

論文の内容の要旨

論文題目 クリニカルシーケンスに向けた神経遺伝性疾患のターゲットリシーケンス解析とエクソーム解析の検討

佐藤 奈穂子

次世代シーケンサー(Next generation sequencer; NGS)を利用したゲノムDNAの配列解析は、標的領域の大きさから全ゲノム解析 (whole-genome sequencing; WGS)、全エクソン解析 (whole-exome sequencing; WES)、ターゲットリシーケエンシングに分類できる。WGSは、非翻訳領域も含めたゲノム全体を解析できる最も網羅的な方法であるが、コスト及びデータ量が多いことが問題である。一方、ヒトゲノムのおよそ1%に当たる全エクソン及びエクソン/イントロン境界に病原性変異が存在することが多いことから、これらの領域を濃縮して解析するWESが、コスト面も考慮され、臨床現場では最も広く使われている。ターゲットリシーケエンシングは、目的とする限られた遺伝子群など特定の領域のみを濃縮するためWESと比較してさらに解析にかかる時間と費用を抑えることができるため、対象の疾患や遺伝子がある程度絞られている場合により実用的である。NGSの実用化に伴い、臨床現場で診断を目的とした遺伝子検査(クリニカルシーケンス)が広く行われつつあるが、臨床面において時に研究面よりも高い精度を担保する必要性と、症例毎に適切な解析方法、変異の解釈の方法を検討する必要性から、ここでクリニカルシーケンスの方法について多方面から検討を行う必要があると考えられる。

今回、効率的かつ精度の高いクリニカルシーケンスへ向けて、神経遺伝性疾患の診断を目的とした解析方法の選択及び実際の臨床応用について検討を行った。はじめに、当科における診断未確定の神経遺伝性疾患5例を対象とし、ヒトゲノムの全エクソンを解析の対象とするWESと、既知のメンデル遺伝病の原因遺伝子のエクソン領域を対象とする2つのターゲットリシーケエンシング (Panel 1, 2)について、解析までの効率と結果の精度を比較した。Panel 1, 2は標的領域が既知のメンデル遺伝病の原因遺伝子のエクソン領域であり、WESと比較すると6-7分の1へ縮小された。1サンプルあたりのインプットに必要なゲノムDNAの量は、WESが3 µgと多く、Panel 1, 2は50 ngと少ない量での解析が可能であった。WESとPanel 1はハイブリダイゼーションにRNAを用いるのに対

し、Panel 2ではDNAが用いられる。また、WESは大規模塩基配列にシーケンサーとしてHiseqを用いる。最大32のサンプルを一度に解析することができるがデータ量が多い分、解析に要する時間は11日間と長い。WESよりも解析対象の領域が小さなPanel 1, 2はMiseqを用いるが、Miseqは解析可能なデータ量がHiseqの100分の1であり解析時間も2日間と短かった。ライブラリ調整システムが自動化されているWESは、ライブラリ調整に要する手間が最も少ないが、症例の蓄積に時間を要するため解析までの時間は最も長い。得られたデータに関して、被覆度は全体にPanel 1が最も高く、次いでWES、Panel 2の順であった。Panel 2では、断片化したDNAライブラリの大きさが300-1000塩基と長いいため、濃縮されたあともエキソンから外れてalignmentされる傾向があり、今回解析対象としたエキソンとスプライス部位の被覆度については低くなる傾向があり、被覆度を高くするためにはPanel 1もしくはWESが適していると考えられた。

塩基の重複率はWESが最も低かった。ゲノムDNAの断片化を機械的にではなく酵素反応やトランスポゾンを用いて行くと、挿入バイアスが増大するため重複率が高くなることが知られている。WESではライブラリ調整の際のゲノムDNAの断片化には超音波を用いて機械的に行う。一方でPanel 1, 2は酵素反応によりゲノムDNAを断片化するため、WESで低い重複率となったと考えられた。

