**DIVERGE CONSORTIUM WORKSHOP REPORT**

Members of the DIVERGE consortium attended a meeting held on Thursday 28th and Friday 29th June 2018 in Edinburgh, UK.

**Attendees:** Marcela Echavarria (Cemic, Buenos Aires, Argentina), Erik Karlsson (Institut Pasteur, Phnom Penh, Cambodia), Gene Tan (JCVI, La Jolla, US), Sandra Jackson (WHO, Geneva, Switzerland), Harry Campbell (University of Edinburgh, UK), Harish Nair (UoE), Andrew Rambaut (UoE), Kate Templeton (UoE), Rachel Reeves (UoE), Elizabeth Wastnedge (UoE), Thomas Williams (UoE).

**Attendees via TC link:** Abdullah Brooks (icddr,b, Dhaka, Bangladesh), Quique Bassat (ISGlobal, Barcelona), Aubree Gordon (Centro de Salud Socrates Flores Vivas/ University of Michigan, Nicaragua), Martha Nelson (Fogarty International Center, Maryland, US), Cecile Viboud (Fogarty), David Spiro (Fogarty).

Harry Campbell (HC) welcomed the participants to the meeting and outlined how the University of Edinburgh has worked in partnership for many years with collaborators in low and middle income countries (LMICs) examining the burden of disease of childhood pneumonia, with a particular focus on Respiratory Syncytial Virus (RSV). This work has become of particular significance in recent years, due to the likely introduction of a vaccine in the near future.

In the course of this work, he and others have become concerned that RSV vaccine development has been largely premised on the lack of RSV F protein variability and that there was a need for a more comprehensive picture of global RSV genome variability. However, available RSV sequences come from very few countries, with majority from the USA. With the UoE focus on global RSV disease burden and prevention, he had felt it important to better understand RSV sequence variability across the entire RSV genome and from a much larger number of global sites, which had led to the establishment of the DIVERGE consortium. It was recognised that these data could also underpin studies of RSV transmission dynamics and phylogenetics.

The purpose of this meeting was to discuss progress within the consortium and discuss and set out the scientific questions to be addressed, and potential future directions.

**The DIVERGE Consortium**

The consortium was set up to represent sites from 5 out of the 6 WHO regions, with a focus on Low and Middle Income Countries (LMICs), where the majority of the morbidity and mortality associated with RSV is concentrated. Aiming for samples over a five year time period, in a paediatric population, it has the following aims:

* Aim 1 (Proof of concept): Demonstrate the feasibility of generating high quality RSV whole genome sequences from multiple sites in Low and Middle Income Countries.
* Aim 2: To describe the population-level diversity of respiratory syncytial virus A or B genomes in paediatric patients at sites with little existing genomic data.
* Aim 3: To relate genetic variability in RSV to clinical presentation in paediatric patients
* Aim 4: To describe the intra-host viral genetic variability of respiratory syncytial A and B virus genomes of pediatric patients.

Longer term, the aims of the consortium moving forwards are to:

* To demonstrate in a high granularity, comprehensive dataset geographical sequence variability in RSV genomes at a global level.
* To establish a baseline picture of RSV genomic variability to map changes subsequent to upcoming immunization programs.
* To relate viral genotype to age and clinical features and determine whether different RSV clades show variable virulence or differences in clinical presentation.

In this pilot phase sites will contribute:

* Nicaragua: 229 samples, 2012-2016
* Banglandesh: 300 samples, 2012-2015
* Cambodia: 41 samples, 2014-2016
* Morocco: 18 samples, 2010
* Mozambique: 42 samples, 2010-2014
* Argentina: 108 samples from 2011-2013 and a proposed further 60 from 2014-2016

**Background talks**

Harish Nair gave a presentation on the global burden on RSV infection, Harry Campbell discussed the current status of vaccines in development for RSV, and Thomas Williams outlined the antigenic consequences of genetic variability in the F protein.

**Talks from LMIC partners**

Abdullah Brooks outlined his work in Kamalapur, Dhaka, Bangladesh, where his team carry out active surveillance for respiratory and other diseases in children. Quique Bassat presented two studies which had collected respiratory samples, one from Morocco looking at the aetiology of severe childhood pneumonia and another in Mozambique looking to identify biomarkers for respiratory infections in children. Erik Karlsson outlined on Severe Acute Respiratory Infection (SARI) monitoring in Cambodia. Marcela Echavarria presented on a number of hospital studies looking at children with respiratory infections. Finally, Aubree Gordon discussed two cohort studies: the Nicaraguan Birth Cohort Study, and the Influenza and Respiratory Syncytial Virus in Infants Study.

**Analysis Plans**

Elizabeth Wastnedge outlined potential hypotheses that could be explored with adequate clinical metadata, including looking at the relationship between genotype and clinical phenotypes suchas disease severity. Given the global nature of the samples, one could examine whether patterns are associated with different centres, or idiosyncratic to particular regions. DIVERGE aims to use a modification of the WHO criteria to grade severity and hope also to identify potential confounders for disease severity.

Gene Tan from JCVI then explained the sequencing pipeline for DIVERGE, and the practicalities of successfully obtaining complete genome sequences from the sequences submitted. He also reviewed progress made with samples from DIVERGE which has been received by JCVI to date. Following this Martha Nelson outlined her plan to break down the analysis of the sequencing data into two different kinds of studies:

1. Examining the local transmission dynamics of RSV in each discrete spatial location, incorporating detailed metadata. This will allow identification of genetic lineages that compose an epidemic, and ask how genetically diverse they are, how they disseminate within a community and whether viruses persist between epidemics.
2. Better understanding the global ecology of RSV- by pooling sequences from all locations an analysis will look at the relationship of viral populations from different/countries/hemispheres, and the location of source populations.

Cecile Viboud followed on from this with a discussion of how one might use genomic data to model RSV transmission dynamics in the context of upcoming immunisation campaigns. Thomas Williams concluded this section by leading discussion on how DIVERGE can help lead on standardisation of RSV nomenclature and archiving of clinical metadata in association with the sequences submitted to GenBank.

**Taking the DIVERGE consortium forwards**

After talks from Kate Templeton on RSV infection in adults, Andrew Rambaut on Project Artic (which aims to bring real time sequencing to epidemic outbreaks) and Sandra Jackson on the WHO RSV surveillance pilot (which is collecting RSV samples from children and adults in 14 countries in all 6 WHO regions) participants focused on ways in which to take the consortium forwards in order to meet its long term goals.

Participants agreed that in order to conduct a prospective study with sample sizes sufficient to answer well designed and important scientific questions external funding would be needed beyond that required for sequencing alone. All the participants reiterated the importance of building capacity and enhancing expertise at partner sites in LMICs, and their commitment to working in an equal partnership between high income and low / middle income country partners to address scientific questions of mutual interest.