SANDIA BIOSCIENCE

SAND2017-11968C

LIBRARY PREPARATION AND ANALYSIS METHODS FOR THE OXFORD MINION NANOPORE SEQUENCER FOR PATHOGEN IDENTIFICATION

Harikrishnan Jayamohan†[1], Anupama Sinha †[2], Raga Krishnakumar †[2], Harrison S. Edwards †[1], Tariq A. Younis[1], Jacob D. Trevithick[1], Kamlesh D. Patel[1], and Michael S. Bartsch[1*]

[1] Advanced Systems Engineering and Deployment, Sandia National Laboratories, Livermore, CA, USA

[2]Department of Systems Biology, Sandia National Laboratories, Livermore, CA, USA

Introduction

The Oxford Nanopore Technologies (ONT) MinION is a long-read, portable DNA sequencer that detects changes in ionic current as single-stranded DNA translocates through a nanopore (Figure 1). The small size and real-time analysis capabilities of the MinION make it particularly well suited for point-of-care and field-based applications like outbreak investigations and environmental metagenomics. Here, we describe our ASPIRE platform for the MinION sequencer that automates and streamlines the preparation of MinION-ready long- fragment DNA libraries for sequencing. We also highlight our work using CRISPR/Cas9 excision of the 16S-23S ribosomal DNA (rDNA) locus for real-time strain-level bacterial identification and analysis with a "Read Until" selective sequencing approach called RUBRIC.

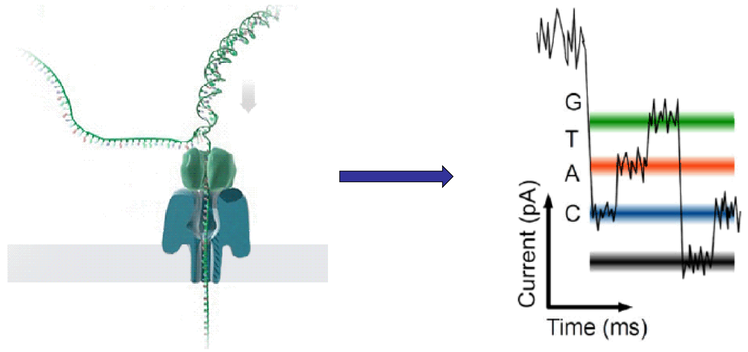
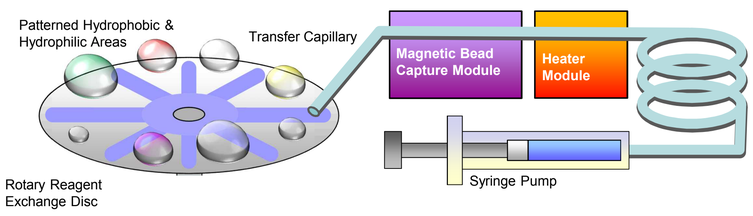
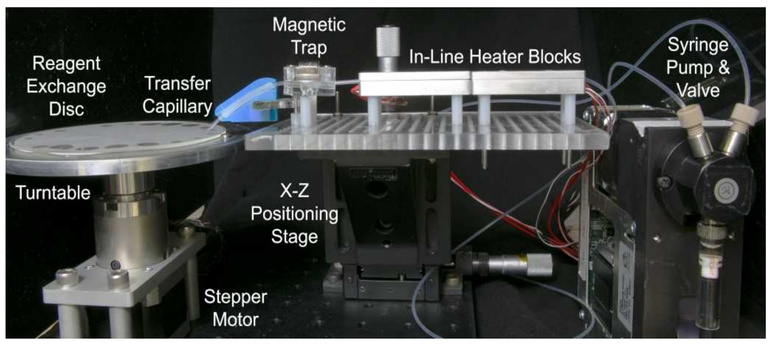
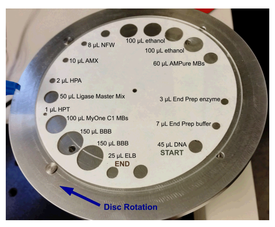


Figure 1 – Nanopore sequencing using ONT's MinION (Oxford Nanopore Technologies, 2016).

ASPIRE: Automated Sample Prep by Indexed Rotary Exchange







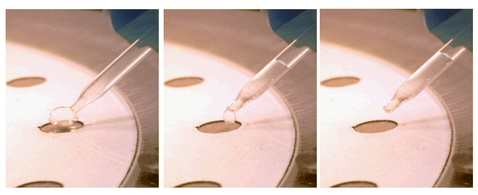


Figure 2 – (Top & Middle) An illustration and image of the first-generation ASPIRE prototype. The user initially pipettes sample and reagent droplets onto the surface of the Teflon AF coated glass disc at locations defined by a template (Bottom-left). Under automated operation, a stepper motor rotates the turntable supporting the disc, bringing reagent droplets one at a time into contact with the thermo-mechanically pulled tip of the transfer capillary (Bottom-right). Droplets are aspirated into the tube as a segmented fluid bolus past a rare-earth magnet bead trap inline with temperature-controlled heater blocks by a syringe pump to carry out basic preparation steps. Integrating the unit operations enable the semi-automated execution of the MinION 2D library preparation protocol—template and complement strands joined by a hairpin loop—including DNA clean-up, end-repair, dA-tailing, and hairpin/adapter ligation reactions.

