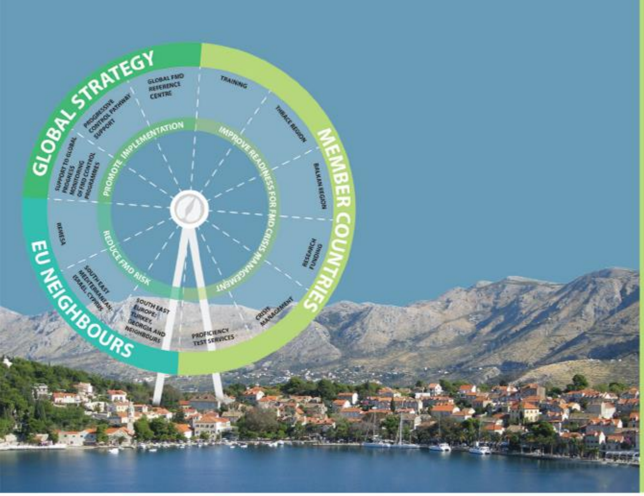


Open session of the Standing Technical
and Research Committees of the EuFMD

Cavtat (Croatia), 29-31 october 2014

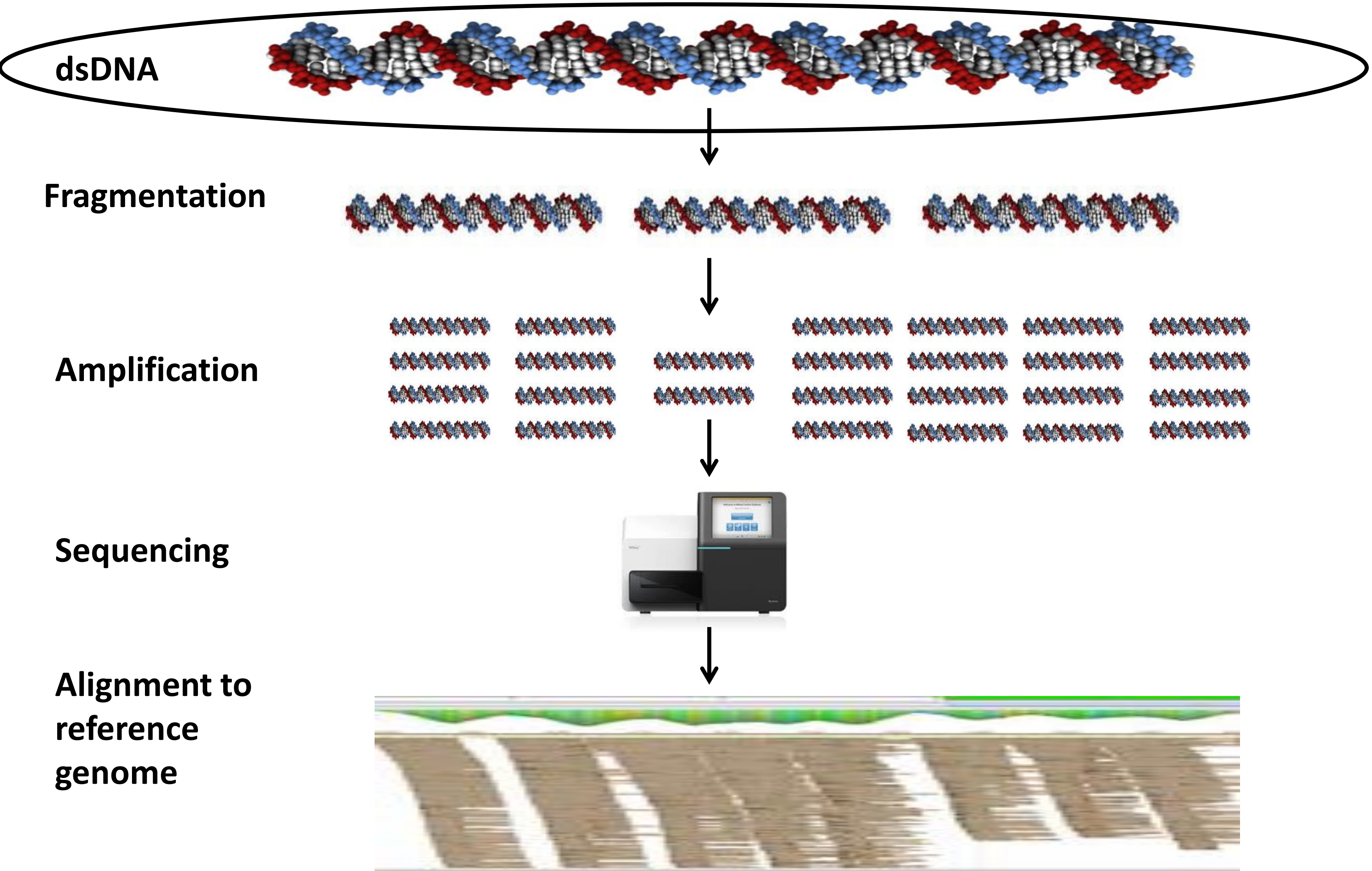
eufmd
european commission for the
control of foot-and-mouth disease

Whole Genome Sequencing of FMDV: Development of protocols and Application to a large outbreak

Graham Freimanis
Non-Vesicular Reference Lab, The Pirbright Institute

EMIDA ERA-NET
SEVENTH FRAMEWORK PROGRAMME
THE Pirbright INSTITUTE 100 YEARS
Epi-SEQ

Next Generation Sequencing (NGS)



dsDNA

Fragmentation

Amplification

Sequencing

Alignment to reference genome

Why Use NGS?

