Sekvenování třetí generace

Petra Vídeňská, Ph.D.

Sekvenování 3. generace

- Nevyužívá amplifikace za účelem zvýšení signálu (měla by být vyšší přesnost –accurancy)
- Produkuje dlouhá čtení
- Dobrá prosekvenovanost GC bohatých oblastí
- Epigenetika
- Zatím dvě dostupné technologie – PacBio a Nanopore (MinION)
- I Illumina chystá nový systém

Pacific Biosciences (Roche)

	RS II (P6-C4)	Sequel
Run time	up to 240 min	up to 240 min
Total output	~500 Mb - 1 Gb	5 Gb - 10 Gb
Output/day	~2 Gb	20 Gb
Mean read length	10 -15 kb	10 -15 kb
Single pass accuracy	~86%	~86%
Consensus (30X) accuracy	>99.999%	>99.999%
# of reads	~50k	~500k
Instrument price	~$700k	$350k
Run price	~$400	~$850

Sequel system

PacBio RS II

- Variety of sample types: Genome DNA, Amplicons, cDNA
- Low input sample amounts from 10 ng to 1 µg

- Complete template preparation in as few as 6 hrs
- Accepts inserts from 250 bp to 40 kb for flexible assay design
- Multiplexing and barcoding solutions available

- Rapid sequence time (0.5 to 4 hrs)
- Serially process up to 16 SMRT Cells in a single run with walkaway automation

- Variety of analysis methods available through SMRT Software Suite and community
- Open source software
- Advanced data visualization and mining

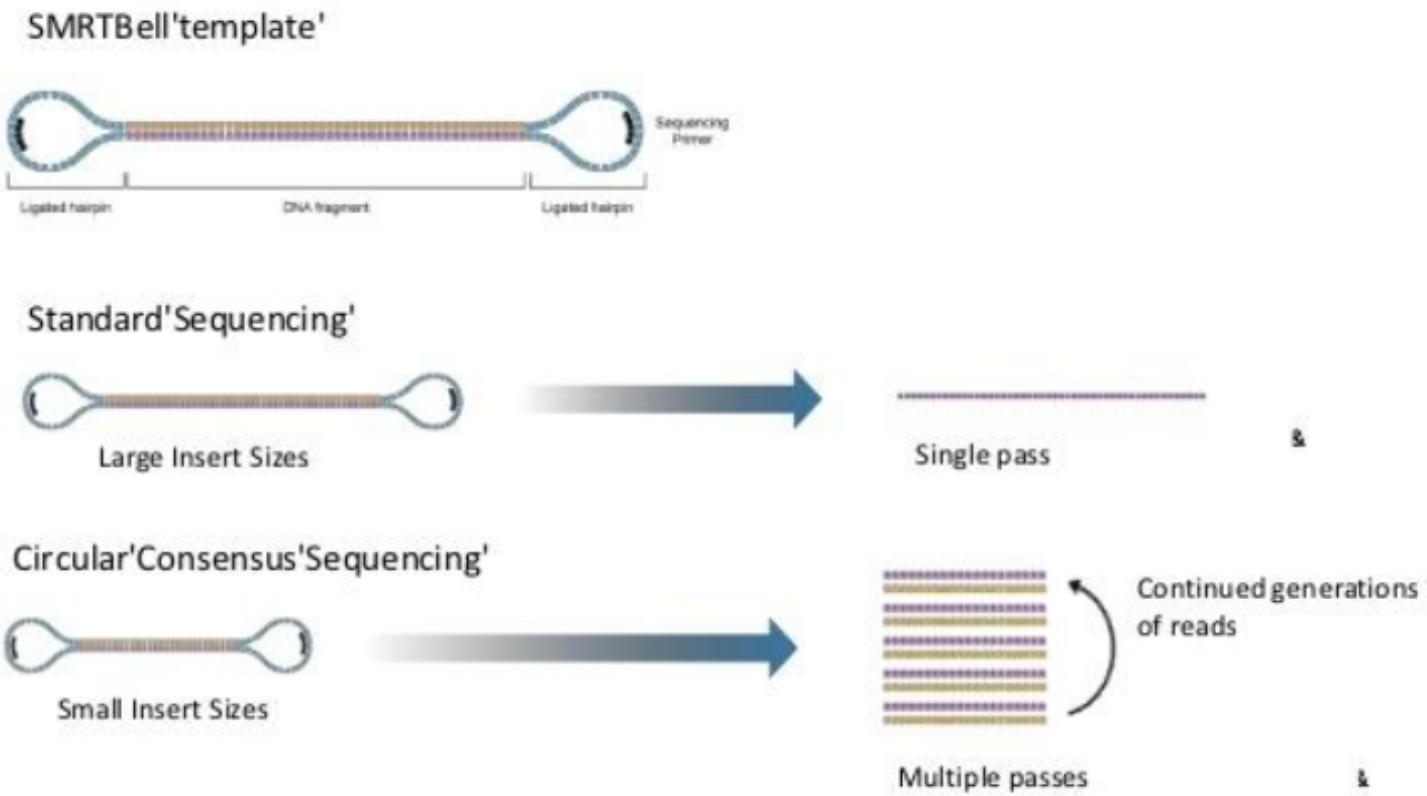
Comprehensive *de novo* assemblies
Full-length isoform transcripts
Phased SNPs & minor variants
Methylation profiles

