SAND2018-2584C

READ UNTIL WITH BASECALL- AND REFERENCE-INFORMED CRITERIA (RUBRIC)

Michael S. Bartsch, Ph.D.

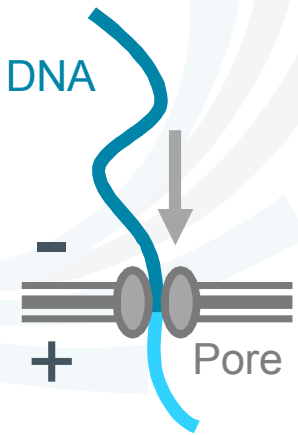
Sandia National Laboratories, Livermore, CA, USA

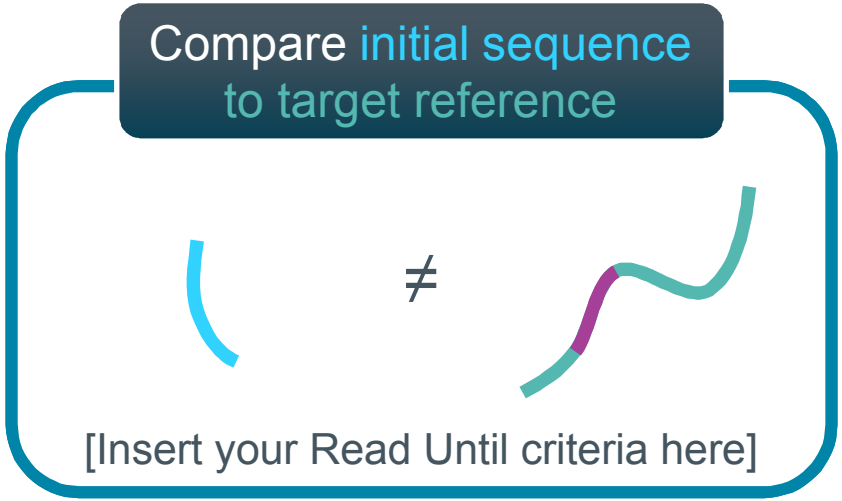
“READ UNTIL” SELECTIVE SEQUENCING

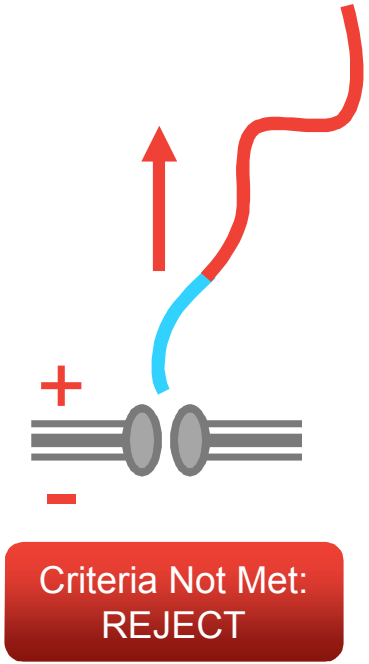
An emerging nanopore sequencing capability