Previous strategies (i.e. Sanger sequencing) subject to:

- limited throughput/labour intensive
- Subject to *a priori* knowledge for sequencing i.e. primers, variation

Criteria for inclusion in protocol development:

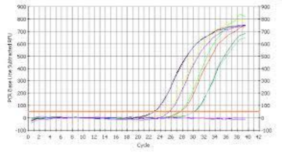
1. **Reduce Bias:** 'No-PCR'
2. **Whole genome coverage**: S and L fragments (polyC tract)
3. Applicable to other viruses
4. **Consensus-level** or deeper
5. Potential for high-throughput

Protocol

Lesion

Tissue Homogenisation

RNA extraction



qRT-PCR

Genomic DNA digest

1st Strand Synthesis

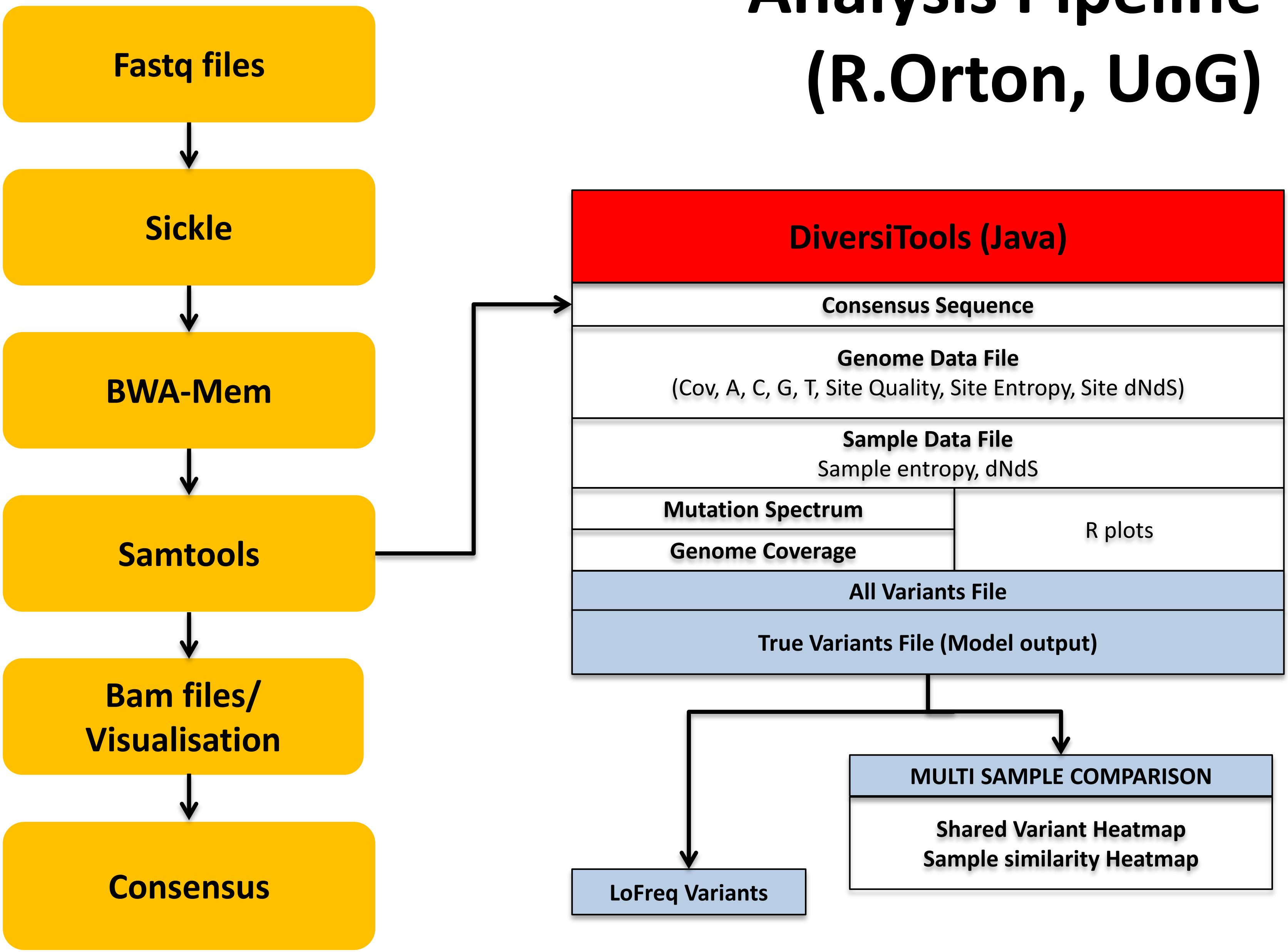
2nd Strand Synthesis

Purify dscDNA

Library Preparation

Sequence on illumina MiSeq

Analysis Pipeline (R.Orton, UoG)



Fastq files

Sickle

BWA-Mem

Samtools

Bam files/ Visualisation

Consensus

DiversiTools (Java)

Consensus Sequence

Genome Data File
(Cov, A, C, G, T, Site Quality, Site Entropy, Site dNdS)

