Development of long-amplicon Nanopore sequencing for metabarcoding of Norovirus genogroups I and II in wastewater

Richard Hill| Genomics Facility Manager| richard.hill@cefas.gov.uk
George Scott, Mary Buckley, Frederico Batista, David Ryder, David Walker
Centre for Environment, Fisheries and Aquaculture Science, Weymouth, United Kingdom

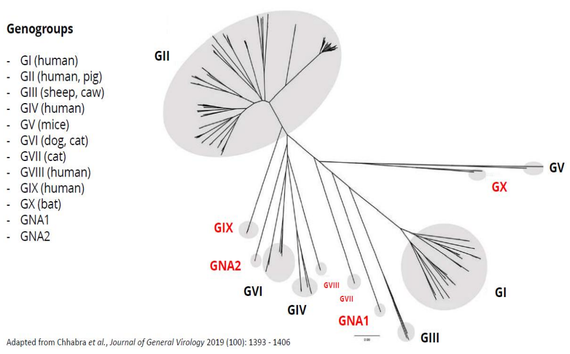
We developed a Nanopore-based method for typing norovirus genogroups I (GI) and II (GII) by sequencing the RNA-dependent RNA polymerase *RdRp* and major structural protein *VP1* genes present in RNA extracted from wastewater samples and amplified by long-range PCR. Reverse transcription, PCR and library preparation were optimised for wastewater RNA and a consensus-based bioinformatics pipeline used for genogroup typing. The resulting assay was used to analyse >3,000 wastewater RNA samples extracted across England to pilot a more cost-effective, non-invasive, national surveillance system for norovirus.

INTRODUCTION

Wastewater-based epidemiology (WBE) gained widespread use as a tool for supporting clinical disease surveillance during the COVID-19 pandemic. In PATH-SAFE WS2a we wanted to build on existing networks and infrastructure and utilise the latest capabilities in high-throughput long-read Nanopore sequencing as part of a 'One sample, many analyses' approach to pathogen surveillance. Specifically, application of WBE could improve the resolution of norovirus identification and improve traceability through the population into the environment.

Noroviruses (NoV) are non-enveloped viruses of the Caliciviridae family, with three of ten genogroups (GI, GII and GIV) causing gastroenteritis in humans. NoV genomes comprise of single-stranded RNA of ~7.5 kb in length. However sequencing NoV genomes is problematic due to the many genogroups, high nucleotide diversity and recombination, making primer design and data interpretation difficult. Furthermore, the chemical composition and low viral load of wastewater makes it a challenging sample matrix.

We assessed influence of sample matrix and reagents on data quality to optimise an effective semi-nested long-amplicon PCR assay for high-throughput Nanopore sequencing and typing of Norovirus GI & GII genogroups.



Adapted from Chhabra *et al.*, *Journal of General Virology* 2019 (100): 1393 - 1406

OBJECTIVES

We used wastewater RNA extracts pooled based on geographic region to optimise:

- **Inhibitor removal**
Mag-Bind® TotalPure NGS (*Omega Bio-Tek*) bead method of Child *et al.* (2023)
- **Reverse transcription**
LunaScript® (*New England Biolabs*) vs SuperScript IV™ (*Invitrogen*)
- **Polymerase enzyme & primers**
Compared performance of commercial polymerase mastermixes in semi-nested PCRs in their sensitivity, inhibitor susceptibility & non-specific amplification.

New England Biolabs: LongAmp®, Q5U®, UltraII Q5®

Invitrogen: PhusionII™, Platinum SuperFi™ and Platinum™ *Taq*

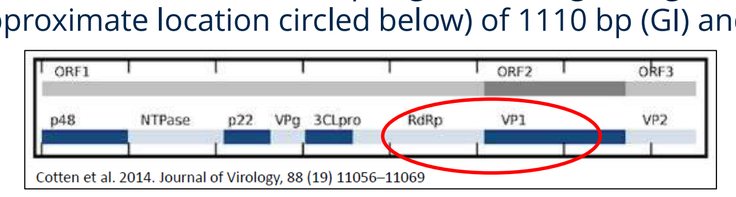
METHODS

Inhibitor removal (IR) and reverse transcription (RT)

- Samples and water control spiked with External Control RNA (~4000 gc/µL) to measure inhibition effect on performance of bead purification and RT kit
- TaqMan assay + inhibitor sensitive mastermix used for GI amplicon (RNA UltraSense™ One-Step Quantitative RT-PCR System + QuantStudio™ 3)

PCR

- Semi-nested PCR of *VP1* and *RdRp* regions in single long amplicon of interest (AOI, approximate location circled below) of 1110 bp (GI) and 971 bp (GII)



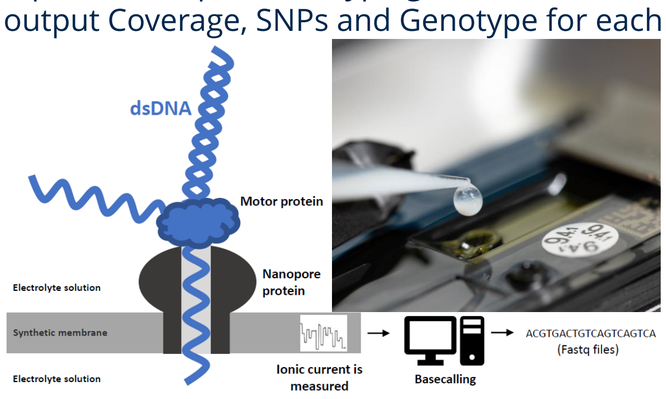
Cotten et al. 2014. Journal of Virology, 88 (19) 11056–11069

