

Title	ロングリードRNA-seqを用いた融合遺伝子検出手法に関する研究
Author(s)	増田, 圭吾
Citation	大阪大学, 2024, 博士論文
Version Type	VoR
URL	<a href="https://doi.org/10.18910/96227">https://doi.org/10.18910/96227</a>
rights	
Note	

*Osaka University Knowledge Archive : OUKA*

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

Osaka University

## 論文内容の要旨

氏名 ( 増田 圭吾 )

論文題名

ロングリードRNA-seqを用いた融合遺伝子検出手法に関する研究

## 論文内容の要旨

融合遺伝子は2つの正常な遺伝子が融合した異常な遺伝子である。融合遺伝子は腫瘍形成において原因的な役割を果たし、ヒトのがんの罹患原因の約20%を占めると言われている。また、融合遺伝子は医薬品の標的となると同時に薬剤の効果を予想するためにも利用され、融合遺伝子を正確に検出手法は、臨床現場や医薬品開発現場で求められている。

近年、個々の転写産物の全長の配列を読み取るロングリードRNA-Seqが可能となり、融合遺伝子の検出に活用され始めている。ロングリードRNA-Seqで読み取られた塩基配列（リード）を参照ゲノムの相同な塩基配列に整列（アライメント）させて融合遺伝子を検出する方法が一般的である。しかし、ロングリードRNA-Seqで得られるリードは塩基の読み取り精度が低いという欠点がある。この欠点のため、アライメントされない領域（ギャップ）がリードの融合点近傍に発生し、先行手法では融合遺伝子および融合点が正しく検出できないという問題点がある。

本研究では、新規の融合遺伝子検出手法を提案することで、先行手法の問題点を解決し、新規の融合遺伝子の発見につなげることを目的とした。提案手法は(1)融合点のエキソン境界固定、(2)ギャップの再アライメント、(3)融合点のクラスタリングの3つ処理を加えて拡張した手法である。(1)と(2)は融合点をエキソンの境界に合わせて補正することで、融合点を正確に特定する狙いがあり、(3)はギャップによって発生する融合点のずれを吸収することで、融合遺伝子の誤検出を防ぐ狙いがある。

本論文は4章から構成される。第1章では、本研究における背景と目的について述べた。第2章では、ロングリードRNA-Seqを用いた新規の融合遺伝子検出アルゴリズムを提案し、シミュレーションデータセットを用いて提案手法の融合遺伝子の検出性能を評価した。その結果、提案手法は先行手法よりも高い融合遺伝子の検出性能を有していることを示した。また、前述の(1)と(2)により検出できる融合遺伝子が増加すること、(3)により融合遺伝子の誤検出が減少することを示した。第3章では、培養がん細胞のシーケンスデータから融合遺伝子を検出するために提案手法の拡張を行ない、複数の培養がん細胞を用いて提案手法の融合遺伝子の検出性能を評価した。その結果、提案手法は先行手法よりも高い融合遺伝子の検出性能を有していることを示した。これにより、新規の融合遺伝子の候補を提示できることを示した。第4章では、本研究の結論と将来の展望について述べた。

## 論文審査の結果の要旨及び担当者

氏 名 ( 増 田 圭 吾 )			
	(職)		氏 名
論文審査担当者	主 査	教 授	松田 秀雄
	副 査	教 授	松田 史生
	副 査	准教授	小蔵 正輝
	副 査	准教授	瀬尾 茂人
	副 査	准教授	下田 雅史 (大阪大学大学院医学系研究科)

## 論文審査の結果の要旨

融合遺伝子は2つの正常な遺伝子が融合した異常な遺伝子である。融合遺伝子は腫瘍形成において原因的な役割を果たし、ヒトのがんの罹患原因の約20%を占めると言われている。また、融合遺伝子は医薬品の標的となると同時に薬剤の効果を予想するためにも利用され、融合遺伝子を正確に検出する手法は、臨床現場や医薬品開発現場で求められている。

近年、個々の転写産物の全長の配列を読み取るロングリードRNA-seqが可能となり、融合遺伝子の検出に活用され始めている。ロングリードRNA-seqで読み取られた塩基配列（リード）を参照ゲノムの相同な塩基配列に整列（アライメント）させ、複数の遺伝子にアライメントされるリードを融合遺伝子として検出することが広く行われている。しかし、ロングリードRNA-seqで得られるリードは塩基の読み取り精度が低いという欠点がある。この欠点のため、アライメントされない領域（ギャップ）がリードの融合点近傍に発生し、先行手法では融合遺伝子を正しく検出できないという問題点があった。

本研究では、先行手法の問題点を解決した融合遺伝子検出手法を提案することで、新規の融合遺伝子の発見につながることを目的とした。提案手法は、(1)融合点のエキソン境界固定、(2)ギャップの再アライメント、(3)融合点のクラスタリングの3つの処理を先行手法に加えて拡張した手法である。(1)と(2)は融合点をエキソンの境界に合わせて補正することで融合点を正確に特定する狙いがあり、(3)はギャップによって発生する融合点のずれを吸収することで、融合遺伝子の誤検出を防ぐ狙いがある。

本論文は4章から構成される。第1章では、本研究における背景と目的について述べている。第2章では、ロングリードRNA-seqを用いた新規の融合遺伝子検出アルゴリズムを提案し、シミュレーションデータセットを用いて提案手法の融合遺伝子の検出性能を評価している。その結果、提案手法は先行手法よりも高い融合遺伝子の検出性能を有していることを示している。また、前述の(1)と(2)により検出できる融合遺伝子が増加すること、(3)により融合遺伝子の誤検出が減少することを示している。第3章では、培養がん細胞のシーケンズデータから融合遺伝子を検出するために提案手法の拡張を行ない、複数の培養がん細胞を用いて提案手法の融合遺伝子の検出性能を評価している。その結果、提案手法は先行手法よりも高い融合遺伝子の検出性能を有するとともに、新規の融合遺伝子の候補を提示できることが示されている。第4章では、本研究の結論として、融合遺伝子を正確に検出し、新規の融合遺伝子の候補を提示することで医学・薬学分野に貢献できる可能性を示唆している。

よって、博士（情報科学）の学位論文として価値のあるものと認める。