

令和 5 年 5 月 31 日現在

機関番号：12601

研究種目：若手研究

研究期間：2021～2022

課題番号：21K15566

研究課題名（和文）肺がん組織における空間的遺伝子発現および局所変異解析手法の構築

研究課題名（英文）Development of spatial genome and transcriptome analytical methods in local regions of lung cancer tissues

研究代表者

鈴木 絢子（Suzuki, Ayako）

東京大学・大学院新領域創成科学研究科・准教授

研究者番号：00770348

交付決定額（研究期間全体）：（直接経費） 3,500,000円

研究成果の概要（和文）：本研究開発は、肺がん組織の局所的な進展の分子メカニズムを明らかにするために、空間トランスクリプトーム解析技術とロングリードシーケンス技術を駆使して、肺がん組織局所のゲノム変異およびトランスクリプトームステータスを解析する手法を開発するものである。本研究では、空間トランスクリプトームVisiumを実施した計8症例の肺腺がん凍結手術検体を対象に、Visiumライブラリ調製の過程で得られる全長cDNAを増幅し、ナノポアシーケンサーで解読した。点変異やスプライシングパターンを抽出することができ、位置バーコードを通じて遺伝子発現クラスターや空間情報と結びつけることができた。

研究成果の学術的意義や社会的意義

本研究で解析対象としている肺がん組織は、不均一かつ複雑ながん微小環境状態を有しており、その分子レベルでの理解には、空間情報・病理組織学情報を考慮したオミクス計測技術が重要な役割を果たすと考えられる。本研究では、空間トランスクリプトーム技術およびロングリードシーケンス技術といった新規ゲノム解析技術を駆使して空間オミクス解析のための実験・情報解析手法の開発を実施している。近年急速に普及が進む空間トランスクリプトーム解析技術をより発展的にがん研究に応用するための基盤整備を行うことができている、学術的意義や社会的意義は十分であると考えている。

研究成果の概要（英文）：In this study, to understand molecular mechanisms of local progression in lung cancer tissues at multilayered levels, we developed basic experimental and computational procedures for detecting genomic mutations and aberrant transcriptome statuses in local regions of cancer tissues by utilizing spatial transcriptome sequencing and long read sequencing technologies. We used fresh frozen specimens of eight lung adenocarcinoma cases which had been provided to spatial transcriptome analysis Visium. We amplified the full-length cDNA libraries with Visium spatial barcodes and sequenced them using a nanopore-type long read sequencer PromethION. We extracted point mutations and splicing patterns from the obtained long reads and integrated them with expression clusters and spatial information via spatial barcodes.

研究分野：がん多層オミクス解析

キーワード：空間トランスクリプトーム解析 ロングリードシーケンス技術 がん

1. 研究開始当初の背景

がん組織は、がん細胞、免疫細胞、間質細胞などのさまざまな細胞種が混ざり合った細胞社会である。近年、これら細胞集団をいわゆるバルクのみではなく、1細胞ごとに分離し、そのオミクス状態を計測するシングルセル解析が広く行われてきた。しかし、シングルセル解析手法は、細胞を単離して解析を行うため、空間情報や細胞間の相互作用の情報が失われてしまい、がん組織の局所領域におけるオミクス状態の理解をするには不十分な点が多い。

空間トランスクリプトーム解析技術は、従来のシングルセル解析では見逃されてきた、細胞間の位置関係、相互作用、および、病理組織学・形態学的特徴を保持したまま、組織の局所領域におけるトランスクリプトーム解析が可能な技術である。空間トランスクリプトーム技術である Visium (10x Genomics 社) は、直径 55 μm (数~数十細胞) のスポットごとに、空間的に遺伝子発現量の計測を行うことができる技術である。約 5000 のスポットが配置されたスライド上に凍結組織切片を置き、スライド上で組織の透過、mRNA の遊離、および、逆転写反応を実施する。各スポットにはそれぞれ固有の位置バーコードが付加されており、それらバーコードを含む cDNA を合成、シークエンスすることにより、由来となるスポットが判別できる。研究代表者らのグループでは、Visium を用いた乳がん (非浸潤性乳管がん; DCIS) の解析を実施し、悪性化の母地となる局所進化におけるオミクスステータスの異常について明らかにし、報告している (Nagasawa S, Kuze Y et al. 2021 *Commun Biol*)。

