

厚生労働科学研究費補助金
難病・がん等の疾患分野の医療の実用化研究事業（再生医療関係研究分野）
研究分担報告書

腎・泌尿器系の稀少遺伝性疾患の原因遺伝子群の網羅的解析系の開発

研究分担者 小崎 健次郎 慶應義塾大学医学部 臨床遺伝学センター 教授

研究要旨

本研究全体の目標は、神経疾患特異的iPS細胞を用いた効率的な創薬システムの構築と高スループット化を行うことである。iPS細胞については継代に応じて細胞に徐々に変異が蓄積することが知られており、創薬研究対象とする患者細胞の性質の明確化・均質化はプロジェクトの成功のために必須のプロセスといえる。遺伝子変異部位の確認を行うために、次世代シーケンサー等による網羅的なスクリーニングのあと、異常の疑われる部位についてサンガー法で確認を行う。今年度は網羅的なスクリーニングのための方法として、体細胞変異を検出するための手法を確立した。

A . 研究目的

本研究全体の目標は、神経疾患特異的iPS細胞を用いた効率的な創薬システムの構築と高スループット化を行うことである。iPS細胞の作成時・継代時・分化誘導時に細胞に徐々に遺伝子変異が蓄積することが知られている。創薬研究を行う場合、患者細胞への遺伝子変異の蓄積の有無および程度を明確化することは、臨床応用を前提として均質な細胞を得るためには必須のプロセスといえる。特にiPS細胞はがん化により大きく性質を変えると懸念されるため、ガン関連遺伝子の蓄積についてプロファイリングが必要と考えられる。

本年は次世代シーケンサーを用いて、培養時の遺伝子変異の蓄積を網羅的に把握するための手法を開発した。

B . 研究方法

一般に入手可能な神経線維腫症患者由来の細胞株XとX由来にて著明な増殖能を獲得した細胞株Yをヌードマウスに移植し、増殖した細胞のゲノムDNAを抽出し、4800の既知ヒト疾患関連遺伝子の変異を網羅的に同定した。Yに存在し、Xに存在しない遺伝子変異を抽出することのためのプログラム群の導入と比較をおこなった。

ゲノムDNAを断片化したのち、TruSightOneキット（イルミナ社）を用いてヒト疾患との関連が報告されている全4800遺伝子の翻訳領域のゲノムDNAを回収した。次世代シーケンサーMiSEQ（イルミナ社）を用いて得られた粗配列を、プログラムBWAを用いて、ヒト参照配

列hg19に整列させた。Picardを用いて重複配列を除去し、GATKで再配列してbam形式・vcf形式の出力ファイルを得た。vcf形式のファイルに対してSnEffおよびannovarでアノテーションを行った。

体細胞変異の有無を検出するため、bam形式の出力ファイルを2種のプログラム（MuTect, Strelka）により比較検討した。MuTectはBroad Institute, StrelkaはStrelka Instituteが開発したプログラムである。

MuTect

Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol. 31:213-9, 2013. <http://www.broadinstitute.org/cancer/cga/mutect>

Strelka

Saunders CT, Wong WS, Swamy S, Becq J, Murray LJ, Cheetham RK. Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. Bioinformatics. 28:1811-7, 2012. <ftp://strelka@ftp.illumina.com>

Linuxオペレーティングシステム的环境下で32コアCPU・メインメモリ200GB超を有するサーバーを用いてパフォーマンスを検討した。得られた変異について、variant toolsを用いてアノテーションを行い、包括的に比較した。San Lucas FA1, Wang G, Scheet P, Peng B.

Integrated annotation and analysis of genetic variants from next-generation sequencing studies with variant tools. San Lucas FA1, Wang G, Scheet P, Peng B. Bioinformatics. 28:421-2, 2012.
<http://varianttools.sourceforge.net>

得られた遺伝子変異のリストのうち、タンパク翻訳領域内の遺伝子変異のみについて体細胞遺伝子変異の有無と質を評価した。アノテーションにあたっては、COSMICデータベース Catalogue Of Somatic Mutations In Cancer <http://cancer.sanger.ac.uk/> を用いた。

(倫理面への配慮)

入手可能なヒト細胞株を用いた検討のため、ゲノム指針の適応とならない。患者検体を用いた解析はヒトゲノム指針に従い、慶應義塾大学医学部倫理委員会の承認を得て解析を行った