Sample Data File
Sample entropy, dNdS

Mutation Spectrum

Genome Coverage

R plots

All Variants File

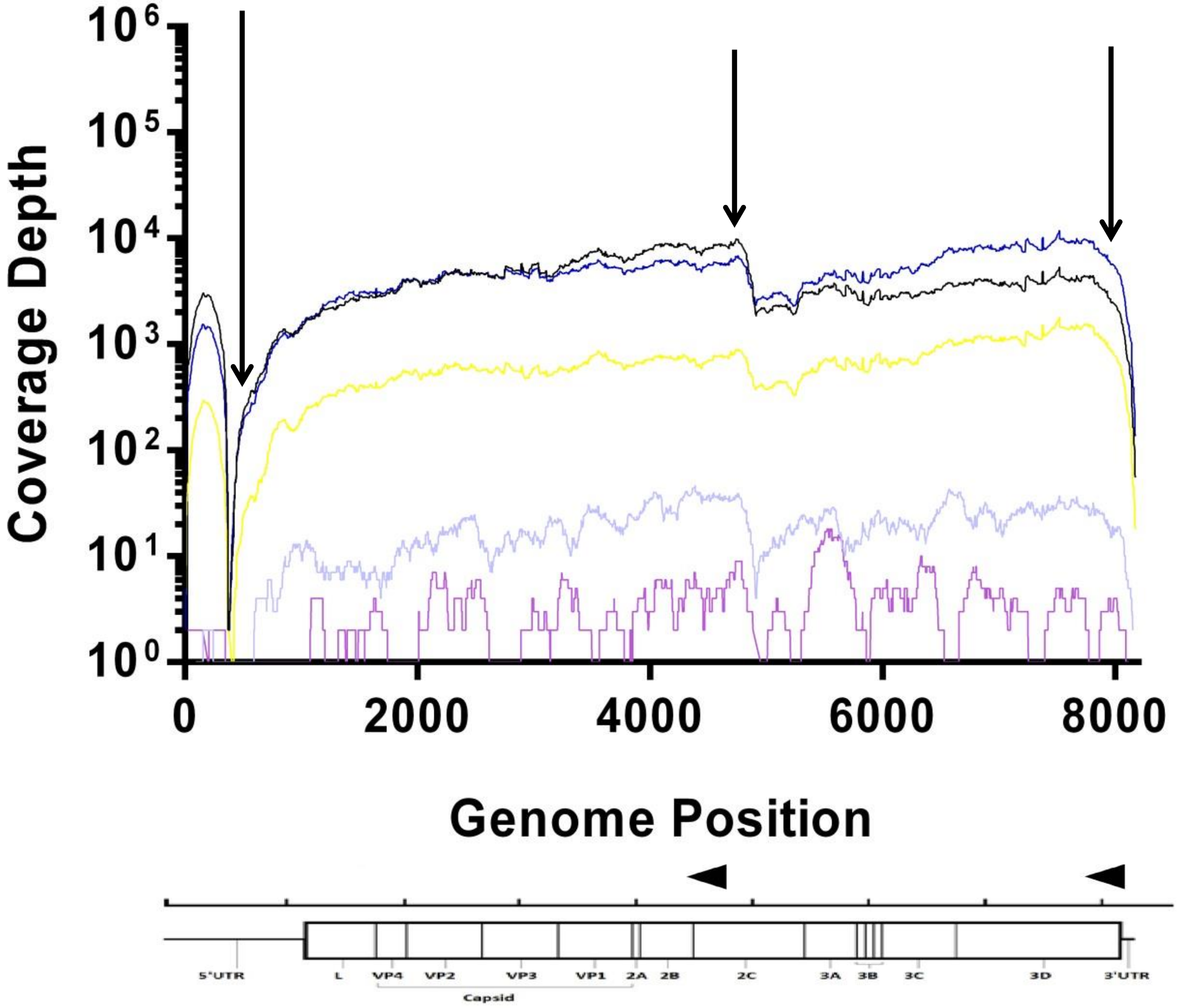
True Variants File (Model output)

MULTI SAMPLE COMPARISON

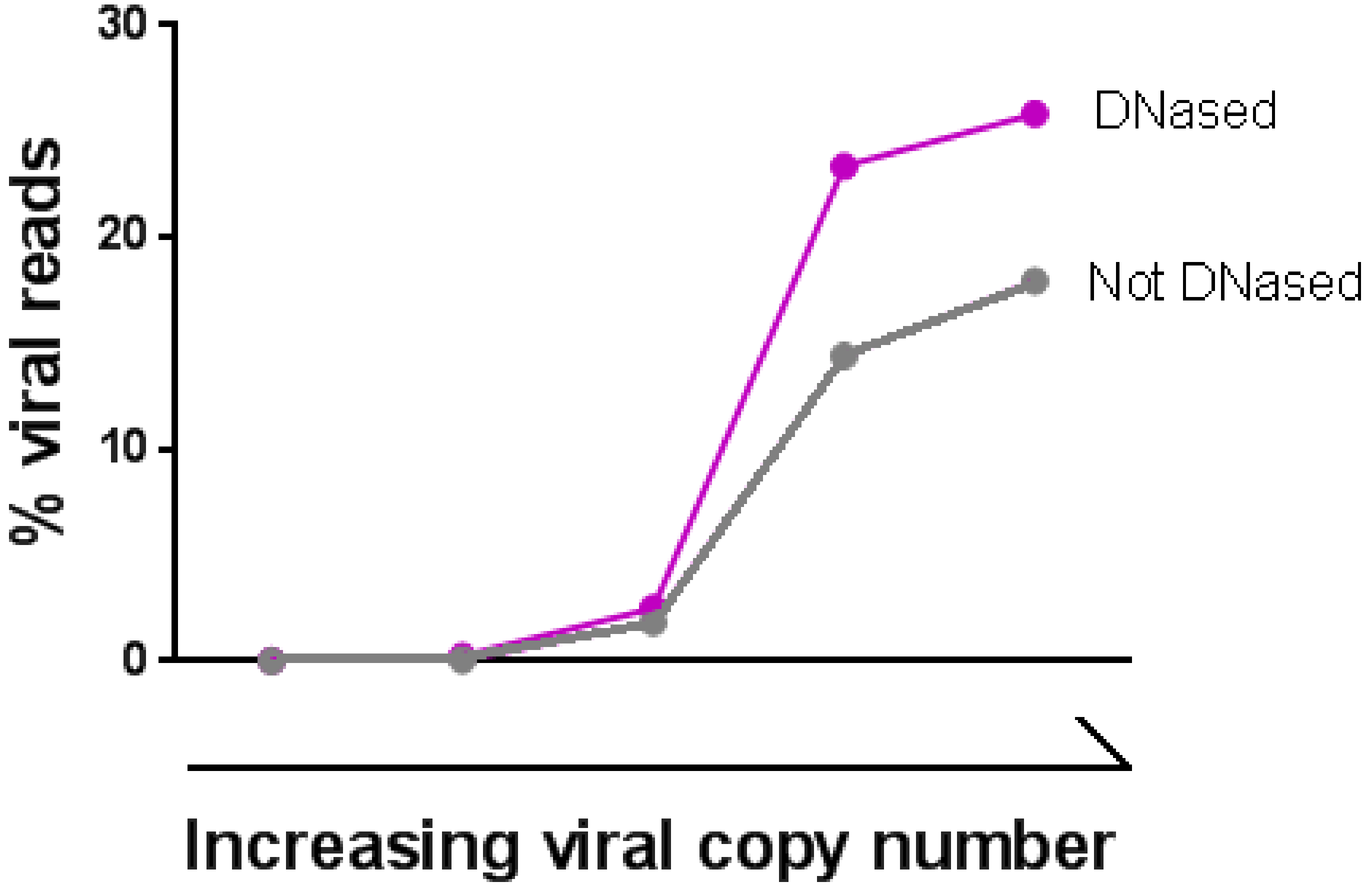
Shared Variant Heatmap
Sample similarity Heatmap