本研究で着目する肺腺がんは、ドライバー遺伝子による層別化が最も進んだがんの一つである。各遺伝子のゲノム変異を指標として患者層別化が行われる。EGFR 遺伝子の変異は日本人肺腺がんの約半数で見られ、EGFR キナーゼ阻害薬の投与の対象となる。この阻害薬は EGFR 変異陽性例には強い治療効果を示すが、治療の過程で一定の割合で薬剤耐性クローンが出現する。このような治療抵抗性は、EGFR 遺伝子そのものに生じた耐性変異や、他のドライバー遺伝子における変異、または、小細胞肺がんといった別のサブタイプへの転化などが原因で生じるとされる。

肺がんの進展や治療抵抗性獲得の背景には、がん組織における腫瘍内不均一性および微小環境が関係していると考えられる。しかし、いまだその全貌は明らかではない。肺がん組織そのものの不均一性、腫瘍内微小環境の状態、および、その相互関係性をオミクスレベルで解明することが、肺がん組織の分子状態、さらには、表現型のコントロールにとって極めて重要であると考えた。肺腺がんには、EGFR 以外にも、さまざまなドライバー遺伝子が報告されており、また、ドライバー不明例も 30%ほど存在する。タンパク質コード領域の変異はその多くが mRNA の中央部に存在するため、現在、Visium にて取得する 3' 端シークエンスでは解読が不可能である。本研究ではこれらの同定にナノポア型長鎖シークエンサーを用いる。先行した科研費若手研究においては、ナノポアシークエンサーを用いて、がん細胞の転写産物全長解析が可能であることを示してきた。本手法では、空間トランスクリプトーム解析とロングリードシークエンス解析を融合し、ゲノム変異と遺伝子発現量の双方について空間多様性の解析を可能にする。ゲノム・トランスクリプトーム上における肺がん組織の多様性を解析することができれば、これまで大きな問題であった治療抵抗性獲得機序の解明、および、適切な層別マーカーのなかった症例に対する新たな治療法の開発につながるのではないかと考えている。

2. 研究の目的

本研究では、空間トランスクリプトーム解析において、空間的遺伝子発現情報を取得すると同時に、局所領域のゲノム変異ステータスを解析する基盤技術を開発することを目指す。また、実際に肺腺がん組織におけるゲノム・トランスクリプトーム情報を空間的に取得し、空間オミクス統合解析を実施する。特に、数十 kb におよぶ長い配列を一本のシークエンスリードで読み切ることができるナノポア型長鎖シークエンサーを用いて、Visium より得られる位置バーコード付き cDNA を全長にわたって解読し、cDNA 上の点変異を検出する手法の基盤整備を行う。

3. 研究の方法

(1) 空間トランスクリプトーム Visium 上での変異解析手法の検討と実践

まず、Visium の cDNA 増幅産物を、ナノポアシークエンサーで解読するための実験条件を検討する。EGFR や KRAS 変異などのドライバー変異に加えて、がん抑制遺伝子などの変異にも着目する。がん抑制遺伝子の変異は機能喪失型であることも多く、mRNA の品質管理機構による分解を受けて発現量が低くなり検出が難しい場合があることに注意する。

まずは、マウス組織 (腎) を用いて条件検討を行う。クライオスタットにて、10 μm 厚の凍結切片を薄切後、Visium 用スライド上で組織の透過、逆転写を行い、得られた cDNA を増幅する。増幅した cDNA は遺伝子発現解析のためのライブラリ調製に供し、残りの一部を、ナノポアシークエンサーを用いた全長 cDNA のタイピングに用いる。