C . 研究結果

1検体あたりの計算時間はMuTectで、約8時間で終了した。Strelkaの計算速度は、使用するCPUコア数に完全に依存した。28コア使用時には5分程度で計算を終了することができた。

一塩基置換について検討したところ、MuTectとStrelkaの両方で検出しえた変異が15個認められた。そのうちの1個は、NF1原因遺伝子と同一のパスウェイに属する遺伝子で、細胞の形質の変化との関連が強く示唆された。MuTectのみで検出し得た変異が1個、Strelkaのみで検出し得た変異が5個であった。Strelkaのみで検出された5個の変異のうち2個について次世代シーケンサーのリードを目視で確認したところ擬陽性と考えられた。

Strelkaの使用により、4800遺伝子中、1遺伝子について欠失を同定することができた。われわれは昨年、MuTectを用いて、Simpson-Golabi症候群患者の末梢血ゲノムDNAと当該患者に発症した肝芽腫のゲノムDNAを比較し、肝細胞ガンの発症に関与する カテニン体細胞変異を同定・報告している。今回、Strelkaで再解析し、カテニンの変異を再確認し得た。Strelkaによって欠失の有無をスクリーニングしたが、欠失を認めなかった。

D . 考察

前述のごとく、MuTectは一塩基置換の検出に特化したプログラムであり、欠失・重複(indel)の検出は不能である。iPS細胞に発生

する体細胞変異にはindelも含まれている。このためMuTect単体による検討では不十分なスクリーニングとなる。

一塩基置換については、MuTectよりもStrelkaの方が多くの変異が同定された。すなわち、今回使用したパラメータでは、MuTectの特異度が高く、Strelkaの感度が高いともいえる。Strelkaのみで検出された2個の変異部位については、Xにもごく少数、変異アレルと同じアレルが存在していた。MuTectはデフォルトで「正常対照」に設定したサンプルに変異アレルが認められた場合は、少量('alternate allele in normal')であっても、その変異を無視する仕様になっている。Strelkaのみで検出された2個の変異についても、alternate allele in normalが認められた。

文献上、iPS細胞を作成するオリジナルの細胞のゲノムDNAに変異がわずかに(数パーセント)存在し、同じ変異が作成後のiPS細胞にヘテロ接合体(50%)として検出される場合が報告されている。オリジナルの細胞群に既に体細胞変異を有する細胞がモザイクの状態が存在し、正の淘汰を受けた可能性が推測される。このような状況を検出するためには、MuTectのデフォルトの 'alternate allele in normal' オプションを外すがまたはStrelkaを用いた解析が必要と考えられる。

本研究の目的は、前述のごとくiPS細胞のハイスループットスクリーニングである。Strelkaは、マルチコア環境下での並列処理を可能としたプログラム構成となっており、32コアCPU下では、30ジョブ(15細胞の解析に相当)を同時に分析可能であり、3プログラムでは1細胞あたりの計算時間は最短となる。スループットを重視するスクリーニング環境下では、Strelkaの単体使用を、さらに精度を重視する場合は、StrelkaとMuTectの併用が進められると考えられた。

一般にMuTectは体細胞変異を有する細胞の割合が低い場合にも、感度・特異度が維持されるとの利点の報告がある。iPS細胞の細胞塊についても変異を有する細胞と変異を有しない細胞がモザイクとして混在する状況は十分想定される。したがって、変異を有する細胞の割合が低い場合にはMuTectの仕様が望まれる。今回の研究では、変異を有する細胞の割合が低い場合のMuTectとStrelkaのパフォーマンスの比較を行っておらず、今後の検討課題と考えられた。

E . 結論

iPS化時に発生する点突然変異の検出には、計算時間と感度の両方からは計算プログラム Strelkaの使用が最も適していると考えられた。計算資源や時間に余裕があれば、MuTectを併用すべきと考えられた。本研究の実施を通じて、ハイスループットスクリーニングのためのパイプラインを構築することができた。スループットに応じて、使用するプログラムを適宜組み合わせ使用して使用する計画である。

F．研究発表

1．論文発表

Kosaki R, Takenouchi T, Takeda N, Kagami M, Nakabayashi K, Hata K, Kosaki K. Somatic CTNNB1 mutation in hepatoblastoma from a patient with Simpson-Golabi-Behmel syndrome and germline GPC3 mutation. Am J Med Genet, in press

2．学会発表

なし

G．知的財産権の出願・登録状況

1．特許取得

なし

2．実用新案登録

なし

3．その他

なし