Metagenomics for diagnosis and surveillance of viruses

Sarah Buddle
Professor Judy Breuer's Group
University College London Great Ormond Street Institute of Child Health

Supervisor and slides credit: Dr Sofia Morfopoulou

NHS
Great Ormond Street Hospital for Children
NHS Foundation Trust

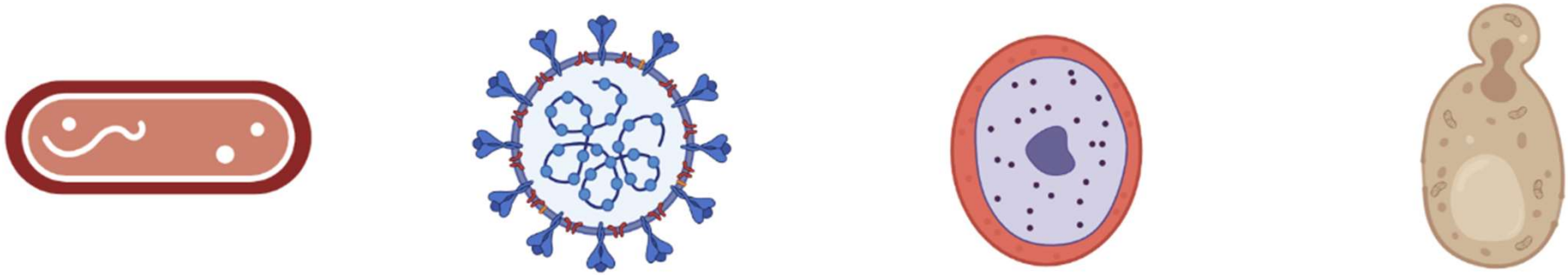
Talk outline

Sterile site metagenomics

- Clinical metagenomics service at Great Ormond Street Hospital, London
- Metagenomics methods evaluation

Future work in clinical metagenomics

Metagenomics



Metagenomics: sequencing all the nucleic acids in a sample.
Allows detection of bacteria, viruses and eukaryotic microbes.

No prior assumptions about microbes present needed

Clinically relevant sequence information

Composition of microbial community

Encephalitis

Characterized by inflammation of the brain.

4000-6000 cases per year in UK

~ 7−8% mortality

70% long lasting consequences

Fatal cases more likely in immunocompromised patients

Can be caused by:

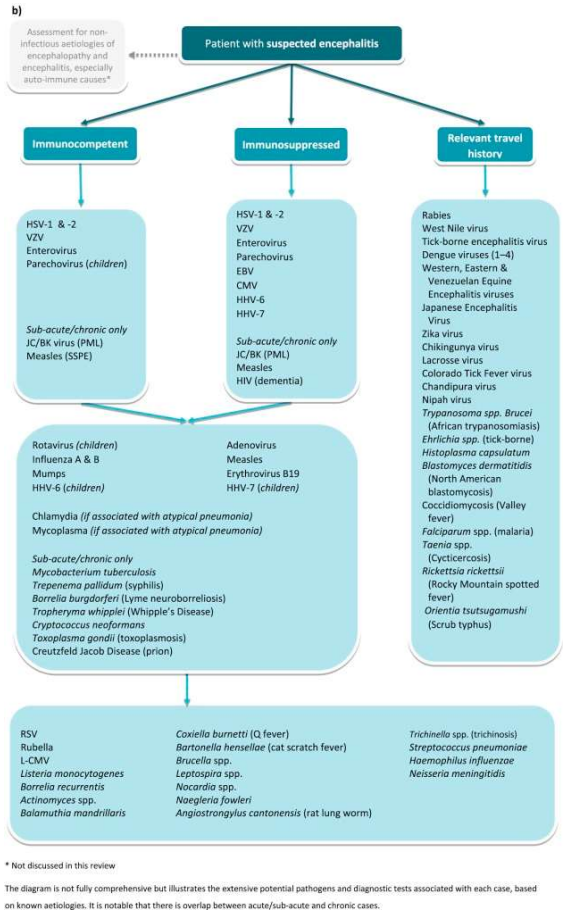
- Infection
- Autoimmune response
- Toxins

50-60% unknown aetiology

Encephalitis diagnosis

Low sample volumes (CSF or brain biopsy)

Traditional tests require prioritisation



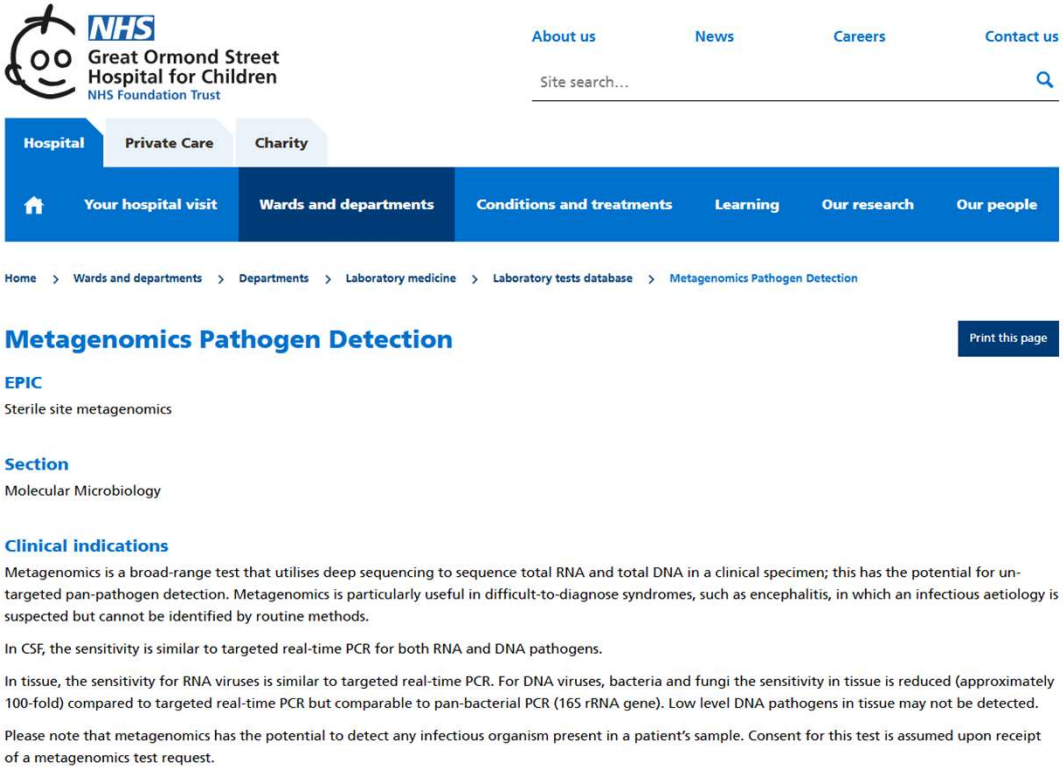
Brown *et al*, Journal of Infection, 2018

Great Ormond Street Hospital Clinical Metagenomics Service

Mainly encephalitis diagnosis using brain biopsy and CSF

Also analyse blood and tissue samples

Great Ormond Street Hospital Clinical Metagenomics Service



NHS Great Ormond Street Hospital for Children NHS Foundation Trust

About us | News | Careers | Contact us

Site search...

Hospital | Private Care | Charity

Your hospital visit | Wards and departments | Conditions and treatments | Learning | Our research | Our people

Home > Wards and departments > Departments > Laboratory medicine > Laboratory tests database > Metagenomics Pathogen Detection

Metagenomics Pathogen Detection

Print this page

EPIC

Sterile site metagenomics

Section

Molecular Microbiology

Clinical indications

Metagenomics is a broad-range test that utilises deep sequencing to sequence total RNA and total DNA in a clinical specimen; this has the potential for un-targeted pan-pathogen detection. Metagenomics is particularly useful in difficult-to-diagnose syndromes, such as encephalitis, in which an infectious aetiology is suspected but cannot be identified by routine methods.

In CSF, the sensitivity is similar to targeted real-time PCR for both RNA and DNA pathogens.

In tissue, the sensitivity for RNA viruses is similar to targeted real-time PCR. For DNA viruses, bacteria and fungi the sensitivity in tissue is reduced (approximately 100-fold) compared to targeted real-time PCR but comparable to pan-bacterial PCR (16S rRNA gene). Low level DNA pathogens in tissue may not be detected.

