

## 課題名(タイトル):精神疾患に対する転移因子の役割

利用者氏名: ○西岡 将基<sup>(1)</sup>、林(上田) 順子<sup>(1)</sup>、数野 安亜<sup>(1)</sup>、加藤 忠史<sup>(1)</sup>

理研における所属研究室名:

(1) 脳神経科学研究センター 精神疾患動態研究チーム

## 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

ヒトゲノムには、ゲノム上をコピー & ペーストないしカット & ペースト形式にて転移する配列があり、これを転移因子という。転移因子はヒトゲノムの約 45%を占めるが、類似した配列が多いため解析困難であり、これまで捨象されてきた。しかし、45%という広大な領域を占めており、また転移の過程で遺伝子を破壊するなど表現型に大きな影響を与えていると予想される。特に遺伝子のコーディング領域ないし近傍への新規挿入は、個体の表現型に対する影響が大きいと考えられる。統合失調症や双極性障害をはじめとする精神疾患は一卵性双生児での診断一致率からゲノムの関与が強いと考えられるが、その遺伝的背景は十分には解明されていない。本計画は、精神疾患罹患者とその両親(トリオ)のシーケンスデータから、通常の一塩基置換や小規模な挿入欠失に加え、転移因子を解析することにより、精神疾患に寄与の大きい変異を検出することが目的である。

## 2. 具体的な利用内容、計算方法

罹患者および両親から得られた試料由来のシーケンスデータを、ヒトリファレンスゲノムにアライメントすることにより、個人ごとの多型・変異を精度よく検出する。シーケンスデータは1個人あたり数十億塩基(通常テキストデータでは文字数に相当する)あり、30億塩基あるリファレンス配列へのアライメントおよびクオリティコントロールには膨大な計算が必要である。BWA (Burrows-Wheeler Aligner; Burrows-Wheeler transformを使用したC++ベースのアライメントソフト)、GATK (Genome Analysis Toolkit; 様々なゲノムデータ操作を行うJAVAベースのソフト)を主に使用し、バッチジョブによる並列の計算を行うことで、現実的な時間内でアライメントを行い、ゲノム上の変異を検出する。転移因子の解析は、MELT (転移因子解析に特化したJAVAベースのソフト)を使用する。リファレンスゲノムによって転移因子の位置が異なるため、GRCh37とhg38の2種類のヒトリファレンスを使用し、2つのパイプラインによる検出を行うことで、網羅的な検出を行う。BWAはメモリ使用効率が高く、コアの並列が大規模に行えるためコア・メモリの使用は比較

的少ない。GATKは計算効率よりも解析パラメーター設定の柔軟性が特徴であり、メモリ使用が大きい。このためGATK利用にあたり、ノード利用を占める割合が多く、また計算時間を長く取る傾向にある。

## 3. 結果

精神疾患罹患者(発端者)にのみ存在し、両親には存在しないデノボ変異(体細胞変異含む)の候補を数百以上検出し、保有試料についてはバリデーション実験にて存在を確認した。このような変異は疾患に対する寄与が大きいと想定され、病態解明につながるリスク変異を含むと期待される。パブリックデータベースからの解析についてはバリデーション実験が行えないため、特異度を上げた解析により、確度の高い変異の選定を行っている。解析結果をまとめて現在投稿準備中である(Nishioka et al. in preparation)。転移因子については、伝達される生殖系列の変異としては複数検出しているものの、デノボ変異としての検出は不完全であり、解析方法を改善している。

## 4. まとめ

精神疾患に寄与が大きいと予想されるデノボ変異を、転移因子含め解析中しており、病態解明につながると期待される変異を複数確認した。転移因子解析はまだ不完全であり、さらなる検出及びバリデーションを進める。

## 5. 今後の計画・展望

精神疾患トリオにて確立したパイプラインを、体細胞変異検出に拡張し一定の成果を得た。転移因子含め、体細胞変異の解析を追加することで、より網羅的な変異検出を行う。転移因子の検出については、解析方法の改善を重ね、次年度にバリデーション実験を行う予定である。精神疾患に対して寄与の大きい変異を網羅的に同定したいと考えている。

2019 年度 利用研究成果リスト

**【ポスター発表】**

西岡将基、「多施設共同トリオ解析による双極性障害の遺伝的構造の解明」、第 115 回日本精神神経学会学術総会、2019 年 6 月 20 日 (朱鷺メッセ:新潟コンベンションセンター)