Full-length Nanopore 16S amplicon analysis in a standardised pipeline

Fredrik Dyrkell[1], Dimitrios Arnellos[1], Oskar Aspelin[1]

[1]1928 Diagnostics AB – Gothenburg (Sweden)

Background

Advances in sequencing technologies have opened up new possibilities for fast, sensitive and cost-efficient applications such as 16S amplicon sequencing. As a result, accurate, standardised and automated workflows are needed for analysing the generated data.

The 1928 platform (1928 Diagnostics) is a cloud-based service for analysing next-generation and third-generation sequencing data, including amplicon sequences for fast pathogen detection in polymicrobial samples.
The platform supports all major sequencing platforms, including a full-length 16S amplicon pipeline for Oxford Nanopore data.

The aim of this project was to comparatively analyse a full-length Nanopore 16S dataset and compare the result with short-read Illumina using the 1928 platform.

Method

One ten-species mock community sample (V1–V9 Nanopore) and six human faecal metagenomic samples (V3–V4 Illumina and V1–V9 Nanopore)[1] were downloaded from ENA (Bioproject PRJDB9744) and uploaded to the 1928 platform.

The sequences were trimmed and filtered on amplicon length (1200–1700 bp for V1–V9).

Reads were compared against each other to determine strain-level representatives which were subsequently mapped against the SILVA (v138.1) reference database for taxonomic assignment.

Results

The relative abundance found in the mock community sample is shown in Figure 1. All expected genera from the mock were found in the Nanopore analysis. Genus level abundances from the faecal metagenomes are shown in Figure 2. Species level alpha diversity (Shannon) (Figure 3) shows a higher diversity in all Nanopore samples compared to their Illumina counterparts.

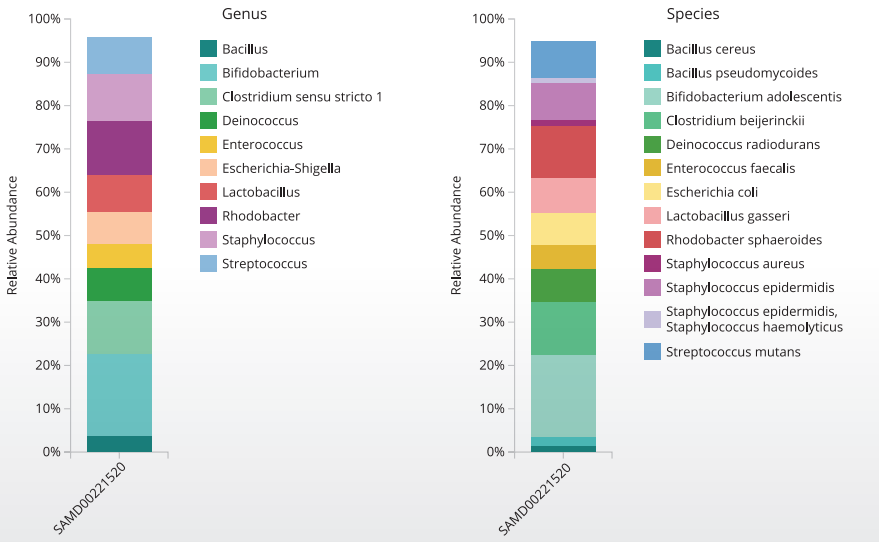

Figure 1. Analysis of the mock community sample – genus and species level, taxa below 1% are not shown.

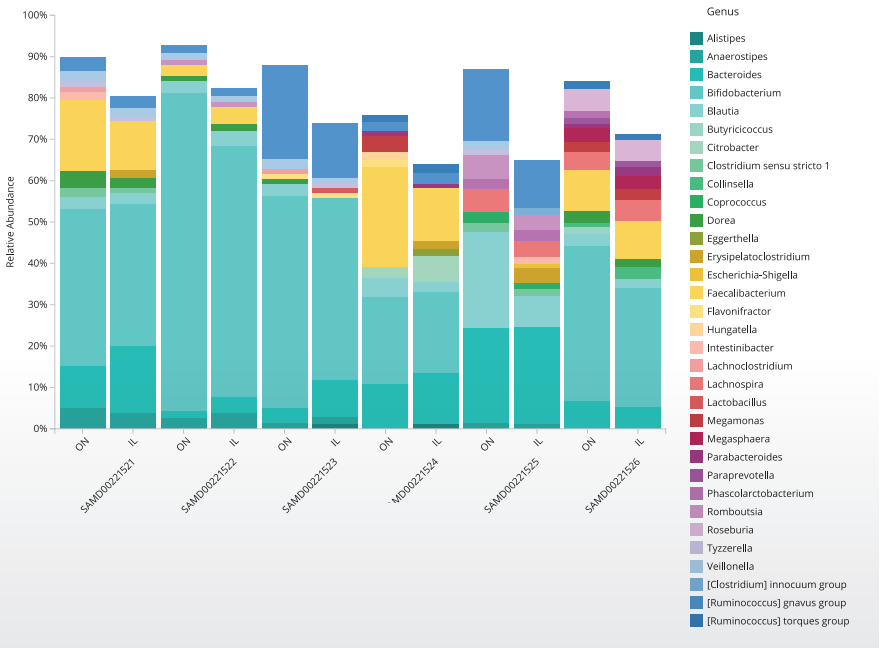

Figure 2. Genus level abundances for Nanopore and Illumina data. Taxa below 1% are not shown.

cont. Results

SAMPLE ACCESSION	RUN ACCESSION	SEQUENCING PLATFORM	SHANNON INDEX	NUMBER OF SEQUENCES	AMPLICON SEQUENCE AVERAGE LENGHT
SAMD00221520	DRR225043-46	Nanopore	5.0	63755	1564
SAMD00221521	DRR225050 DRR225048	Illumina Nanopore	3.4 4.7	132484 104895	401 1589
SAMD00221522	DRR225053 DRR225051	Illumina Nanopore	3.0 3.7	137648 84065	401 1576
SAMD00221523	DRR225056 DRR225054	Illumina Nanopore	2.5 3.6	264114 76968	402 1582
SAMD00221524	DRR225059 DRR225057	Illumina Nanopore	4.6 5.2	207064 114060	401 1593
SAMD00221525	DRR225062 DRR225060	Illumina Nanopore	4.2 5.4	144272 85912	401 1597
SAMD00221526	DRR225065 DRR225063	Illumina Nanopore	3.9 5.0	104364 108938	402 1589

Figure 3. Metrics of analysed samples. Shannon index denotes species level alpha diversity.

Conclusions

Full-length 16S Nanopore analysis shows high concordance to Illumina sequences and has the potential of yielding higher strain level resolution due to the longer region sequenced. The 1928 platform facilitates easy comparison of amplicon data between platforms in a standardised way.

References

1. Matsuo Y et al. Full-length 16S rRNA gene amplicon analysis of human gut microbiota using MinION™ nanopore sequencing confers species-level resolution. BMC Microbiology. Springer Science and Business Media LLC; 2021. doi: 10.1186/s12866-021-02094-5

Conflict of interest statement

F.D., D.A. and O.A. are employees of 1928 Diagnostics AB.

Contact

fredrik.dyrkell@1928diagnostics.com