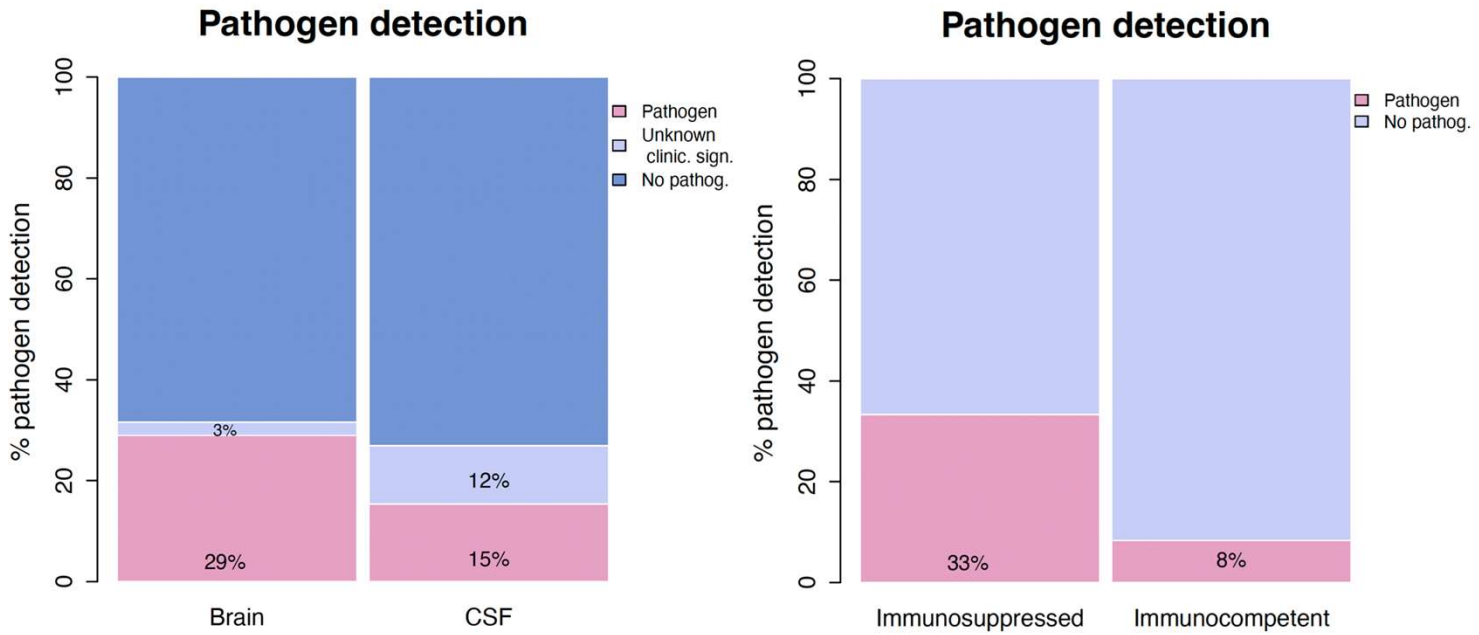
Please note that metagenomics has the potential to detect any infectious organism present in a patient's sample. Consent for this test is assumed upon receipt of a metagenomics test request.

Established 2014

Now a routine clinical service with a weekly sequencing run

Accredited by ISO 15189

Evaluation of clinical service



Penner et al, Journal of Infection, 2023

Changes in clinical management

22% pts CMg uncovered the infectious cause

13 cases: Targeted anti-infective treatment incl. 3 repurposed antivirals

Negative findings still clinically useful for informing immunomodulation

Clinical management changes 74% (n=42/57).

Penner et al, Journal of Infection, 2023

Unexpected pathogens identified

Astrovirus VA1/HMO-C: An Increasingly Recognized Neurotropic Pathogen in Immunocompromised Patients

Julianne R. Brown,[1,2] Sofia Morfopoulou,[3] Jonathan Hubb,[4] Warren A. Emmett,[3] Winnie Ip,[5] Divya Shah,[2] Tony Brooks,[6] Simon M. L. Paine,[7,9] Glenn Anderson,[7] Alex Virasami,[2] C. Y. William Tong,[4] Duncan A. Clark,[4] Vincent Plagnol,[3] Thomas S. Jacques,[7,9] Waseem Qasim,[5] Mike Hubank,[6] and Judith Breuer[1,8]

[1]Virology Department, Great Ormond Street Hospital for Children NHS Foundation Trust, [2]NIHR Biomedical Research Centre, Great Ormond Street Hospital for Children NHS Foundation Trust and University College London, [3]UCL Genetics Institute, University College London, [4]Virology Department, Barts Health NHS Trust, [5]Molecular and Cellular Immunology, [6]Molecular Haematology and Cancer Biology Unit, Institute of Child Health, University College London, [7]Department of Histopathology, Great Ormond Street Hospital for Children NHS Foundation Trust, [8]Department of Infection and Immunity, and [9]Birth Defects Research Centre, Institute of Child Health, University College London, United Kingdom

CORRESPONDENCE

Human Coronavirus OC43 Associated with Fatal Encephalitis

Acta Neuropathol (2017) 133:139–147
DOI 10.1007/s00401-016-1629-y

CrossMark

CASE REPORT

Deep sequencing reveals persistence of cell-associated mumps vaccine virus in chronic encephalitis

Sofia Morfopoulou[1] · Edward T. Mee[2] · Sarah M. Connaughton[2] · Julianne R. Brown[3] · Kimberly Gilmour[4] · WK 'Kling' Chong[5] · W. Paul Duprex[6] · Deborah Ferguson[2] · Mike Hubank[7] · Ciaran Hutchinson[8] · Marios Kaliakatsos[9] · Stephen McQuaid[10,11] · Simon Paine[8,12] · Vincent Plagnol[13] · Christopher Ruis[1] · Alex Virasami[8] · Hong Zhan[14] · Thomas S. Jacques[8,15] · Silke Schepelmann[2] · Waseem Qasim[16,17] · Judith Breuer[1,3]

Pathogens detected

Incomplete list

Viruses	Bacteria	Fungi	Parasites
Adenovirus	Acinetobacter lwoffi	Aspergillus fumigatus	Toxoplasma gondii
Aichi virus	Enterococcus faecalis	Candida glabrata	
Astrovirus*	Mycobacterium avium		
Avian orthoavulavirus*	Mycobacterium Tuberculosis		
CMV			
Coronavirus OC43			
Coxackievirus*			
EBV			
HHV6			
JC virus			
Measles			
Mumps vaccine strain*			
Rotavirus			
Rubella			

Sample types

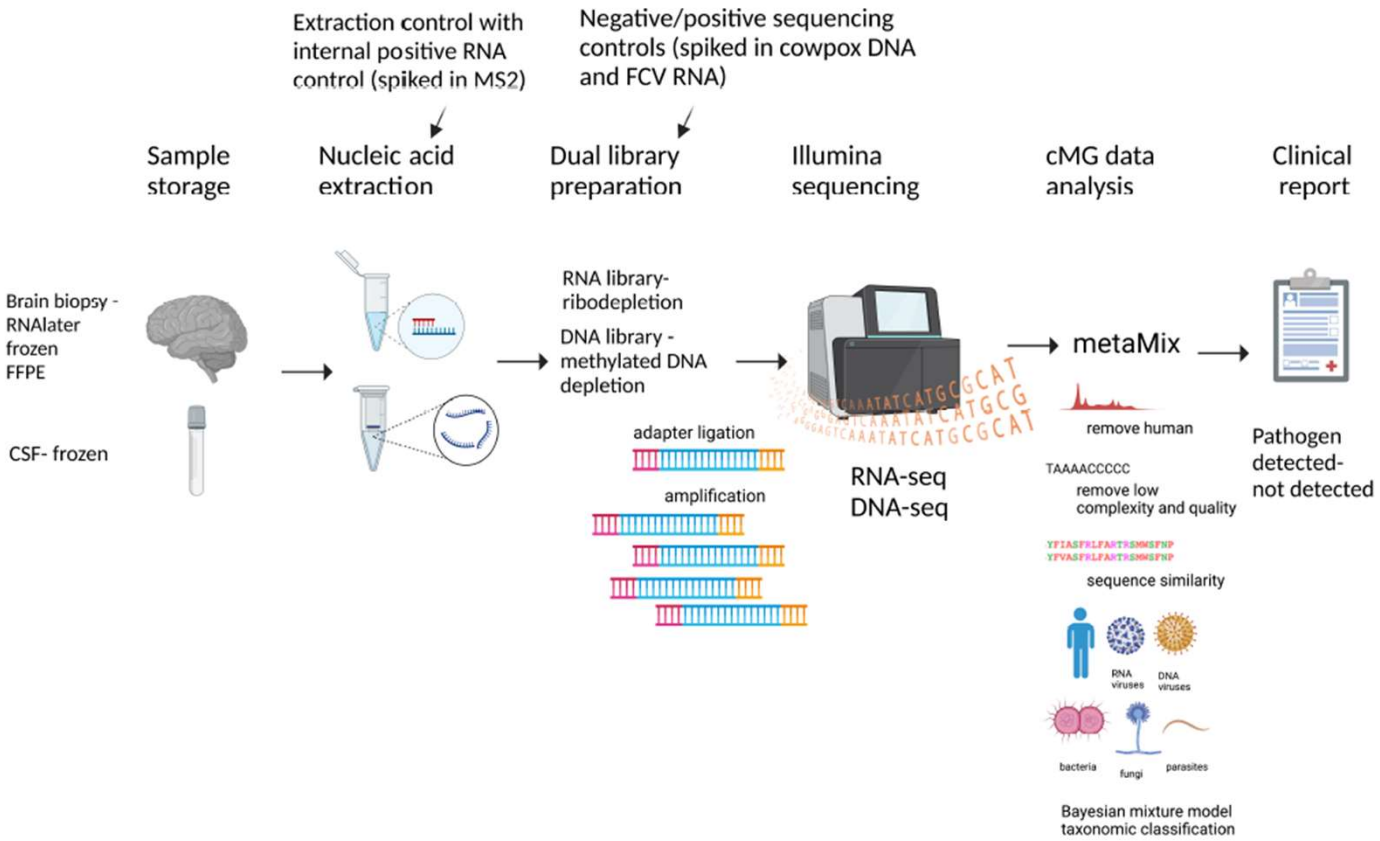
CSF & brain biopsy

Sometime pathogens found in brain but not CSF

High levels of human material

Low microbial diversity

Metagenomics protocol



Penner et al 2023

Why not just sequence RNA?

