

# RNA 結合タンパク質の標的分子探索手法を開発

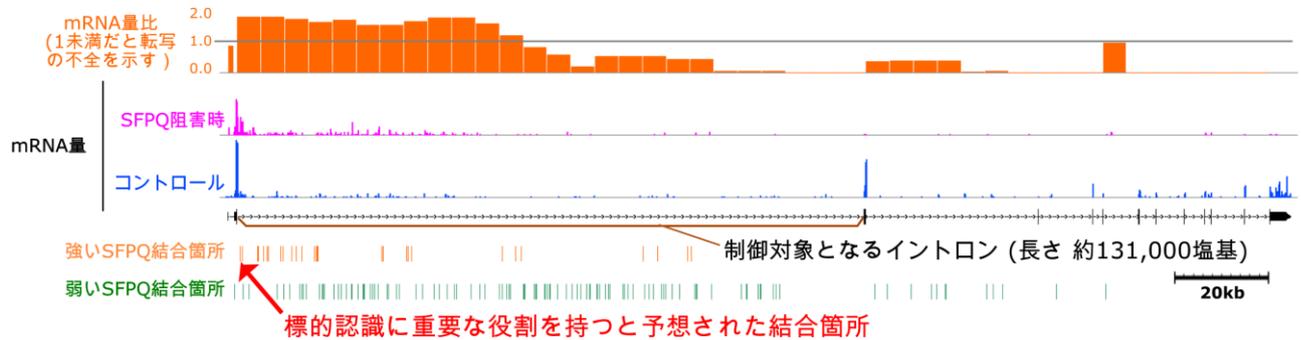
—複雑な RNA 制御のメカニズム解明に期待—

## 概要

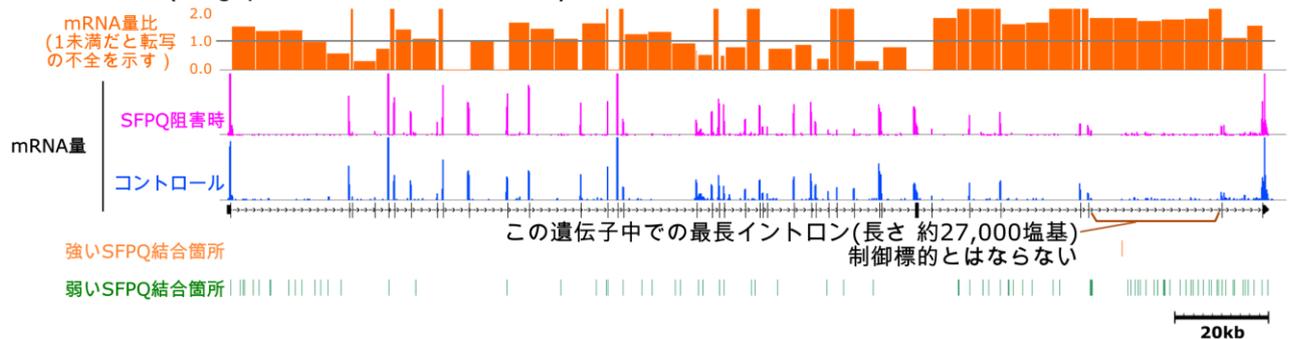
哺乳類の細胞において、遺伝情報は染色体 DNA から RNA に転写され、さらに様々な転写・転写後制御を受けることが知られています。RNA の制御を行う分子は RNA 結合タンパク質と呼ばれ、ヒトで 1500 種類以上が存在することが知られています。RNA の制御は、複雑な遺伝子発現制御を必要とする哺乳類で特に重要であり、その異常がさまざまな疾患を引き起こすことが知られています。一方、それぞれの RNA 結合タンパク質が数万種類の RNA のうちどれをどう制御するのが大きな謎となっています。京都大学大学院医学研究科飯田慶 特定助教、武内章英 同准教授、萩原正敏 同教授らの研究グループは、RNA 結合タンパク質の 1 つである SFPQ(Splicing Factor Proline And Glutamine Rich)をモデルとして、「RNA の機能的制御につながる結合」を探索するバイオインフォマティクス手法を開発しました。これにより、観察されていた 236,000 箇所以上の結合部位のうち、標的分子を決定する際に重要な役割を持つと考えられる約 200 箇所の同定に成功しました。これは、複雑な RNA 制御のメカニズムおよびその包括的制御の全容解明や、RNA 制御異常により引き起こされる疾患の治療法開発につながる成果です。