LoFreq Variants

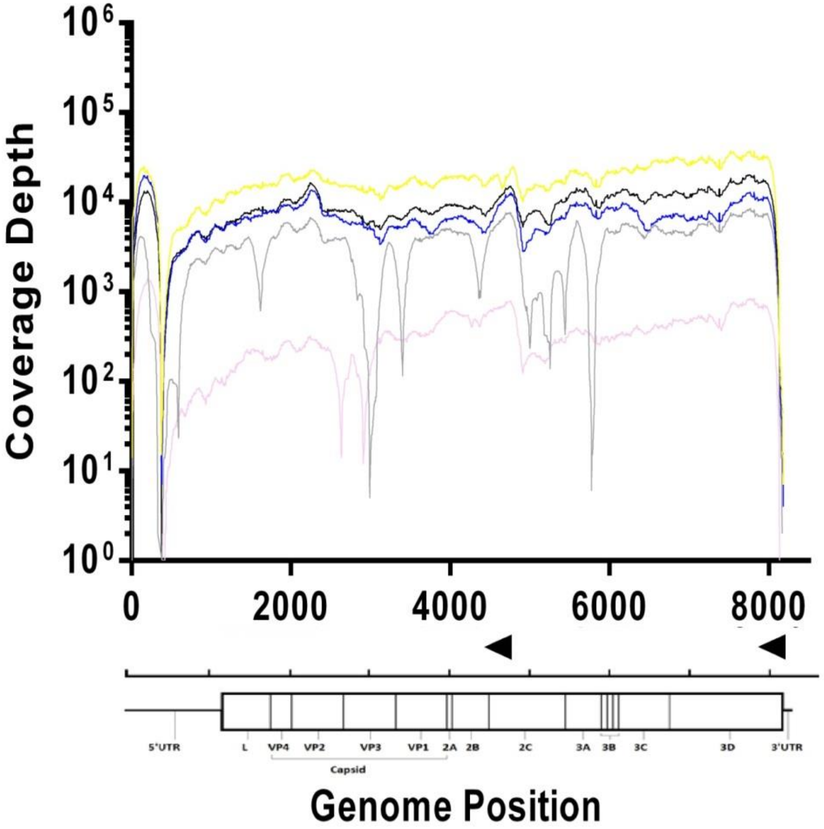
Protocol sensitivity: viral loads >10^6 copies/µl