Sandia National Laboratories

Access to real-time data enables real-time rejection of non-informative DNA







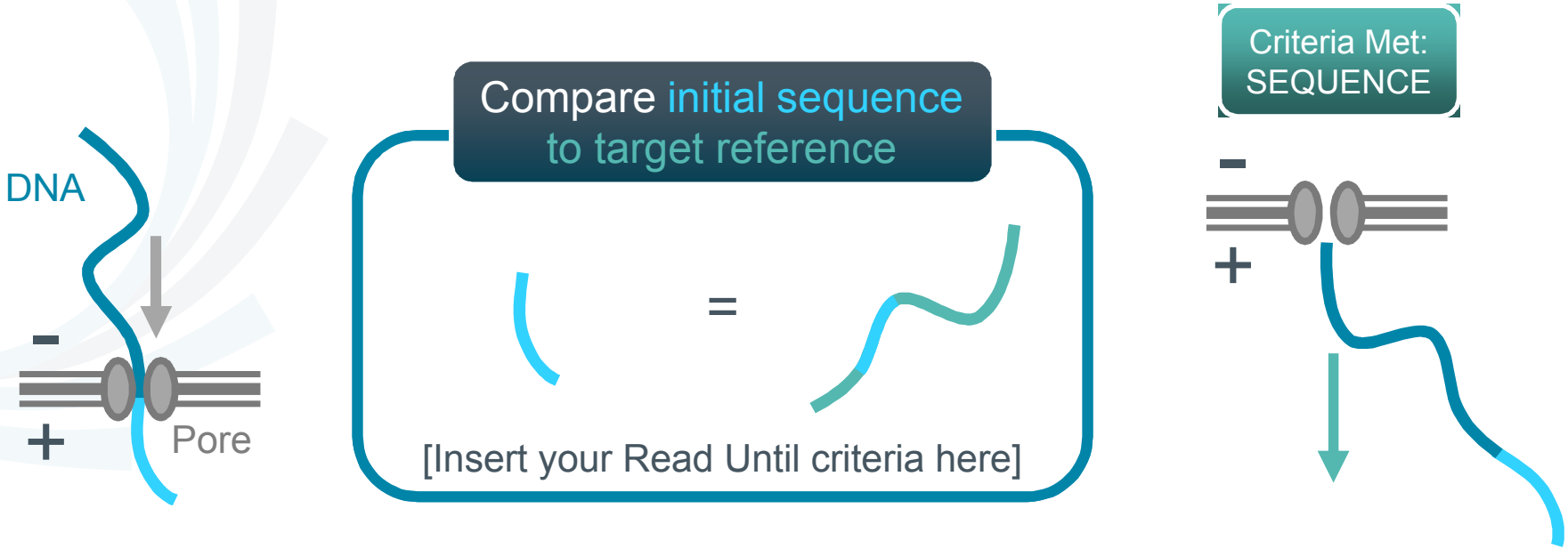
Oxford NANOPORE Technologies

“READ UNTIL” SELECTIVE SEQUENCING

An emerging nanopore sequencing capability

Sandia National Laboratories

Access to real-time data enables real-time rejection of non-informative DNA

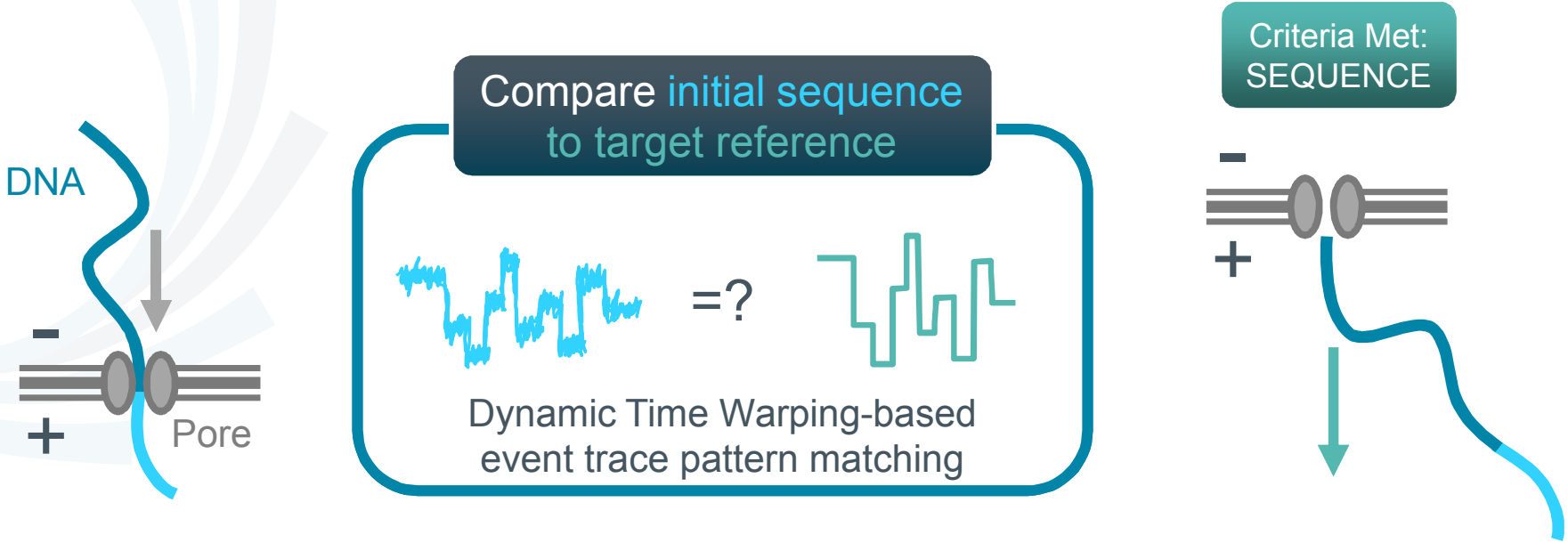


Oxford NANOPORE Technologies

“READ UNTIL” SELECTIVE SEQUENCING

Original pattern-matching approach

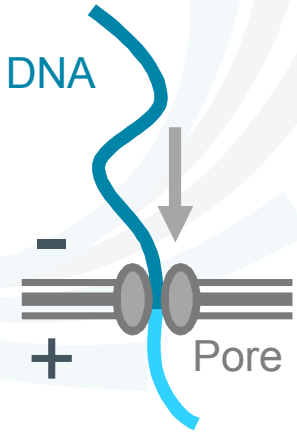
First published implementation of Read Until by Loose, Malla, and Stout, *Nature Methods*, 2016



