

戦略的創造研究推進事業 CREST
研究領域「異分野融合による新型コロナウイルスを
はじめとした感染症との共生に資する技術基盤の
創生」
研究課題「先端ゲノム解析と人工知能によるコロナ
制圧研究」

研究終了報告書

研究期間 2021年02月～2024年03月

研究代表者：井元 清哉
(東京大学 医科学研究所 教授)

§ 1 研究実施の概要

(1)実施概要

- 当初目標とした 5,000 名の日本における COVID-19 患者の情報収集について、目標を大きく超える 6,400 名以上の検体 (DNA、RNA、血漿) を収集し、約 5,500 症例の全ゲノムシーケンス、約 3,000 症例の RNAseq、約 1,400 症例のプロテオーム解析、約 500 症例のメタボローム解析を実施しマルチオミクスのデータベースを構築。
- COVID-19 患者群 GWAS 個人別ジェノタイプデータ 2,400 名、コントロール群 GWAS 個人別ジェノタイプデータ 2,000 名、COVID-19 GWAS 解析サマリー統計量、COVID-19 患者群 RNA-seq データ 460 名、COVID-19 RNA-seq eQTL 統計量、COVID-19 患者群シングルセルシーケンス元データ 75 名、コントロール群シングルセルシーケンス元データ 75 名、シングルセルシーケンスに対応する GWAS データ 150 名を NBDC、EGA データベースにて公開した。
- 重症例を中心とした約 2,400 症例を用いたゲノムワイド関連解析を実施し、DOCK2 遺伝子の領域の遺伝子多型が 65 歳以下の非高齢者における重症化リスクと関連性を示すことを発見した (Nature, 2022)。
- 国際コンソーシアム COVID-19 Host Genetics Initiative にアジア最大の cohorts として参画し、重症化に関連する 15 箇所のバリエーションの発見に貢献した (Nature, 2021, 2023)。
- COVID-19 の重症度とゲノム変異による遺伝子発現制御の相互関係の解析を実施し、CLEC4C 等、重症度に応じた遺伝子発現の変化パターンが特定のヒトゲノム変異の有無で異なる遺伝子を同定した (Nature Communications, 2022)。
- 収集したデータを検体収集に協力頂いた医療従事者に共有し、データのサマリーなどを簡易に取得、解析できるシステムを構築することで共同研究を協力を推進し、COVID-19 の臨床像の解明に関わる 20 報以上の研究論文となって出版されている。
- 2021 年に開催された第 32 回オリンピック競技大会 (2020/東京) 及び東京 2020 パラリンピック競技大会 (以下、東京 2020 大会) の選手村において新型コロナウイルス感染症 (COVID-19) の下水疫学調査を実施し、下水中の新型コロナウイルス RNA 量と陽性確定者数との相関を解析した結果を報告した (Journal of Travel Medicine, 2022)。
- 東京 2020 大会選手村下水検体中の新型コロナウイルス RNA 量は陽性者の存在と統計的に有意な正の相関が認められ、さらに下水中ウイルス RNA 量の増加は個人検査による陽性者の発見に 2 日間先行していたことが示唆された (JAMA Network Open, 2022)。
- ウイルス血症に関して、末梢血細胞に細胞内ウイルス RNA が検出された COVID-19 患者の臨床的特徴を世界で初めて明らかにした (International Journal of Infectious Diseases, 2023)。
- 初診時に得られる臨床情報の任意の組み合わせ、および演算によって新たな変数を自動生成し、重症化症例をバリデーション cohoot (トレーニングを行ったデータとは別時期に取得されたデータ) において AUC 0.906 を達成する高精度予測が可能な説明可能 AI を構築した。

(2)顕著な成果

<優れた基礎研究としての成果>

1.

概要: 120 以上の医療機関の協力の下、日本における 6,400 名以上の COVID-19 患者の検体 (DNA、RNA、血漿) を収集し、それらから得られた全ゲノムシーケンスを含むマルチオミクスデータと精緻な臨床情報を収集し、東アジアにおいては随一、世界的にも一級のデータベースを構築し、その基盤を用いて COVID-19 の臨床像の解明に大きく寄与する多くの研究が実施された。

2.

概要: 収集したコロナ患者のデータから、特に、若年重症者を対象として、日本人集団に特異

的な重症化遺伝子変異と重症化遺伝子 DOCK2 を同定した。RNAseq およびシングルセルシーケンス解析による遺伝子発現解析によって、その発現低下が重症化に寄与し、重症化のバイオマーカーとなる可能性を示した。更に、モデル動物感染実験系を用いて、DOCK2 遺伝子の機能障害が新型コロナウイルス感染症重症化に関わる生物学的メカニズムを明らかにした。

3.

概要: COVID-19 の重症度とゲノム変異による遺伝子発現制御の相互関係の解析を実施し、CLEC4C 等、重症度に応じた遺伝子発現の変化パターンが特定のヒトゲノム変異の有無で異なる遺伝子を同定した。

< 科学技術イノベーションに大きく寄与する成果 >

1.

概要: 発見した DOCK2 の遺伝子バリエーションについて、DOCK2 は、リンパ球の活性化、リンパ球の遊走、I 型インターフェロンの産生に重要な役割を持つことが知られているが、RNA-seq 解析、single cell RNA-seq 解析、一細胞解析、病理解析、細胞実験、動物実験による詳細な解析から、DOCK2 が COVID-19 の重症化のマーカーとなるだけでなく、COVID-19 の治療標的となることを見出した。

2.