また、本研究では、代表的な肺腺がん症例について、凍結手術検体を取得し、解析を実施する。解析対象例には、EGFR 変異陽性および陰性例を含む。また、病理組織学的な観点から、高分化な肺胞上皮置換型から低分化な充実型の組織像を含む不均一な症例、もしくは、それぞれを個別に含む症例を選択することを目指す。

マウスでの検討と同様に、10 μm 厚の凍結切片より、Visium 解析を実施している。その際に得られる位置バーコード付き全長 cDNA 増幅産物の一部をナノポアシークエンサーにより解読する。うまくいかない場合は、ハイブリキャプチャーによる標的がん関連遺伝子の濃縮を試行する。

(2) 肺腺がん組織における空間的遺伝子発現・変異統合解析

得られたデータにおける、点変異、および、遺伝子発現パターンについて解析する。点変異については、ナノポアシークエンサーで解読した cDNA 配列より、分子バーコード (UMI: unique molecular identifier) や位置バーコード (SB: spatial barcode) が抽出できたリードに対して解析を行い、検出する。コピー数異常については、各スポットの遺伝子発現パターンより inferCNV といったツールを用いて予測を行う。また遺伝子発現情報については、すでに先行研究にて報告している肺がん遺伝子発現モジュールを写像することにより、各スポットにおける遺伝子発現状態をスコア化する (Suzuki A, Onodera K et al. 2019 *Sci Rep*)。細胞間相互作用については各スポットの成分分解により細胞種の構成比を算出、またリガンド-受容体の関係性などを参考に、細胞種ペアの関係性そのものをスコア化する。それらを病理組織上に配置し、オミクスの遷移状態を空間情報と重ね合わせる。

また、得られた解析結果をバリデーションするために、局所の微小切片よりオミクス解析を行う手法を開発、実施する。上記で Visium の解析に使用した凍結切片の近傍切片を取得し、ダイセクション装置や微小切片採取装置等を用いて、注目領域ごとに局所組織片を回収する。得られた組織片より、バルクのゲノムシークエンスや RNA-seq を行い、上記の空間的オミクス解析の結果と照らし合わせ、評価を行う。

4. 研究成果

(1) 空間トランスクリプトーム上での変異解析における実験・情報解析手法の基盤構築

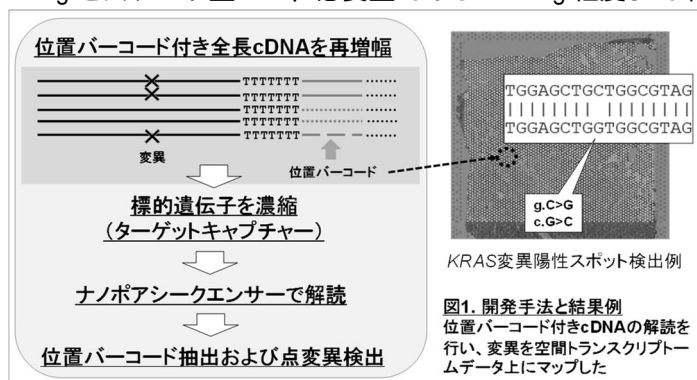
図1(左図)のように、実験・情報解析手法の基盤構築を実施した。まず、Visium の全長 cDNA テンプレートを再増幅し、ナノポアシークエンサーでの解析に必要な量まで増やした。マウス腎および肺がん検体での検討の結果、10 ng をスタート量とし、必要量である 500 ng 程度まで増幅を行うことができた。

次に、ナノポアシークエンサーを用いてシークエンス解析を実施した。計画当初は、MinION を用いる予定であったが、より多くのシークエンスリードを得ることができる大型機 PromethION を用いた。まず、肺がん 2 検体 (LUAD-1 および 4) より、得られた Visium 全長 cDNA を再度増幅した。得られた cDNA 増幅産物に対してライブラリ調製を実施し、PromethION でシークエンス解析を実施した。その結果、それぞれ、69,479,107 および 97,423,791 本の 1d pass リードを取得することができた (WT: whole transcriptome; 表 1A)。

