AF project: Genomic epidemiology for SARS-CoV-2: sooner and later

Caroline Colijn, Christophe Fraser, Marc Lipsitch

We propose an AF project to establish the immediate-term genomic epidemiology for SARS-CoV-2, including its use in real time, and the longer-term integrated modelling and genomic surveillance to detect increased transmission, monitor for emerging potential vaccine escape strains and to monitor for impacts of selection.

In the immediate future we would establish robust and flexible clustering methods to identify sets of cases that are plausibly linked by transmission; cluster-specific epidemiologic parameters and high-resolution (in time and across clusters) transmission dynamics. These include age- and/or location-specific contributions to the force of infection through time. Clustering is challenging because both rapid bursts of transmission and longer-term partially-observed transmission chains contribute to epidemiologically-linked transmission, and because SC2 accrues limited genetic diversity on the time scale of an outbreak. The combination of genomic and epidemiologically-defined clusters can help to determine likely routes of transmission and to identify settings with a high risk of onward transmission, particularly if this can be done in collaboration with public health teams. We will establish a warning system to detect emerging clusters with a higher-than-expected rate of growth, and will develop the appropriate baselines for comparison.

We will establish methods to estimate population dynamics and epidemiological parameters in resolved clusters, using genomes and Ct values, age and demographic information where available, together with state of the art inference methods. Using and adapting insights from the PANGEA2 consortium we will elucidate genomic signatures of transmission sources and sinks over time, answering questions such as “what was the force of infection from university students in the autumn of 2020”. We will extend these to real-time analyses and estimates as the data flows become established. Together, these high-resolution cluster-specific inferences will allow us to model cluster-specific transmission dynamics and to determine which interventions might have the most impact on curbing transmission at which place and time. We anticipate that this will have direct application as vaccines become available, as age-specific forces of infection over time (together with knowledge of the timing of control measures) will help to inform models to simulate vaccination programmes.

In the second phase of the project we will establish real-time genomic surveillance capable of monitoring for sharp increases in transmission, building on the estimation and detection methodology in the first phase. In particular we would seek to detect any potential vaccine escape strains as early as possible, using a combination of the cluster-specific dynamics and knowledge of plausible determinants of vaccine escape, in the context of strong real-time sequencing. We will also establish real-time monitoring of impacts of other forms of selection acting on the virus. The aspect of the project should link to work in international genomic surveillance of zoonotic viruses and their evolution, including evolution in animal hosts.