Depletion of host genomic DNA increases the proportion of viral reads

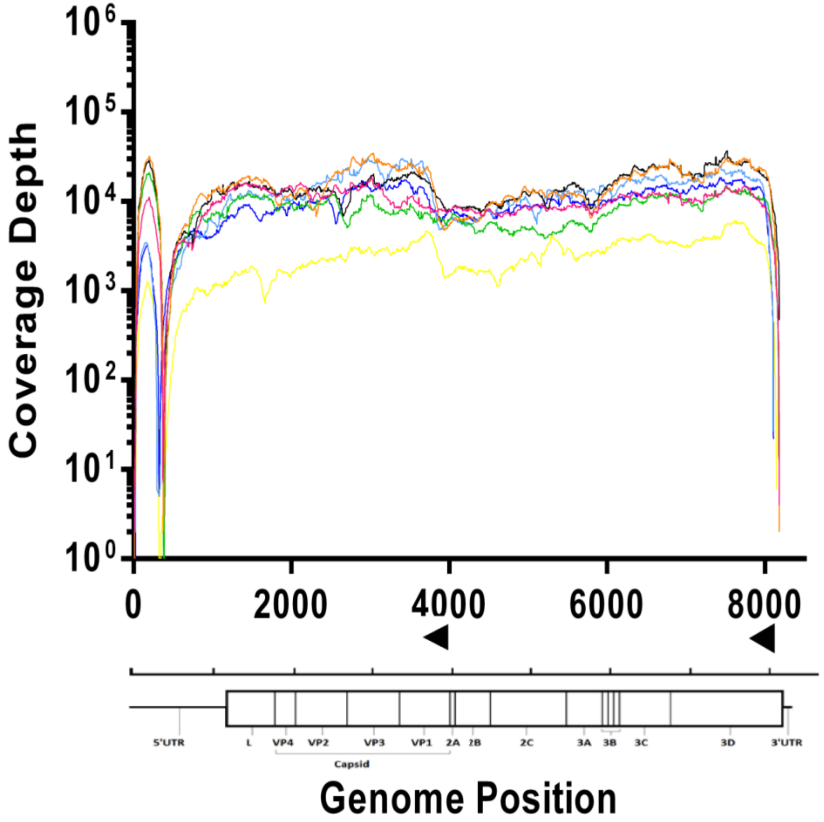


FMDV Serotype Coverage



▲ Genome coverage of 5 UK 2001 field isolates

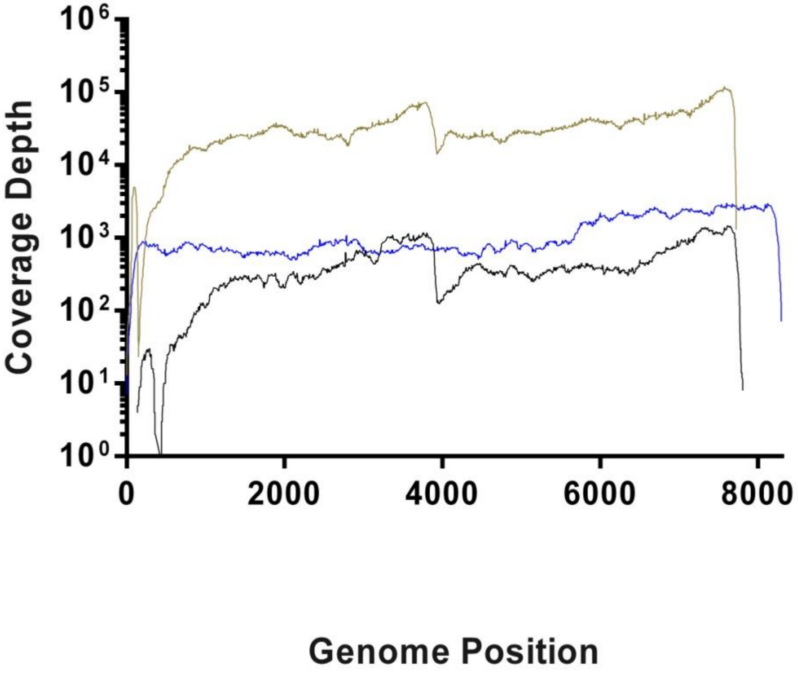
- UKG/1450
- UKG/1558
- UKG/1734
- UKG/14597
- UKG/4998



▲ Genome coverage of 7 FMDV serotypes

- Serotype O
- Serotype A
- Asia 1
- Serotype C
- SAT1
- SAT2
- SAT3

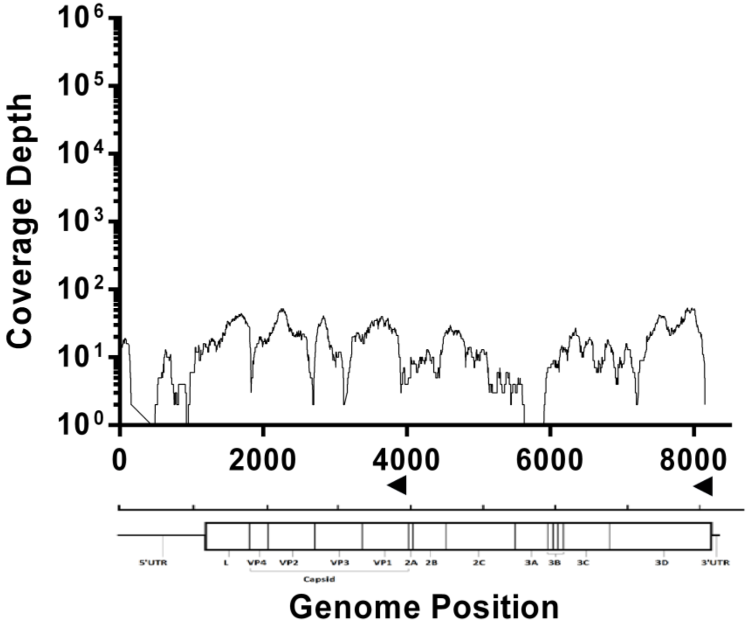
Other Applications



◀

Family	Genus	Species
Picornaviridae	Aphthovirus	Foot-and-mouth disease virus (FMDV)
		Equine rhinitis A virus (EMCV) [Beige]
	Cardiovirus	Encephalomyocarditis virus (ERAV) [Black]
Caliciviridae	Vesivirus	Vesicular exanthema of swine virus (VESV) [Blue]

WGS of viruses unable to grow in culture (FMDV O/ISR/2/2013) ►



Summary I

- ‘No-PCR’ protocol for WGS of FMDV
- Limit of Detection: $>10^6$ copies/µl
- Able to sequence all 7 serotypes of FMDV
- Able to sequence other viruses (non-FMDV)
- Able to sequence viruses without amplification (culture)