被覆度の分布については、Panel 1で最も幅広い分布を呈していたが、最も低い平均被覆度に合わせるようにデータ量を減じた解析では3つのプラットフォーム間の差は縮小し、ほぼ同等の結果であった。共通の標的領域のうち、20X以上の被覆度を得られたエキソンの分布については、大部分が全てのプラットフォームで共通して20X以上の被覆度を得られていたものの、Panel 1のみで20X以上の被覆度を得られたエキソンの割合が2.9 %と、Panel 2のみの0.5 %、WESのみの0.9 %と比べ高かった。Panel 1は先述の通り平均被覆度が他のプラットフォームよりも明らかに高く、全体で高い被覆度の得られた領域が広範囲であると考えられる。20X以上でシーケンスできたエキソンの違いは、プラットフォームによって濃縮、キャプチャーする範囲が異なっており、その違いを反映していると考えられる。

解析結果から得られたdataのうち、3つのプラットフォームで異なる判定がなされた255の塩基について直接塩基配列決定法により確認を行い、プラットフォームによる正

答率を比較したところ、Panel 1が最も正答率が高く、次いでWES、Panel 2の順であり平均被覆度の高い順と一致していた。さらに、被覆度及び塩基のquality value (QV)ごとの正答率は、低い被覆度/QVにおいて、Panel 1、WES、Panel 2の順で高く先述の2つのパラメーターと一致していた。プラットフォームにより長所および短所があり、診断の目的や被覆される領域なども合わせて考慮する必要がある。

なお、対象の5症例のうち、1例でNGSの結果をもとに診断が確定した。patient 5は*FKTN*上に既知のp.R179Tをヘテロ接合性に全てのプラットフォームで共通して認め、直接塩基決定法で同じ変異を確認した。PCR法を用いて同遺伝子の3 kb挿入変異をヘテロ接合に確認したため、Cardiomyopathy, dilated, 1Xの診断となった。patient 3は、*DNM2*にc.2370_2372delGCCを認め、直接塩基決定法で同じ変異を確認した。同変異の報告はないが、コントロールのデータベースに登録がなく、病的変異の可能性が考えられたが、家系内の他の構成員の協力を得ることができなかった。2例とも、考慮されたvariantは全てのプラットフォームで一致していた。

続いて、プラットフォームの比較の結果に基づいて、当科における診断未確定の筋疾患 12 症例に対し、クリニカルシーケンスの応用と変異の解釈に関して検討した。アメリカ臨床遺伝学会 (American College of Medical Genetics; ACMG)の発表したガイドラインでは、変異の解釈を病原性の高い順より pathogenic, likely pathogenic, uncertain significance, likely benign, benign の 5 段階に分類しており、pathogenic または likely pathogenic となった変異が病原性変異として判断される。今回検討した対象 13 個の rare variant のうち、うち 10 個 (77%, 7 例) は病的変異が検出できたと考えられ、3 個 (3 例) については慎重な判断を要すると考えられたが、これらの変異を ACMG によるガイドラインに則り病原性の評価を行うと、病原性変異と考えられる、pathogenic または likely pathogenic に分類された variant は 13 個の候補変異のうち *FKTN* 上の 3kb insertion および p.R179T、*DYSF* 上の p.W999C、*COL6A1* 上の p.G419S、*SCN4A* 上の p.T704M の *CAPN3* 上の p.R440W、*LDB3* 上の p.A147T の 7 つ(54%)であり 77%から低下した。ACMG のガイドラインに則り病原性を判断すると、variant の病原性が既知のものであった場合や、家系内で遺伝子検査可能なメンバーが多数おり、変異の共有または非共有が表現型に矛盾しない場合は病原性ありと判断できる要素となるが、過去に報告のない

変異や、家系内の遺伝子検査可能な構成員が確保できない場合には **uncertain significance** となり、このガイドラインのみの判断では、臨床的に変異と考え得るような **variant** が見出された場合にも **uncertain significance** と判断されてしまうことに注意が必要である。

診断の確定した変異だけでなく、**uncertain significance** に分類された **variant** についても病原性を支持する情報をデータベースに登録することにより蓄積し、分類を更新していくことが重要である。情報が蓄積しデータベースが充実していくことで、確定診断率の上昇が望めるばかりでなく、さらには治療へ向けた研究の発展への利用が期待できる。

今回の検討では、主に **SNVs** と短い挿入・欠失変異に注目したが、診断に向けて、染色体構造変化や **repeat expansion** の検出も含めた解析方法の構築も必要である。今後はクリニカルシーケンスへの応用へ向けて、臨床へのフィードバックの方法についても検討が必要である。