NCPV500
An Electronic Resource for Mining Viral Genomes

Karen Buttigieg[1], Sarah Alexander[1], Lidia Liberty[1], Joseph Taylor[1], Nicholas Grayson[2], Edward Burnett[1], Julian Parkhill[2], Julie E. Russell[1]

[1]Public Health England, Porton Down, UK. [2]Wellcome Trust Sanger Institute, Cambridge, UK.

Public Health England

Protecting and improving the nation's health

INTRODUCTION

NCPV500 is a collaboration between Public Health England and the Wellcome Trust Sanger Institute to produce 500 viral genomes from PHE's National Collection of Pathogenic Viruses (NCPV) using the Illumina sequencing platform. NCPV curates and supplies authenticated human pathogenic viruses for the research community.

NCPV National Collection of Pathogenic Viruses — Operated by Public Health England

Data mining the sequences of individual genomes helps to catalogue the genes encoded in a particular strain and is a vital step for in-depth characterisation studies. Sequencing of multiple isolates, strains or species enables understanding of the factors responsible for varying virulence using comparative genomics.

METHODS

Figure 1. Extraction and sequencing technology: Maxwell 16 and Illumina MiSeq

Viral DNA was extracted from virus strains within NCPV, using a Maxwell 16 (Promega) automated system. Whole genome sequencing was performed on the Illumina MiSeq platform at the Wellcome Trust Sanger Institute. Assembled and annotated genome sequences will be freely publicly accessible via a comprehensive web-based Biological Resource Information Centre (BRIC).

RESULTS

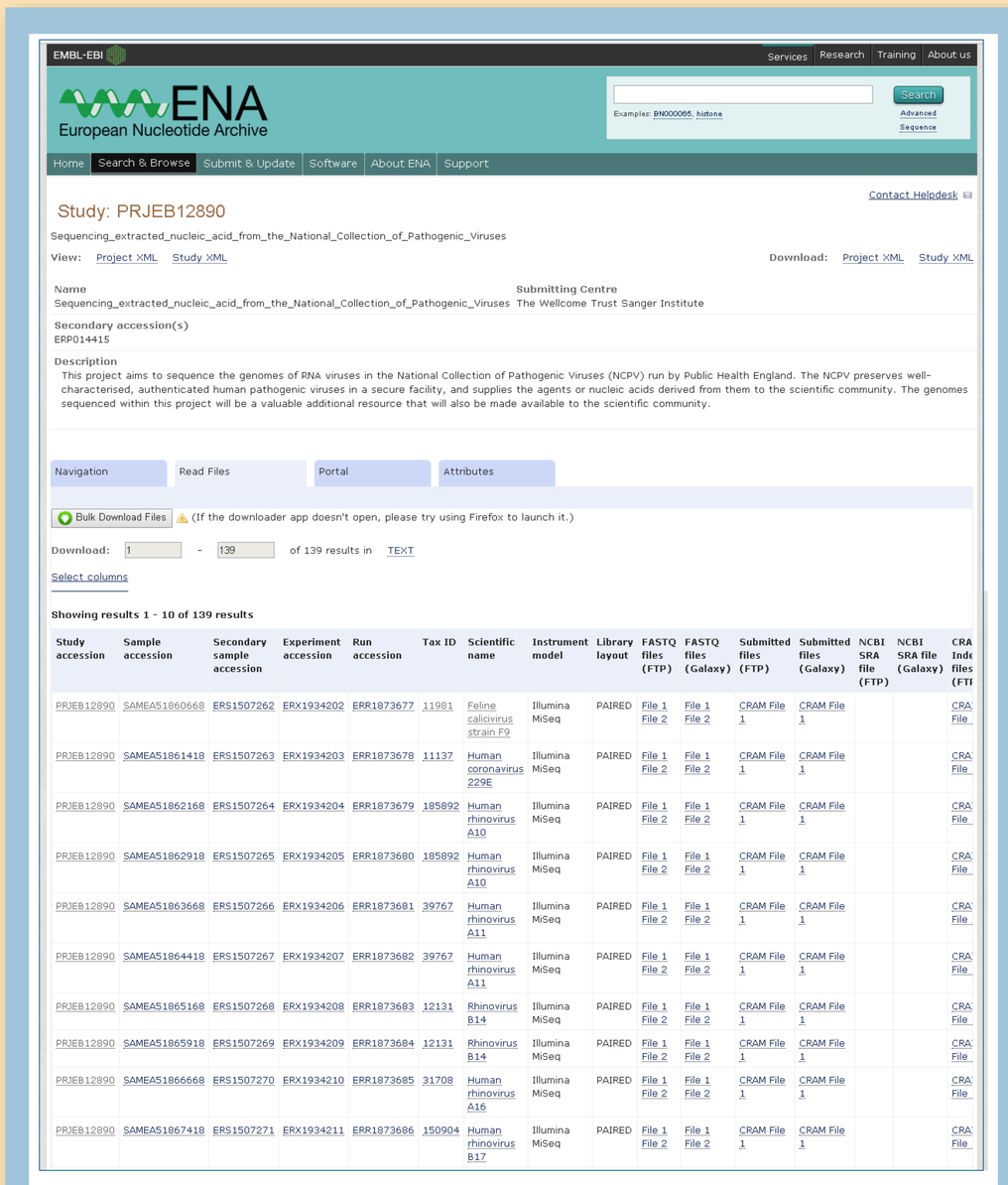

Figure 2. ENA uploaded sequences

Nucleic acid from 196 ACDP Hazard Group 2 viruses has been extracted and sequenced.

