

令和 3 年 5 月 20 日現在

機関番号：12601

研究種目：若手研究

研究期間：2019～2020

課題番号：19K16108

研究課題名（和文）ナノポアシーケンサーを用いたRNA合成及び分解の新規解析手法の開発

研究課題名（英文）Development of novel measurement methods for RNA synthesis and degradation using nanopore sequencing

研究代表者

関 真秀（Seki, Masahide）

東京大学・大学院新領域創成科学研究科・特任准教授

研究者番号：90749326

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：本研究では、人工的に導入したRNA修飾塩基を標的にナノポアシーケンスによる検出の可能性を検証し、転写及び転写後制御の多くの方法論簡略化に向けて基盤技術を創出することを目的とした。人工塩基である4sUをCに塩基変換するTUC-SeqとR2C2法を用いたナノポアシーケンスでの高精度のシーケンスを組み合わせることで、新生RNA及び既存のRNAの全長構造を検出する実験手法、データ解析の開発に成功した。さらに、がん特異的なスプライシングアイソフォームの全長構造の同定を行った。がんにおいて、それらのアイソフォームが蓄積する機構を開発した方法に応用した。

研究成果の学術的意義や社会的意義

細胞の中にあるRNAは、その合成と分解の合計によって総量が決まります。RNAの合成と分解を測定するためのこれまでの多くの方法では、短いRNAの配列を決めるため、どのような全長配列を持ったRNAがどのような合成や分解の制御を受けているのかは十分にわかっていませんでした。今回、合成された、または、元からあるRNAをラベリングして、ナノポアシーケンサーというRNAの全長の配列を解読できるシーケンサーで検出する方法を確立しました。この方法を利用することで、どのような全長配列を持ったRNAがどのような分解や合成制御を受けているのかが明らかになることに貢献できると考えられます。

研究成果の概要（英文）：In this study, we aimed to verify the feasibility of detection of artificially introduced RNA-modified bases by nanopore sequencing and to create a fundamental method for analysis of transcription and post-transcriptional regulation. By combining TUC-Seq, which converts an artificial base, 4sU, to cytosine, with high-accuracy sequencing by nanopore sequencing using the R2C2 method, we successfully developed an experimental method and pipeline for data analysis to detect 4sU-labeled and unlabeled RNA-derived reads. Furthermore, we identified the full-length structures of cancer-specific splicing isoforms. We applied the method to analyze the mechanism of accumulation of these isoforms in cancer.

研究分野：ゲノム科学

キーワード：ナノポアシーケンス トランスクリプトーム

様式 C - 19、F - 19 - 1、Z - 19 (共通)

1. 研究開始当初の背景

ゲノムから mRNA へ、遺伝子発現制御について転写開始制御を中心に以前から多くの研究が行われてきた。その一方で、転写伸長あるいは RNA 修飾、タンパク質への翻訳効率といった転写後の制御過程については、依然として不明な部分が多い。近年盛んにおこなわれる RNA-Seq により測定される発現量は、複数の制御素過程の結果の総和である。遺伝子発現制御の全体像の理解には、mRNA の転写が開始され最終的に分解される、個々の過程を解析することが必須である。近年、次世代シーケンサーを用いて、それぞれの制御過程について網羅的に解析する方法論の開発がなされている。RNA の新生鎖の検出法としては GRO-Seq(Global run-on-sequencing) (Core et al. 2008) など、RNA の分解速度の測定法としては BRIC-Seq (bromouridine immunoprecipitation chase-sequencing) (Tani et al. 2012) などの方法が開発されている。これらの 2 つの手法に代表されるように、多くの RNA ダイナミクスの解析には核酸アナログを取り込ませるメタボリックラベリング法が用いられる。例えば、これら 2 つの方法では、プロモウリジン (BrU ; bromo-uridine) を取り込ませて、BrU に対する抗体で BrU を取り込んだ RNA を免疫沈降 (IP) することで、ラベルされた RNA 量を検出する。これらは核酸アナログに対する IP を介した方法では、煩雑な多段階に実験工程を要する。その過程での試料ロスも大きく、要求される RNA インプット量も多い。また、含有される BrU が多い RNA 分子ほど高率で回収される傾向がある等、いくつかの重要な技術上の困難を有するために、多くの研究室で日常的に実施できるまでにはその方法論は成熟していない。近年、SLAM-Seq (thiol-linked alkylation for the Metabolic sequencing) と呼ばれる IP を介さない方法が開発されてきているが、細胞毒性の高い 4sU (4-thiouridine) を使用する必要があるなどの問題点が存在する (Herzog et al. 2017)。

