

論文審査の結果の要旨

氏名 武田 淳一

本論文は4章からなり、第1章は完全長 cDNA を用いたヒトの選択的スプライシング (AS) バリエーションの同定とタンパク機能アノテーション解析、第2章はヒトとマウスの完全長 cDNA を用いた AS バリエーションの比較ゲノム解析、第3章は AS のデータベースである H-DBAS の構築、第4章は次世代シーケンサーから得た RNA-Seq タグを用いた AS バリエーションの翻訳検証について述べられている。ヒトの AS 解析は、これまで EST を用いた転写物の一部、遺伝子予測の情報を含む転写産物モデル、あるいはエクソン-エクソンジャンクションのプローブを用いることによって行われてきた。しかし、これらは転写物の 5'末端 (転写開始点) の情報が不十分なため、エクソンの数や順番に正確性を欠くという欠点がある。この欠点を克服したのが、5'末端が転写開始点である完全長 cDNA である。本論文では完全長 cDNA を用いることにより、転写物上のエクソンの位置が明確な AS バリエーションを初めてゲノムワイドに同定し、その特徴を解析している。さらに、EST、転写産物モデル、およびマイクロアレイのデータでは不可能であった、精度の高い AS バリエーションのタンパク機能アノテーション解析および比較ゲノム解析を行っている。これらの解析結果から、タンパク機能に影響を及ぼすヒトの AS 遺伝子が多いこと、一方でマウスとの種間保存 AS 遺伝子は非常に少ないこと、また、非保存 AS 遺伝子に濃縮して現れるタンパク機能は生殖に関係していることを明らかにした。ヒトの AS バリエーションは、計算機による自動アノテーションだけではなくマニュアルでもアノテーションを行っている。これにより、典型的な AS パターンに合致しない複雑な AS パターンを見出している。本論文ではこの他に、次世代シーケンサーを使用した RNA-Seq 解析による AS バリエーションの翻訳検証を行っている。ヒト DLD-1 細胞のポリソームおよび核画分由来 RNA-Seq タグのマッピング結果により、翻訳している (ポリソーム由来) または翻訳していない (核由来) と考えられる AS バリエーションを区分した。また、これらの解析結果を網羅的に検索、そしてインタラクティブに閲覧できるデータベースの構築を行った。

本論文は、5'末端が転写開始点である完全長 cDNA を用いることにより、転写物上のエクソンの位置が明確な AS バリエーションのゲノムワイドな同定および解析を初めて行った。さらに、次世代シーケンサーの膨大なデータを用いることにより、細胞単位で AS バリエーションに翻訳情報を付加することを可能にした。このような精度および網羅性の高い情報を用いた解析は、未だ不明な点の多い AS が寄与する細胞内遺伝子機能における役割の解明に向けて大きく貢献するものであると考えられたために、博士 (生命科学) を授与するのに適当であると判断された。