



Collaborative Management Platform for
detection and Analyses of (Re-) emerging
and foodborne outbreaks in Europe

A global platform for sequence-based rapid identification of pathogens

Prof. Frank M. Aarestrup, coordinator
Technical University of Denmark

Prof. Marion Koopmans, deputy coordinator
Erasmus Medical Center, the Netherlands

Dr. Matthew Cotten
Erasmus Medical Center, the Netherlands



COMPARE Goals

- A. Use risk assessment models to drive risk-based sampling**
- B. Collect and store sequence data in standardised structures**
- C. Identify or create validated workflows to extract**

Actionable information

Pathogen identification

Outbreak prediction and detection

Outbreak investigation

- D. Identify or create tools for communicating risk**
- E. Be cost-effective**

Compare members (Scientists from 30 Institutes)

5 years funding EU Horizon 2020 program

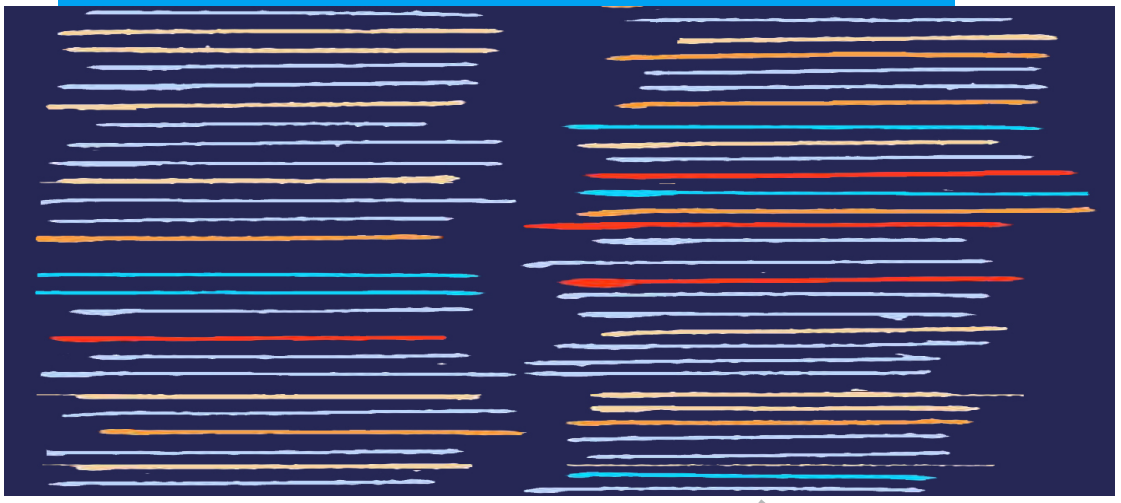
Technical University of Denmark
Erasmus University Medical Center
Statens Serum Institute
Friedrich Loeffler Institute
Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail
Robert Koch-Institut
Istituto Superiore di Sanita
European Molecular Biology Laboratory
Animal and Plant Health Agency/DEFRA
Rijksinstituut voor Volksgezondheid en Milieu
Universitaetsklinikum Bonn
University of Edinburgh
Universiteit Antwerpen
Academisch Medisch Centrum Universiteit van Amsterdam

University of Cambridge
Universidad de Castilla- la Mancha
Artemis One Health Research by Aristotelio Panepistimio Thessalonikis
Tierärztliche Hochschule Hannover
Erasmus Universiteit Rotterdam
Fondation Mérieux
Magyar Tudományos Akademia Wigner Fizikai Kutatóközpont
Institut Français de Recherche pour l'Exploitation de la Mer
Responsible Technology
The Australian National University
Leibniz Institut Deutsche Sammlung von Mikroorganismen und Zellkulturen
Civic Consulting
University of Bologna
Wellcome Trust Sanger Institute



