

## 平成 27 年度 全体研究開発報告書

1. 補助事業名：創薬等ライフサイエンス研究支援基盤事業（創薬等支援技術基盤プラットフォーム事業）
2. 補助事業課題名：創薬等支援のための大規模シーケンスによるゲノミクス解析支援と高度化（創薬等支援のための大規模シーケンスデータのためのゲノミクス解析支援と高度化）
3. 研究開発分担者：大学共同利用機関法人情報・システム研究機構 国立遺伝学研究所  
生命情報研究センター 准教授 池尾 一穂

### 4. 研究開発の成果

今年度は、

創薬プロセス又は医療等のライフサイエンス研究に活用可能な研究基盤の整備の一環として、次世代シーケンサを用いた大規模シーケンスによるデータ解析のためのゲノミクス技術基盤を整備し、幅広いライフサイエンス研究者に対して、次世代シーケンサを用いたゲノミクス解析を提供し総合的に支援することを目標に、

支援に関しては、

1) 一次解析から高度なデータ解析までをトータルで支援するとともに最先端のゲノムシーケンスおよびデータ解析技術を提供する。

高度化としては、

- 2) 統合的解析を可能とする解析手法及び
- 3) 統合的エピゲノミクスデータベースの開発を行う。

以上の項目に関して研究開発を進めた。

具体的には、統合的解析を可能とする解析手法の開発においては、データを処理する統計手法、データ解析の自動化、複数のエピゲノムデータの統合解析、新規発見を成し得るような統合データ解析ツールの開発提供を目指し、統合的エピゲノミクスデータベースの開発においては、創薬プラットフォームにおいてゲノム情報を利用するために必要な要素情報の収集と可視化技術の開発をおこない提供する。加えて、情報拠点データクラウドとの連携を進める。また、解析に関するデータベース及び解析パイプラインの利用法も随時公開するほか、年 1 回程度講習会を開催して、利用者との情報共有を図ることを目指した。各項目に関しては、

#### ・支援

（コンサルテーション）

公募支援に関し、領域A支援 2 件、特別支援1件を実施した。利用申請者、代表機関、分担機関の三者が代表機関に集まり話し合いを行い、支援内容に最も適していると考えられる実験方法、サンプル数、シーケンス法、データ解析方法について意見交換した。

また、データ解析法方に関するコンサルティングを都合 5 件行った。このうち 2 件は、厚労省難治疾患対策事業と肝炎対策事業におけるNGSデータ解析実施に関するものである。

それ以外に、wikiを利用した解析関連技術情報の公開、公開解析システム利用に関する技術的相談に対応した。

（データ解析支援）

平成 26 年度分の解析公募支援を引き続き 39 件について行った。その内訳は、67 種類の異なる解析を行い、プログラム総実行数 2873 回に至った。また、支援先との打ち合わせ回数 57 回に成る。

その内訳は、ゲノミクス A 2 2 件、ゲノミクス B 1 7 件（橋本グループ 1 0 件、白髭グループ 5 件、二階堂グループ 1 件）となる。

上記の支援は、当初予定の解析を終了し先方にレポートとして報告済であるが、その後、成果の検証取りまとめに必要な追加解析の希望があり、2 1 件が進行中である。

27 年度分の解析に関しては、現在、シーケンシングが終了しつつあり、7 件が解析進行中で、第一回目の報告を終了もしくは打ち合わせ予定の調整中である。残りのデータに関しても、月当たり 5 件の解析が可能であり本年度分のシーケンスサンプルが揃う秋までに終了する予定でいる。

特別支援としては、解析拠点・解析領域が支援する竹内恒博士（産業技術総合研究所）より課題申請があり、平成 27 年 9 月に関係者一同が集まって、事前相談により情報共有を行った後、シーケンス解析が 2 月末に終了し、データ解析が進行中である。

高度化支援として、ゲノミクス B の単一細胞 RNAseq 技術開発に協力して、検証のためのデータ解析を橋本グループと協力して行った。

#### ・高度化

（統合的解析を可能とする解析手法）

我々のデータ解析システムは、現在、4 6 2 種類の解析パイプラインが実行可能となっており、その対象は CAGE, SAGE, ChIP-seq を中心にゲノム解析を含めた NGS アプリケーション全体をカバーしている。また、ゲノムビューワーや総合アノテーションパイプラインの開発により、当初目標としていた統合解析システムの構築も順調に進んでいる。実際の支援において、本システムを利用して、幾つかの疾患マーカー候補の抽出、疾患機序モデルの構築などに貢献した。これらの成果は、支援機関が中心となって投稿準備中である。

（統合的エピゲノミクスデータベースの開発）

上の統合的解析とも関連するが、遺伝子発現、及びその制御に関するデータを扱うためのバイオロジカルネットワークビューワーは、ほぼ開発を終了しており、支援での利用と情報拠点への提供を開始した。