ナノポアシーケンサー MinION や PromethION は、100kbp 以上の長鎖核酸のシーケンスや、直接 RNA をシーケンスする direct RNA-Seq (dRNA-Seq) が行えるなど、これまでのシーケンサーにない特徴を有している。核酸がナノスケールの穴であるナノポアを通過する際に起こる電流のかく乱 (squiggle) パターンから塩基配列を推測することでシーケンスを行う。そのパターンはナノポアに取り込まれる塩基の種類によって異なるため、性質の異なる修飾塩基がナノポアに導入された場合、通常の塩基とは異なるパターンが生じるため修飾塩基の検出することが可能である。これまでに、DNA の m5C、hm5C や m6A、RNA の m6A や m5G の各種のメチル化の検出を行えることが示されている (Garalde et al. 2016, Simpson et al. 2017)。申請者らは、ヒト肺腺がん細胞株の mRNA の dRNA-Seq を行った (Seki et al. 2019)。イルミナ社の RNA-Seq で測定した遺伝子発現との比較を行った。相関係数が 0.92 と非常に高い相関を示しており、定量性の高いトランスクリプトーム解析を行えることを示した。また、1本のリードで mRNA の全長をカバーでき、そのリードが 70 万本以上得られることを確認した。

2. 研究の目的

本研究では、人工的に導入した RNA 修飾塩基を標的にナノポアシーケンスによる検出の可能性を検証し、転写及び転写後制御の多くの方法論簡略化に向けて基盤技術を創出することを目的とした。具体的には、以下の 2 点を目的とした。

・人工修飾塩基のナノポアシーケンスデータからの検出

dRNA-Seq、あるいは、全長 cDNA をシーケンスする cDNA-Seq データから、人工核酸アナログの検出法の確立を目指した。さらに、それらの検出法の感度について、評価した。

・ナノポアシーケンスを利用した RNA 合成と RNA 分解速度の測定方法の確立と応用

BrU を利用した既存の手法である GRO-Seq と BRIC-Seq に、上記の核酸アナログの検出方法を応用し、IP を必要としない RNA 合成と分解の測定法の開発を行う。がん特異的に発現する転写産物についてこれらの方法で解析し、それらの合成・分解制御を明らかにすることを目的とした。

dRNA-Seq や cDNA-Seq の利点は、全長の配列決定が行えることがあげられる。そのため、RNA スプライシングパターンと RNA 合成や分解との関係性を明らかにできると考えた。また、ナノポアのデータからラベルされた mRNA 由来のリードを検出できれば、IP を必要せず、実際の値に近い合成量や分解速度を測定できると考えた。

3. 研究の方法

当初は、dRNA-Seq データからの核酸アナログを直接検出することで、ラベルされた RNA を検出することを試みていたが、研究開始後に本研究と同様のコンセプトの手法開発の論文 (Drexler et al. 2020; Maier et al. 2020) が公開された。また、dRNA-Seq には 100 µg 程度と大量の total RNA が必要になることや、核酸アナログは読み取り精度への悪影響が大きいことが

ら、より実用的な手法を確立するために、微量の RNA から実施可能で、読み取り精度も高い cDNA-Seq を利用することとした。cDNA-Seq では、逆転写により、核酸アナログは通常の塩基に置き換えられるため、4sU をシトシンに変換する幾つかの既存の手法について検討を行った。

