

Content archived on 2024-06-18



Bridging the Evolution and Epidemiology pean Research Council of HIV in Europe

Reporting

Project Information

BEEHIVE

Grant agreement ID: 339251

Project website 🗹

Project closed

Start date 1 April 2014

End date 31 March 2019

Funded under

Specific programme: "Ideas" implementing the Seventh Framework Programme of the European Community for research, technological development and demonstration activities (2007 to 2013)

Total cost € 2 499 739,35

EU contribution € 2 499 739,35

Coordinated by THE CHANCELLOR, MASTERS AND SCHOLARS OF THE UNIVERSITY OF OXFORD United Kingdom

This project is featured in...



Final Report Summary - BEEHIVE (Bridging the Evolution and Epidemiology of HIV in Europe)

Our primary aim in the BEEHIVE project was to investigate the viral molecular basis of HIV virulence: linking the virus' genetic sequence to severity of disease. We define this as the viral load, the amount of virus in the blood, which is roughly stable over time, and which varies by orders of magnitude between patients. We also look at the rate at which patients' CD4 cells decline during untreated infection.

To do this, we studied patients with well-defined time of infection and detailed clinical follow up, from eight European countries, and Uganda. We recruited the patients into this study from national cohorts, and obtained plasma samples, from which we genetically sequenced a large number of viruses. Within a sample, these are typically closely genetically related, but not identical. We developed new high throughput sequencing and bioinformatic methods to process this type of data, and generated viral genomes from 3,193 participants. For each sample, we typically have 1 million viral sequence fragments.

We quantified the fraction of variation in viral load and in CD4 decline that is attributable to viral genetics, obtaining a result of around a third and a tenth respectively.

We developed mathematical models to explain the variation in viral loads in terms of basic virological, immunological and evolutionary mechanisms, showing that the structure of the viral population in an infected person is important to explaining variation.

Using a genome-wide association study, we have identified specific mutations associated with altered virulence. These include changes in the most common base, but also kmer (set of k bases in a row) or insertion or deletion, as well as the presence of minor viral variants; they include mutations novel and known. The strongest unexpected effects concern minority variants at sites that are conserved in most patients. We also showed the the total number of rare variants in the genome explains ~5% of viral load variation.

We also found a linked set of mutations, present together in a small group of patients, associated with hypervirulence: viral load (concentration of virus in the blood) 3.5 times higher, and a decline in CD4 immune system cells 2 times faster, independently confirmed in ~100 patients in an expansion of the

2 of 3

BEEHIVE data set. Despite the widespread availability of treatment, this hypervirulent lineage will likely be associated with increased mortality amongst those infected with it.

To address the project's secondary aim, we identified cases of transmission in our (carefully anonymized) data set. These allowed us to quantify transmission patterns, such as those between different cities, and identify demographic characteristics associated with increased risk of transmission.

We also focussed on a detailed exploration of one of the cohorts, in the Netherlands, where the national coverage of sequence data is highest. We showed the epidemic is maintained by a long persisting clusters of transmission, and that most transmission events occur during early HIV, before individuals become aware of their infected status. Based on these data, we modelled different epidemic control strategies, and concluded that it would be difficult to send the epidemic into dtermined decline without offering pre-exposure prophylaxis.

For the project's tertiary aim, we determined which of the patients were infected with two distinct viral strains, and found strong evidence for these individuals having slightly raised viral load.

To process and analyse the BEEHIVE viral sequence data, we developed the computational methods, 'shiver' and 'phyloscanner'. 'shiver' accurately assembles whole viral genomes from raw sequencing experiment output. 'phyloscanner' considers within- and between-host pathogen diversity to make inferences about transmission from one individual to another, identify multiply infected individuals, and detect low-level contamination in sequencing data.

We released both methods as accessible documented software to facilitate their uptake by the wider community. Coupling these two methods to each other and a sequencing protocol completes a chain starting with virus-containing samples and ending with a quantitative picture of pathogen flow in populations. Using this output in mathematical models of epidemics and interventions leads directly to recommendations for targeted public health interventions of maximum effectiveness.

There have been several succesful spin-offs of the BEEHIVE project focussed on HIV pathogen sequencing for public health in sub Saharan Africa, where much less is known about the structure of the epidemics, and the need for better prevention is acute.

Last update: 16 May 2017

Permalink: https://cordis.europa.eu/project/id/339251/reporting

European Union, 2025