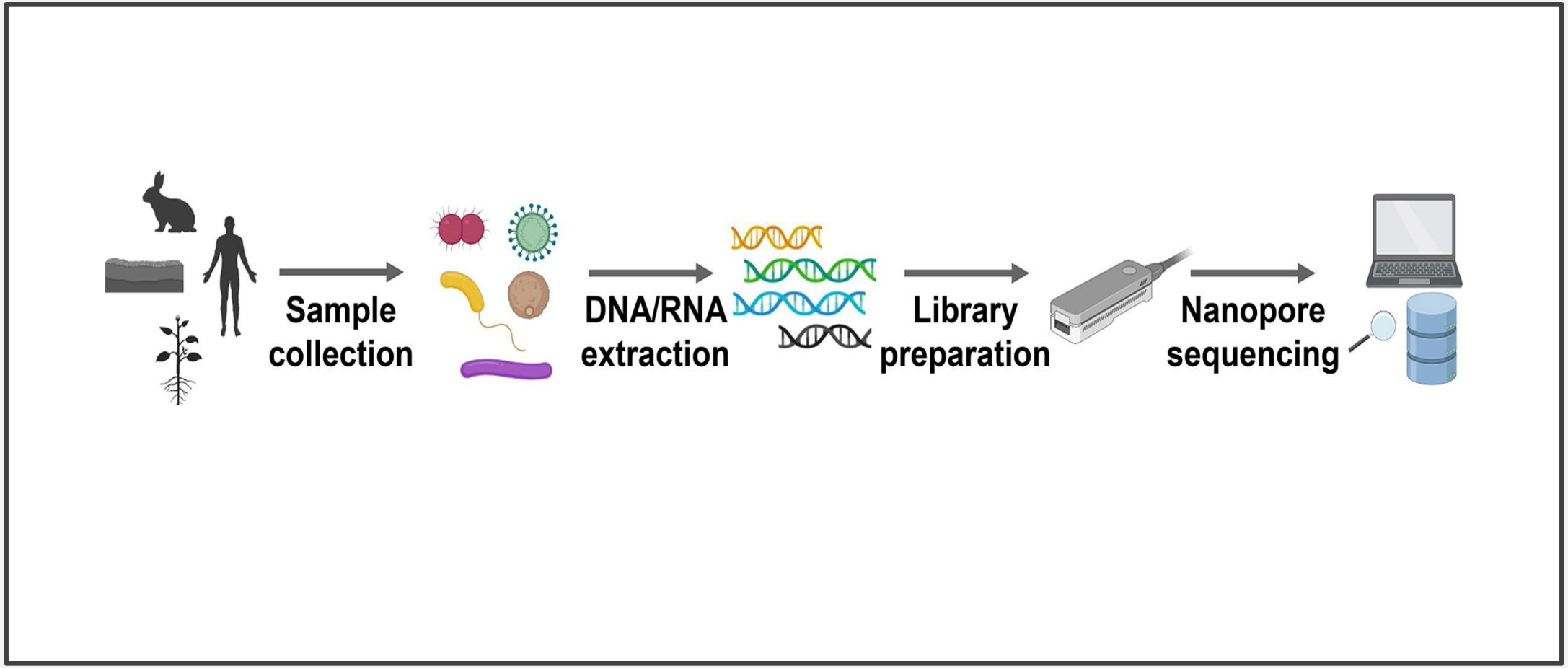
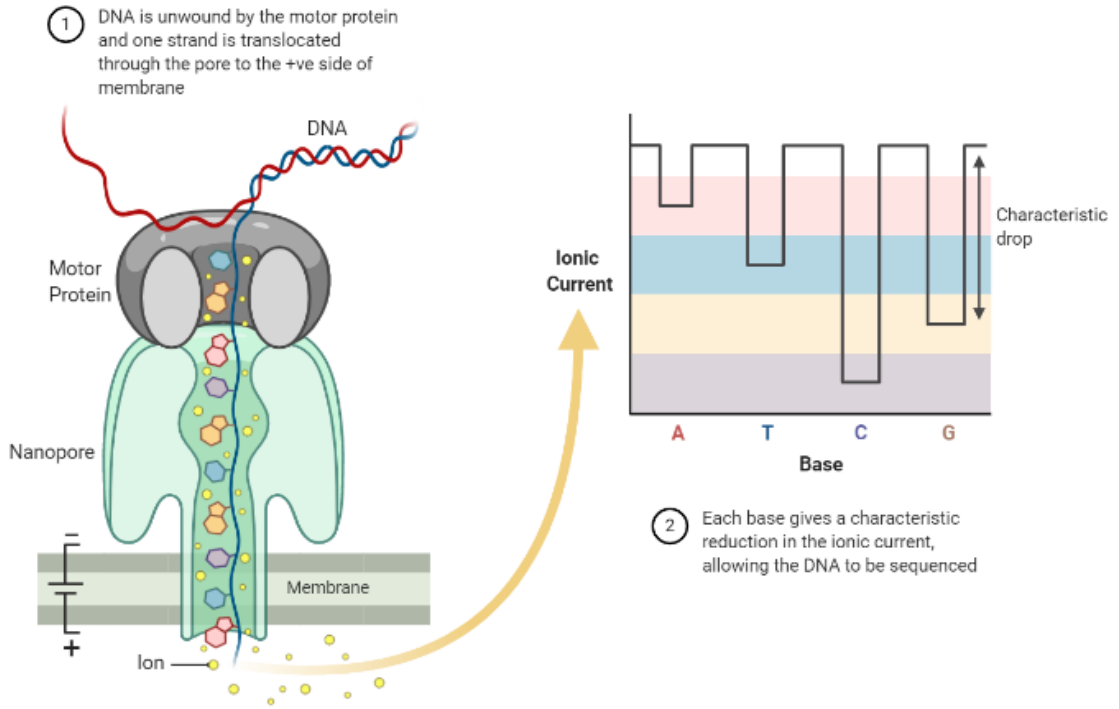
