OUKA
The University of Osaka
Institutional Knowledge Archive

Title	RNAIndel: discovering somatic coding indels from tumor RNA-Seq data
Author(s)	萩原, 浩平
Citation	大阪大学, 2021, 博士論文
Version Type	
URL	https://hdl.handle.net/11094/85267
rights	
Note	やむを得ない事由があると学位審査研究科が承認したため、全文に代えてその内容の要約を公開しています。全文のご利用をご希望の場合は、大阪大学の博士論文についてをご参照ください。

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

論　文　内　容　の　要　旨
Synopsis of Thesis

氏　名 Name	萩原浩平
論文題名 Title	RNAIndel: discovering somatic coding indels from tumor RNA-Seq data（RNAIndel：癌トランスクリプトームにおける体性欠損挿入変異の発見アルゴリズム）

論文内容の要旨

〔目　的(Purpose)〕

トランスクリプトームシークエンシングは遺伝子発現や融合転写産物の解析に用いられるが、変異検出の目的に用いることは困難である。その理由としてライブラリーを調整する際に用いるPCRで導入される変異や、スプライス部位周囲のマッピングが不完全なために生じる変異、正常組織のトランスクリプトームは癌組織とは遺伝子発現パターンが異なるためにコントロールして用いることができないことがあげられる。しかし、トランスクリプトームはシークエンスとデータのストレージにかかるコストが比較的低いことと、発現している遺伝子を見ているため、表現型への解釈がしやすい利点がある。本研究では特に解析が困難な挿入・欠損変異をトランスクリプトームで解析する手法を開発した。

〔方法ならびに成績(Methods/Results)〕

挿入・欠損変異は常に一意に配列されるとは限らないため、変異配列を再配列し、同じ変異ならば同じ配列パターンとして認識される再配列の工程を開発し実装した。この再配列の結果にもとづき、がんゲノムに付随する体性変異に特徴的な塩基配列のパターンや遺伝子産物への効果を抽出し、それら特性量を学習させることでトランスクリプトーム中の挿入・欠損変異を体性・胚性・アーティファクトに分類した。小児がんを含む５種類の癌トランスクリプトームデータで性能を評価し、８８−１００％の感度を偽陽性数が一塩基挿入・欠損変異と多塩基挿入・欠損変異ともに１０個以下で得た。現在、トランスクリプトームでの変異検出で広く使われる手法にくらべても１４倍低い偽陽性数で１．５倍以上感度よく検出できた。さらに、非小細胞肺がんデータにおいて、同一症例から採取された肺正常組織をコントロールとして用いて体性変異検出に特化したアルゴリズムと比較したところ、本手法は正常組織をコントロールとして用いないにもかかわらず、一塩基挿入・欠損変異で１０倍低い偽陽性数で２倍の感度、多塩基挿入・欠損変異で２倍低い偽陽性数で１．５倍の感度を得た。大腸がんのデータを例に、マイクロサイト不安定性が存在する場合は本手法では効率よく検出できないことが示された。

〔総　括(Conclusion)〕

本研究は挿入・欠損変異を分類する特性量を定義し、生物学的に興味深い。特に、挿入・欠損配列と周辺配列間の編集距離として定義された量は体性変異に特異的であることが示された。それら特性量の計算と学習のアルゴリズムはソフトウエアとしてまとめられた。

論文審査の結果の要旨及び担当者

（申請者氏名）萩原浩平

	（職）		氏　名
論文審査担当者	主　査	大阪大学教授	高島成二
	副　査	大阪大学教授	岡田随象
	副　査	大阪大学教授	谷内田真一

論文審査の結果の要旨

がんにおける体性変異の同定法として、正常細胞とがん細胞のDNAを比較する方法は現在もっとも広く用いられている。コストや、表現型への解釈の容易さという面ではRNAの利用は従来から興味をもたれていたが、RNAシークエンシング中に発生する大量のアーチファクト等の問題のため、一塩基置換の検出に関する報告が散見されるにとどまっていた。本研究は、RNAから検出する変異の範囲を、一塩基置換に次いで多い変異であるインデルに拡張する点で高く評価できる。この拡張は、がんに付随するインデルが特徴的な配列を持ちやすいことを利用しており、生物学的な根拠に基づくものであり興味深い。またこの工夫により、本手法はがん組織のRNAのみから体性インデルを同定し、様々な制約からコントロールとしての正常組織が入手できないことがあるという、実際のがんゲノムシークエンスにおける運用上の問題も解決している。

これらの発見は、博士（医学）の学位授与に値するものと認める。