領域) における数値データを抽出する処理を RICC によっておこなう。

3. 結果

今年度は、NCBI (National Center for Biotechnology Information) に登録されたデータをもちいて解析の検討をおこなった。二つの近交系統間の雑種マウス (遺伝的背景の異なる二系統のマウス同士を交配させたマウス) における RNA-seq データをもちいることによって、アレル特異的な遺伝子発現の同定を試みた (アレル特

異的な遺伝子発現とは、父親由来、母親由来 2 コピーずつ遺伝子群のなかでもその発現がどちらか一方に偏っている遺伝子を意味する)。その結果、これまで論文で報告のあるようなゲノム刷り込み遺伝子を同定することに成功し、その組織間比較などが可能となった。これらの解析はすでに論文報告されたデータをもちいているため、新たな知見を得ることはできなかったが、この検討によって解析パイプラインの構築ができ、新規の RNA-seq データへの対応が可能となった。

4. 今後の計画・展望

今年度確立した解析パイプラインを、新規のデータに適用していく。現在、マウス亜種間雑種をもちいた新規の RNA-seq データを取得中であるため、そのデータをもちいて同様の解析をおこない、上記結果との比較・検討をおこなう。

平成 24 年度 RICC 利用報告書

平成 24 年度 RICC 利用研究成果リスト

【国際会議、学会などでの口頭発表】

沼田 興治「亜種間ハイブリッドマウスにおける遺伝子発現のアレル間差異」 日本進化学会第 14 回大会（2012 年 8 月東京、招待講演）