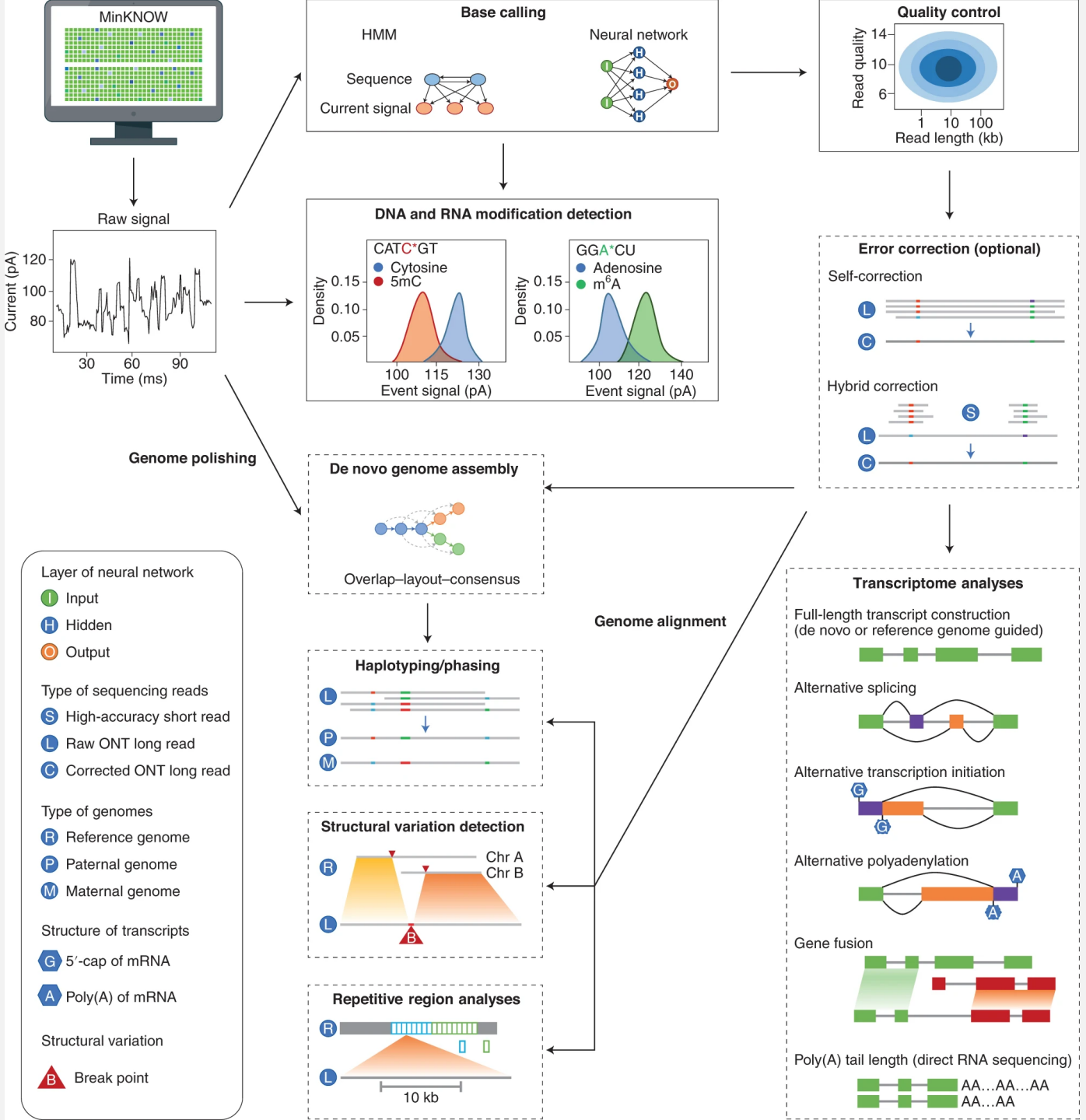
REVIEW ABOUT AVAILABLE SOFTWARE FOR NANOPORE SEQUENCING DATA ANALYSIS