Application: UK 2001 Outbreak

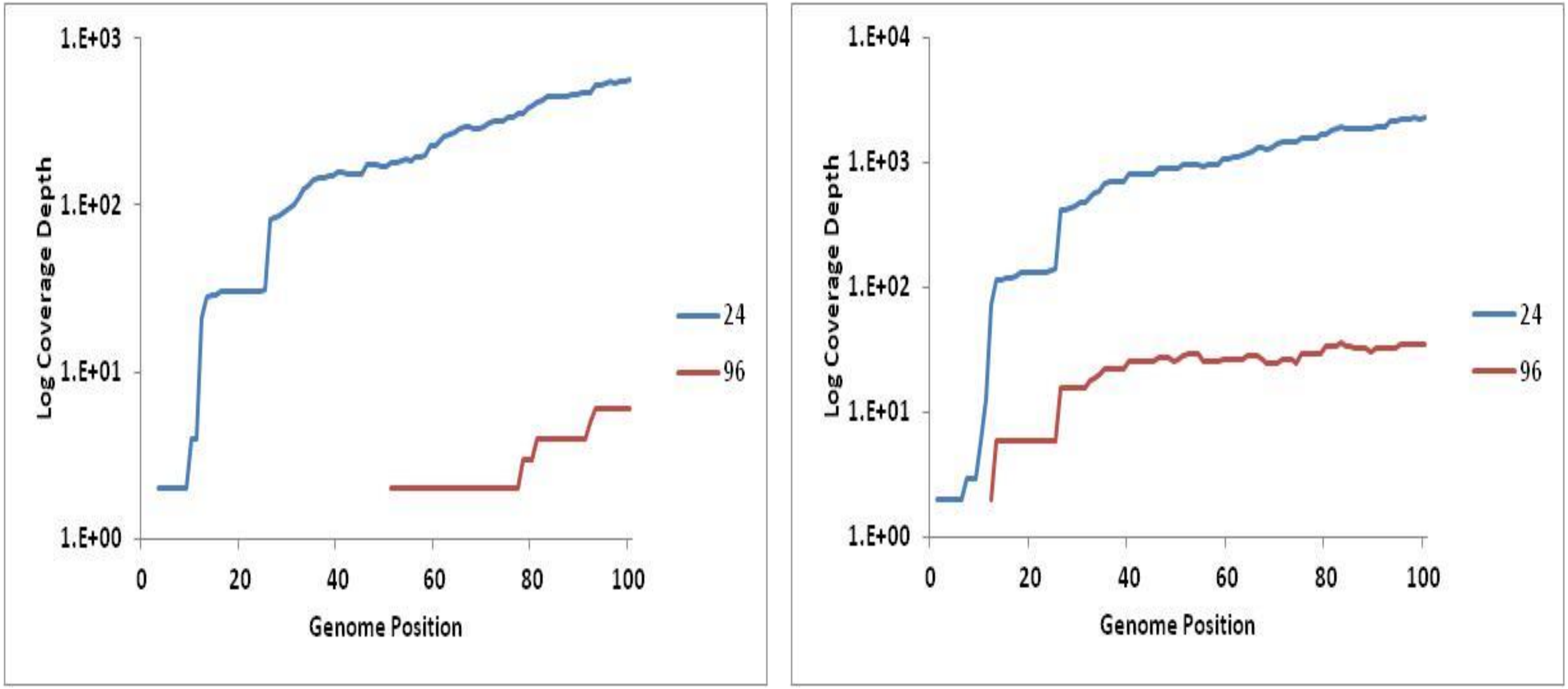
- **2026 IP's confirmed/1300 samples**
- **Economic impact**: > £2billion
- **Aim**: Apply WGS to give a fine point resolution to outbreak tracing

Adapt protocol to high-throughput format (i.e. 96 samples) @consensus level

  - 88 samples, 8 controls
  - 24 samples repeated – 24 format
  - **FMDV**: SAT1/2/3/UKG/Prev Pos
  - **Non-FMDV** : ERAV/EMCV/VESV