	DNA-seq	RNA-seq
Total reads	31,052,510	68,105,008
HSV1	44,627	20
HSV2	47,344	8

Removing contaminants and false positives

Lab

Negative and positive controls

Physical separation
- Pre and post PCR
- Sterile site metagenomics from high microbial load samples

PCR confirmation of anything unusual

Analysis and reporting

Specific tool: metaMix

Confirmatory mapping

Mini-MDT discussion before reporting

Confirmatory mapping

Probable contaminant

Probably real

Removing contaminants and false positives

Lab

Negative and positive controls

Physical separation
- Pre and post PCR
- Sterile site metagenomics from high microbial load samples

PCR confirmation of anything unusual

Analysis and reporting

Specific tool: metaMix

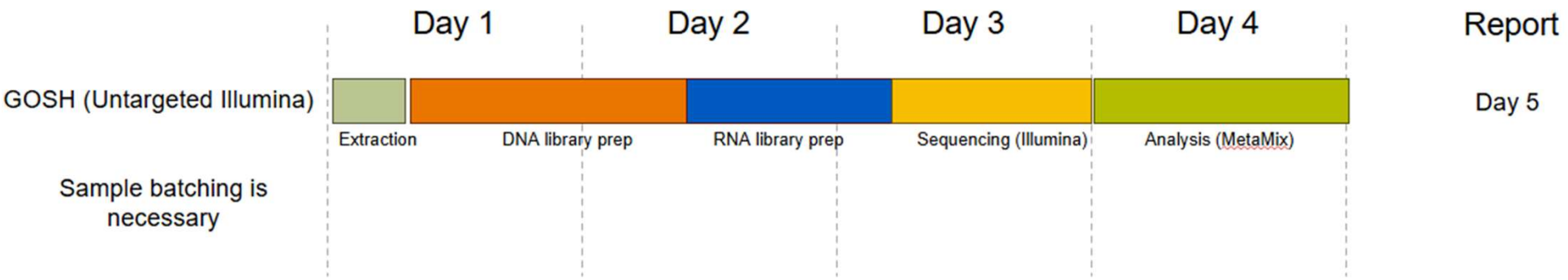
Confirmatory mapping

Mini-MDT discussion before reporting

Potential improvements

Decrease
turnaround time

Turnaround times



Oscar Torres

Potential improvements

Decrease turnaround time

Oxford Nanopore Technologies (ONT) Sequencing

Increase sensitivity

Validation

Spike in dilution series

- RNA-seq ≈ qPCR - tissue and CSF.
- DNA seq ≈ qPCR for CSF
- DNA-seq 100-fold less sensitive qPCR for tissue ≈16S pan-bacterial PCR

Mock Tissue – spiked dilution series results with model organisms

RNA virus (feline calicivirus)		DNA virus (cowpox virus)	
PCR Ct value	**Detection by mNGS (No of reads)**	**PCR Ct value**	**Detection by mNGS (No of reads)**
29	Detected (3,095)	28	Detected (676 reads)
32	Detected (5,412)	31	Detected (81 reads)
39	Detected (344)	34	Not detected
		36	Not detected

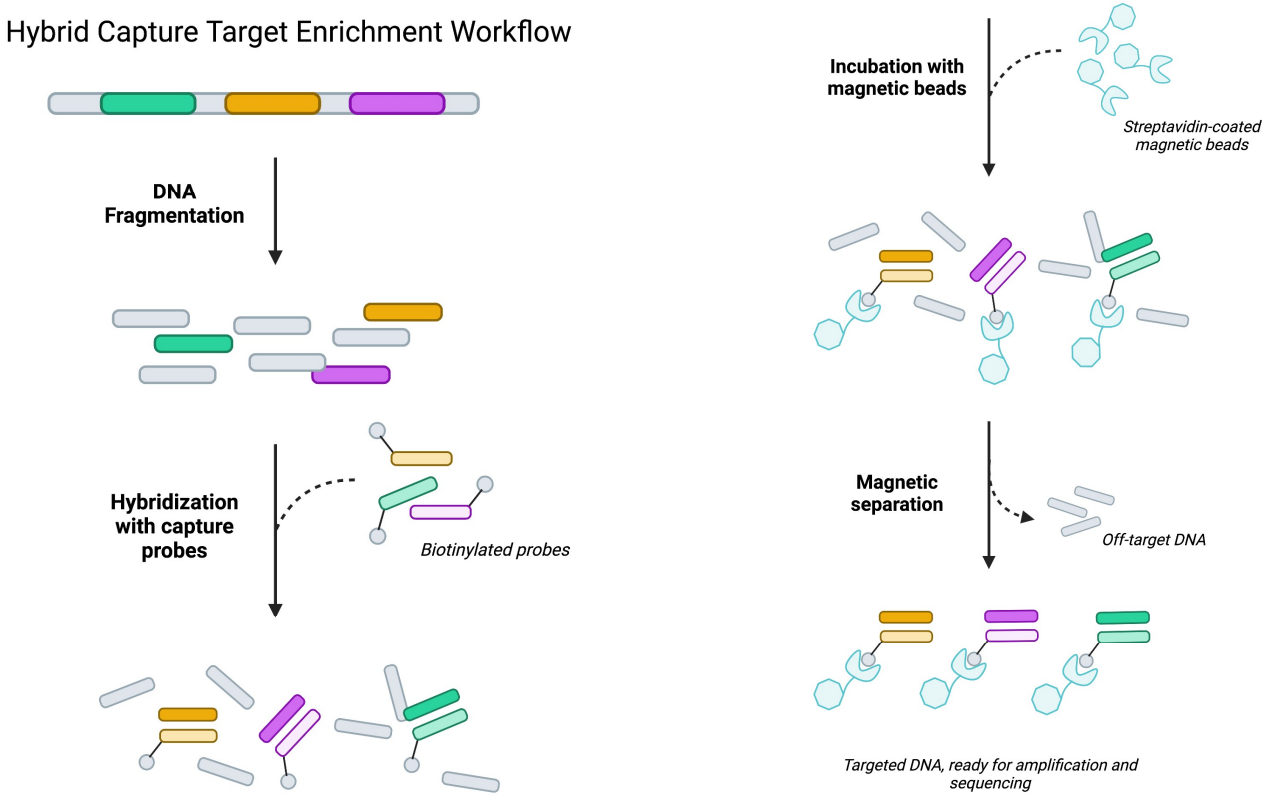
Mock CSF – spiked dilution series with model organisms

RNA virus (feline calicivirus)		DNA virus (cowpox virus)	
PCR Ct value	**Detection by mNGS (No of reads)**	**PCR Ct value**	**Detection by mNGS (No of reads)**
30	Detected (436 reads)	28	Detected (1329 reads)
33	Detected (1,090 reads)	31	Detected (127 reads)
37	Detected (87 reads)	34	Detected (32 reads)
41	Detected (20 reads)	37	Detected (11 reads)

Potential improvements

Decrease turnaround time		Oxford Nanopore Technologies (ONT) Sequencing
Increase sensitivity		Targeted approaches

Hybridisation capture panels



Twist Bioscience

Potential improvements

Decrease turnaround time	→	Oxford Nanopore Technologies (ONT) Sequencing
Increase sensitivity	→	Targeted approaches
Confirm negative results and exclude infection	→	Host transcriptomics

Methods comparison

Untargeted Illumina Sequencing

Separate DNA & RNA Seq

Human genomic DNA depletion through CpG methylation

Ribodepletion

Untargeted ONT Sequencing

Separate DNA & RNA Seq

Human genomic DNA depletion through CpG methylation

No ribodepletion

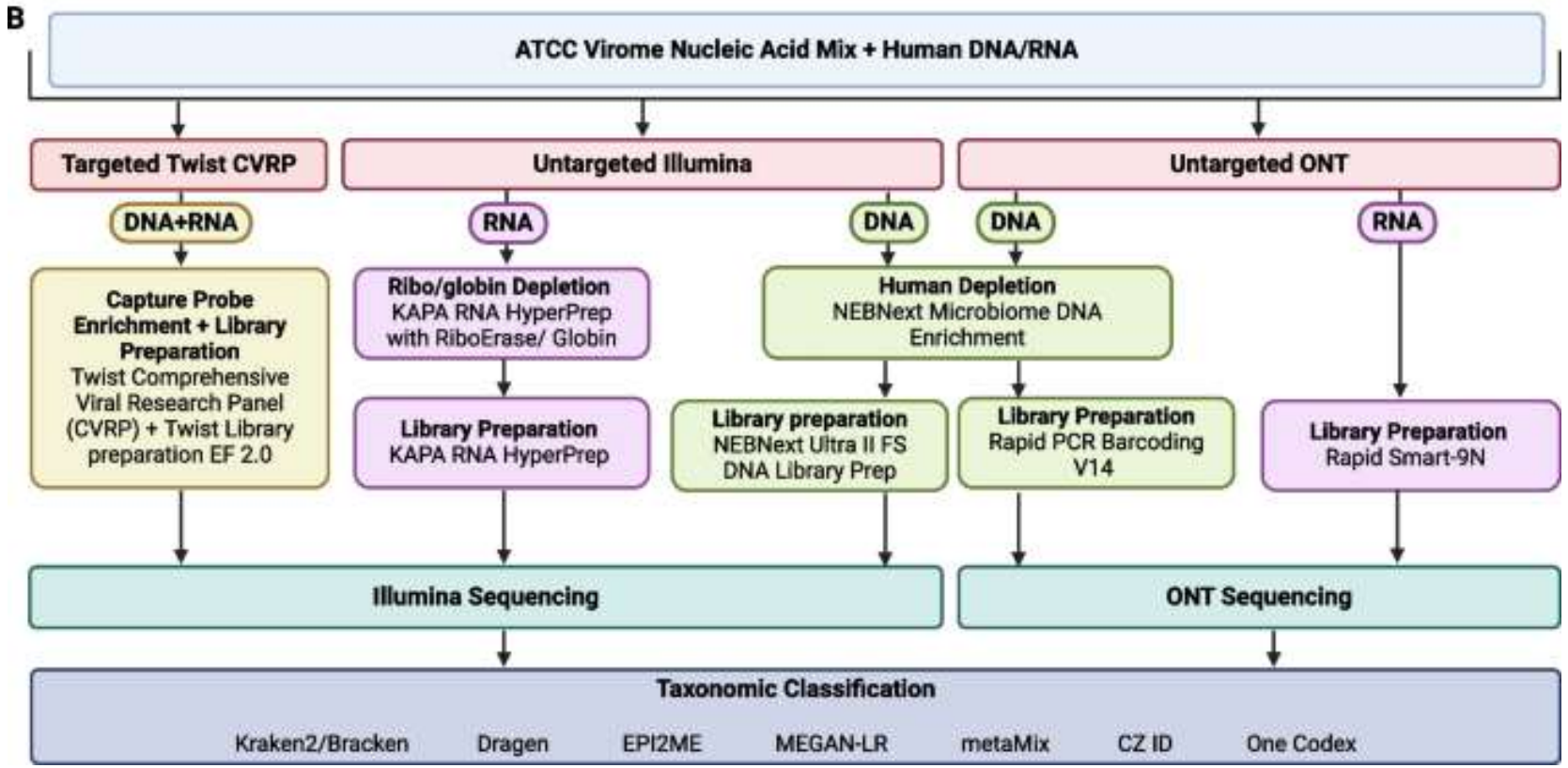
Capture Probe Enrichment + Illumina

Twist Biosciences Viral Research Panel (3153 species)