Barbara Popławska


Sample collection
DNA/RNA extraction
Library preparation
Nanopore sequencing

Nanopore Sequencing



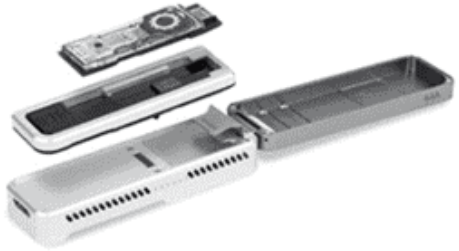

MinKNOW
Raw signal
Current (pA)
120
100
80
30
60
90
Time (ms)
Base calling
HMM
Sequence
Current signal
Neural network
Quality control
Read quality
14
10
6
1
10
100
Read length (kb)
DNA and RNA modification detection
CATC*GT
Cytosine
5mC
Density
0.15
0.10
0.05
100
115
130
Event signal (pA)
GGA*CU
Adenosine
m6A
100
120
140
Event signal (pA)
Error correction (optional)
Self-correction
Hybrid correction
Genome polishing
De novo genome assembly
Overlap–layout–consensus
Genome alignment
Haplotyping/phasing
Structural variation detection
Chr A
Chr B
Repetitive region analyses
10 kb
Transcriptome analyses
Full-length transcript construction
(de novo or reference genome guided)
Alternative splicing
Alternative transcription initiation
Alternative polyadenylation
Gene fusion
Poly(A) tail length (direct RNA sequencing)
AA…AA…AA
AA…AA
Layer of neural network
Input
Hidden
Output
Type of sequencing reads
High-accuracy short read
Raw ONT long read
Corrected ONT long read
Type of genomes
Reference genome
Paternal genome
Maternal genome
Structure of transcripts
5′-cap of mRNA
Poly(A) of mRNA
Structural variation
Break point

