

令和 5年 9月 15日

## 「離れ合う転写」「近づき合う転写」を同定する ソフトウェア「CCIVR2」を開発

### <研究成果のポイント>

- 「離れ合う転写」ディバージェント転写、「近づき合う転写」コンバージェント転写による転写産物を網羅的に同定できるソフトウェア「CCIVR2」を開発しました。
- 11種類の生物のゲノムデータをもとに、転写産物のスタート・ゴール地点の特徴を同定しました。
- 特定の生命現象の特徴的な発現変動を示すディバージェント転写産物を同定しました。
- 本研究成果は医学・生物学分野のさらなる発展につながることを期待されます。

※本研究成果は、英国電子版科学誌「**Scientific Reports** (サイエンティフィック・リポーツ)」に日本時間9月8日に公表されました。

### <概要>

浜松医科大学分子生物学講座の大畑樹也学内講師、医学科6年の鈴木麻耶らのグループは、同大学先進機器共用推進部および器官組織解剖学講座との共同研究により、同グループが以前発表したソフトウェア「CCIVR」に、ディバージェント転写やコンバージェント転写による転写産物<sup>\*1</sup>を同定する機能を追加した「CCIVR2」を新しく開発しました。この成果は、英国電子版科学誌「**Scientific Reports** (サイエンティフィック・リポーツ)」に日本時間9月8日に公表されました。

### <研究の背景>

DNAは+鎖と-鎖からなる2本鎖の形で存在し、それぞれのDNA鎖上にたくさんの遺伝子が存在しています。+鎖と-鎖ではDNAの並びが反対方向になっているため、遺伝子が転写される方向も反対になります。したがって、ある2つの遺伝子がそれぞれ+鎖と-鎖に存在するとき、この2つの遺伝子は別々の方向に転写されます。このとき、+鎖と-鎖の2つの遺伝子が両方向に離れ合うように進む転写を「ディバージェント転写」、両方向から近づき合うように進む転写を「コンバージェント転写」と呼びます。このような両方向性に起こる転写の特徴や機能についての情報は限られていました。

### <研究手法・成果>

私たちは以前、転写のスタート地点またはゴール地点において、別の遺伝子と重なり合っている遺伝子を探し出すことができるソフトウェア「CCIVR」を開発していました。そこで今回は「CCIVR」の技術を応用し、転写のスタート地点またはゴール地点において“重なり合う遺伝子”だけでなく、“重なり合っていないが、ある程度近接している遺伝子”も同時に探し出すことができる機能を追加した、進化版ソフトウェア「CCIVR2」を新たに開発しまし

た。すなわち、ひとつの生物の遺伝子データから、ディバージェント転写やコンバージェント転写による転写産物を網羅的に抽出できるようになりました。この機能を用いて行った解析により、新たな知見が得られました。

まず、CCIVR2 および異なる 11 種の生物のゲノムデータを用いて、ディバージェント転写、コンバージェント転写が起きる遺伝子の位置関係を解析しました。その結果、転写される遺伝子がタンパク質コード遺伝子<sup>\*2</sup> 同士の組合せでは、転写のスタート地点、ゴール地点の付近で特定の配列（プロモーター<sup>\*3</sup>やポリ A シグナル<sup>\*4</sup>）を共有しているものが多くあることを示唆する結果が示されました。

さらに、ES 細胞<sup>\*5</sup>の分化と上皮間葉転換（EMT）<sup>\*6</sup>という 2 つの生命現象の RNA-seq<sup>\*7</sup> データを用いて CCIVR2 解析を行いました。その結果、これらの生命現象に関わりのある遺伝子群の中に、相関、逆相関の発現変動を示すディバージェント転写産物が複数見つかりました。この遺伝子群の中には、ES 細胞の分化や EMT といった特定の生命現象に関与するものも含まれていました。このことは、ディバージェント転写の一部は遺伝子発現調節を介して特定の生命現象に関与する可能性が示唆されました。

