

Nanopore sequencing reveals TACC2 locus complexity and diversity of isoforms transcribed from an intronic promoter

メタデータ	言語: English 出版者: 公開日: 2019-03-20 キーワード (Ja): キーワード (En): 作成者: 伊藤, 陽介 メールアドレス: 所属:
URL	https://jair.repo.nii.ac.jp/records/2002294

授与機関名 順天堂大学

学位記番号 甲第 2072 号

Complexity of biomarker TACC2 for predicting endometrial cancer lymph node metastasis

(子宮体癌リンパ節転移予測バイオマーカーTACC2 の複雑性)

伊藤 陽介 (いとう ようすけ)

博士 (医学)

論文内容の要旨

先行研究で子宮体癌のリンパ節転移予測マーカーの一つとして新規転写開始点(p10)から転写される p10 TACC2 を報告した。さらに PCR 法による p10 TACC2 の増幅産物(アンプリコン)を大腸菌を用いてサブクローニングし、プライマーウォーク法およびサンガーシーケンスによって配列決定を行った結果、非常に多様な p10 TACC2 バリエントが存在することが明らかになった。多様なエクソン構造および数十~百塩基前後の短い塩基配列の挿入・欠失が示唆される複雑なバリエントの網羅的探索において、ショートリードシーケンサーを用いた従来法ではそれぞれの全長配列を再現することは不可能であった。そこで本研究では、全長配列を保持したままのシーケンスを可能とする一分子長鎖シーケンサー

(MinION) を用いて全長配列の網羅的探索を実行した。既報の TACC2 遺伝子のバリエント頻度 (variant usage) を基準として p10 TACC2 の新規バリエントを決定し、p10 TACC2 のエクソン選択の特徴を確認することを目的とした。そこで子宮体癌細胞株 (JHUEM-1)、子宮体癌原発腫瘍組織、正常子宮内膜組織から p10 TACC2 アンプリコンを作成し、それぞれの配列解析を行い比較した。JHUEM-1 のシーケンス結果において、新たに 889 の新規 p10 TACC2 バリエントを同定した。各バリエントのリード数は、総リード数で除した usage (頻度) として正規化した。先行研究で配列決定された既報の 8 バリエントの usage を閾値に設定すると、閾値以上の 41 バリエントが p10 TACC2 バリエント全体の 80% 以上を占めることが明らかになった。そこで既報の 8 バリエントを除く 33 バリエントを usage (頻度) の高い新規 p10 TACC2 バリエントとして同定した。エクソン usage は、各エクソンのリード数を総リード数で除して算出した。細胞株、癌組織、正常内膜の各検体におけるエクソン usage を比較すると、核移行シグナルと TACC domain 領域のエクソン usage はどの検体においても高く保持され、転写開始点近傍のエクソン 1 とエクソン 3 の usage も同等に高く保持されていた。さらに新規エクソン候補領域を 3 箇所同定した。

MinION シーケンスによる高感度で得られた新規 TACC2 の多様なバリエント結果は、先行研究における特異性の高いサンガーシーケンスの結果と組み合わせることによって、高い特異度で配列を決定することができた。新規 TACC2 の転写バリエントはその由来に関わらず多様性を有しており、転写機構は従来考えられてきたものより複雑であることが明らかとなった。一方で機能ドメインは高い保存性を示しており、翻訳産物であるタンパク質は、既知の塩基配列に基づくタンパク質と機能は変わらないと考えられた。ショートリードシーケンスではリードをつなぐアセンブリが必要であり、非常に多様なバリエントが存在する場合は断片化前の配列は推定であり完全復元は不可能となる。これに対し MinION を用いたアンプリコンシーケンスは、ロングリードシーケンスが可能であり、1つのアンプリコンの配列決定を連続した配列から行うことで網羅的探索が可能となった。ショートリードとロングリードのシーケンス結果を組み合わせる配列決定を行う方法はまだ確立されていないが、本研究は 1つの手法を提案する非常に有意義なものであったといえる。