Sequences from 139 RNA virus strains have been uploaded to the European Nucleotide Archive, under Study PRJEB12890 (Figure 2).

Files can be downloaded and analysed by users in FASTQ or Galaxy formats.

The BRIC will capture all the information regarding the virus strains, including isolation and deposit details (where known), safety data sheets, production and quality control results. This will be combined with data from external sources, such as sequences from ENA and NCBI, and publications.

NCPV National Collection of Pathogenic Viruses — Operated by Public Health England

ECACC European Collection of Authenticated Cell Cultures — Operated by Public Health England

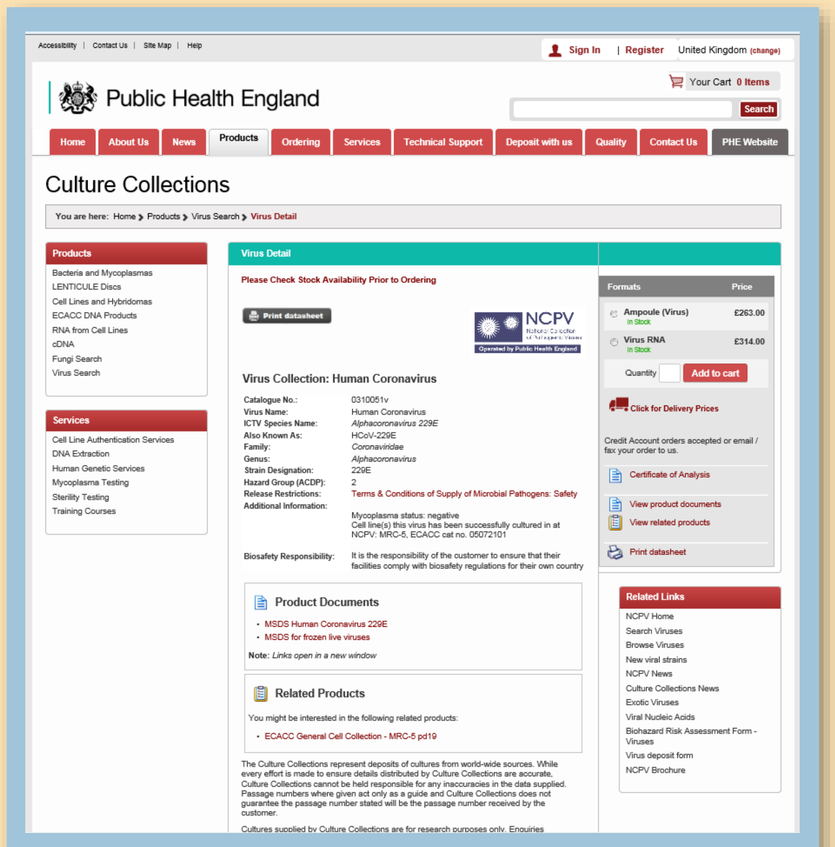

Figure 3. Culture Collections catalogue listing

All the sequenced strains are publically available from the National Collection of Pathogenic Viruses (Figure 3).

Cell lines for culturing viruses can be obtained from the European Collection of Authenticated Cell Cultures (ECACC), another Culture Collection of Public Health England.

DISCUSSION

Next Generation Sequencing is an emerging technology that is becoming more widely used in pathogen microbiology. applications have become more diverse, including real-time epidemiology, clinical diagnosis, microbe discovery, taxonomic classification, quality control of vaccines, tracking adaptation and evolution and understanding the roles of viral genes in infection.

Online availability of the viral genomes, coupled with the biological availability of the virus strains from NCPV, will enable investigation of the interactions between genotype and phenotype of known, emerging and novel viral pathogens.

Specific research questions can be addressed using the historical depth of the Collection:

- compare isolates from before and after vaccine introductions or changes in uptake
- examine changes in vector biology for vector-borne viruses
- examine the effect of serial passage, as an analogy to transmission during an outbreak
- identify co-infections
- examine the effect of culture in alternative cell lines

Future work will focus on sequencing a further 304 viruses, including Hazard Group 3 organisms.

Researchers are invited to deposit newly isolated virus strains into NCPV free of charge, and have the viral genomes sequenced and added to the BRIC.

ACKNOWLEDGEMENTS

The authors are grateful to ECACC for the provision of quality cell cultures for propagation of virus cultures.

The Culture Collections IT team are leading on the design of the BRIC.

CONTACTS

www.phe-culturecollections.org.uk

https://www.ebi.ac.uk/ena/data/view/PRJEB12890

@NCPV