Identify sequences in metagenomic samples

SLIM wrapper



contigs/reads

ublast +

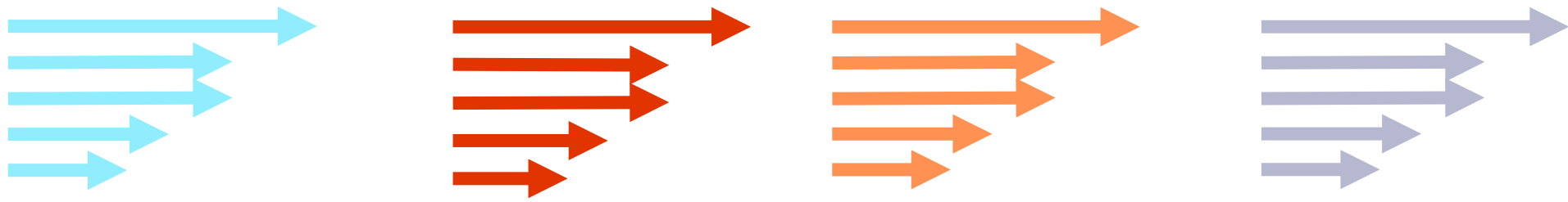
..... output.....etc.

GenBank

taxon ID

- virus_1_protein_database
- virus_2_protein_database
- virus_3_protein_database
- virus_4_protein_database

1. fasta files, contigs sorted, labeled, correct orientation

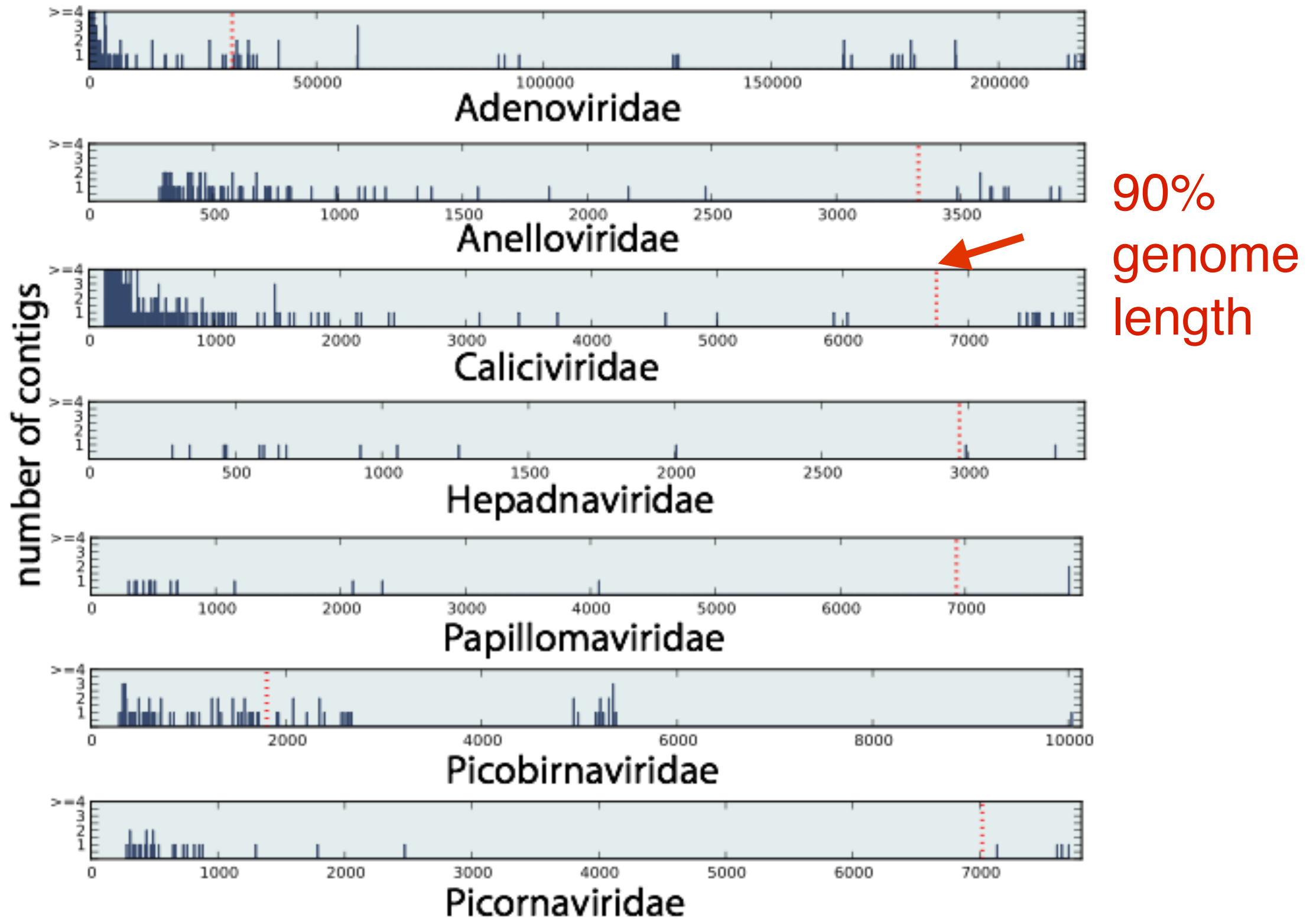


2. simple summary: hit, length, ID etc.

