

2021 年度卒業論文

新型コロナウイルスの RNA 解析

龍谷大学 理工学部 数理情報学科

T170058 中川 祐輔

指導教員 佐野 彰

概要

現在新型コロナウイルスが世界中を脅かしている。最初に中国の武漢で確認された新型コロナウイルスであるが、その当初のウイルスと現在のウイルスでは主にスパイク部分がかなり違ってきている。既に新型コロナウイルスにおいての変異株系統は α 株 β 株 $\dots$ $\omicron$ 株と 12 種ほどあるが、今現在の感染株は $\omicron$ 株が感染地域と感染者を急激に増やしており、メインの株となりつつある。

私は現在日本で流行している新型コロナウイルスは、初めて武漢で発見された新型コロナウイルスからどの程度変異しているのかという点に興味を抱いた。

本研究では、R 言語を使用し解析を行った。新型コロナウイルスのデータは NCBI (アメリカ国立生物工学センター) 取得した。本研究では、中国の武漢で最初に流行りだした新型コロナウイルスのデータと日本の第 3 波 (2021 年 2 月) 第 4 波 (2021 年 6 月) と第 5 波 (2021 年 9 月) の時期のデータも採取した。これらを R 言語を用いて、アラインメントしたのち、系統樹を作ること株間の遺伝的距離を比較した。第 4 波では主にアルファ株が第 5 波では主にデルタ株が流行っていた。元の原点とされる武漢のデータからの共通祖先などを見てアルファ株は原点からどれほど距離があるのか、デルタ株はアルファ株に近いものなのかを見ていく。

さらに新型コロナウイルスは、その性質上ヒト細胞の受容体結合部分であるスパイク部分に大きな変異が起こることが分かっている。そこでスパイク部分のデータから出した系統樹には全体のデータとどれほどの差が出るのかなどを系統樹から見ていく。

その結果、原点と比べると第 3 波、第 4 波、第 5 波、現在の日本 (11 月から 12 月) の 3 つのデータには距離があることが見られ、現在の日本 (11 月から 12 月) のデータは第 5 波のデルタ株との共通祖先が近いことからデルタ株の特徴である感染力と重篤性の高さを持っているものが現在の日本では流行していることが考えられる。第 3 波は他のデータと比べて距離が遠い結果であった。第 3 波がその後に変異したとは考えづらく、それ以前のものか海外からウイルスが持ち込まれ感染が広がったと考えられる。

今後の課題として、原点と日本のデータで比較したが、その比較対象は数が多いほど変化は見られやすい。時間の都合で可能にはならなかったが、複数の PC を使うなどをしてより多くのデータから解析できることが挙げられる。