

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

2020 年度採択研究代表者

2021 年度 年次報告書

井元 清哉

東京大学 医科学研究所
教授

先端ゲノム解析と人工知能によるコロナ制圧研究

§ 1. 研究成果の概要

全国の病院、大学機関からなるコロナ制圧タスクフォースが中心となり COVID-19 臨床データベースを構築してきた。2022 年 3 月時点で COVID-19 患者 5,600 例以上から臨床情報、血液検体を採取し、4,700 例でゲノムワイド関連解析 (GWAS)、1,200 例以上で全ゲノムシーケンス、RNAseq、およびプロテオミクス解析、500 例以上でメタボローム解析を施行し臨床情報とマルチオミクスデータを蓄積している。現在、臨床情報のクリーニングを継続しながら、遺伝情報、血中たんぱく情報などと統合したデータベースの整備・構築を進めている。また、これらのデータベースを用いて、国際コンソーシアム COVID-19 Host Genetic Initiative に参画し、国際連携のもと重症化に関連する遺伝子多型解析に貢献した (Nature, 2021)。さらに、若年重症者を対象とした解析を通じて、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。収集したデータの基本的な統計や簡単なクロス集計について、データ提供者が簡単に解析できるツールを備えたエクセルファイルを配布した。この効果もあり、慶應義塾大学の若手研究者を中心に次々に臨床研究が立ち上がり、データベースを活用した研究が進められている。

ウイルスゲノムについては、GISAID に登録されている SARS-CoV-2 ゲノムシーケンスデータを SHIROKANE 上にダウンロードし、その情報を可視化するアプリケーション「HGC SARS-CoV-2 Variant Browser」を開発し、システムの運用を開始した。現在 500 万配列を超える情報を格納している。このシステムを用いることで、新型コロナウイルスのゲノム情報から「どういった変異を持っており」「いつどの国から流入し」「どのように感染拡大してきたのか」を迅速に把握することができるようになった。

約 3,300 名のクリーニング済み臨床情報を用い、複数の項目を組み合わせ人が解釈可能な変換量を新たな項目として人工知能によって自動的に生成し、重症例と非重症例を判別する機械学習技術を構築した。

§ 2. 研究実施体制

(1) 井元グループ

- ① 研究代表者:井元 清哉 (東京大学 医科学研究所 教授)
- ② 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ージメント

(2) 福永グループ

- ① 主たる共同研究者:福永 興壱 (慶應義塾大学 医学部 教授)
- ② 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

- ① 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)
- ② 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ージメント

(4) 小川グループ

- ① 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)
- ② 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ージメント

(5) 岡田グループ

- ① 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)
- ② 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

【代表的な原著論文情報】

- 1) “Japan COVID-19 Task Force: a nation-wide consortium to elucidate host genetics of COVID-19 pandemic in Japan”, medrxiv, doi: <https://doi.org/10.1101/2021.05.17.21256513>
- 2) “Mapping the human genetic architecture of COVID-19”. *Nature*, 2021 Dec;600(7889):472-477. doi: [10.1038/s41586-021-03767-x](https://doi.org/10.1038/s41586-021-03767-x)
- 3) “Direct derivation of human alveolospheres for SARS-CoV-2 infection modeling and drug screening”. *Cell Rep.* 2021 Jun 8;35(10):109218. doi: [10.1016/j.celrep.2021.109218](https://doi.org/10.1016/j.celrep.2021.109218)
- 4) “Clinical characteristics of patients with coronavirus disease (COVID-19): preliminary baseline report of Japan COVID-19 Task Force, a nation-wide consortium to investigate host genetics of COVID-19”. *International Journal of Infectious Diseases*, 2021 Sep 30;S1201-9712(21)00775-X. doi: [10.1016/j.ijid.2021.09.070](https://doi.org/10.1016/j.ijid.2021.09.070)
- 5) “Viral loads and profile of the patients infected with SARS-CoV-2 Delta, Alpha or R.1 variants in Tokyo”. *J Med Virol* 2022 Apr 94(4):1707-1710. doi: [10.1002/jmv.27479](https://doi.org/10.1002/jmv.27479)