3. alignment file



Identification



Cotten, et al., 2014 PLoS ONE, 9(4), e93269–15. doi:10.1371/journal.pone.0093269

Differentiate **deliberate release** outbreak from **natural cause** outbreak.

Phylogenetic clues

1. Virus derived from cloned source may lack diversity, **limited quasi-species**.

2. **Phylogeography** may show inappropriate virus movement.

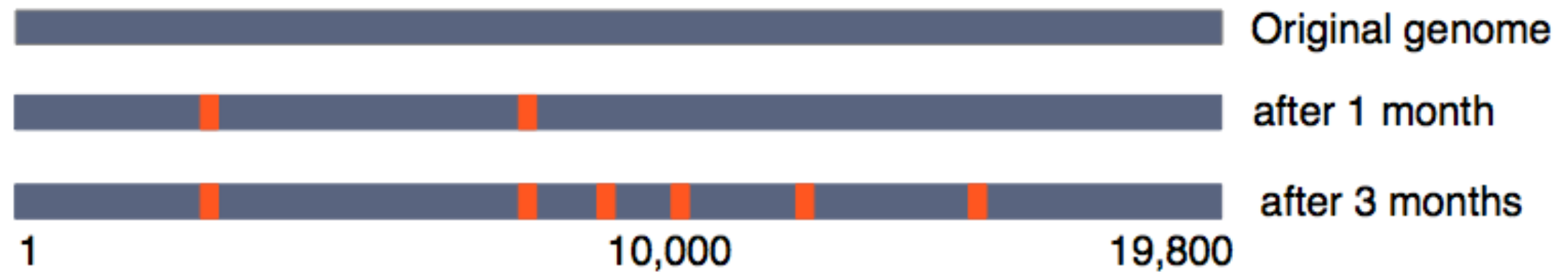
3. **Too Rapid changes**: phylogenetic distance may indicate engineering, cell culture selection.

4. **Frozen evolution**.

RNA viruses (except retroviruses) must continue to infect/replicate/evolve.

On dated tree, paused evolution = frozen sample = deliberate release?

Ebola virus RNA genome evolves at a defined rate



The virus genome sequence can reveal close virus ancestors and potential sources of the infection

