

洪 実

慶應義塾大学医学部坂口記念システム医学講座・教授

動的遺伝子ネットワークの多次元構造解析による
高精度な細胞分化制御技術の開発

§1. 研究実施体制

(1) 洪 グループ

- ①研究代表者: 洪 実(慶應義塾大学医学部坂口記念システム医学講座、教授)
- ②研究項目: 転写調節因子導入ヒト ES 細胞株の樹立、培養、トランスクリプトーム解析、
遺伝子発現調節ネットワークの数理解的解析

(2) 阿久津 グループ

- ①主たる共同研究者: 阿久津 英憲(国立成育医療研究センター生殖・細胞医療研究部、室長)
- ②研究項目: ヒト ES 細胞株の提供、ヒト ES 細胞培養のクオリティコントロール

(3) 小原 グループ

- ①主たる共同研究者: 小原 収(かずさ DNA 研究所、副所長)
- ②研究項目: 遺伝子クローンの提供、トランスクリプトーム解析

(4) 的場 グループ

- ①主たる共同研究者: 的場 亮
- ②研究項目: トランスクリプトーム解析

(5) 西村 グループ

- ①主たる共同研究者: 西村 邦裕
- ②研究項目: データベース・ウェブサイト構築、データ解析、データの可視化、
遺伝子発現調節ネットワークの数理解析

§2. 研究実施内容

【研究の背景】

本 CREST では、「生命現象の統合的理解や安全で有効性の高い治療の実現に向けた *in silico/in vitro* での細胞動態の再現化による細胞と細胞集団を自在に操る技術の創出」を戦略目標とし、さらに「将来実現しうる重要課題の達成ビジョンの一つとして、各種疾患の原因となる細胞の老化や、再生医療において重要な幹細胞の分化など、複合要因で制御される現象の再現」を達成目標としている。本研究課題では、これらの戦略目標に沿うプログラムとして、ヒト胚性幹細胞 (ES 細胞) の動的遺伝子ネットワークの構造とその動態解析を行う。

そこで平成 24 年度は、特定の転写調節因子 (TF) を発現誘導できるようなヒト ES 細胞 (TF-hESC) の作製を開始した。

(1) 転写調節因子 (TF) の選択 (洪、小原)

本プロジェクトでは、ヒトで約 1,000 から 1,500 個程度存在すると考えられている TF の内、500 個を選択し、それぞれの TF を選択的に薬剤で発現誘導をかけることができる細胞株の樹立を、第一目標とする。そこで、情報解析 (Bioinformatics) ツールを用いて、我々の過去の TF 誘導マウス ES 細胞株を用いたデータを含め、これまで公開されてきた様々な生物医学情報を解析することで、ヒトの遺伝子の中で TF と考えられる遺伝子のリストアップを行った。さらに、本年度は最終的に使用する 500 遺伝子の内、入手可能と考えられる約 400 遺伝子をリストアップした。

(2) TF を Dox で誘導できる基本ベクターの構築 (洪)

ヒト ES 細胞株で、テトラサイクリン的一种であるドキシサイクリン (Dox) によって外来性の遺伝子発現誘導が可能な系を樹立するため、今年度は piggyBac ベクターに gateway システムを用いて簡便自在に外来遺伝子を導入できる基本ベクターを作成した。

(3) TF 発現誘導可能なベクターの作成 (洪、小原)

リストアップした 400 種類の TF 遺伝子の内、今年度は日本国内の DNA バンク (AIST, BRC) と研究分担機関であるかずさ DNA 研究所 (小原) より、約 100 遺伝子を購入または譲渡を受けた。本年度は入手した遺伝子の内、約 50 種類の TF を Dox 誘導可能な基本ベクターに導入した。

(4) TF 発現を自在に誘導できるヒト ES 細胞バンクの作成 (洪、阿久津)