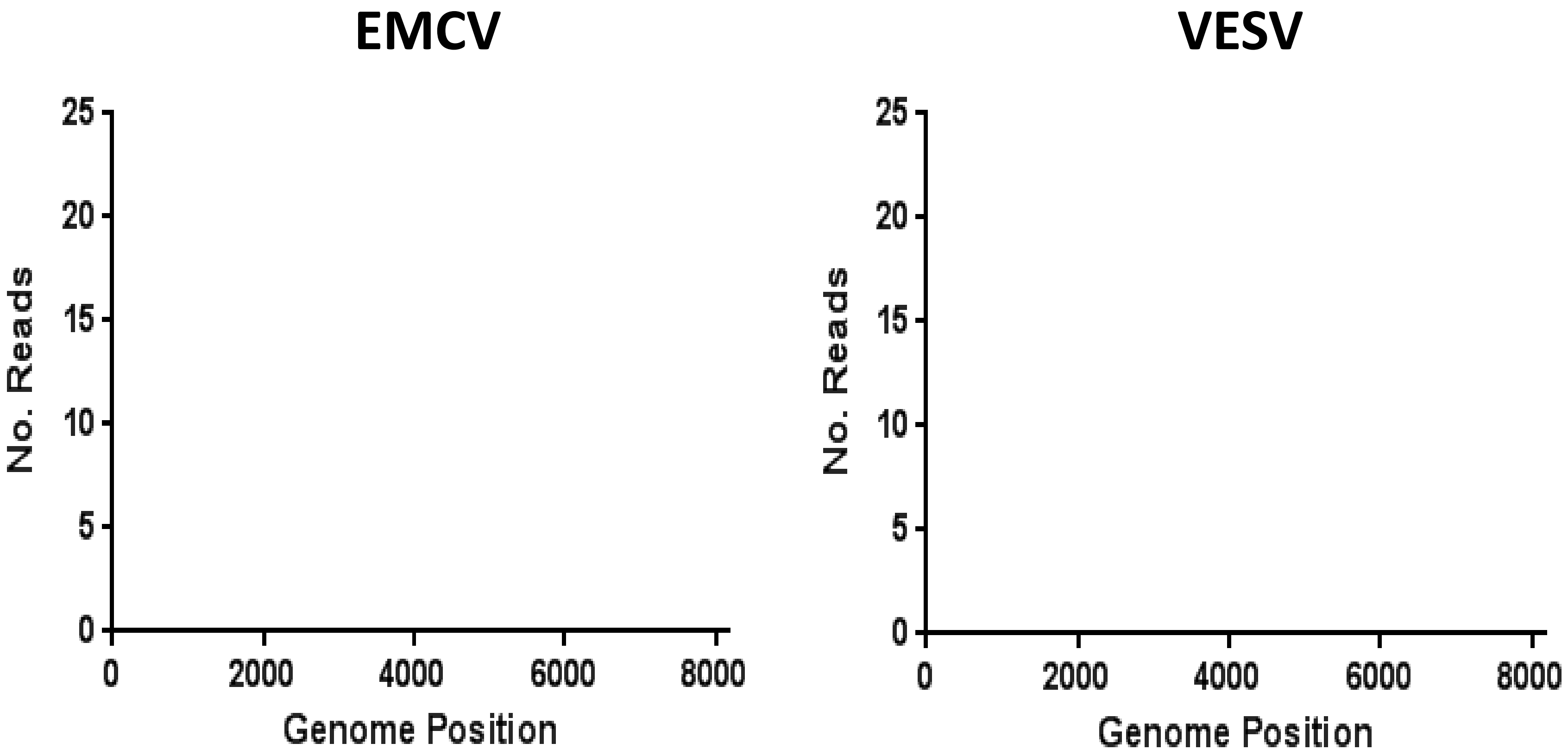
Picture source: Telegraph: http://www.telegraph.co.uk/earth/agriculture/8335705/Farmers-rise-from-the-ashes-of-foot-and-mouth.html

Recovery of 5’ genomic termini



5’ genomic termini recovery - 24 and 96 sample format

Cross-reactivity between samples



▲ Cross reactivity between FMDV samples against EMCV and VESV reference genome

Validation with Sanger Sequencing

4 Samples previously sequenced using Sanger sequencing (gold standard):

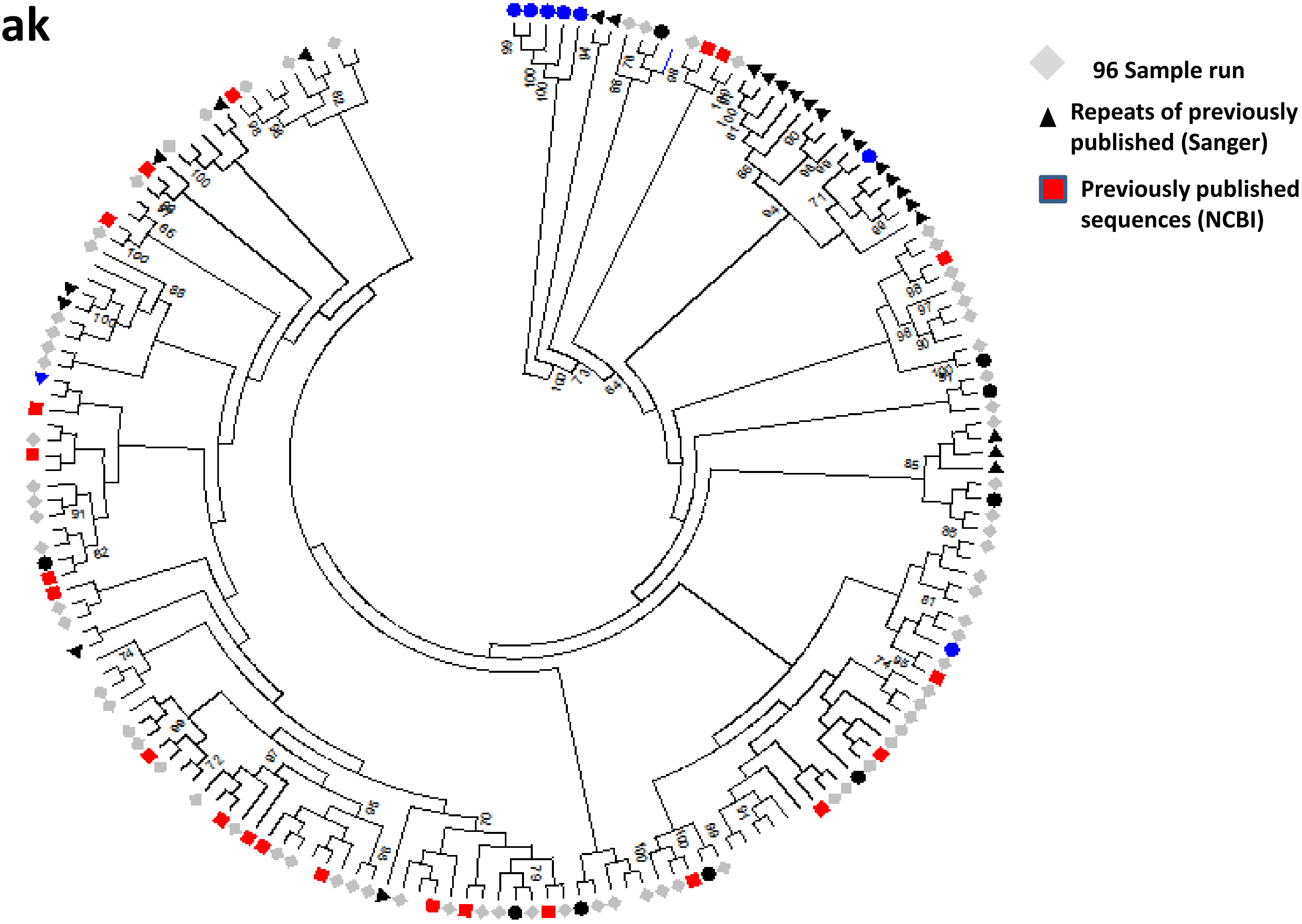
- 3/4 consensus identical with Sanger sequences
- 1/4 differed by one position (ambiguous in Sanger)

