



東北大学



東北メディカル・メガバンク機構
TOHOKU MEDICAL MEGABANK ORGANIZATION



2016年8月25日

報道機関各位

東北大学 東北メディカル・メガバンク機構
国立研究開発法人日本医療研究開発機構

日本人の基準ゲノム配列（JRG）を公開

～長鎖読みとり型次世代シーケンサーを用いた日本人のもつゲノム構造解析～

<公開のポイント>

- 長鎖読みとり型次世代シーケンサーを用いて、全ゲノム解読を完了。従来の技術では正確な解読が困難だった挿入配列の解読に成功。
- 国際ヒトゲノム参照配列に903個の挿入配列（約100万塩基配列）を新規追加した、日本人の基準ゲノム配列（JRGv1）を公開。
- 日本人の基準ゲノムを作成することで日本人ゲノム解析の精度向上と、結果として疾患関連遺伝的多様体の高精度検出・同定が期待できる。

【概要】

東北大学東北メディカル・メガバンク機構（以下、ToMMo）は、コホート調査^{*1}の参加者から提供されたDNAをもとに、長鎖読みとり型の次世代シーケンサー^{*2}PacBio RS II（Pacific Biosciences社製）を用いて、ヒトゲノム全長の100倍に相当する3,000億塩基のシーケンシングを行い、全ゲノム解読しました。本シーケンス解析解読の結果、国際ヒトゲノム参照配列^{*3}に対して、日本人が保有しこれまで報告されてこなかった約3,700箇所の新たな挿入配列、約250万塩基の同定に成功したことを4月23日付で発表^{*4}し、それらの配列を国際ヒトゲノム参照配列上にマップしない状態でデコイ配列^{*5}として7月15日から公開しています。今回、ToMMoでは、得られた挿入配列のうち日本人集団で存在を検証した挿入配列903個（約100万塩基配列）を国際ヒトゲノム参照配列上に配したものを、日本人の基準ゲノム配列JRG v1（Japanese Reference Genome version 1）として、公開することとなりました。

なお、デコイ配列 decoyJRG v1は日本人に高頻度でありながら国際ヒトゲノム参照配列には含まれていない配列をまとめたもので、ゲノム解読時の読み取り精度の向上に活用されます。

両配列の公開は、国際ヒトゲノム参照配列だけを用いている際には正確に読みとることのできなかった領域の研究に大きく寄与し、日本のゲノム研究全体を底上げ、加速させるものと期待されます。また今回の成果は日本人に特徴的なゲノム構造を明らかにする成果であり、今後、日本の医学研究の大きな基盤となる成果と考えられます。

【公開詳細】

公開URL：<http://jrg.megabank.tohoku.ac.jp/>

公開方式：fasta形式のファイルによるダウンロード

配列数：約30億塩基分（国際ヒトゲノム標準配列中に、挿入配列を配したもの）

条件：
・研究目的であればどなたでも自由にダウンロードできます。
・ダウンロードにあたっては、研究機関名・氏名等の登録をお願いします。
・登録にあたって審査等はありません。

【今後の計画】

日本人基準ゲノムJRGv1は、今後の研究の進展により、より精細な配列情報が得られた場合は、バージョンアップする予定です。

【用語解説】

- *1. コホート調査：ある特定の人々の集団を一定期間にわたって追跡し、生活習慣などの環境要因・遺伝的要因などと疾病の関係を解明するための調査のこと。
- *2. 次世代シーケンサー：主に2000年代半ば以降に登場した、DNA配列を高速で読み取る（シーケンス）機器の総称。主に、ランダムに切断されたDNA断片の塩基配列を1塩基ずつ決定する解析過程を、数百万以上ものDNA断片に対して同時並列的に処理することが可能。それまでの解析方法と比較して、精度の高いデータが大量かつ低価格、短時間で得られる。
- *3. 国際ヒトゲノム参照配列：国際的な学術組織The Genome Reference Consortiumが継続的に改訂を行っているヒトゲノムの全染色体の塩基配列。同配列は主に欧米の複数のヒトゲノムを読むことで構築されている。事実上、ヒトゲノムのデファクトスタンダードの塩基配列として全世界のヒトゲノム研究に利用されている。平成28年4月現在、最もよく使われている最新の国際ヒトゲノム参照配列はGRCh38である。
- *4. 4月23日付けで日本人基準ゲノム配列を解明したことを報告。
<プレスリリース>日本人の基準ゲノム配列（JRG）を決定～長鎖読みとり型次世代シーケンサーを用いて日本人のもつゲノム構造を解明～
<http://www.megabank.tohoku.ac.jp/news/15143>
- *5. デコイ配列：繰り返し配列などの、短鎖型次世代シーケンサーでの難読領域等を仮想配列として統合した配列。難読配列は、短鎖型次世代シーケンサーによって読みとりはされるが、その読みとり結果を、国際ヒトゲノム参照配列等に照らして並べ（マップシ）ようとすると、適合しなかったり、特定箇所に過度に集中するなどして、うまくマップすることができない。そうした領域を、（デコイ＝おとりのようにして）人為的に集める仮想配列をつくと、短鎖型次世代シーケンサーの結果の解析に対して有用である。

図1 日本人基準ゲノムのロゴ



【参考】

東北大学東北メディカル・メガバンク機構 <http://www.megabank.tohoku.ac.jp/>

東北メディカル・メガバンク計画は、平成 27 年度より、国立研究開発法人 日本医療研究開発機構（AMED）が本計画の研究支援担当機関の役割を果たしています。

本計画の事業の実施は、東北大学東北メディカル・メガバンク機構と岩手医科大学いわて東北メディカル・メガバンク機構とが連携して行っています。

【お問い合わせ先】

（研究に関すること）

東北大学東北メディカル・メガバンク機構

シーケンス解析室長

教授 安田 純（やすだ じゅん）

電話番号：022-272-3102

E メール：jyasuda@megabank.tohoku.ac.jp

東北大学東北メディカル・メガバンク機構

インシリコ解析室長

教授 長崎 正朗（ながさき まさお）

電話番号：022-273-6051

E メール：nagasaki@megabank.tohoku.ac.jp

（報道に関すること）

東北大学東北メディカル・メガバンク機構

広報戦略室長

長神 風二（ながみ ふうじ）

電話番号：022-717-7908

ファックス：022-717-7923

E メール：fnagami@med.tohoku.ac.jp

（AMED 事業に関すること）

日本医療研究開発機構（AMED）

バイオバンク事業部 基盤研究課

電話番号：03-6870-2228

E メール：kiban-kenkyu@amed.go.jp