検討には、4sU でのラベリングを実施実績のある大腸がん細胞株 DLD-1 株を用いた。まず、SLAM-Seq での塩基変換を行ったところ、SLAM-Seq の反応で用いるアルキル化の副反応により、全長の cDNA 合成が阻害されることが判明した。そこで、TUC-Seq(Riml et al. 2017)と呼ばれる塩基変換法に採用し、条件検討を進めた。塩基変換した RNA から、全長 cDNA を合成してナノポアシーケンサー PromethION でシーケンスを行った。得られたリードをゲノム上にマッピングしたところ、T から C への塩基置換の頻度は数%程度であった。この塩基置換の頻度でラベルされた RNA とラベルされていない RNA 由来のリードを区別することは、通常のナノポアシーケンサーの読み取り精度では困難であった。そこで、全長 cDNA を環状化し、RCA 法で増幅し、ナノポアシーケンスする R2C2 法(Volden et al. 2018)を組み合わせた。R2C2 法では同一リード上に同じ cDNA 由来の配列がタンデムに存在することになるため、コンセンサス配列を取ることで高精度の配列を得ることができる。予備検討として、DLD-1 の RNA について R2C2 法を実施し、コンセンサスの回数ごとに読み取り精度について集計を行った(図 1)。4 個以上の配列からコンセンサスが取られた cDNA 配列を使用することにより、99%以上の精度のシーケンスが得られることが確認できた。次に、非ラベルと 8 時間もしくは 24 時間の 4sU でラベリングした DLD-1 から RNA を抽出して、TUC 変換後、R2C2 法を用いて解析した。4 個以上の配列から取得されたコンセンサス配列の T から C への置換率を算出したところ、ラベル時間に依存して T->C 塩基置換を持ったリードの割合が上昇することが確認できた(図 2)。

データ解析パイプラインについては、T->C 置換の高いリードを分類するパイプラインと RNA スプライシング構造を同定・集計するパイプライン(Oka ... *Seki, *Genome Biol.* 2021; Seki et al., *Methods Mol. Biol.* 2021)を構築を行った。よって、新生と既存の RNA ごとの mRNA の全長構造を同定・発現解析を行う実験的・情報学的パイプラインの開発に成功した。

肺がん細胞株と臨床検体について cDNA-Seq を行い、開発した RNA のスプライシング構造を同定するパイプラインを用いて解析を行ったところ、がんの特異的な RNA のスプライシング構造を持った 2000 個以上のスプライシングアイソフォームを同定することに成功した(Oka ... *Seki, *Genome Biol.* 2021)。また、がん特異的なアイソフォームの蓄積にスプライシング因子や RNA 品質維持機構の一つである NMD 因子の発現や変異の影響が示唆された。しかし、がん特異的なアイソフォームの生成メカニズムが不明な点が多い。がん特異的なアイソフォームが多かった肺がん細胞株 VMRC-LCD 株と中程度の A549 株について、開発した方法を応用してそれらのアイソフォームの生成及び蓄積機構について解析を進めている。

4. 研究成果

TUC-Seq と R2C2 法を組み合わせることで、新生 RNA と既存 RNA を分類し、スプライシングアイソフォームごとの RNA の合成と分解を測定する手法の開発に成功した。cDNA-Seq データから、T->C 置換率によりリードを分類するパイプラインやスプライシング構造を同定するためのパイプラインの開発を行った。肺がんにおいて、cDNA-Seq 法と開発したパイプラインを用いて、がん特異的な転写産物の全長構造のデータの収集を行い、投稿論文として発表した(Oka ... *Seki, *Genome Biol.* 2021)。本研究の過程で情報収集したナノポアシーケンサーを用いた DNA 及び RNA の修飾塩基の検出の技術的な進展について、Review 論文として発表した(Xu and *Seki, *J Hum Genet.* 2020)。また、direct RNA-seq や cDNA-Seq のデータから転写産物の全長構造同定のためのデータ解析について検討を行った結果を英語総説として発表した(Seki et al., *Methods Mol. Biol.* 2021)。さらに、本研究で得られた塩基変換についての知見を DNA エピゲノム解析に応用することで、新規の長鎖 DNA エピゲノム解析法の開発に成功し、投稿論文として発表した(Sakamoto ... *Seki, *Nucleic Acids Res.* 2021)。共同研究において、2 種類人工核酸を用いて短鎖シーケンサーによる RNA の合成と分解の同時測定法 Dyrec-seq の開発を行った(Kawata et al. *Genome Res.* 2020)。