Combined DNA & RNA Seq

No human genomic DNA depletion or ribodepletion

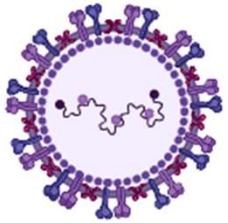
Methods comparison



Mock clinical samples

Human mastadenovirus F (HAdV-F)
DNA
34392

Human betaherpesvirus 5 (HHV5) (Cytomegalovirus, CMV)
DNA
229354



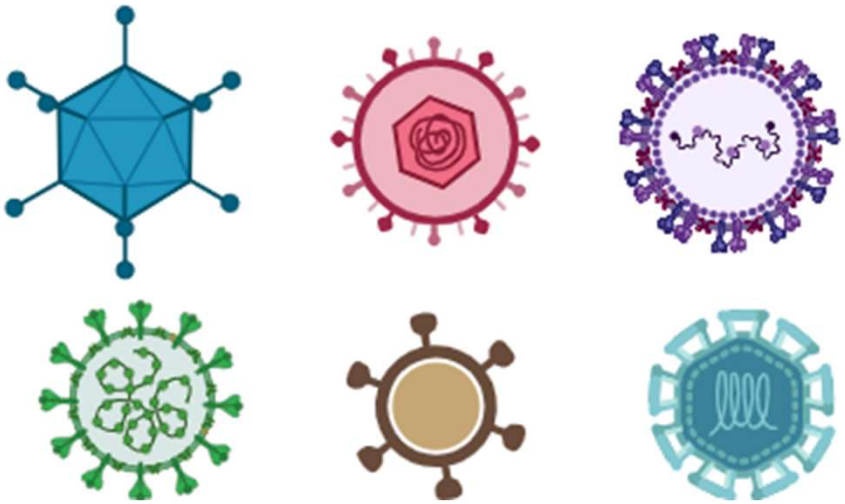
Human orthopneumovirus (Respiratory syncytial virus, RSV)
-ve RNA
15228

Influenza B virus
-ve RNA
18527

Mammalian orthoreovirus
ds RNA
23416

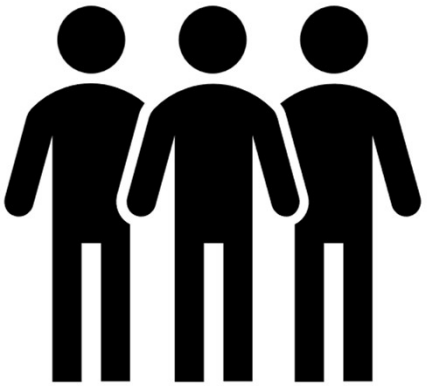
Zika virus
+ve RNA
10952

Mock clinical samples



Viral mock community
(4 dilutions: ~ 6 x 10^1 – 6 x 10^4 genome copies per ml)