得られたリードは SiCeLoRe toolkit (Lebrigand K et al. 2020 *Nat Commun*) を用いた解析に供した。このツールを用いた主な解析内容は以下のとおりである。まず、通常の Visium プロトコルにより、cDNA をショートリードシークエンス解析し、解析ツール Space Ranger (10x Genomics 社) で一次情報解析したデータを用意した。これらデータより、各リードの SB、UMI、および、遺伝子の組み合わせについてリファレンス情報を作成した。次に、PromethION より得られたロングリードより、アダプターおよび poly-A 配列を有するリード (想定通りのライブラリ構造を有するリード) を抽出した (表 1B)。さらにヒトゲノムリファレンス配列にマップし、遺伝子領域にマップできたリードより UMI および SB を抽出した。ショートリードデータで作成したリファレンス情報と照らし合わせ、最終的に、それぞれ 7,975,872 および 10,344,456 リードの cDNA 配列について、正しい UMI/SB 情報を抽出できたことが確認できた。

UMI/SB 情報が紐づいた cDNA 配列を解析することで、各スポットのゲノム変異の抽出を試みた。あらかじめバルクのゲノムシークエンス解析により、該当症例のゲノム変異をリスト化しておき、SAMtools 等を用いて、各変異を有する cDNA リードの抽出を行った。しかし、転写産物を網羅的にシークエンス解析したため、ハウスキーピング遺伝子等の高発現遺伝子にリードの多くがアサインされてしまい、がん関連遺伝子をカバーするリードの数は十分得ることができなかった。

そこで、がん関連遺伝子のリードを濃縮する手法を試みた。がん関連遺伝子を搭載した NCC Oncopanel を用いて、ハイブリダイゼーションによるターゲットキャプチャーを実施し、再増幅



した Visium 全長 cDNA ライブラリからがん関連遺伝子の cDNA を濃縮した。濃縮された cDNA 長をバイオアナライザで確認したところ、元の cDNA ライブラリ長の分布と遜色ない長さの cDNA がキャプチャーできていることが分かった。これらを同様に PromethION でのシーケンス解析に供したところ、それぞれの症例について、137,670,943 および 115,533,728 リードが得られた (Cap: target captured; 表 1A)。また、最終的に、34,870,111 および 19,998,169 リードにおいて、UMI・SB を正しく抽出できた。ターゲットキャプチャーを行ったほうが正しいライブラリ構造を有する cDNA 配列を効率よくシーケンスできることが分かった。また、9 割以上のリードが標的遺伝子にマップされることも確認できた。

表1. ナノポアシーケンサーでのVisium cDNA解析結果

(A). シーケンス統計

	LUAD-1		LUAD-4	
	WT	Cap	WT	Cap
1d reads	82,494,773	137,670,943	113,178,332	115,533,728
Total bps	43,897,736,596	107,490,130,503	75,234,809,773	99,299,371,068
1d average length (bp)	532	781	665	859
1d average qscore	9.8	10.3	10.1	9.9
1d pass reads	69,479,107	118,201,042	97,423,791	99,903,559
1d pass average length (bp)	550	790	681	868
1d pass average qscore	10.6	11.2	10.9	10.7

(B). リードのフィルタリングとUMI/SB抽出結果

	LUAD-1		LUAD-4	
	WT	Cap	WT	Cap
1d pass reads	69,479,107	118,201,042	97,423,791	99,903,559
Passed (poly-A/adapters found)	29,381,110	77,012,530	43,665,259	52,946,938
Passed & mapped	25,147,921	76,207,693	34,992,627	52,442,008
UMI/SB found	7,975,872 [27.1%]	34,870,111 [45.3%]	10,344,456 [23.7%]	19,998,169 [37.8%]

(2) 肺腺がん組織における空間的変異解析・遺伝子発現解析

本研究グループでは、肺腺がん計 8 症例の凍結手術検体の Visium を実施している。うち 6 症例が *EGFR* 変異陽性、2 症例が *KRAS* 変異陽性例である。ドライバー変異不明例は含まれていない。本研究では、ライブラリ調製過程で得られる Visium 全長 cDNA ライブラリを PromethION でのシーケンス解析に供した。