Mateneh Ebola Treatment Centre, Makeni Sierra Leone

Ian Goodfellow
Univ of Cambridge
Sequencing
Tent

PHE
Diagnostic
Lab

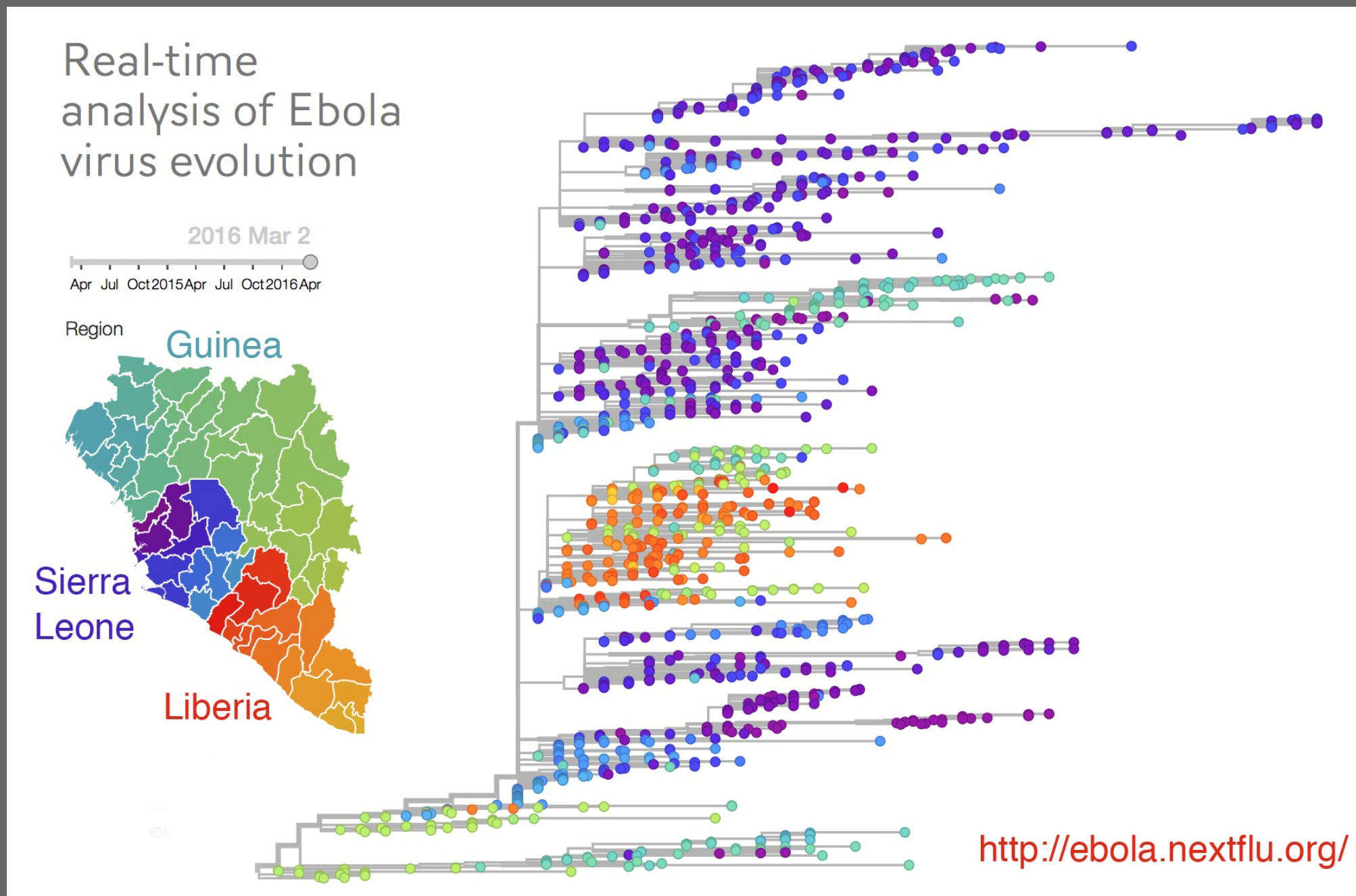


1 of 6 UK sponsored
treatment centres,
Funded by DfID, built by
Royal Engineers
1 of 3 diagnostic labs set up
by Public Health England

Without an open database of contemporary genomes placement of a genome from a new case is not possible

