Nanopore Sequencing Technology and Tools:  
Computational Analysis of the Current State, Bottlenecks and Future Directions

Damla Senol¹, Jeremie Kim¹,³, Saugata Ghose¹, Can Alkan² and Onur Mutlu³,¹

¹Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, USA  
²Department of Computer Engineering, Bilkent University, Bilkent, Ankara, Turkey  
³Department of Computer Science, Systems Group, ETH Zürich, Zürich, Switzerland

Nanopore Sequencing

Nanopore sequencers rely solely on the electrochemical structure of the different nucleotides for identification and measure the change in the ionic current as long strands of DNA (ssDNA) pass through the nano-scale protein pores.

Biological Nanopores for DNA Sequencing

- first proposed in the 1990s,
- recently made commercially available in May 2014 by Oxford Nanopore Technologies (ONT).

MinION

- first commercial nanopore sequencing device
- high-throughput sequencing apparatus
- produces real-time data
- inexpensive
- pocket-sized / portable

Pipeline and Current Tools

Step 1. Basecalling

- Translates raw signal output of MinION to generate DNA sequences.  

Step 2. Genome Assembly for noisy long reads

- Using only the basecalled DNA reads, generates longer contiguous fragments called draft assemblies.  
- canu [5], miniasm [6]

Step 3. Polishing (Optional)

- Generates an improved consensus sequence from the draft assembly  
- nanoplot [7], racon [8]

Our Goal

- In order to take advantage of nanopore sequencing, it is important to increase the accuracy and the speed of the whole pipeline.
- Although new nanopore chemistry R9 improves the data accuracy, the tools used for nanopore sequence analysis are of critical importance as they should overcome the high error rates of the technology.
- Our goal in this work is to comprehensively analyze tools for nanopore sequence analysis, with a focus on understanding the advantages, disadvantages, and bottlenecks of the various tools.

Results

- ONTs cloud-based basecaller, Metrichor and local basecaller, nanonet perform similarly with high accuracy. However, another local basecaller nanocall is not suitable for R9 data.
- Canu, the assembler with error correction, produces high-quality assemblies but is relatively slow compared to Miniasm, the assembler without error correction.
- Miniasm is not as accurate as canu, but it is suitable for fast initial analysis and the quality of the assembly can be increased with an additional polishing step.