

## Śledzenie pojawiania się i rozprzestrzeniania COVID-19 przy użyciu bazy danych sekwencji

Podsumowanie prezentacji Sama Lycetta na warsztatach SARS-CoV-2 / COVID-19.

Analizy filogenetyczne sekwencji SARS-CoV-2 pokazują, że wirus jest bardzo podobny do SARS, ale także do wielu innych beta-koronawirusów wcześniej zidentyfikowanych u nietoperzy. Jesteśmy w stanie wywnioskować te relacje dzięki wirusowym danym sekwencjonowania, które są udostępniane globalnie w ogólnodostępnej bazie danych GISAID (<https://www.gisaid.org>). W styczniu i lutym większość zdeponowanych sekwencji była izolowana w Azji, ale w ostatnich tygodniach obserwujemy rosnącą liczbę z Europy i Ameryki Północnej.

Sekwencje SARS-CoV-2 są do siebie bardzo podobne, ale widzimy kilka mutacji między izolatami. Chociaż nie oczekujemy, że mutacje te będą miały wpływ na zjadliwość, możemy je wykorzystać do śledzenia rozprzestrzeniania się epidemii.

Budując skalowane czasowo drzewo filogenetyczne ze wszystkich dostępnych sekwencji, możemy oszacować pochodzenie wirusa na listopad 2019 r. To samo drzewo wskazuje, że wirus został wprowadzony do Wielkiej Brytanii wiele razy, a klastry rozprzestrzeniają się z każdego wprowadzenia. Możemy również zobaczyć wyraźne skupienia w sekwencjach pochodzących z różnych kontynentów; na przykład sekwencje z Europy są o wiele bardziej podobne do siebie niż sekwencje z Oceanii.

Dane te można nałożyć na mapę świata, która pokazuje, że w styczniu i lutym epicentrum rozprzestrzeniania się wirusa była Azja, w kierunku Europy, Australii i Ameryki Północnej. Jednak w miarę upływu czasu epicentrum przeniosło się do Europy, rozprzestrzeniając się w Ameryce Południowej i innych krajach.

Gdy zbieramy więcej danych o sekwencji, możemy być w stanie wyciągnąć więcej szczegółów na temat rozprzestrzeniania się wirusa, takich jak oszacowanie różnic w  $R_0$  między różnymi krajami i kontynentami. To z kolei może pomóc pokazać, które strategie interwencji były najbardziej skuteczne.

Data publikacji 13 kwietnia 2020.