本研究成果は、2020 年 6 月 27 日に米国の国際学術誌「iScience」に掲載されました。

## 標的遺伝子 (*Samd4*, 長さ 約222,000塩基)



## 非標的遺伝子 (*Ralgapa1*, 長さ 約218,000塩基)



## 1. 背景

遺伝情報はゲノム DNA から RNA に転写され、さらにそこからタンパク質に翻訳されます。RNA 結合タンパク質 (RBP) はこの RNA に結合し転写および転写後調節(注1)を綿密に制御します。ヒトやマウスなどの哺乳類のゲノム上には約 1,500 個(全遺伝子の約 7%) の RBP が存在しており、これらが遺伝子発現制御の多様性を生み出していることが明らかになりつつあります。高速シーケンサーの発展や、実験技術の発達に伴い、細胞中で RBP が結合している RNA 結合箇所を網羅的に解読することが可能となってきましたが、これが解読されてもなお、どの結合箇所が機能的に重要なのか？どのように RBP の標的分子が認識されるのか？などの多くの重要な問題が未解明のまま残されています。

SFPQ は、2018 年に当研究グループが、神経発生の際に必要な不可欠な役割を持つことを見出した RNA 結合タンパク質の 1 つです(Takeuchi, Iida *et al.* 2018 *Cell Rep.*)。SFPQ は  $10^6$  塩基以上の大きさを持つ遺伝子 (遺伝子の長さの中央値は約  $10^5$  塩基) の遺伝子発現に必要であることが分かっていたのですが、制御の標的となる遺伝子とそうでない遺伝子がどのように区別されているのかは明らかにされていませんでした。

## 2. 研究手法・成果

本研究プロジェクトでは、SFPQ の標的特異性を生み出すメカニズムを明らかにするため、RNA-seq データと CLIP-seq データ(注2)という複数種類の大規模データを組み合わせてバイオインフォマティクス(注3)解析をするというアプローチを取りました (Multilateral bioinformatics analyses, 多層バイオインフォマティクス分析)。[データの観察]に基づき、[仮説構築]を行い、その[仮説を検証]する、というプロセスを繰り返すことで、SFPQ の標的領域や特異性に重要な結合箇所を絞り込んでいくという点を特徴としています(参考図)。SFPQ は  $10^6$  塩基以上の大きさを持つ遺伝子の転写に必要であることがわかっていましたが、本研究では、同程度の大きさを持つ遺伝子でも SFPQ を必要とする遺伝子と必要としない遺伝子が存在することを明確にしました(図1)。これを最初の手がかりとして、SFPQ を必要とする遺伝子は、大きなイントロン(注4)を持つこと、またイントロンの最上流に近い位置に特徴的な SFPQ 結合箇所があることを見出しました(図1)。同定された重要な結合箇所の情報から、さらに SFPQ の結合様式や制御機構、標的 RNA の予測ができるようになりました。

この解析により、SFPQ の標的分子がより正確に解明され、SFPQ が関わる転写後制御がどのように神経発生に寄与するのかについて理解が進むことが期待されます。また、SFPQ における遺伝子の変異は認知症等の疾患を引き起こすことが知られています。本研究成果はこれらの疾患メカニズムの解明に寄与することが期待されます。

