Evaluation of long-read vs. short-read platforms for the genomic epidemiology study of SARS-CoV-2 outbreaks

Nanopore Community Meeting 2022

Panisello Yagüe D[1], Pelegrin AC[1], Soler L[1], Bordoy AE[1], Martínez-Puchol S[1,8], Clara G[1], González-Gómez S[1], Paris de León A[1], Noguera M[3], Armengol MP[4], Francino O[5], Blanco A[1], Cardona PJ[1,7], Blanco I[6], Saludes V[1,2], Martró E[1,2,*]

1. Microbiology Department, Laboratori Clínic Metropolitana Nord (LCMN), Germans Trias i Pujol University Hospital, Germans Trias i Pujol Research Institute (IGTP), Badalona, Spain
2. CIBER in Epidemiology and Public Health (CIBERESP), Instituto de Salud Carlos III, Madrid, Spain
3. IrsiCaixa, Badalona, Spain
4. Translational Genomics Unit, IGTP, Badalona, Spain
5. Facultat de Veterinària, Universitat Autònoma de Barcelona (UAB), Bellaterra, Spain
6. Clinical Genetics Department, LCMN, Hospital Universitari Germans Trias i Pujol, Badalona, Spain
7. CIBER in Respiratory Diseases (CIBERES), Instituto de Salud Carlos III, Madrid, Spain
8. Vicerectorat de recerca, Universitat de Barcelona (UB), Barcelona, Spain

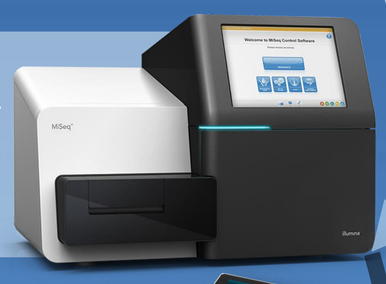
Background and Aims:

The use of SARS-CoV-2 whole-genome sequencing (WGS) as a high-resolution epidemiological marker eases the interpretation of transmission chains, so that tailored control interventions can be promptly implemented. Given the rapid dissemination of SARS-CoV-2, a short time to result is needed. In cases with a monophyletic origin and an epidemiological link, reliably determining the number of single nucleotide polymorphisms (SNPs) between pairs of sequences is important to identify recent transmission events. **We aimed to evaluate the applicability of Oxford Nanopore MinION Mk1C for SARS-CoV-2 WGS in outbreak characterization, by comparing its technical and practical differences with the Illumina MiSeq platform.**

Methods:

38 nasopharyngeal swab samples
Ct (RdRp gene) <30

Four outbreaks: A(*n*=8), B(*n*=11), C(*n*=13), D(*n*=6)



SuperScript IV RT → ARTIC v4 tiled PCR (IDT) → Illumina DNA prep kit → MiSeq (v2, 2x150 cycles, 27 h)

Midnight RT-PCR Expansion kit (v6 +28_Right alternative for BA.1) → Rapid Barcoding kit → Mk1C (R9.4 flow cell). Time was adjusted to the run size (12-26 samples/run) and flow cell use status, from 2h to overnight sequencing.

First, **Nf-core/viralrecon pipeline (v2.4.1)** generates Fasta consensus sequences and analyzes them with Nextclade (v1.10.2) and Pango (v3.1.20).

Second, **homemade Nexflow pipeline**:

1. Downloaded a subset of samples from the same (sub)lineages from GISAID and then queried our sequences against them with BLAST (v2.12.0).
2. Closely related sequences were aligned with ours, and the Wuhan reference, using MAFFT (v7.490).
3. A phylogenetic Maximum Likelihood tree was inferred with IQTREE (v2.2.0) and viewed with iTOL (v6).
4. A pairwise SNP distance matrix for all samples was created with SNPS_DIST (v0.8.2).

Results:

The SARS-CoV-2 genomic sequence was obtained by both sequencing platforms for 97.4% (37/38) of samples, but two of them had suboptimal genome coverage (<90%).

Comparative results for good quality samples (*n*=35) are shown below:

	Illumina MiSeq	Nanopore MinION Mk1C
Mean genome coverage at ≥10×	99.2% (IQR: 2%)	98.9% (IQR: 0%)
Mean depth of coverage	691× (IQR: 216×)	407× (IQR: 623×)

- **Pango lineage** assignments were **100% concordant** between pairs of sequences, including several Delta and Omicron lineages (AY.4, AY.43, AY.121, AY.122, BA.1, BA.1.1, BA.1.1.1, BA.1.7).
- No **differences in identified SNPs** were observed between pairs of sequences (**100% concordance**).

The Nanopore method had an 1.5 to 5-fold shorter lab turn-around time (TAT; from extracted RNA to completed sequencing), and an approximate 3-fold cheaper sequencing process.

Illumina
Sequencing time: 27h
Total lab TAT: 35h
Cost: $$$

Nanopore
Sequencing time: Flexible (2-19h)
Total lab TAT: 7-24h
Cost: $

Conclusions:

These results support the use of MinION Mk1C for a rapid and reliable response to nosocomial SARS-CoV-2 outbreaks.

Contact:

emartro@igtp.cat
dpanisello@igtp.cat

This study was carried out with the support from "Fundació La Marató de TV3" (Project Ref. 342/C/2021)