† **Co-authors contributed equally**
* **Direct correspondence & inquiries to Michael Bartsch (mbarts@sandia.gov)**

Table 1. Comparison of ASPIRE and bench library prep results

Metrics	λ DNA		*E. coli* gDNA	
	ASPIRE	Bench	ASPIRE	Bench
Post-End Prep Yield (ng DNA)	680 ± 187	642 ± 204	866	672†
Post-Ligation Yield (ng DNA)	118 ± 12.7	67.3 ± 12.6	100	60†
Total Reads	66,583	-	70,051	43,873
Mean Aligned Read (bp)	7,886	-	5,244	4,936
Longest Aligned Read (bp)	47,581	-	75,642	76,946
% Aligned Reads	88.0%	-	85.7%	68.0%

† Quantitation and sequence results are from two different preps

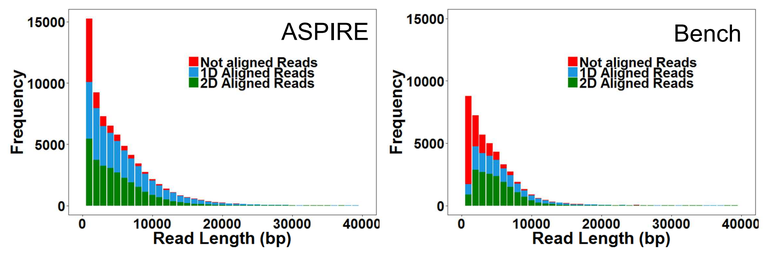


Table 1 – key figures of merit for Lambda phage and *E. coli* library preparation and sequencing. DNA yields following end-repair and dA-tailing were found to be comparable, while ASPIRE yielded slightly more DNA with less variation following adapter ligation, perhaps a benefit of automation. ASPIRE and bench methods produced a usable number of reads and showed essentially equivalent size distributions, comparable mean read sizes, and maximum read lengths in excess of 75 kbp. In these experiments, ASPIRE also produced a significantly higher proportion of genome-mapped reads (both 2D template/complement and 1D template-only) than the bench prep, particularly for short reads.