## 3. 波及効果、今後の予定

RNA 結合タンパク質の研究において、実験的に結合箇所を決めた後であっても、機能に重要な結合部位や標的認識メカニズムがうまく解明できないケースは多くあると考えられています。本研究で示された解析コンセプトはこのようなケースの解決に役立ち、転写後制御によって生じる細胞の多様性の理解に貢献することが期待されます。我々研究グループも、さらに他の RNA 結合タンパク質を対象として、多層バイオインフォマティクス解析の実行、および手法の改良を行なっていく予定です。

#### 4. 研究プロジェクトについて

本研究は、下記機関より資金的支援を受けて実施されました。

- 日本学術振興会 (JSPS)・文部科学省 科学研究費補助金科学研究 費基盤研究 (A)、基盤研究 (C)
- 文部科学省 革新的細胞解析研究プログラム「セルイノベーション」
- 科学技術振興機構(JST)・戦略的創造研究推進事業(CREST)
- 厚生労働省厚生労働科学研究費
- 日本医療研究開発機構(AMED)

#### <用語解説>

注1) **転写および転写後調節**：DNA を鋳型として合成(転写)される RNA が受ける制御の総称。RNA の転写、成熟過程や細胞内局在の制御、安定性の制御や、タンパク質の翻訳量の調整などを含む。

注2) **RNA-Seq, CLIP-Seq**：高速シーケンサーを用いて細胞の遺伝暗号の読み出しやその制御の状態を、全遺伝暗号を対象に網羅的に読み解く解析方法。

注3) **バイオインフォマティクス**：コンピューターを用いて、細胞や生物由来のデータの処理を行ったり、データ間の関連性を見出したりすることで、生命科学の問題を解くことを目指す研究分野。

注4) **イントロン**：ゲノム DNA から RNA に転写されるものの、RNA の成熟過程において切り取られる RNA 領域のこと。成熟 RNA に残る領域はエクソンという。

#### <研究者のコメント>

RNA は細胞内の情報分子:DNA と機能分子:タンパク質をつなぐ非常に重要な分子です。また RNA そのものが機能を持つケースも数多くわかっています。本研究は RNA の転写後制御の多様性や制御様式の解明を目的としたバイオインフォマティクス解析手法の一端となります。本手法やさらなる解析手法の研究・開発を通じて、RNA 制御が作り出す細胞の多様性の解明や、疾患メカニズムの解明に寄与していきたいと考えています。

#### <論文タイトルと著者>

タイトル：Multilateral bioinformatics analyses reveal the function-oriented target specificities and recognition of the RNA-binding protein SFPQ (多層バイオインフォマティクス分析による、機能指向的な SFPQ の標的特異性と認識機構の解明)

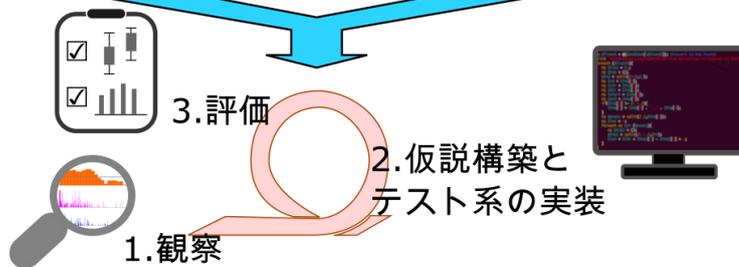
著者：飯田 慶、武内 章英、萩原 正敏

掲載誌：iScience DOI：https://doi.org/10.1016/j.isci.2020.101325

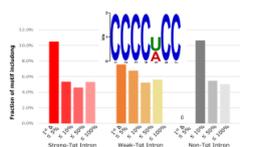
< 参考図表 >



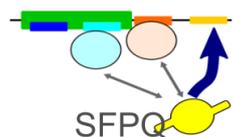
CLIP-Seq RNA-Seq



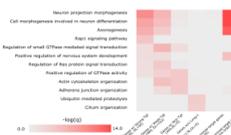
多層バイオインフォマティクス解析



機能に重要な結合箇所の解析



メカニズム解析



制御を受ける遺伝子の特徴解析