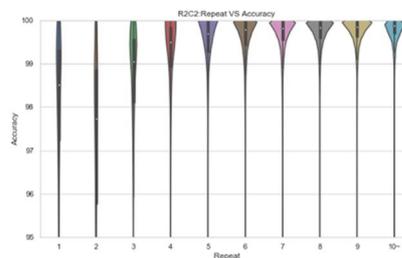


図 1. R2C2 リードのコンセンサス回数と塩基読み取り精度の関係

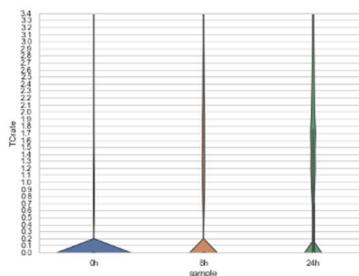


図 2. ラベリング時間ごとの T から C への置換率の比較

5. 主な発表論文等

〔雑誌論文〕 計5件（うち査読付論文 4件/うち国際共著 0件/うちオープンアクセス 4件）

1. 著者名 Xu Liu, Seki Masahide	4. 巻 65
2. 論文標題 Recent advances in the detection of base modifications using the Nanopore sequencer	5. 発行年 2019年
3. 雑誌名 Journal of Human Genetics	6. 最初と最後の頁 25 ~ 33
掲載論文のDOI（デジタルオブジェクト識別子） 10.1038/s10038-019-0679-0	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Oka Miho, Xu Liu, Suzuki Toshihiro, Yoshikawa Toshiaki, Sakamoto Hiromi, Uemura Hayato, Yoshizawa Akiyasu C., Suzuki Yutaka, Nakatsura Tetsuya, Ishihama Yasushi, Suzuki Ayako, Seki Masahide	4. 巻 22
2. 論文標題 Aberrant splicing isoforms detected by full-length transcriptome sequencing as transcripts of potential neoantigens in non-small cell lung cancer	5. 発行年 2021年
3. 雑誌名 Genome Biology	6. 最初と最後の頁 9
掲載論文のDOI（デジタルオブジェクト識別子） 10.1186/s13059-020-02240-8	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -
1. 著者名 Seki Masahide, Oka Miho, Xu Liu, Suzuki Ayako, Suzuki Yutaka	4. 巻 2284
2. 論文標題 Transcript Identification Through Long-Read Sequencing	5. 発行年 2021年
3. 雑誌名 Methods in Molecular Biology	6. 最初と最後の頁 531 ~ 541
掲載論文のDOI（デジタルオブジェクト識別子） 10.1007/978-1-0716-1307-8_29	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Sakamoto Yoshitaka, Zaha Suzuko, Nagasawa Sato, Miyake Shuhei, Kojima Yasuyuki, Suzuki Ayako, Suzuki Yutaka, Seki Masahide	4. 巻 49
2. 論文標題 Long-read whole-genome methylation patterning using enzymatic base conversion and nanopore sequencing	5. 発行年 2021年
3. 雑誌名 Nucleic Acids Research	6. 最初と最後の頁 in press
掲載論文のDOI（デジタルオブジェクト識別子） 10.1093/nar/gkab397	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Kawata Kentaro, Wakida Hiroyasu, Yamada Toshimichi, Taniue Kenzui, Han Han, Seki Masahide, Suzuki Yutaka, Akimitsu Nobuyoshi	4. 巻 30
2. 論文標題 Metabolic labeling of RNA using multiple ribonucleoside analogs enables the simultaneous evaluation of RNA synthesis and degradation rates	5. 発行年 2020年
3. 雑誌名 Genome Research	6. 最初と最後の頁 1481 ~ 1491
掲載論文のDOI (デジタルオブジェクト識別子) 10.1101/gr.264408.120	査読の有無 有
オープンアクセス オープンアクセスとしている(また、その予定である)	国際共著 -

〔学会発表〕 計0件

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

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
---------	---------