<http://virological.org/c/ebolavirus>, Andrew Rambaut

<http://ebola.nextflu.org/>, Richard Neher, Trevor Bedford

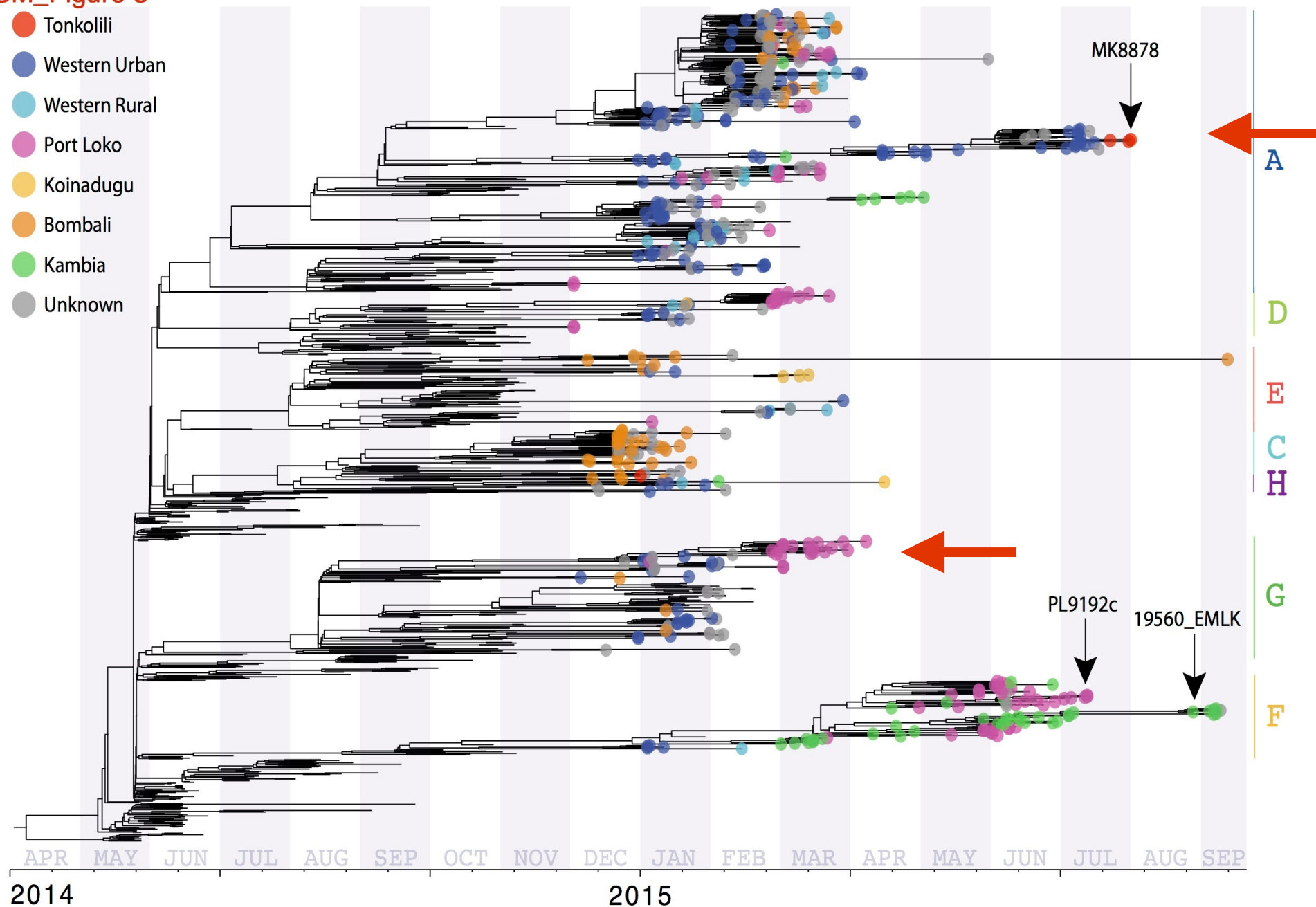


2. Phylogeography may show rapid virus movement

Example from Ebola virus

SM_Figure 3

- Tonkolili
- Western Urban
- Western Rural
- Port Loko
- Koinadugu
- Bombali
- Kambia
- Unknown



Movement of virus from Freetown to Tonkolili

Movement of virus from Freetown to Port Loko

Rapid outbreak sequencing of Ebola virus in Sierra Leone identifies transmission chains linked to sporadic cases
Arias et al. *Virus Evolution* Jan 2016, 2 (1) vew016; DOI: 10.1093/ve/vew016

