

作成日 2026 年 3 月 27 日
(最終更新日 2026 年 6 月 6 日)

「情報公開文書」 (Web ページ掲載用)

受付番号 : 2026-1-181

課題名 : ゲノム言語モデルを用いた希少難病家系の原因遺伝子探索研究

1. 研究の対象

研究課題「NGS を用いた希少難病家系の網羅的ゲノム解析の追加研究」に参加された方

2. 研究期間

2026 年 6 月 (研究実施許可日) ~ 2032 年 3 月

3. 試料・情報の利用及び提供を開始する予定日

当機関で試料・情報の利用を開始する予定日及び外部への提供を開始する予定日は以下の通りです。

利用開始予定日 : 2026 年 7 月 1 日

提供開始予定日 : 2026 年 7 月 1 日

4. 研究目的

近年のゲノム解析技術 (全ゲノムシーケンス) の進歩により、私たちの体をつくる設計図である DNA 配列のすべてを読み取ることが可能になりました。しかし、病気の原因を特定しようとしても、解析で見つかる膨大な配列の変化 (バリエーション) の中には、現在の医学の知識だけでは、それが病気に直接関係しているのか、それとも単なる個人の特徴なのか判断できないものが数多く存在します。これらの病的な意義が不明なバリエーションは、診断を確定させる上での大きな壁となっています。

本研究の目的は、最新の AI 技術である「ゲノム言語モデル」を用いて、これまで解釈が困難だったバリエーションの意味を正しく読み解くことです。ゲノム言語モデルは、膨大なゲノム配列を言葉の並びとして学習した AI です。人間が気づけないような複雑な配列の法則性を理解することで、見つかった変化が生命の設計図として「意味が通じなくなるような重大な変化」なのか、それとも「影響のない変化」なのかを高精度に予測します。この技術により、これまで解決できなかった希少難病の新たな原因遺伝子を発見できることが期待されます。

5. 研究方法

この研究は、以下の流れで進められます。

1) 試料および情報の収集

本研究では、先行研究で取得済みの既存試料・データに加え、研究対象者から新たに採取する生体試料 (血液、唾液、尿沈渣、組織等) およびその加工物 (抽出 DNA/RNA 等) を用います。次世代シーケンシング (NGS) により新規の配列情報を創出し、年齢、性別、病歴、家族歴等の臨床情報と統合して解析に使用します。

2) ゲノム言語モデルの学習 (AI にゲノムの文法を教える)

まず、公開されている膨大なヒトゲノムデータ (特定の個人を識別できない情報) を AI に読み込ませ、ゲノム配列の規則性を学ばせます。これは、AI が膨大な文章を読んで言

葉の並びや文脈を理解するプロセスに似ています。この「学習」によって、AI は DNA 配列における正常な文法や、重要な領域がどこであるかを判断する能力を身につけます。

3) ゲノム言語モデルによる推論（患者さんのバリエーションを解析する）

次に、学習を終えた AI を用いて、患者さんのゲノムデータに見つかっている病原性が不明な配列変化を一つずつ評価します。AI は、学習した知識をもとにこの配列の変化は、設計図の意味を大きく変えてしまうものを高度な計算によって予測します。これらと患者さんの症状やご家族の情報を照らし合わせることで、これまで見逃されていた病気の真の原因である可能性が高い変化を、新たな有力候補として特定します。

6. 研究に用いる試料・情報の種類

試料：先行研究ですでにいただいた検体（血液、唾液、その他組織[頬粘膜、臍帯血、毛根、爪、歯牙、尿、皮膚、骨髓血、固形腫瘍政権、手術摘除残余標本の一部]、あるいはそれらから抽出された DNA/RNA、樹立された細胞株を含む）

情報：ゲノム情報（全エクソームシーケンスデータ、全ゲノムシーケンスデータ、候補バリエーション、ヒト多型データ）、年齢、性別、病歴、家族歴

7. 外部への試料・情報の提供

情報は個人が特定できないよう氏名等を削除し、記録媒体、郵送、電子的配信等により共同研究機関へ提供します。

対応表は、当院の研究責任者が保管・管理します。

本研究では、理化学研究所計算科学研究センターの富岳クラウドプラットフォームを利用し、あなたのデータを保存し解析します。この理化学研究所のクラウドサーバの設置場所については以下のページをご覧ください。<https://www.r-ccs.riken.jp/fugaku/fcp/>

8. 研究組織

研究代表機関

東北大学大学院医学系研究科 田宮 元

共同研究機関

理化学研究所革新知能統合研究センター 成田 暁

成育医療研究センター 要 匡

東京慈恵会医科大学 大石 公彦

9. 利益相反（企業等との利害関係）について

本学では、研究責任者のグループが公正性を保つことを目的に、情報公開文書において企業等との利害関係の開示を行っています。

使用する研究費は大学運営費交付金です。

外部との経済的な利益関係等によって、研究で必要とされる公正かつ適正な判断が損なわれる、または損なわれるのではないかと第三者から懸念が表明されかねない事態を「利益相反」と言います。

本研究は、研究責任者のグループにより公正に行われます。本研究の利害関係については、現在のところありません。今後生じた場合には、所属機関において利益相反の管理を受けたうえで研究を継続し、本研究の企業等との利害関係について公正性を保ちます。

この研究の結果により特許権等が生じた場合は、その帰属先は研究機関及び研究者等になります。あなたには帰属しません。

10. お問い合わせ先

本研究に関するご質問等がありましたら下記の連絡先までお問い合わせください。

ご希望があれば、他の研究対象者の個人情報及び知的財産の保護に支障がない範囲内で、研究計画書及び関連資料を閲覧することが出来ますのでお申出ください。

また、試料・情報が当該研究に用いられることについて研究対象者もしくは研究対象者の代理人の方にご了承いただけない場合には研究対象としますので、下記の連絡先までお申出ください。その場合でも研究対象者に不利益が生じることはありません。

当機関における照会先および研究への利用を拒否する場合の連絡先：

担当者の所属・氏名：東北大学大学院医学系研究科 高山 順
住所：宮城県仙台市青葉区星陵町 2-1 宮城県仙台市青葉区星陵町 2-1
連絡先：022-728-3071

当機関の研究責任者：東北大学大学院医学系研究科 田宮 元

研究代表者：東北大学大学院医学系研究科 田宮 元

◆個人情報の開示等に関する手続

本学が保有する個人情報のうち、本人の情報について、開示、訂正及び利用停止を請求することができます。

保有個人情報とは、本学の役員又は職員が職務上作成し、又は取得した個人情報です。

- 1) 診療情報に関する保有個人情報については、東北大学病院事務部医事課が相談窓口となります。詳しくは、下記ホームページ「配布物 患者さまの個人情報に関するお知らせ」をご覧ください。（※手数料が必要です。）

【東北大学病院個人情報保護方針】

<http://www.hosp.tohoku.ac.jp/privacy.html>

- 2) 1)以外の保有する個人情報については、所定の請求用紙に必要事項を記入し情報公開室受付窓口へ提出するか又は郵送願います。詳しくは請求手続きのホームページをご覧ください。（※手数料が必要です。）

【東北大学情報公開室】<http://www.bureau.tohoku.ac.jp/kokai/disclosure/index.html>

※注意事項

以下に該当する場合には全部若しくは一部についてお応えできないことがあります。

- ①研究対象者等又は第三者の生命、身体、財産その他の権利利益を害するおそれがある場合
- ②研究機関の研究業務の適正な実施に著しい支障を及ぼすおそれがある場合
- ③法令に違反することとなる場合