“READ UNTIL” SELECTIVE SEQUENCING

Read Until with Basecall- and Reference-Informed Criteria (RUBRIC)

Sandia RUBRIC method by Edwards, Krishnakumar, Sinha, Patel, and Bartsch, 2017



Compare initial sequence to target reference

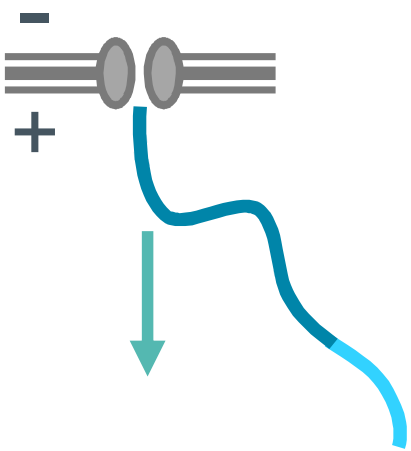
Basecall:
AACTGGGCTA

=?

GGTACATAAC
TGGGCTATTC
GACCAATCTA

Alignment to conventional ACTG reference sequence

Criteria Met:
SEQUENCE



THE RUBRIC WORKFLOW

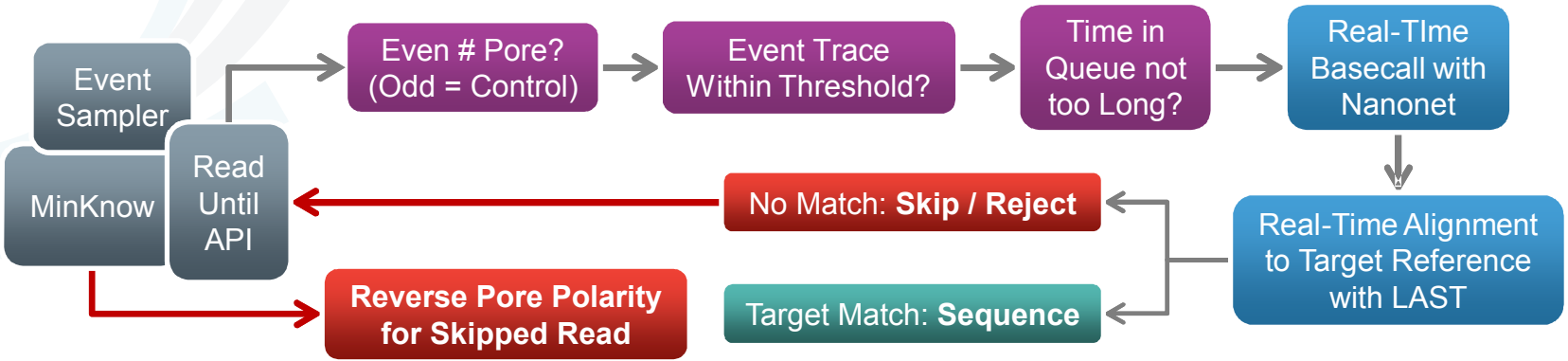
An exercise in efficient use of limited computing resources

Sandia National Laboratories

Key Elements of the Method

- Laptop runs MinKnow while ethernet-linked desktop runs RUBRIC
- Event trace threshold filters out reads unlikely to become mappable fast5s
- Decision queue screens out reads that are too “old” for a timely decision
- **IN REAL-TIME**: nanonet basecalls 300 events (~80 bases) and LAST aligns to reference

