Developing a global genomics partnership to inform pneumococcal vaccine strategy

Christine Boinett[1], Kate Mellor[1], Lesley McGee[2], Narender Kumar[1], Stephanie Lo[1], Stephen Bentley[1]

[1]Parasites and Microbes, The Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridge CB10 1SA, UK

[2]Respiratory Diseases Branch, Centers for Disease Control and Prevention, Atlanta, GA 30333, USA

Hundreds of thousands of infant deaths are attributed to *Streptococcus pneumoniae* each year. Vaccines targeting 7 (PCV7), 10 (PCV10) and 13 (PCV13) of the most prevalent serotypes were introduced in 2000, 2009 and 2010, respectively, with great success. However, in many countries' infant mortality due to pneumococcal diseases remains high, partly due to serotype replacement driving vaccine evasion in the circulating pathogen population. Starting in 2011, the Global Pneumococcal Project (GPS) set out to record the global diversity and evolution of *S. pneumoniae* using whole-genome sequencing. In collaboration with partners from >50 countries, we sequenced the genomes of >26,000 pneumococcal isolates around the world and GPS was the largest study of its kind at the time. This enabled the development of a new standard for designating *S. pneumoniae* strains based on total genome data, that can be consistently applied across national and international genomic surveillance studies. These global population genomic data reveal extensive differences in population structure between countries and such variations can partly be explained by selective pressure due to vaccines and local antibiotic usage. Using mathematical models, genome data can even be used to forecast the likely outcome of future vaccine strategies.

In 2020, we began a new phase of the GPS project. As well as continuing genomic sampling in established partner countries, we aim to address the lack of data for countries with a very high burden of pneumococcal disease such as India, Nigeria, DRC, and Pakistan. Furthermore, we are shifting to a decentralised model of working where the majority of genome sequencing, analysis and interpretation is conducted locally, with appropriate support from the core GPS team. Accordingly, we will be supporting the development of local sequencing capacity and delivering a portfolio of bioinformatics training and portable software for GPS partner labs, thus building a robust pneumococcal genomics network that will support local decision-making and global vaccine development strategies. Through outreach activities we are raising awareness of the benefits of genomic pathogen surveillance and engaging with policy makers to advocate for integration in national infrastructures. Here, we will present an overview of the GPS network, discussing developments over time, lessons learned and potential future paths.