## <今後の展開>

我々が開発したソフトウェア CCIVR2 はディバージェント・コンバージェント転写産物を同定するだけでなく、遺伝子の位置情報に関するマクロな解析や、発現量の情報を結び付けた機能的な解析にも応用できる有用なオープンソースのツールです。CCIVR2 により新たに発見されるディバージェント・コンバージェント転写産物の中には、疾患の発生病序に関与し新規の治療ターゲットになるものも期待されます。私たちは多様なデータを CCIVR2 解析することでさらなるディバージェント・コンバージェント転写産物の探索を進め、生命現象の実態を明らかにするとともに医学の発展に貢献していきたいと考えています。

## <用語解説>

- \*1 転写産物：RNA ポリメラーゼによりゲノム DNA を鋳型として合成された RNA の総称。
- \*2 タンパク質コード遺伝子：ゲノム上に存在する遺伝子のうち、タンパク質のアミノ酸配列を指定している遺伝子。RNA ポリメラーゼによって転写されたのち、いくつかのプロセスを経て mRNA となり、タンパク質へ翻訳される。
- \*3 プロモーター：遺伝子の転写のスタート地点の上流にある領域。ここに基本転写因子や RNA ポリメラーゼが結合し、転写複合体を形成することで遺伝子の転写が開始する。
- \*4 ポリ A シグナル：遺伝子の転写領域のゴール地点付近には、ポリ A シグナルと呼ばれる特定の配列が存在する。タンパク質コード遺伝子では、転写されたポリ A シグナルの下流にポリ A 鎖が付加されることで mRNA が安定化される。
- \*5 ES 細胞：哺乳類の発生の初期段階である胚盤胞から、内部細胞塊と呼ばれる部分を取り出し培養した細胞。全身のあらゆる組織の細胞に分化できる能力「多能性」をもつ。生物学において多くの実験に用いられる。
- \*6 上皮間葉転換（EMT）：細胞の形態は大きく上皮系と間葉系に分けられる。もともと上皮系であった細胞が間葉系に変化する現象を EMT と呼ぶ。EMT は発生、創傷治癒、癌の転移などに関わるとされている。
- \*7 RNA-seq：遺伝子発現は DNA の配列が RNA に転写されることによって起こる。RNA-seq は、次世代シーケンサーによって細胞内にある RNA の配列を網羅的に読み、その情報からどの遺伝子がどの程度転写されたかを算出することで、発現量を調べる解析方法である。

### <発表雑誌>

Scientific Reports (サイエンティフィック・リポーツ)  
(DOI: 10.1038/s41598-023-42044-x)

### <論文タイトル>

CCIVR2 facilitates comprehensive identification of both overlapping and non-overlapping antisense transcripts within specified regions

### <著者>

Maya Suzuki, Satoshi Sakai, Kosuke Ota, Yuki Bando, Chiharu Uchida, Hiroyuki Niida, Masatoshi Kitagawa & Tatsuya Ohhata

### <研究グループ>

本研究は、浜松医科大学分子生物学講座を中心に、同大学先進機器共用推進部および器官組織解剖学講座との共同研究として行われました。

### <研究支援>

本研究は、日本学術振興会科学研究費補助金「科研費番号JP20K06541」（研究代表者・大畑樹也）および浜松医科大学 研究支援事業の支援によって行われました。

### <本件に関するお問い合わせ先>

国立大学法人 浜松医科大学 分子生物学講座  
〒431-3192 浜松市東区半田山 1-20-1  
学内講師 大畑 樹也

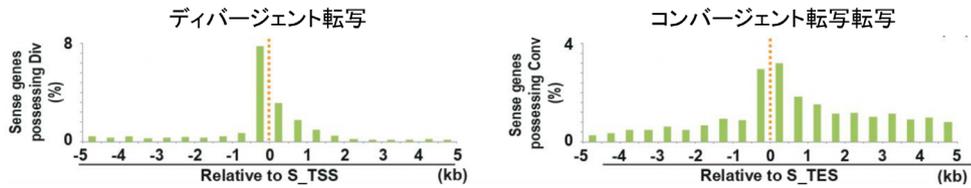
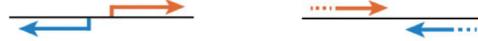
Tel/Fax: 053-435-2323

E-mail: ohhata@hama-med.ac.jp

<参考図>



遺伝子の位置情報を用いて、  
 ディバージェント転写を同定    コンバージェント転写を同定



ヒトのタンパク質コード遺伝子における、ディバージェント転写とコンバージェント転写の位置関係分布