Dox で誘導可能なベクターが機能するためには、リバーステトラサイクリン制御性トランス活性化因子 (rtTA) を発現するプラスミドを、予め、ヒト ES 細胞に導入しておく必要がある。そこで、本年度は、rtTA ユニートを発現するプラスミドの構築と、それを均質に安定に発現できる、ヒト ES 細胞の樹立を行った。ヒト ES 細胞株は、成育医療研究センターで樹立されたヒト ES 細胞株 (SEES1, SEES2, SEES3, SEES4) の分与をうけ入手した。この ES 細胞株に、複数の遺伝子導入方法、複数のヒト ES 細胞培養方法を用いて、rtTA 発現ヒト ES 細胞株を合計 9 株樹立した。これらの細胞株に、緑色蛍光蛋白 Emerald を Dox で誘導可能な基本ベクターを導入したところ、蛍光顕微鏡下で Dox の濃度依存的な Emerald 蛍光が観察され、樹立された rtTA 発現ヒト ES 細胞株が、本研究課題の目的にかなったものであることが明らかとなった。

(5)トランスクリプトーム計測(洪、的場)

本年度は、予備実験として、次世代シーケンサーを用いたトランスクリプトーム解析のコスト削減と高効率化を試みた。ヒト iPS 細胞から抽出した RNA サンプルを材料とし、複数のプロトコルにより次世代シーケンス用ライブラリを作成し、その再現性の確認、さらに DNA マイクロアレイデータとの比較検討を行った。

(i) ライブラリ作成時の PCR (Polymerase Chain Reaction) の影響を調べるために、RNA サンプル増幅過程の PCR cycle 数を 2 cycles と、12 cycles で別々にライブラリを作成した。データ解析の結果、12 cycles プロトコルが 2 cycles プロトコルと比較して約 8 倍のデータ量が得られ、かつデータ再現性がよいことがわかった。さらに、12 cycles プロトコルにおいて、2つの異なるサンプル調製試薬から得られたデータを、スクアタープロットにより比較したところ、ほぼ同等の結果が得られることが確認された。

(ii) マイクロアレイデータ(Agilent 社のアレイを用いたデータ)との比較検討を行った。次世代シーケンスデータにおいて Read Count 10 以上の遺伝子数は平均 15,000 程度であり、一方、マイクロアレイにおいて、シグナル値 200 以上の遺伝子数は平均 11,000 程度であった。これらのうち共通に発現が見られた遺伝子数は平均 10,000 程度であり、これらの相関を見るためにスクアタープロットを作図した結果、発現量について強い相関が得られた。またデータ解析により、遺伝子上の各エクソン毎の発現情報が得られ、個々の遺伝子発現状態の詳細情報を得ることができた。

以上の解析により、次世代シーケンサーによるトランスクリプトーム解析が、これまでのマイクロアレイと同等以上のデータが得られることが確認された。

(6)データベース・ウェブサイト構築とデータ解析(西村)

当該 CREST のホームページを作成し、公開すると共に、今後の次世代シーケンサーデータの解析のためのシステムおよびデータベースのプロトタイプ構築として、次世代シーケンサーから出力されたデータを前処理した SAM(Sequence Alignment/Map)ファイルや BAM ファイルを対象として、管理・解析のためのシステムを構築した。(i)データ自体のアップロードやダウンロードできる仕組みを Web ベースで実装した。(ii)データのソートを行い、リード数のカウントし、データベースに格納することを分散型(マルチプロセス)のシステムで実現した。(iii)次世代シーケンサーのデータ自体を整理するために「タグ」付けする仕組みを導入し、「タグ」をデータベースに格納し、Web インタフェースを整えることで、ファイル自体の整理や検索を容易にする仕組みを開発した。

この仕組みを利用すると、データをデータベースに格納すると自動的に「タグ」による整理が出来るため、今後 high throughput の解析が可能となる。来年度は、TF ネットワークのモデル化を行い遺伝子発現調節ネットワークが描けるようになった際に、データベースと接続して解析パイプラインとしてシステムを構築し、研究を進めやすくしていく予定である。