NANOPORE SEQUENCING PLATFORMS

MinION™

GridION™ X5

PromethION™

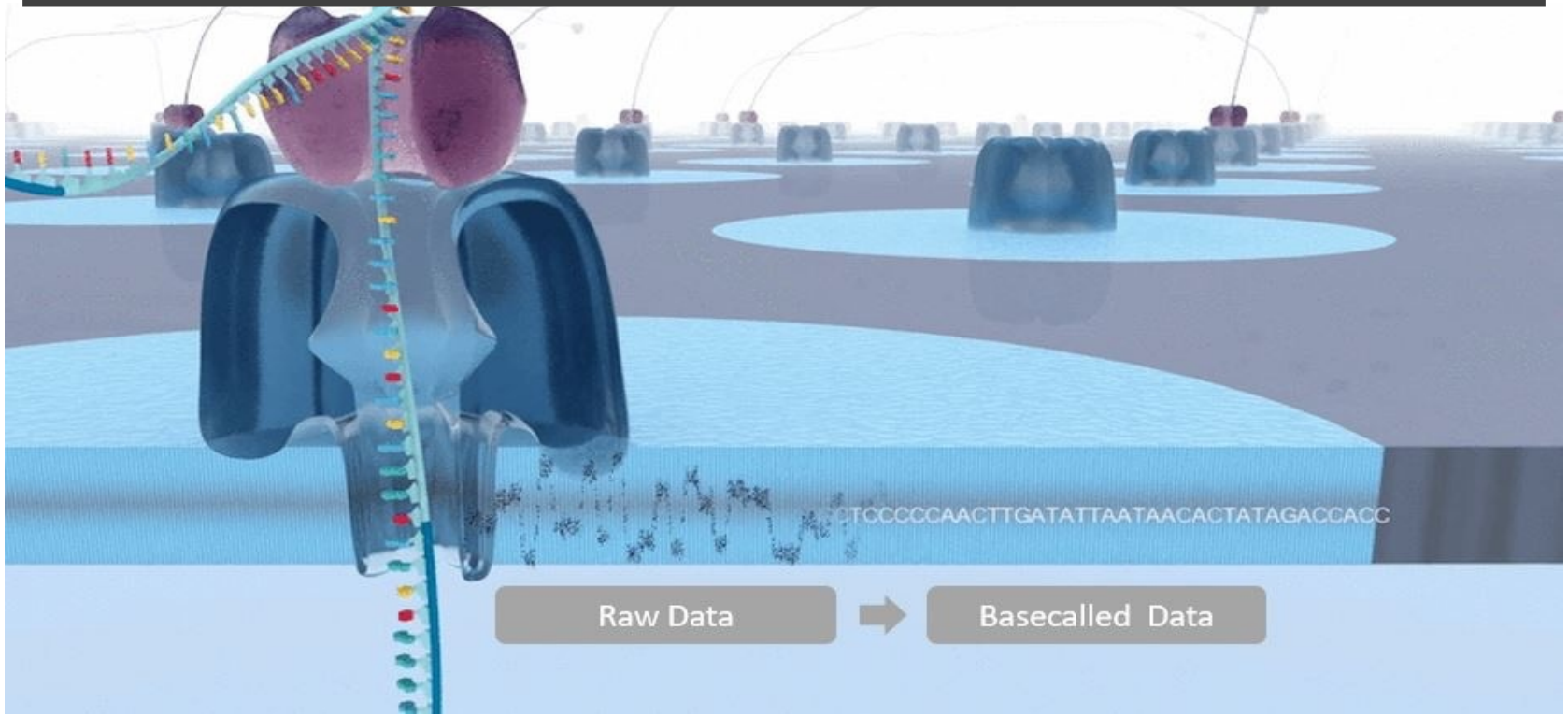


Flongle™

MINKNOW

Channel states panel
Run state: sequencing
368
Sequencing
41
Pore
10
Recovering
81
Inactive
12
Unclassified
+ Show detailed