Real-time strain-level ribotyping

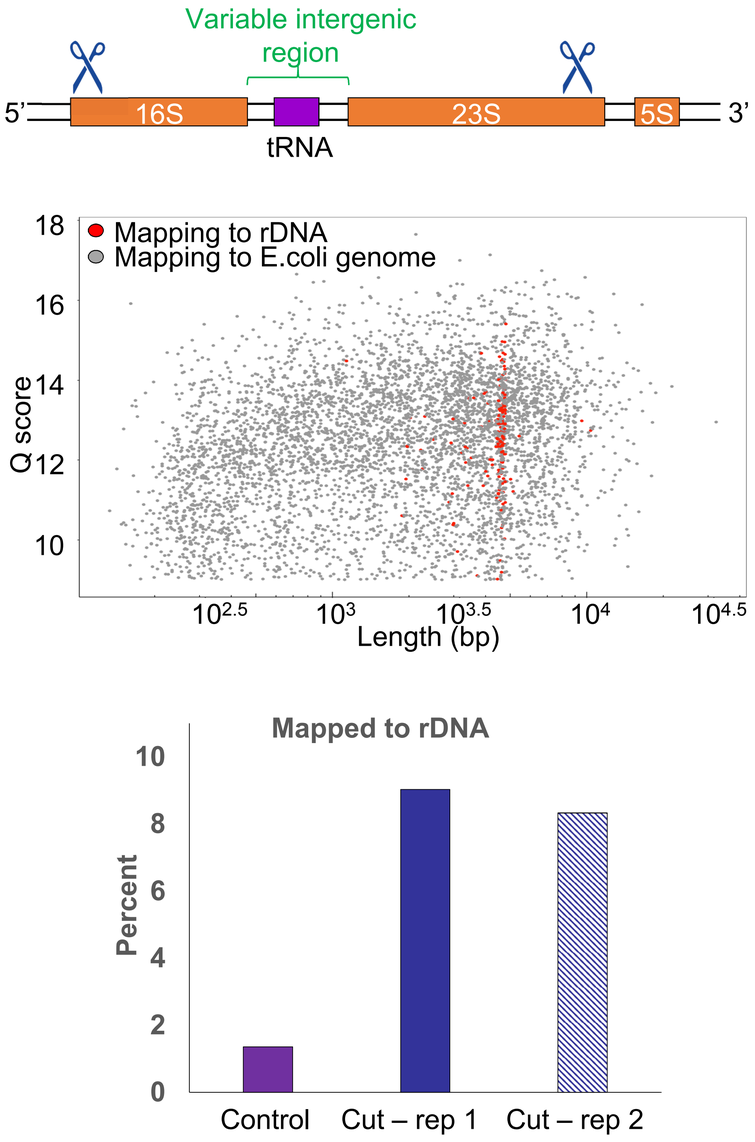


Figure 4 – Cas9-based excision of the ~5kb *E.coli* rDNA locus and downstream MinION sequencing. (Top) – Cas9 cut locations used to excise the rDNA locus. (Middle) – Q-score vs. length scatter plot of Cas9-cut *E.coli* showing the expected clustering of cut rDNA fragments in a band near 5kb. (Bottom) – Enrichment of rDNA mapping reads in Cas9 cut *E.coli* genomic DNA versus uncut control.

RUBRIC: Read Until with Basecall- and Reference-Informed Criteria

First published by Loose, et al. (*Nature Methods*, 2016), "Read Until" is a selective sequencing method unique to the MinION. As DNA transits a nanopore, it can be matched in real-time against a target reference sequence. DNA of interest is allowed to continue to sequence while non-matching DNA is physically rejected from the pore by reversing its polarity. The Sandia RUBRIC system takes this method a step further by enabling the use of conventional "ACTG" reference genomes for sequence target selection, dramatically expanding the accessibility, flexibility, and utility of MinION selective sequencing.

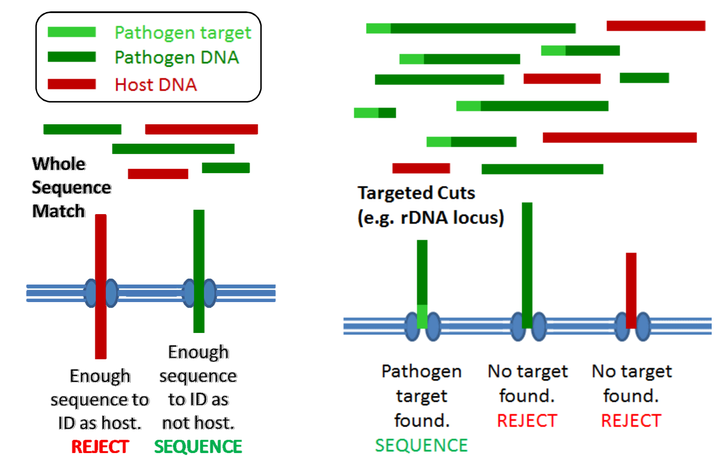


Figure 5 – Different scenarios for applying Read Until selective sequencing to a host/pathogen sample. (Left) – Whole sequence matching is computationally intensive. (Right) – RUBRIC leverages targeted Cas9 cuts and short leading target sequences to reduce computational load.

Figure 6 – To test the RUBRIC software, we prepared a MinION library with bacteriophage lambda genome digested using Eag1 restriction endonuclease to generate fragments of 11.7, 16.7 and 19.9 kb. The 16.7 kb fragment was chosen as the RUBRIC sequencing target.

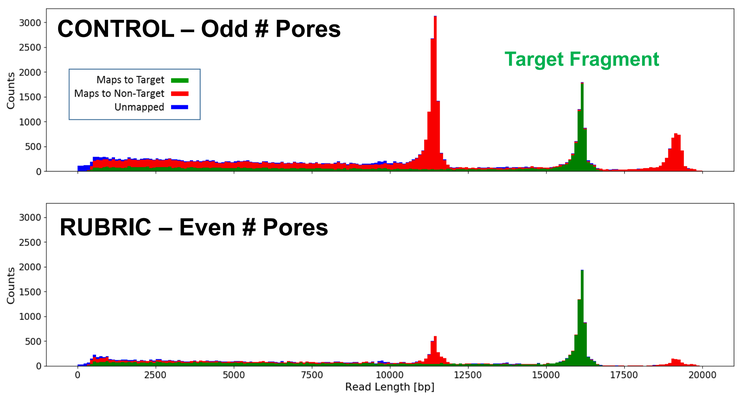


Figure 7 – Comparison of DNA fragment size and disposition for reads obtained by MinION pores implementing RUBRIC selection (even numbered pores) and those with no selection (odd numbered). Depletion of non-target DNA by RUBRIC is clearly shown.

Conclusions & Future Work

Target-based selective sequencing of long read samples on the MinION represents a significant advance on the path toward adapting nanopore sequencing to rapid, real-time pathogen diagnostics and organism-independent biodetection. Ongoing work includes investigated mixed host/pathogen and targeted CRISPER/Cas9 to facilitate target recognition and selection.

Acknowledgements

This work was funded as part of the Real-time Pathogen Identification by Enhanced Ribotyping (RAPIER) project through Sandia's Laboratory-directed research and development (LDRD) program.