まず、Visium のショートリードデータを用いて、空間的遺伝子発現解析を実施した。Space Ranger により一次解析を行い、出力結果を R パッケージ Seurat にて解析した。クラスタリング解析を実施し、各クラスターで特徴的な発現パターンを有する遺伝子群を抽出した。特に、分化度やストレス応答に関わる遺伝子に着目して、クラスターのアノテーションを行い、高分化で悪性度の低い領域と、ストレス応答や脱分化がみられる比較的悪性度の高い領域にクラス分けした。また、研究代表者らがすでに構築済みの肺がんを特徴づける遺伝子発現ネットワークモジュール (Suzuki A, Onodera K et al. 2019 *Sci Rep*) を用いたスコアリングも実施した。組織局所のがん細胞における細胞系譜・分化度・ストレス応答ステータスを詳細に定義づけることができ、また酸化ストレス応答モジュールが全体的に亢進している症例を見出すこともできた。

次に、PromethION での全長 cDNA シーケンスデータを取得し、変異解析を実施した。上記検討にて、ターゲットキャプチャーによる濃縮が有効であると明らかとなったため、ターゲットキャプチャーを引き続き用いることとした。得られた PromethION リードは、上述と同様の情報解析手法を用いて解析し、UMI・SB が抽出できたリード上における変異の有無を判定した。肺がん関連遺伝子の多くは 3' UTR の長い転写産物を有することが報告されており、3' 端から標的の変異部位までがかなり長い。そのため、増幅ができない、リード長が足りないといった問題点が想定されていたが、結果として変異領域をカバーするリードが取得できた。poly-A の位置の違いや internal poly-A priming 等による影響と考えられる。ドライバー変異である *EGFR* および *KRAS* のホットスポット変異についてそれぞれの症例において確認したところ、図 1 (右図: LUAD-14 の *KRAS* 変異) のように検出することができた。また、各スポットのスプライシングパターンも確認することができた。

本研究では、さらに各領域における変異ステータスの確認のために、Visium に供した組織切片の近傍より切片を取得し、微小切片採取装置等を用いて局所の微小組織切片を採取して、シーケンス解析を実施するための実験検討も実施した。シーケンス解析に供するのに十分な質・量のゲノム、エピゲノム (DNA メチル化)、RNA シーケンスライブラリを調製することができた。

本研究では、不均一かつ複雑な微小環境を有するがん組織において、空間情報を考慮した空間オミクス解析の実験・情報解析手法の開発を実施した。また実際に、肺腺がん検体を用いた解析を行った。近年急速に普及が進む空間トランスクリプトーム解析技術をより発展的にがん研究に応用するための基盤整備を行うことができたと考えている。本研究成果については、関連する日本語総説にて紹介し、また学会のセミナーにて発表した。

5. 主な発表論文等

〔雑誌論文〕 計0件

〔学会発表〕 計2件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 鈴木絢子
2. 発表標題 ロングリード解析とシングルセル・空間解析の統合に向けて
3. 学会等名 第45回日本分子生物学会年会（バイオテクノロジーセミナー）
4. 発表年 2022年

1. 発表者名 Suzuki A, Araki T, Suzuki J, Zenkoh J, Suzuki Y.
2. 発表標題 Elucidation of diverse transcriptome statuses and their spatial transition in lung adenocarcinoma tissues
3. 学会等名 第44回日本分子生物学会年会
4. 発表年 2021年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

<p>本研究に関連した日本語総説を執筆した。 ・藤井元人，鈴木絢子，水野裕介，関真秀，鈴木稜「ロングリードシーケンシングを用いたscRNA-seqおよびVisiumの全長cDNA解析」空間オミクス解析 スタートアップ実践ガイド，実験医学別冊 最強のステップUPシリーズ，鈴木稜 / 編，羊土社，pp.47-55</p>

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------