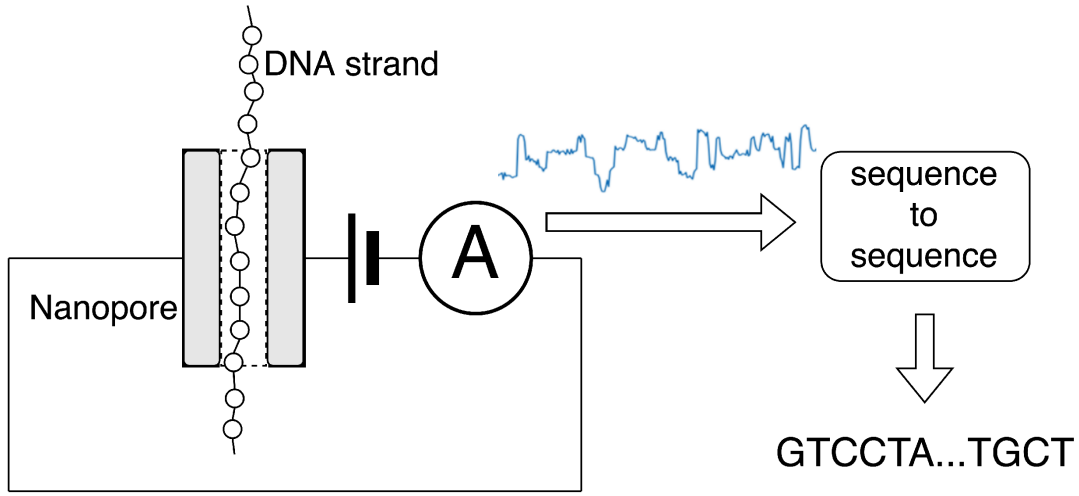
BASECALLING



BONITO SOFTWARE

Three basecaller models are offered:

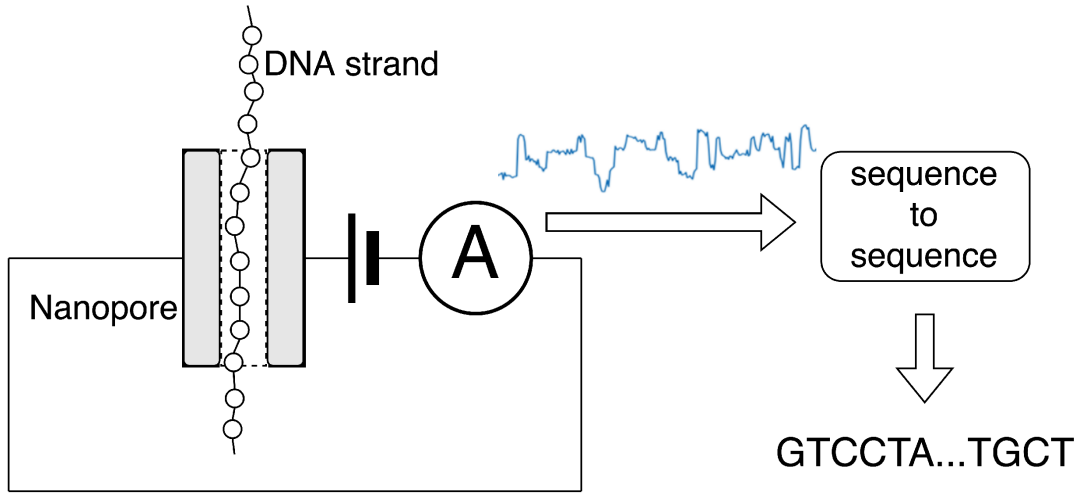
- Fast Model
- High Accuracy Model (HAC)
- Super Accurate Model (SUP)



GUPPY

The Guppy toolkit contains:

- Basecaller
  - Calibration strand detection
  - Adapter trimming
- Barcoding/demultiplexing
- Alignment
- Modified basecalling



DORADO

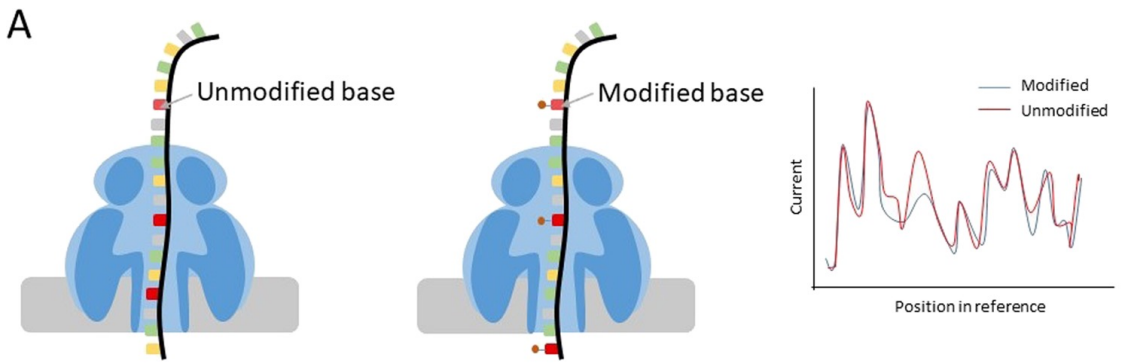
- Basecalling
- Alignment
- Modified basecalling

