

Workshop on Next Generation Sequencing of Viruses
Institut Pasteur, Paris, France
20 & 21 May 2015



© Institut Pasteur

Analysis and interpretation of Next Generation Sequencing (NGS) data of influenza and other viruses, in particular as regards the significance of minor variants within a virus population, is in its infancy. There is a level of urgency in addressing this subject with respect to formats for reporting of data, what data to be stored in databases, issues peculiar to segmented viruses such as influenza, interpretation of sequence variation and its usefulness, in terms of significance for potential emergence and impact of minority variants detected in a virus population.

The objective of the workshop is to bring together scientists with expertise in NGS technologies for the analysis of genetic variation in virus populations, and intra-host diversity, and the emergence of minor variants in their evolution and in response to immune and antiviral pressure, and those with an interest in using this knowledge to inform public health decision making.

The 2-day expert workshop will include 'state of the art' presentations, reflecting current and future developments, and discussions relating to sequencing technologies, data processing (quality control, mapping), assembly and analysis of data, database format, interpretation and the significance and use of NGS data for public and animal health. Lessons from studies of different viruses will be included in the programme, together with short talks (and posters) selected from submitted abstracts. It is intended that the outcome of the expert discussion will form the basis of guidance on the generation and interpretation of NGS data on viruses.

The workshop is being organised by the isirv-Antiviral Group (AVG) in collaboration with GISAID and partners of the EU PREDEMICS Consortium.

Organising Committee

Martin Beer - Friedrich-Loeffler-Institut, Riems, Germany
Ron Fouchier - Erasmus MC, Rotterdam, The Netherlands
Monica Galiano - PHE, London, UK
Maria Giovanni - NIAID, Maryland, USA
Alan Hay (Co-Chair) - NIMR, London, UK
Aeron Hurt - WHO CC, Melbourne, Australia
Paul Kellam - The Sanger Institute, Hinxton, UK
Philippe Lemey - Rega Institute, Leuven, Belgium
John McCauley - NIMR, London, UK
Andrew Rambaut - Institute of Evolutionary Biology, Edinburgh, UK
Gavin Smith - Duke NUS, Singapore
David Spiro - NIAID, Maryland, USA
Sylvie van der Werf (Co-Chair) - Institut Pasteur, Paris, France
Marco Vignuzzi - Institut Pasteur, Paris, France
Richard Webby - St Jude Children's Research Hospital, Memphis, USA
David Wentworth - CDC, Atlanta, USA



Preliminary Programme

Day 1 – Wednesday 20th May

Session 1. Objectives of NGS/Deep Sequencing – Possibilities & Limitations

- Population diversity (evolution/minor variants; diagnosis)
Andrew Rambaut, Institute of Evolutionary Biology, Edinburgh, UK
- Virus (pathogen) discovery
Christian Drosten, University of Bonn, Bonn, Germany (tbc)
- Outbreak investigation, transmission
Martin Beer, Friedrich Loeffler Institute, Isle of Riems, Germany

Session 2. Sequencing Technologies - 'State of the Art' Current & Future Developments

- Current platforms (advantages/disadvantages)
David Wentworth, CDC, Atlanta, USA
- Future developments & improvements
Chad Nusbaum, Broad Institute, Cambridge, MA, USA (tbc)

Session 3. Sample Preparation, Data processing, Assembly, Analysis

- Sample preparation & analysis
Saskia Smits, Erasmus MC, Rotterdam, The Netherlands (tbc)
- Data processing, assembly, analysis
Paul Kellam, Sanger Institute, Hinxton, UK

Session 4. Examples of Different Pipelines/Viruses (shorter talks)

- *Influenza viruses* - Simon Watson, Sanger Institute, Hinxton, UK
- *Ebola/Lyssaviruses* - Pardis Sabeti, Broad Institute, Cambridge, MA, USA (tbc)
- *Arboviruses* - Yong-Zhen Zhang, China-CDC, Beijing, China
- *HIV* – Christophe Fraser, Imperial College, London, UK
- *MersCoV/SARS* - Leo Poon, Hong Kong, SAR China
- *RSV + Other viruses* - Matthew Cotten, Sanger Institute, Hinxton, UK
- *Hepatitis E viruses* - Marc Eloit, Institut Pasteur, Paris, France (tbc)

Day 2 - Thursday 21st May

Session 5. Data Handling, Storage and Database Format

- Representation of data & integration of other data sets
Maria Giovanni, NIAID, Rockville, MD, USA
- Data storage and format
Guy Cochrane, EMBL-European Bioinformatics Institute, Cambridge, UK (tbc)

Session 6. Interpretation of Results (Sequencing Depth, Relevant Minor Variants)

- Bio-informatics (virus discovery/diversity)
Elodie Ghedin, New York University, New York, USA
- Linking SNPs (circle sequencing)
Raul Andino, University of California, San Francisco, CA, USA (tbc)
- Genetic diversity (intra-host/inter-host)
Gavin Smith, Duke NUS, Singapore
- Minority variant emergence - evolution/adaptation, antigenic drift
Trevor Bedford, Hutchinson Cancer Research Center, Seattle, USA (tbc)
- Minority variant emergence - antiviral resistance
Aeron Hurt, WHO CC, Melbourne, Australia

Session 7. Significance and Communication of Information

- Timely data sharing
Nancy Cox, GISAID SAC
- Public Health perspective (surveillance; zoonotic risk)
Maria Zambon, Public Health England, London, UK
- Animal Health perspective (outbreak spread)
David Swayne, SEPRL, Athens, GA, USA (tbc)