Example

July 2015, new virus infections in region free of Ebola for previous 130 days

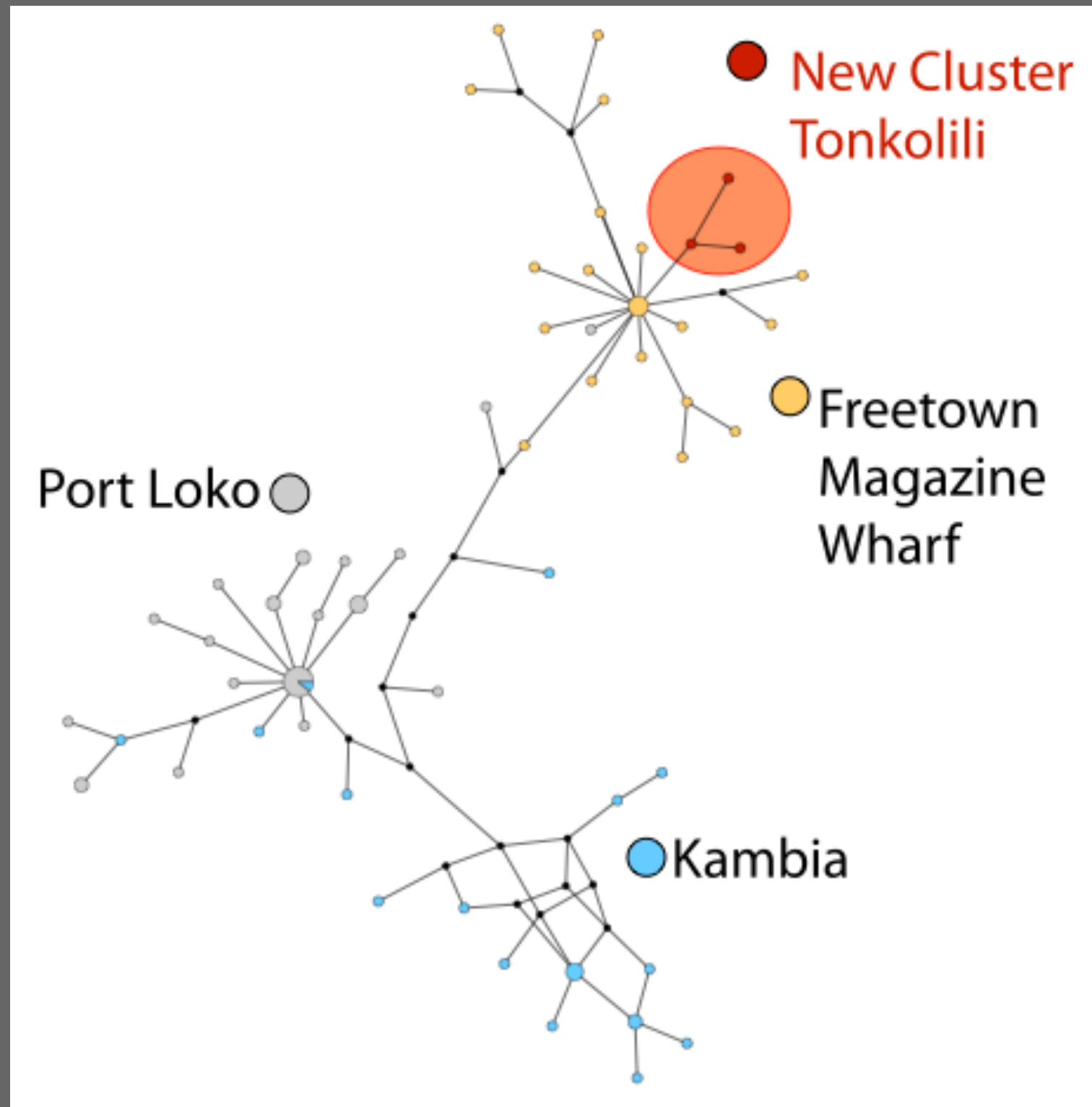
Virus sequenced within 48 hours

Clustered with viruses from Freetown

Patient had recently travelled from Freetown

Excluded:

1. unknown transmission chain,
2. movement from Guinea,
3. new zoonosis

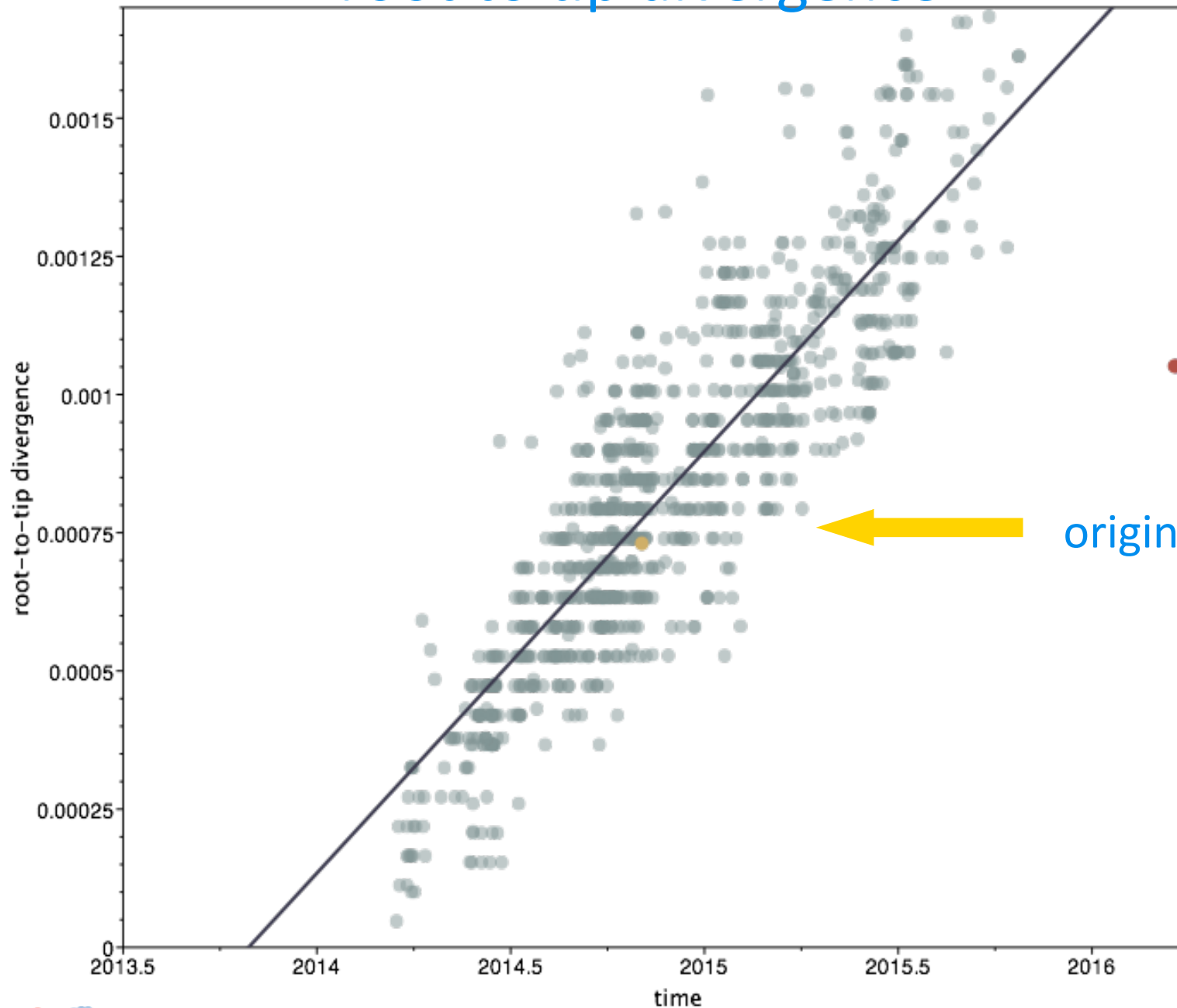


Point 4. Reduced virus evolutionary rate

EBOV persisted in the survivor for 531 days.

The virus in blood from 3Nov14 differed from that in semen from 21Mar16 by 6 nt changes
Evolutionary rate: 0.24×10^{-3} substitutions per site per year = 5X slower than rate seen during human-to-human transmission in this outbreak (Fig. S2).

root to tip divergence



virus from semen
504 days later

original virus (acute)

Analysis from Gytis Dudas,
Andrew Rambaut

From Resurgence of Ebola virus disease in
Guinea linked to a survivor with virus
persistence in seminal fluid for more than 500
days.

**Diallo, et al. Clinical Infectious Diseases
09/2016; DOI:10.1093/cid/ciw601**

Challenges

1. Validation of sample handling and sequencing platforms
2. Validation of genome assembly, data analysis and virus discovery methods
3. Databases: is GenBank enough?
4. New algorithms for virus detection

University of Cambridge

Ian Goodfellow

Jia Lu Armando Arias

Luke Meredith Lucy Thorne

Sarah Caddy

University of Makeni

Umaru Jah

Raoul Emeric

Alimamy Tarawalie

University of Edinburgh

Andrew Rambaut

Gytis Dudas

Erasmus Medical Center/

Dutch Mobile Labs

Marion Koopmans

Suzan Pas

Saskia Smits

WHO

Dhamari Naidoo
Sierra Leone MOH, European
Mobile Labs, Public Health
England, DFID, CDC

Wellcome Trust Sanger Institute

Matthew Cotten now EMC

My Phan soon EMC

Bas Oude Munnink now EMC

Simon Watson

Paul Kellam

Clinicians, nurses, scientists, PHE
volunteers who helped run
diagnostic and care facilities