nanoporetech/
dorado

Oxford Nanopore's Basecaller

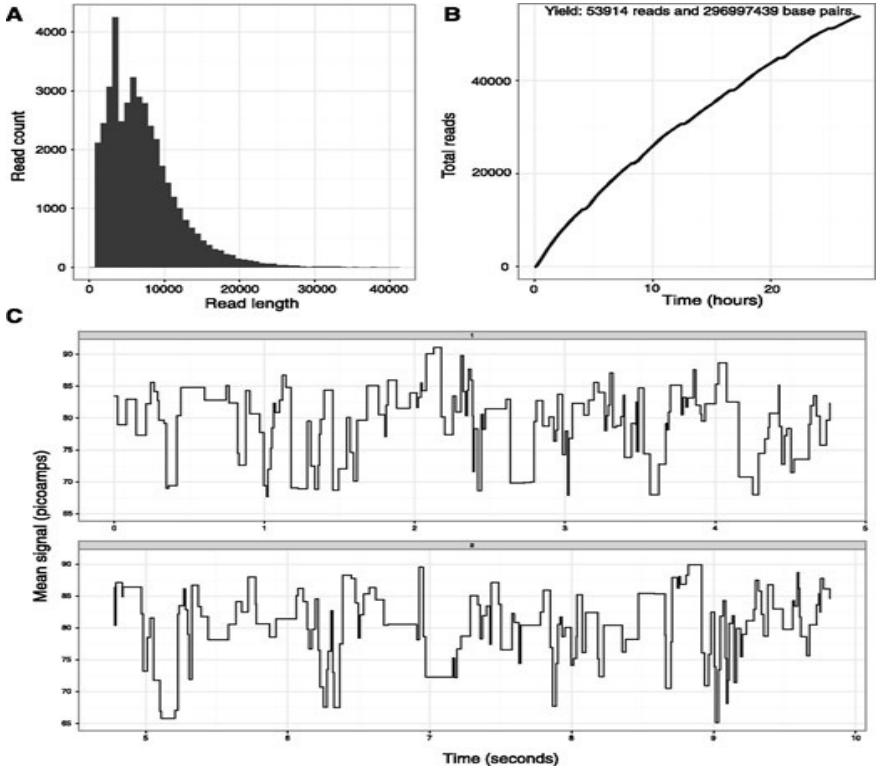
NANOPOLISH

- Detecting DNA and RNA modifications



PORETOOLS

- Data Conversion
- Quality Control
- Data Visualization
- Subsetting and Filtering
- Custom Analysis Pipelines



EPI2ME

- Basic signal processing
- Assembly
- Metagenomic
- Direct RNA sequencing
- Quality control

BIBLIOGRAPHY

- Alberto Magi, Roberto Semeraro, Alessandra Mingrino, Betti Giusti, Romina D'Aurizio, Nanopore sequencing data analysis: state of the art, applications and challenges, Briefings in Bioinformatics, Volume 19, Issue 6, November 2018, Pages 1256–1272, https://doi.org/10.1093/bib/bbx062
- Nanopore sequencing and its application to the study of microbial communities, Ciuffreda, Laura et al., Computational and Structural Biotechnology Journal, Volume 19, 1497 – 1511
- Xu, L., Seki, M. Recent advances in the detection of base modifications using the Nanopore sequencer. J Hum Genet 65, 25–33 (2020). https://doi.org/10.1038/s10038-019-0679-0
- Nicholas J. Loman, Aaron R. Quinlan, Poretools: a toolkit for analyzing nanopore sequence data, Bioinformatics, Volume 30, Issue 23, December 2014, Pages 3399–3401, https://doi.org/10.1093/bioinformatics/btu555
- https://community.nanoporetech.com/technical_documents/data-analysis/v/datd_5000_v1_revr_22aug2016/basecalling-algorithms
- https://community.nanoporetech.com/docs/prepare/library_prep_protocols/Guppy-protocol/v/gpb_2003_v1_revax_14dec2018/guppy-software-overview
- https://labs.epi2me.io/wfindex/
- https://www.cd-genomics.com/nanopore-sequencing-principles-platforms-and-advantages.html