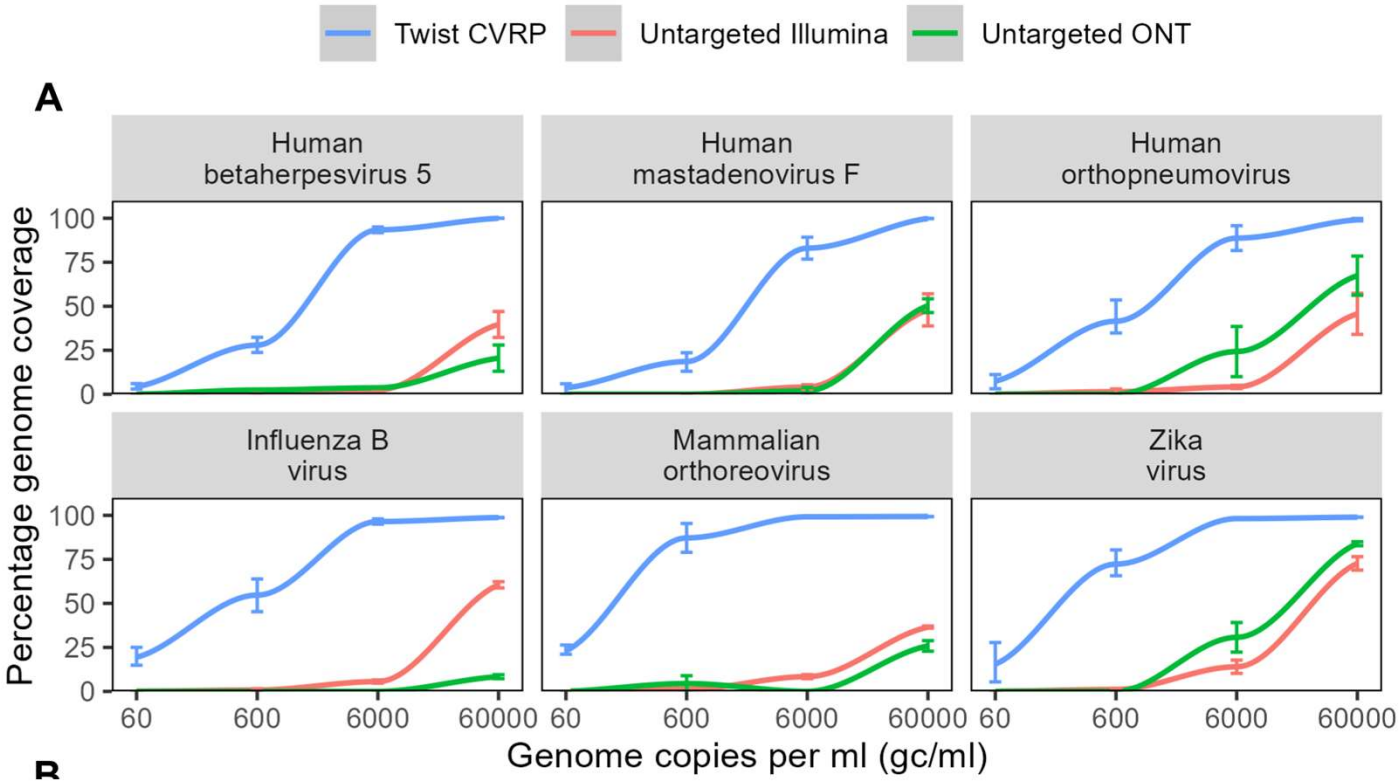
+



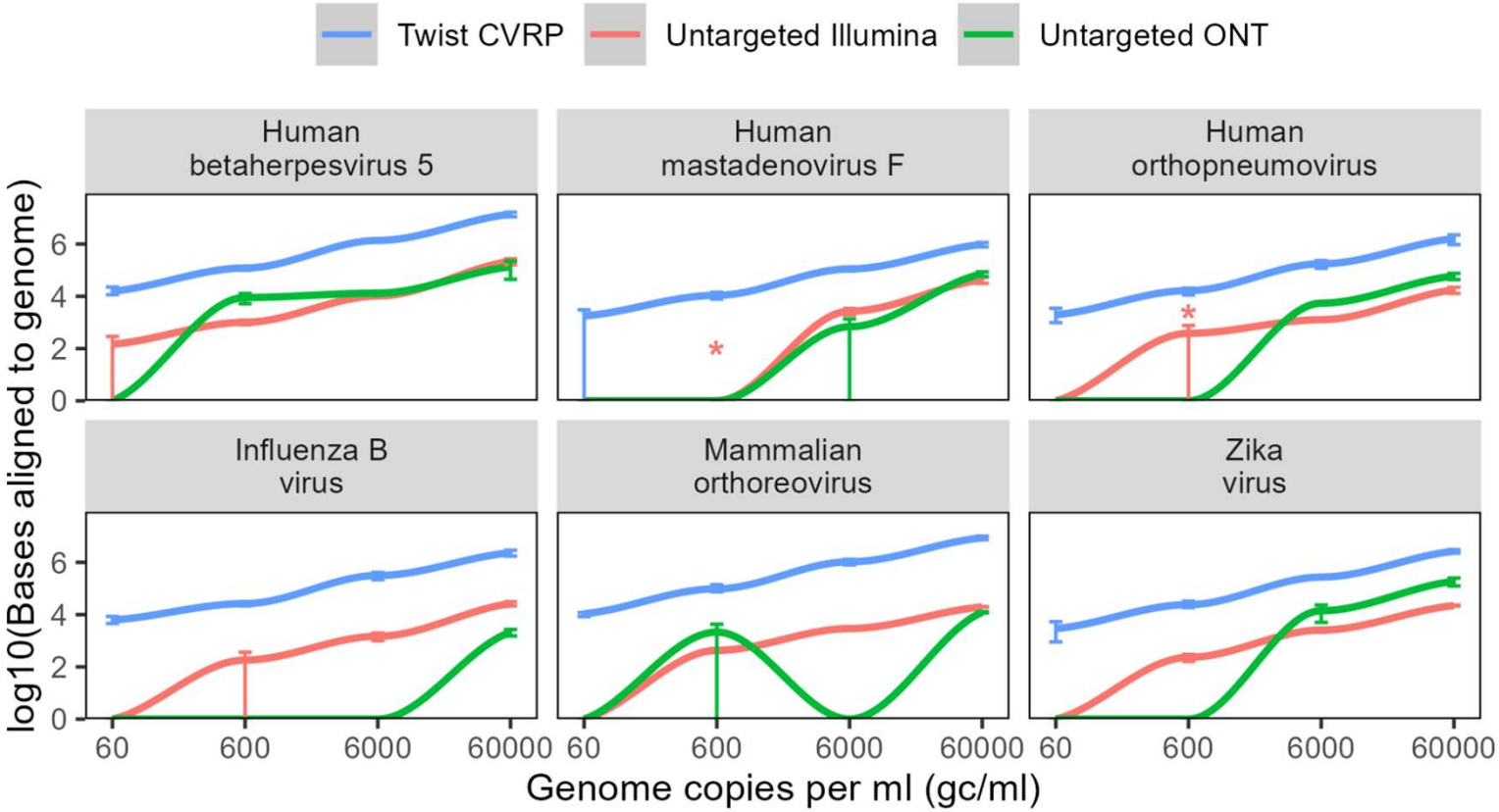
Human DNA + RNA

Negative controls – human DNA and RNA only

Sensitivity



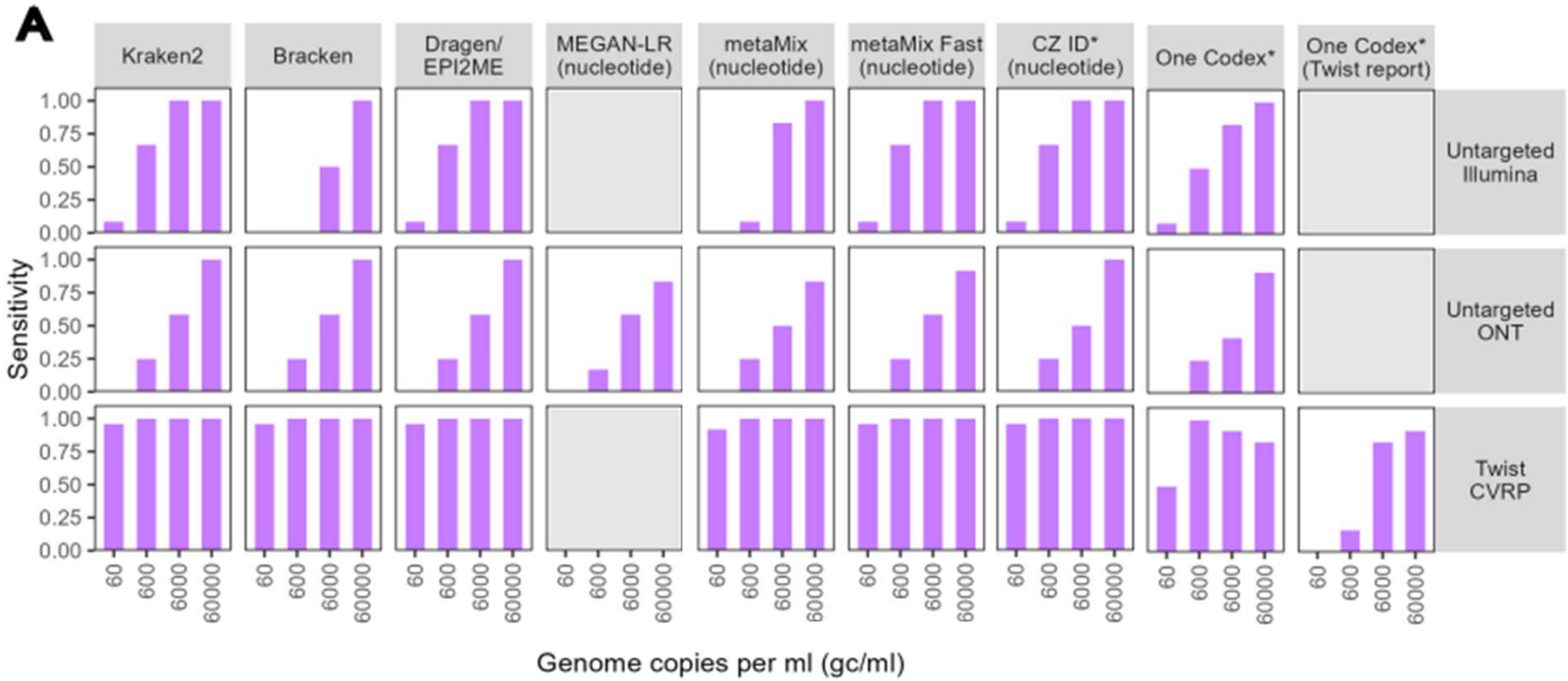
Sensitivity



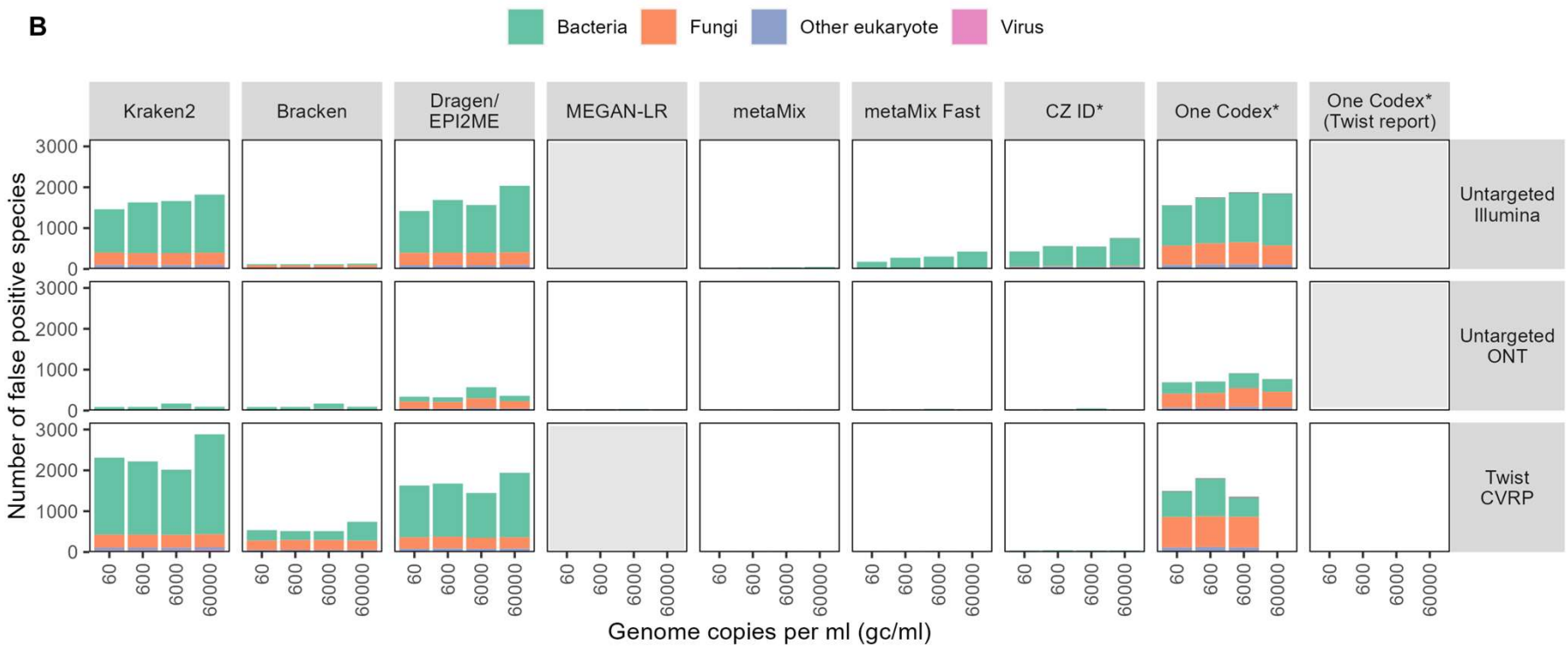
Taxonomic classifiers

Classifier	Method	Reason included	Platform	GUI or CLI	Local or cloud	Database size	Approximate time taken (hours)*
Kraken2 & Bracken	Kmer-based, lowest common ancestor	Very widely used	Illumina & ONT	CLI	Local	124 GB	1-3
DRAGEN Metagenomics (Kraken2)	See Kraken2	Illumina's platform	Illumina & ONT	CLI & GUI	Cloud	124 GB	1.5-3
EPI2ME Labs wf-metagenomics (Kraken2 & Bracken)	See Kraken2	ONT's platform	ONT	CLI & GUI	Local. Cloud in development	124 GB	0.5-2
MEGAN(-LR)	Lowest common ancestor	Good performance in benchmarking study	Illumina & ONT	CLI required for preprocessing. GUI (free), CLI (paid for short reads)	Local	327 GB (Minimap2) 88 GB (DIAMOND)	5-8
metaMix	Bayesian mixture models	Good performance in benchmarking study, used clinically	Illumina & ONT	CLI	Local	148 GB (BLAST) 88 GB (DIAMOND)	5-12+
CZ ID	Alignment and assembly	Free, cloud-based platform	Illumina & ONT	CLI & GUI	Cloud	NA - inbuilt online database	0.5-2
One Codex	Kmer-based	Recommended platform for use with Twist CVRP	Illumina & ONT	CLI & GUI	Cloud	NA – inbuilt online database	~0.5-2