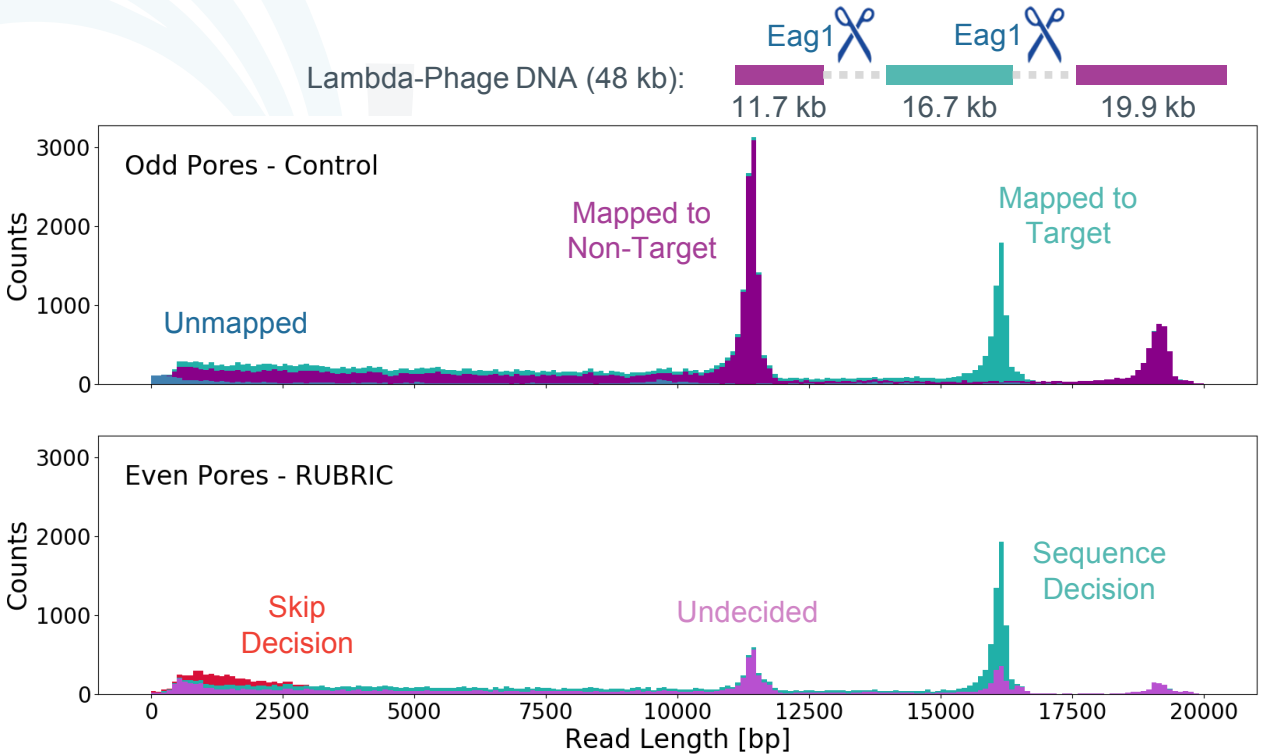
Average Decision Time: 0.65 sec

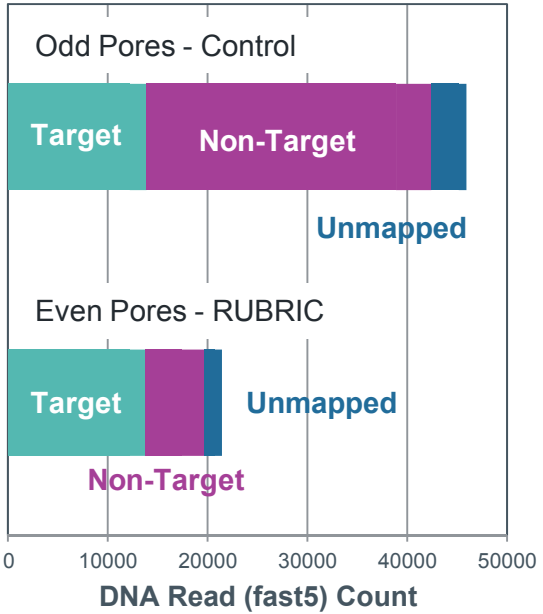


Oxford NANOPORE Technologies

SUCCESSFUL SELECTION WITH RUBRIC

Proof of concept: Eag1-digested Lambda DNA





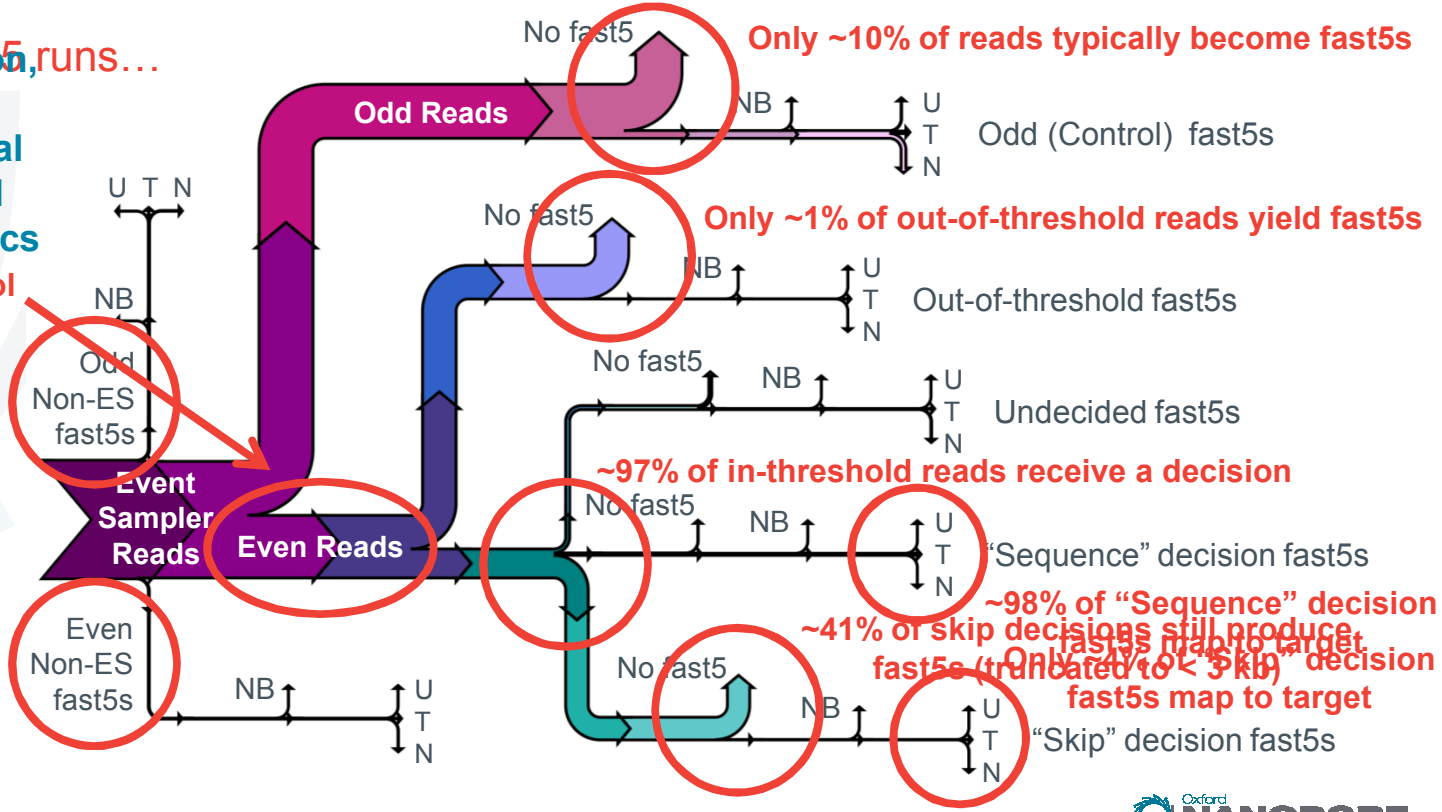
EVALUATING RUBRIC PERFORMANCE

Detailed "data flow" analysis provides critical insights

Testing, optimization, and analysis are ongoing with a goal of RUBRIC-based pathogen diagnostics

~10% more RUBRIC reads/pore vs. control

~16% of all fast5s are never reported by the Event Sampler



U – Unmapped
T – Maps Target
N – Maps Non-Target

NB – No Basecall

ACKNOWLEDGEMENTS

Sandia National Laboratories

Name	Division / Location
Harrison S. Edwards	Advanced Systems Engineering & Deployment Dept., Sandia National Labs (CA)…now University of Toronto
Raga Krishnakumar, Ph.D.	Systems Biology Dept., Sandia National Labs (CA)
Anupama Sinha	Systems Biology Dept., Sandia National Labs (CA)
Kamlesh D. Patel, Ph.D.	Advanced Systems Engineering & Deployment Dept., Sandia National Labs (CA)
Michael S. Bartsch, Ph.D. (mbarts@sandia.gov)	Advanced Systems Engineering & Deployment Dept., Sandia National Labs (CA)

Special thanks to Matt Loose (U. Nottingham), and Julian Atienza & Richard Ronen (ONT)

This work was supported by the Laboratory Directed Research and Development program at Sandia National Laboratories, a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

THANK YOU

ADDITIONAL OBSERVATIONS ON RUBRIC OPERATION

Fugitives from RUBRIC selection