概要: シングルセルを含むマルチオミクスデータ解析によって、COVID-19 の重症度に応じた遺伝子発現の変化パターンが特定ヒトゲノム変異の有無で異なる遺伝子で生じていることを明らかにし、また、その仕組みが、いくつかの細胞種に特異的な遺伝子発現により説明できることを突き止めた。これらの結果は、感染する可能性を、特定遺伝子の細胞種に特異的な制御の仕組みの違いとして理解するための基盤的情報となり得る成果である。

3.

概要: 第 32 回オリンピック競技大会(2020/東京)及び東京 2020 パラリンピック競技大会の選手村において新型コロナウイルス感染症の下水疫学調査を実施し、下水中の新型コロナウイルス RNA 量と陽性確定者数との相関を解析した結果を報告した。下水検体中の新型コロナウイルス RNA 量は陽性者の存在と統計的に有意な正の相関が認められ、さらに下水中ウイルス RNA 量の増加は個人検査による陽性者の発見に 2 日間先行していたことが示唆された。

< 代表的な論文 >

1.

概要: Namkoong H, Edahiro R, et al. DOCK2 is involved in the host genetics and biology of severe COVID-19. *Nature*, 2022 Sep;609(7928):754-760. 免疫機能での重要な役割が知られる DOCK2 と呼ばれる遺伝子の領域の遺伝子多型が、65 歳以下の非高齢者における重症化リスクと関連性を示すことを発見した。また、RNA-seq 解析、single cell RNA-seq 解析一細胞解析、病理解析、細胞実験、動物実験による詳細な解析から、DOCK2 が COVID-19 の重症化のマーカーとなるだけでなく、COVID-19 の治療標的となることを見出し、今後の新しい治療戦略につながる成果を得た。

2.

概要: Wang QS, et al. The whole blood transcriptional regulation landscape in 465 COVID-19 infected samples from Japan COVID-19 Task Force. *Nature Communications*, 2022 Aug 22;13(1):4830. COVID-19 の重症度とゲノム変異による遺伝子発現制御の相互関係の解析 (ieQTL 解析) により、重症度に応じた遺伝子発現の変化パターンが特定のヒトゲノム変異の有無で異なる遺伝子が存在することを明らかにした。さらに、その仕組みが、いくつかの細胞種

に特異的な遺伝子発現により説明でき、感染する可能性を、特定遺伝子の細胞種に特異的な制御の仕組みの違いとして理解するためのさらなる考察につながる成果を得た。

3.

概要: Kitajima M, et al. Association of SARS-CoV-2 load in wastewater with reported COVID-19 cases in the Tokyo 2020 Olympic and Paralympic Village from July to September 2021. JAMA Network Open. 2022 Aug 1;5(8):e2226822. 東京 2020 大会開催期間を含む 2021 年 7 月 14 日から 9 月 8 日にかけて、選手村にて下水疫学調査を実施した。下水検体中の新型コロナウイルス RNA 量は陽性者の存在と統計的に有意な正の相関が認められ、さらに下水中ウイルス RNA 量の増加は個人検査による陽性者の発見に 2 日間先行していた。この下水疫学調査結果は、他の感染モニタリング指標と併せて、パラリンピック期間中に更なる感染防止対策の判断材料の一部として活用された。

§2 研究実施体制

(1) 研究チームの体制について

① 井元グループ

研究代表者: 井元 清哉 (東京大学医科学研究所 教授)

研究項目

- ・重症化を予測する説明可能な AI の構築
- ・ウイルスゲノム解析ツール開発
- ・シーケンスデータマネジメント

② 福永グループ

主たる共同研究者: 福永 興彦 (慶應義塾大学医学部 教授)

研究項目

- ・臨床情報の収集／臨床情報データベースの構築
- ・DNA、RNA、血漿等検体の収集
- ・ウイルス血症解析の実施

③ 宮野グループ

主たる共同研究者: 宮野 悟 (東京医科歯科大学・M&D データ科学センター 特任教授)

研究項目

- ・全ゲノム、RNA、ウイルスゲノムの解析による重症化因子の探索
- ・臨床情報の収集／DNA、RNA、血漿等検体の収集
- ・ホストゲノム(全ゲノム・RNA)ウイルスゲノムシーケンス
- ・シーケンスデータマネジメント

④ 小川グループ

主たる共同研究者: 小川 誠司 (京都大学大学院医学研究科 教授)

研究項目

- ・COVID-19 の重症度に関するクローン性造血の関連解析
- ・プロテオーム解析
- ・ホストゲノム(全ゲノム・RNA)ウイルスゲノムシーケンス
- ・シーケンスデータマネジメント

⑤ 岡田グループ

主たる共同研究者: 岡田 随象 (大阪大学大学院医学系研究科 教授)

研究項目

- ホストゲノム関連解析の実施
- シングルセル RNA 解析
- ホストゲノム SNP 解析

(2)国内外の研究者や産業界等との連携によるネットワーク形成の状況について

- コロナ感染者の臨床情報、検体(DNA、RNA、血漿)の収集に協力頂いた 120 の医療機関とのネットワークは極めて重要である。
- 国際コンソーシアム COVID-19 Host Genetics Initiative と連携し、GWAS データの提供(東アジアで first data fix までに重症例を含むデータ提供したのは我々のみ)、WGS からの genotype の提供を実施し、重症化などの臨床像と宿主ゲノムの関連の解明に貢献している。
- 片山チームの片山浩之教授や北海道大学の北島正章准教授と下水疫学調査について連携し、下水疫学調査の社会実装を進めている。
- 有田チームと連携し、ウイルスゲノム解析のパイプラインやデータベース構築、可視化ツールについて意見交換を行っている。