Sensitivity



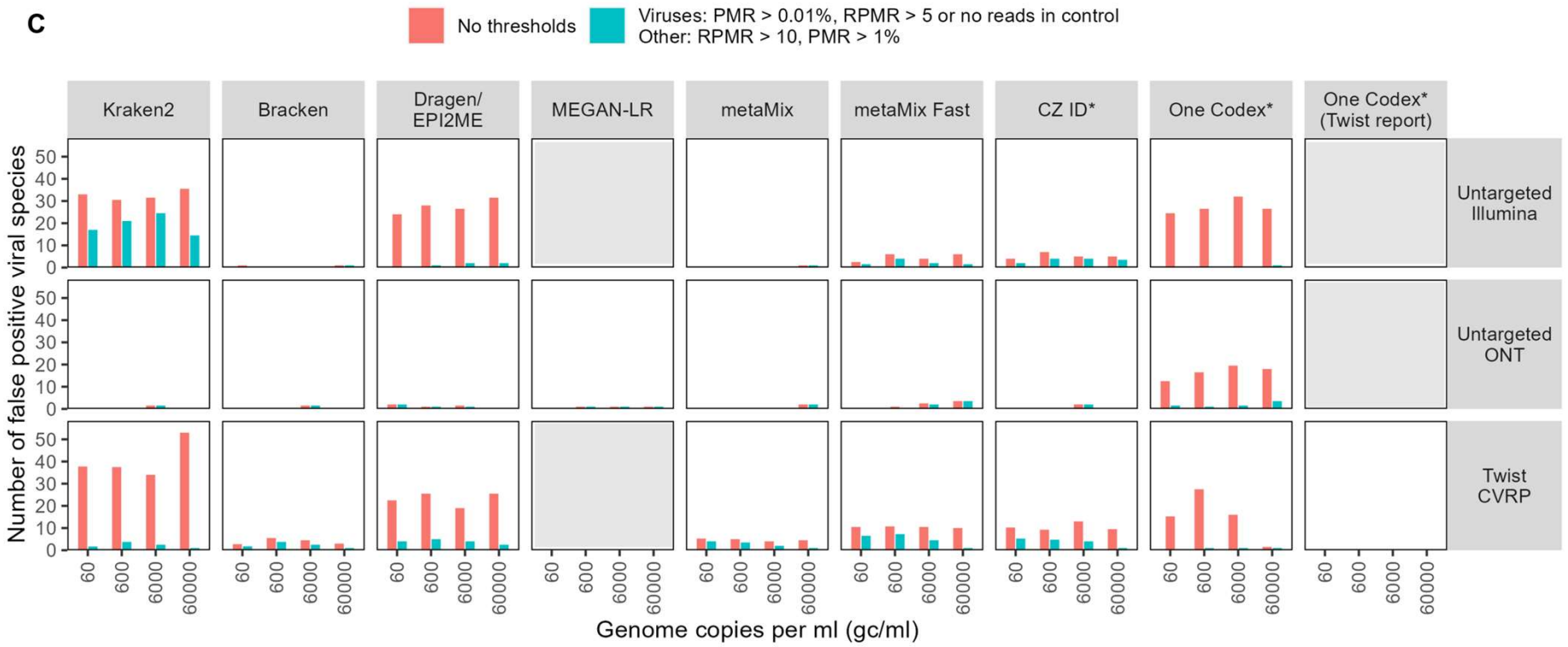
Specificity



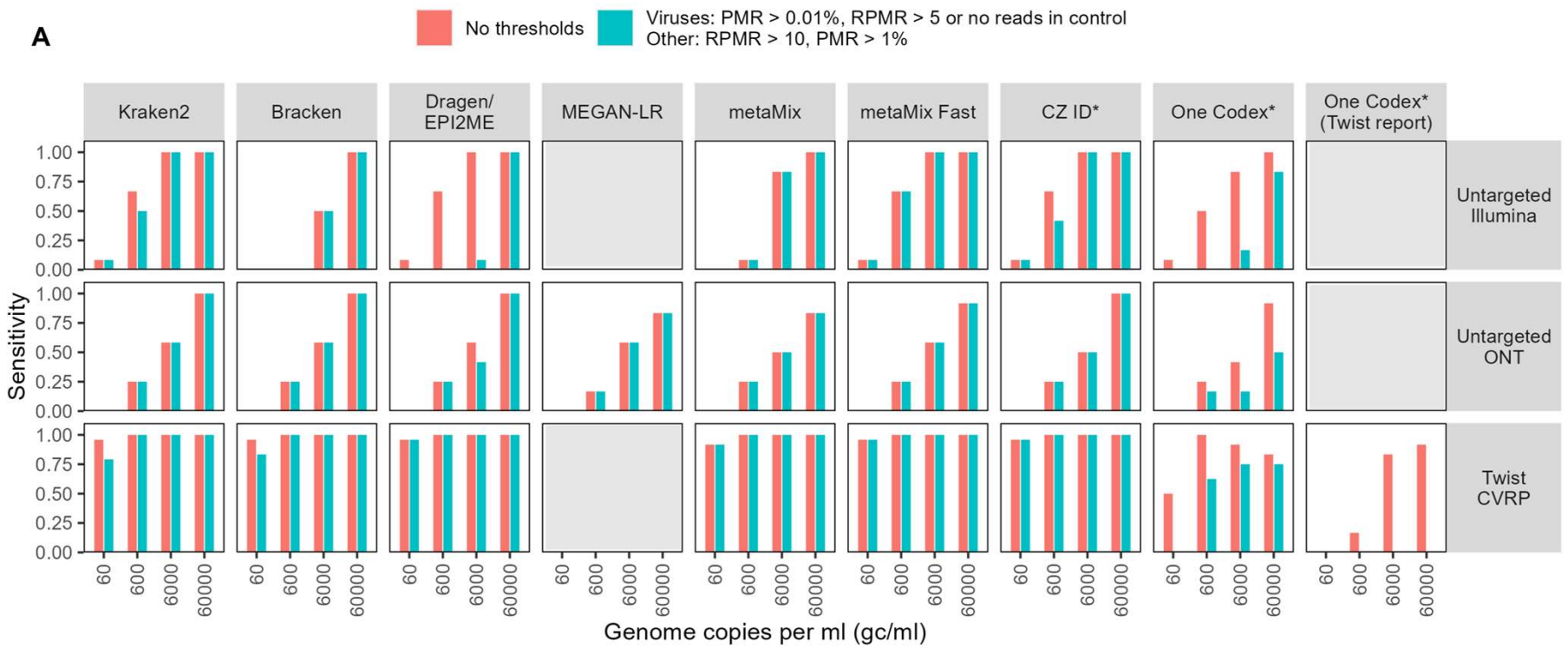
Types of threshold

Raw reads	reads assigned to taxon
Reads per million (rpm)	$\text{reads per million (rpm)} = \frac{\text{reads assigned to taxon} * 10^6}{\text{total raw reads in sample}}$
Reads per million ratio	$\text{rpm ratio} = \frac{\text{sample rpm}}{\text{control rpm}}$
Proportion of microbial reads	$\text{proportion microbial reads} = \frac{\text{reads assigned to taxon}}{\text{total nonhuman, classified reads in sample}}$