8 controls on run validated

24 and 96 sample formats:

- Both gave identical consensus

Applications I: Outbreak Tracing

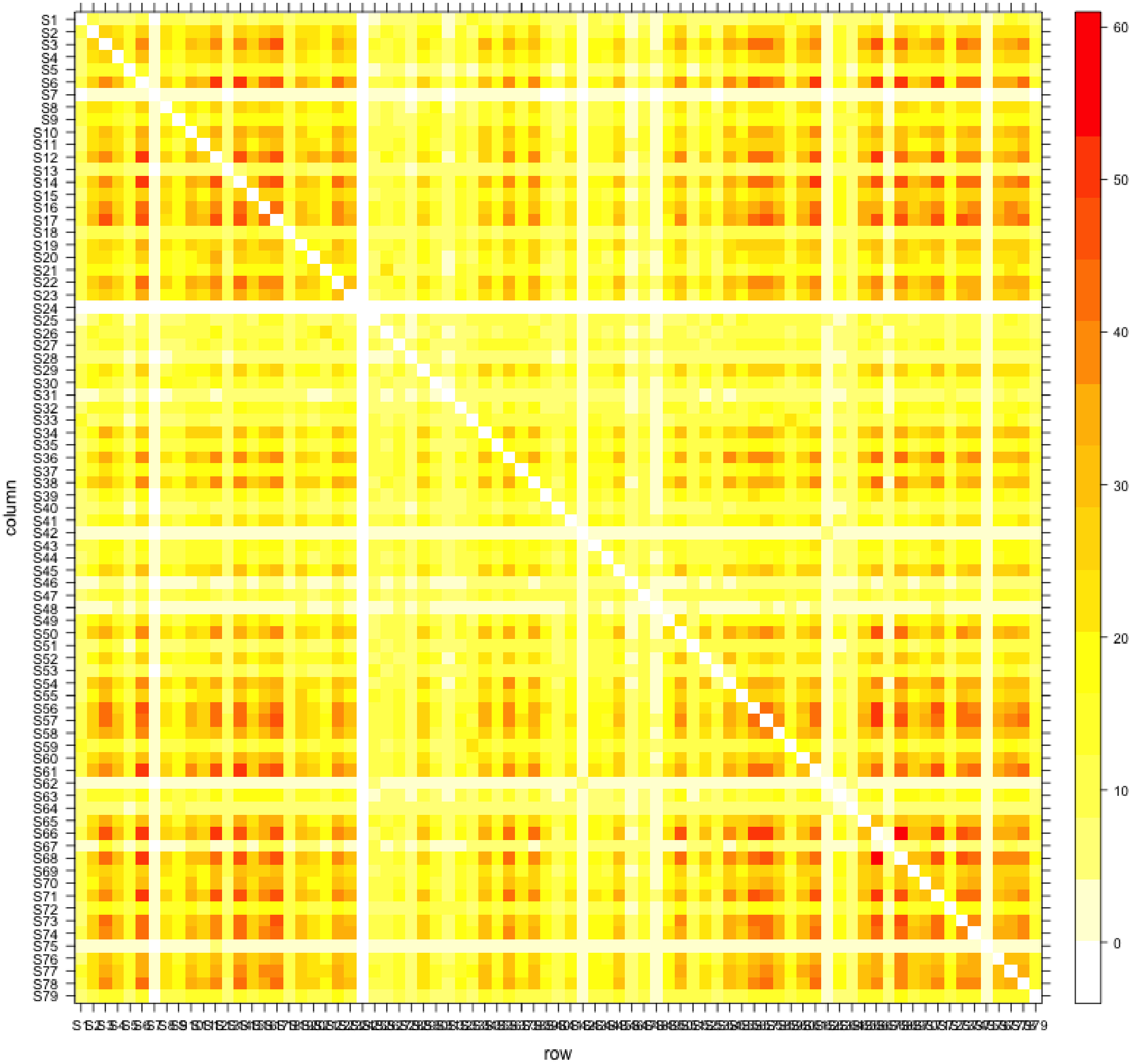


◆ 96 Sample run

▲ Repeats of previously published (Sanger)

🟥 Previously published sequences (NCBI)

Application II:
Consensus Similarity



Summary II

- Protocol successfully generates consensus level (or deeper) sequence for whole FMDV genomes (BMC Genomics, 2014)
- 5’ genomic termini dependent upon viral load
- No detectable cross-reactivity
- Adaptable for high-throughput format (96 samples) – <u>application for larger outbreaks</u>
- Sequencing of UK 2001 outbreak ongoing

Acknowledgements

The Pirbright Institute:

- **Grace Logan**
- **David King**
- **Kasia Bankowska**
- **Begona Valdazo-Gonzalez**
- Nick Sanderson
- Nick Knowles
- Donald King
- Eleanor Cottam
- Valerie Mioulet & WRLFMD

University of Glasgow:

- **Richard Orton**

The Epi-Seq Consortium