Sequencing

- Nanopore (ONT) sequencing using Native Barcoding Kit V14 + GridION
- Number of aligned reads used as a metric for RT performance

Bioinformatics

Guppy, Cut Adapt and NGSpeciesID, typing with Norovirus Typing Tool Version 2.0 (RIVM), to output Coverage, SNPs and Genotype for each sample



RESULTS

Inhibitor removal (IR) and reverse transcription (RT)

- RNA purification with Mag-Bind® beads was extremely effective, reducing average inhibition from **90.6 %** to **13.2 %** in the GI qPCR
- Resulted in **6.4-fold** increase in NoV quantification
- The long-range PCR was successful when IR combined with *NEB* LunaScript® RT

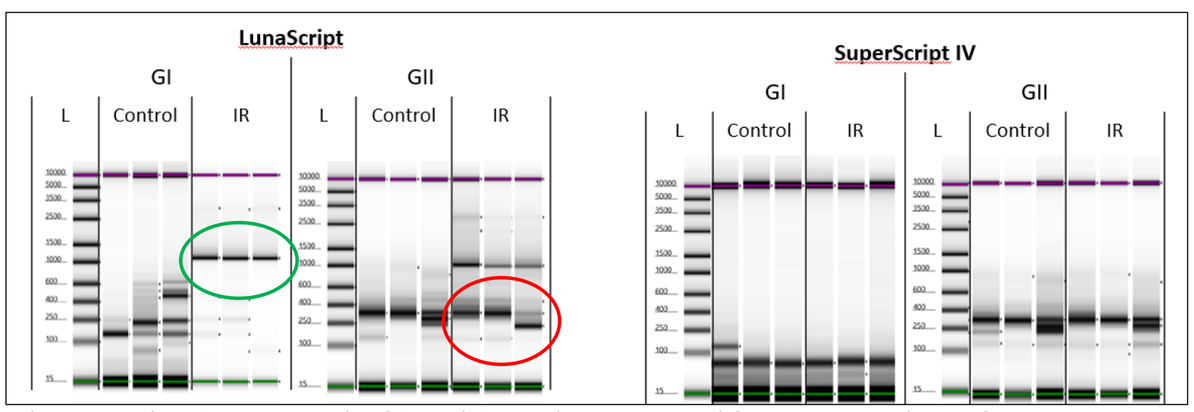


Figure 1. Agilent Tapestation gels of GI and GII amplicons generated from semi-nested PCR after RNA treatment with *MagBind* beads (IR) in conjunction with reverse transcription using *Lunascript®* or *Superscript IV™* reagents. The red and green circled bands highlight respectively the amplicon of interest at ~1 kB and non-specific amplification.

PCR

Platinum™ *Taq* gave best overall performance in long-range PCR in the metrics outlined in our objectives. The PCR conditions were further optimised, increasing annealing temperatures, and resulting in (data from Scott *et al.* (in prep)):

- Significantly increased percentage of AOI in the PCR library, from **77.6 %** to **89.1 %** for GI and **33.4 %** to **61.5 %** for GII
- Significantly increased (**57.7%**) taxa richness for GII
- Using data from 130 wastewater samples we applied an optimal pooling ratio of 85.3 nM GI + 114.7 nM GII amplicons for ONT library preparation and sequencing, to reduce dilution of GII reads caused by non-specific amplification

APPLICATIONS

A subset of 3,232 RNA samples extracted from 152 sewage treatment works across England were subsequently analysed using our optimised long-range PCR protocol and ONT sequencing, resulting in an overall **85 %** success rate across the samples. The sequence data this study will ultimately provide will greatly improve our understanding of NoV diversity and prevalence within the population, whilst also demonstrating our ability to track outbreaks and gain valuable insights into the genetic diversity of this important pathogen.

ACKNOWLEDGEMENTS Cefas thanks the UKHSA and Environment Agency for their collaboration in collecting and extracting the wastewater RNA and the support and expertise of Cefas colleagues in Weymouth and Exeter laboratories for the sample processing and analysis.

REFERENCES
Cotten *et al.* (2014) 'Deep sequencing of norovirus genomes defines evolutionary patterns in an urban tropical setting'. J Virol, Oct;88(19), pp. 11056-69.
Child, H. T. *et al.* (2023) 'Optimised protocol for monitoring SARS-CoV- 2 in wastewater using reverse complement PCR-based whole-genome sequencing', PLoS ONE, 18(4 APRIL), pp.1–19.
Chhabra, P. *et al.* (2019) 'Updated classification of norovirus genogroups and genotypes', Journal of General Virology, 100(10), pp. 1393–1406.
Scott *et al.* (in prep) 'High-Throughput Nanopore Sequencing for Dual-Typing of Norovirus Genogroups I and II in Wastewater'.

Cefas