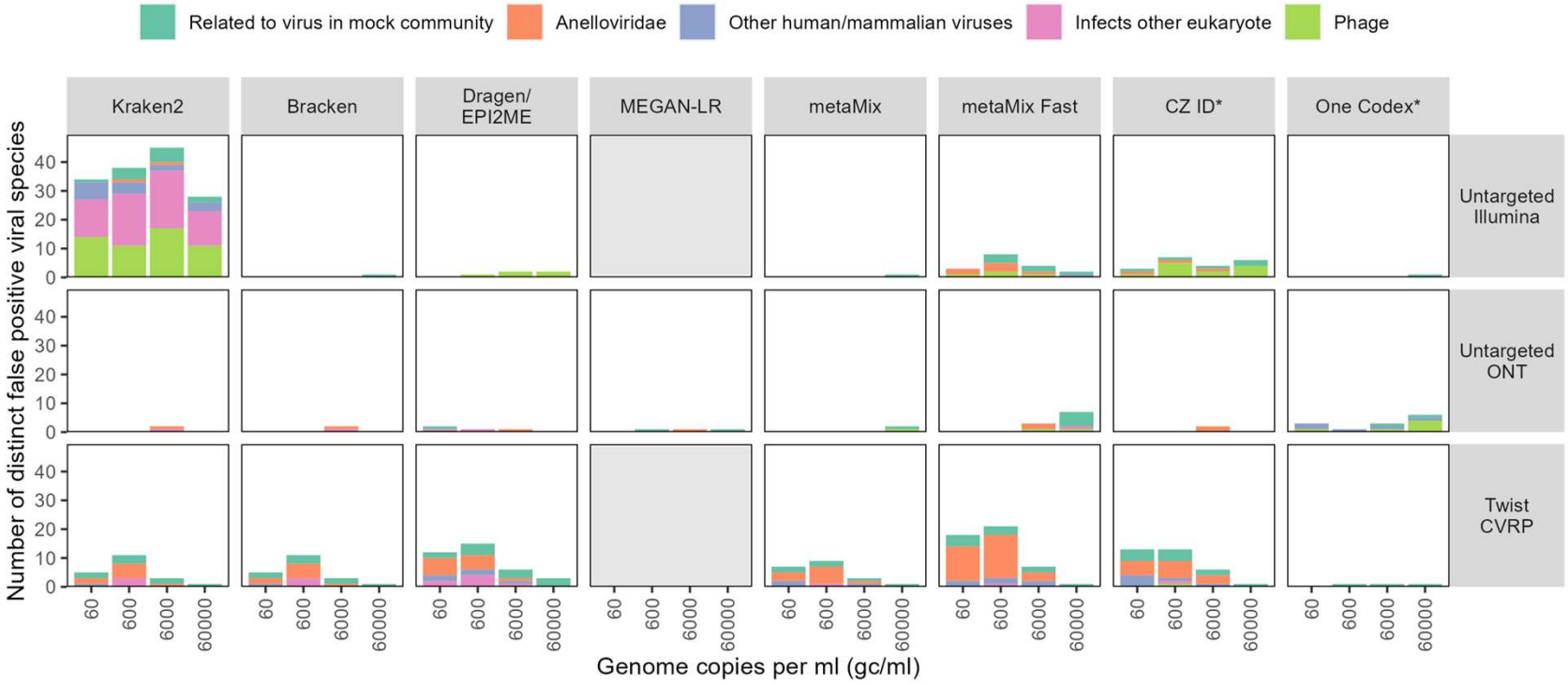
False positive species



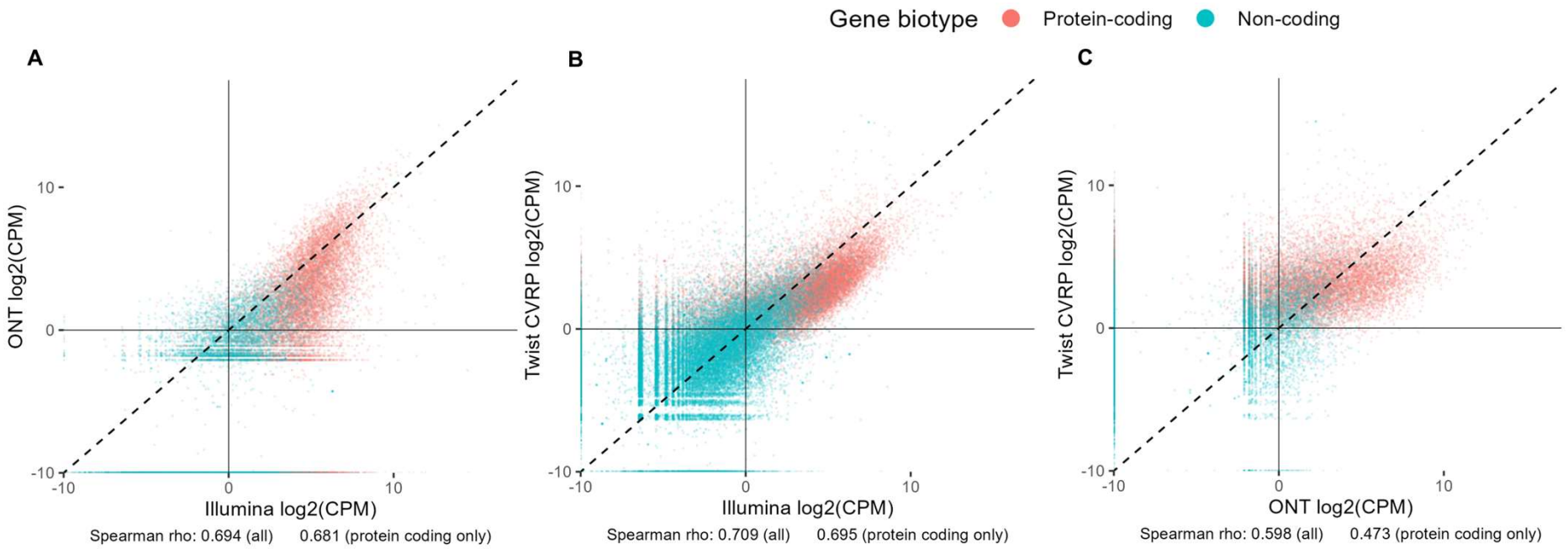
Sensitivity with thresholds



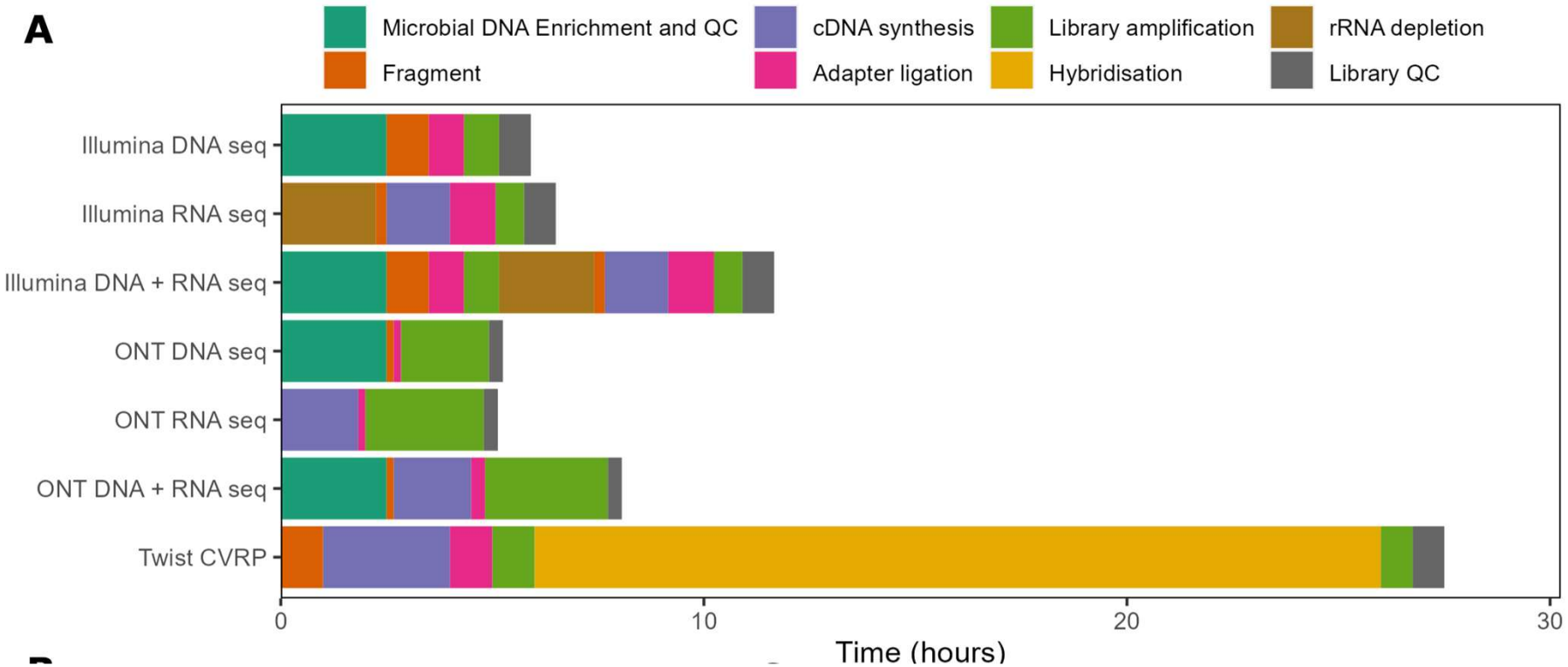
False positive species



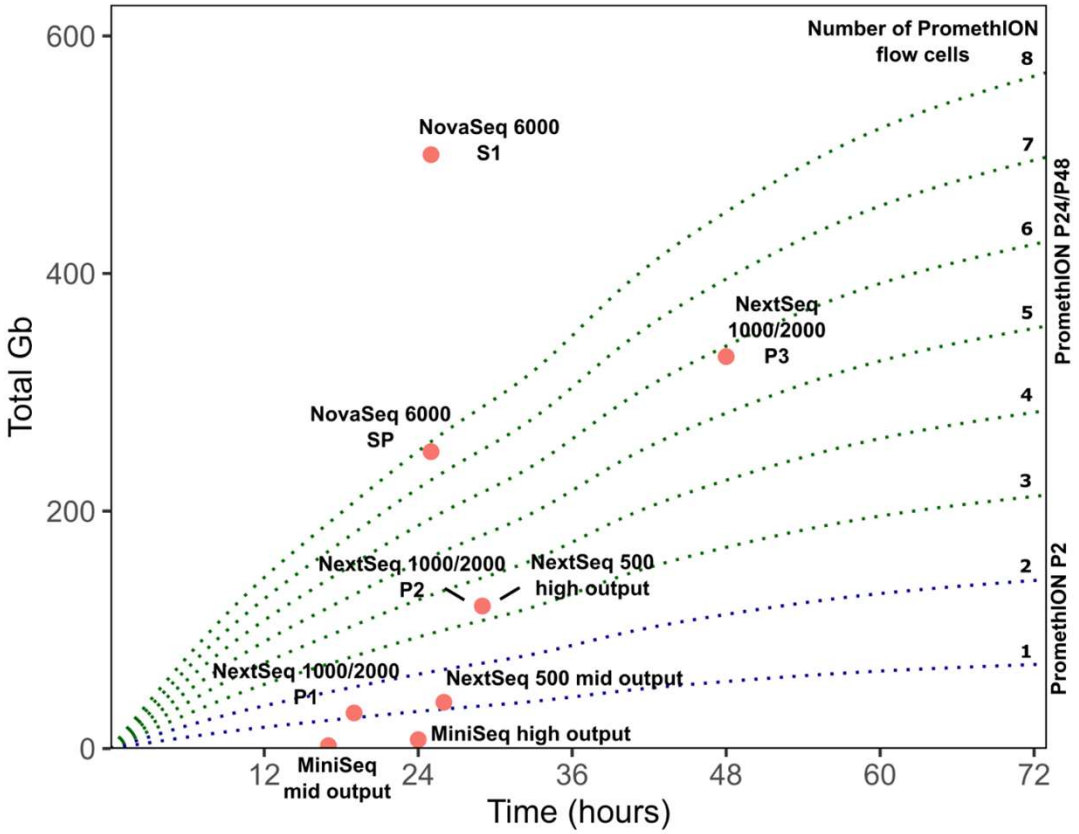
Host transcriptomics



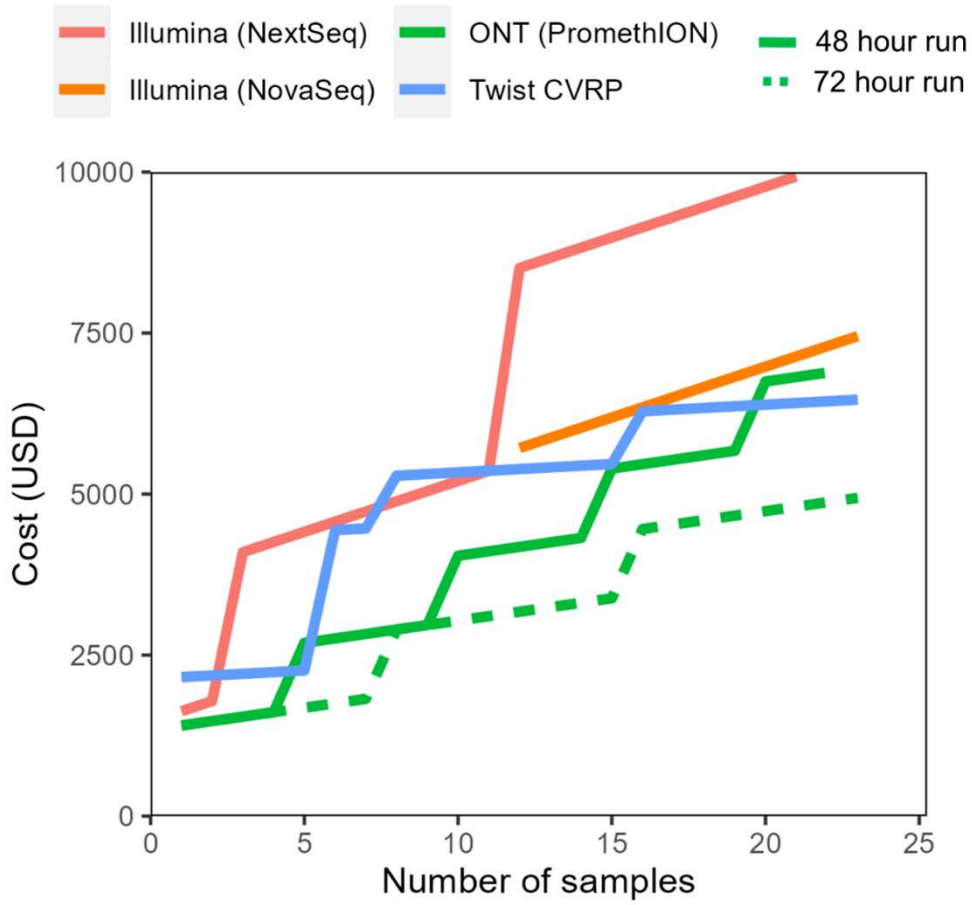
Turnaround times



Turnaround times



Costs


Illumina (NextSeq)
Illumina (NovaSeq)
ONT (PromethION)
Twist CVRP
48 hour run
72 hour run
10000
7500
5000
2500
0
Cost (USD)
0
5
10
15
20
25
Number of samples

Conclusions

Sensitivity: Twist CVRP >> Untargeted Illumina > Untargeted ONT

ONT was the most specific approach

Use of thresholds can standardise results across classifiers

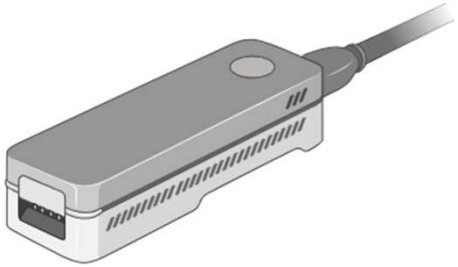
Host analysis is possible with all platforms.

Cost and turnaround times vary with sample number, but ONT is quickest and cheapest for small sample numbers.

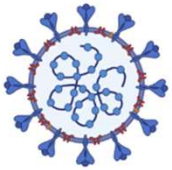
Paper: Buddle & Forrest et al, Genome Medicine, 2024

Future directions

Clinical samples



Improving sensitivity of ONT



Targeted panels for encephalitis

Other uses of blood and tissue metagenomics

Sepsis and bloodstream infection

Fever of unknown origin

Transplanted organs

Respiratory metagenomics

Diagnosis of pneumonia in intensive care units

More complex microbial community

Faster turnaround times required

Surveillance potential

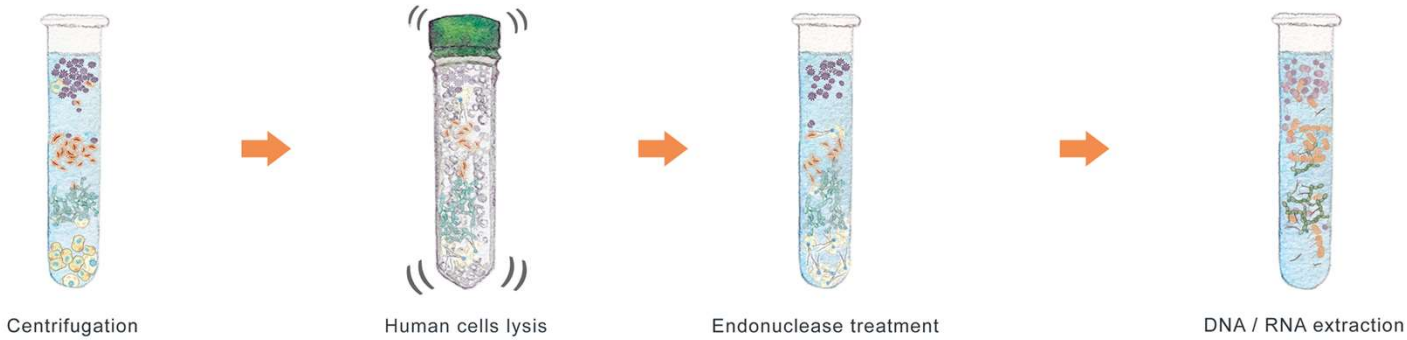
Respiratory metagenomics with ONT

Routine Metagenomics Service for ICU Patients with Respiratory Infection

Themoula Charalampous[1]*, Adela Alcolea-Medina[1,3]*, Luke B. Snell[1,4]*, Christopher Alder[1,4], Mark Tan[1], Tom G. S. Williams[4], Noor Al-Yaakoubi[1], Gul Humayun[1], Christopher I. S. Meadows[2,5], Duncan L. A. Wyncoll[5], Richard Paul[5], Carolyn J. Hemsley[4], Dakshika Jeyaratnam[4], William Newsholme[4], Simon Goldenberg[4], Amita Patel[1,4], Fearghal Tucker[3], Gaia Nebbia[4], Mark Wilks[6], Meera Chand[7], Penelope R. Cliff[3], Rahul Batra[1,4], Justin O'Grady[8], Nicholas A. Barrett[5], and Jonathan D. Edgeworth[1,4‡]

