

【目的】 遺伝性疾患の原因変異として、これまで主にコード領域内でアミノ酸置換やフレームシフトを起こすものやスプライス部位を壊すものが同定されてきた。しかし、近年、新たなスプライス部位を形成する変異により通常用いられるエクソンとは異なるエクソン（偽エクソン）が活性化されることで有害な転写産物がもたらされるために疾患となる事例が報告されるようになった。このような変異は、機能を持っていない領域で起こる変異であることから、これまで見過ごされることが多かった。そこで本研究は、スプライス部位を形成することで遺伝性の希少疾患の原因となるレアバリエントの網羅的な同定を目的とした。本研究は、遺伝性疾患について、これまであまり注目されてこなかったスプライス部位形成変異に着目することで新たな原因変異を同定し、確定診断が困難なことが多いそれらの疾患の診断率の向上に貢献することを目指すものである。

【方法】 対象とする疾患は特定のものに絞らず、遺伝性疾患全般について原因変異を探索する疾患横断的なアプローチをとった。方法としては、バイオインフォマティクス解析を中心とし、公共データベースに蓄積しているヒトの変異データに対し、スプライス部位の予測プログラムを適用することで、スプライス部位形成能を持った有害な変異を探索した。得られた候補変異については、その変異を有する個人のトランスクリプトームデータを解析し、実際にその変異がスプライスに変化をもたらしているかを検証した。これらを通し、様々な遺伝性の希少疾患について、その原因となる新たな変異を同定した。

【結果】 遺伝性疾患の原因となることが知られている約 4,000 遺伝子を対象に、ヒトの変異データベースに登録されている約 7 万人分の変異データの中から、既知エクソン内、あるいはその周辺に存在する変異を抽出した。それらの変異について、統計的手法および人工知能を応用したスプライス部位予測法を適用することで、各変異のスプライス部位形成能を評価した。その結果、5,656 個のスプライス部位形成変異を得た。さらに、これらの変異について、コドン読み枠のズレや終止コドンの挿入、タンパク質ドメイン構造内の挿入・欠失などにより、病原性が考えられるかを評価したところ、3,942 個の変異を病原性があると判断された。実際の変異の機能について、その変異を持つ個人の RNA-seq データで検証したところ、実際にスプライス部位を形成する変異であることが確認された。今回の手法を、原因変異が未同定の織毛病患者から得られた変異データに適用したところ、新規原因候補となるスプライス部位形成変異の同定に成功した。

本研究で解析対象とした変異とそれらにより生成される異常なエクソン