[1]Centre for Clinical Infection and Diagnostics Research, Department of Infectious Diseases, School of Immunology and Microbial Sciences and [2]Faculty of Life Sciences and Medicine, King's College London, London, United Kingdom; [3]Infection Sciences, Synnovis, London, United Kingdom; [4]Department of Infectious Diseases and [5]Critical Care Directorate, Guy's and St Thomas' NHS Foundation Trust, London, England; [6]London School of Medicine and Dentistry, Queen Mary University, London, United Kingdom; [7]UK Health Security Agency, London, United Kingdom; and [8]Oxford Nanopore Technologies, Oxford, United Kingdom.

GSTT Respiratory Metagenomics Protocol



Alcolea-Medina *et al*, Communications Medicine, 2024

Surveillance

May require distinct bioinformatics pipelines and databases

Requires data/results sharing

Key challenges (in addition to general challenges of metagenomics)

- Sharing sensitive data, possibly internationally
- Validating pipelines for unknown organisms

Acknowledgements

UCL

Oscar Torres

Naomi Akinsuyi

Cristina Venturini

Leticia Scalioni

Ines Ringue

Sofia Morfopoulou

Judy Breuer

UCL Genomics

Leysa Forrest

Luz Marina Martin Bernal

Tony Brooks

Sunando Roy

Sian Goldsworthy

Sergi Castellano

Rachel Williams

Great Ormond Street

Julianne Brown

Nathaniel Storey

Angelika Kopec

Laura Atkinson

Charles Miller

Robbie Hammond

BTRU-GEMS

Tanya Golubchik

Heli Harvala

Peter Simmonds

Funders

NIHR | Blood and Transplant Research Unit in Genomics to Enhance Microbiology Screening at University of Oxford

NIHR | Great Ormond Street Hospital Biomedical Research Centre

Some figures produced using Biorender.com