PacBio

High-throughput sequencing

PACIFIC BIOSCIENCES™

Library preparation



A

PacBio

B

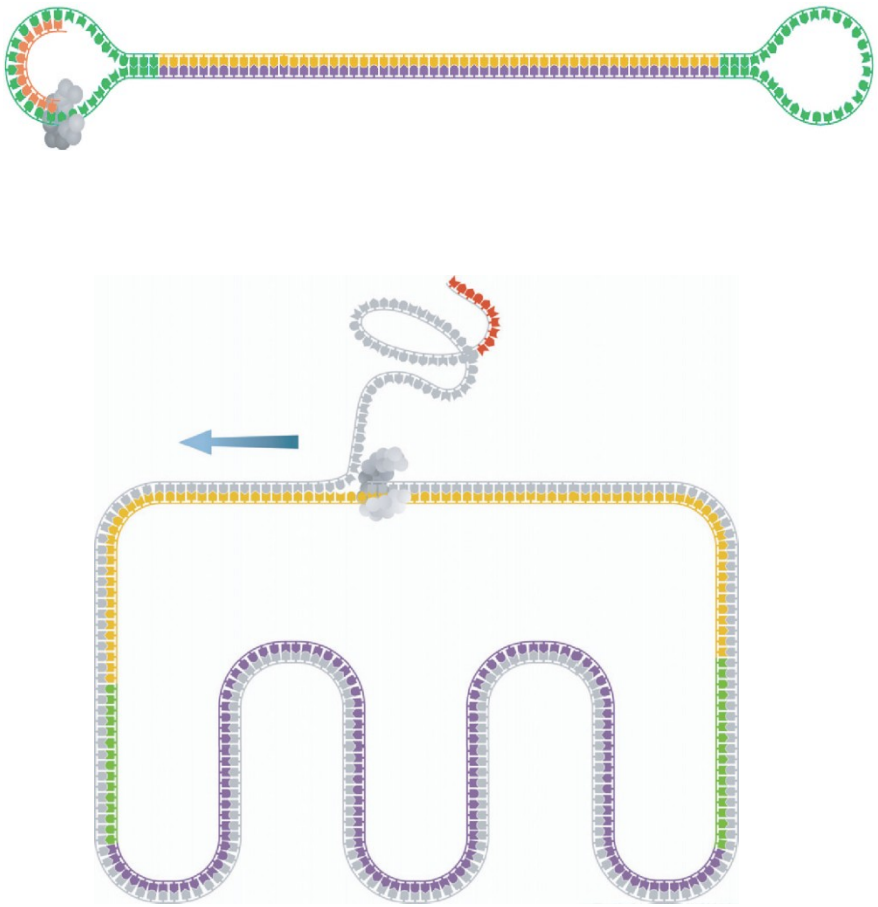
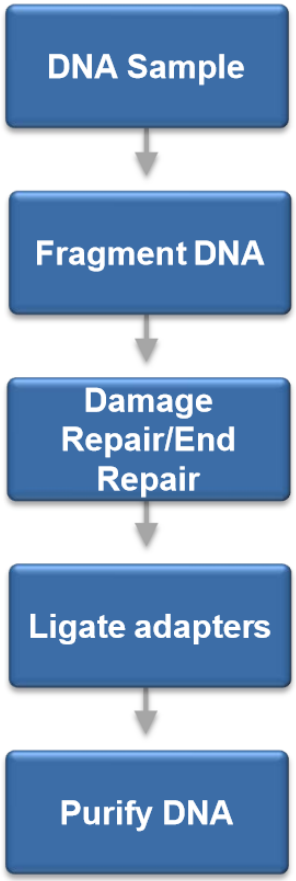
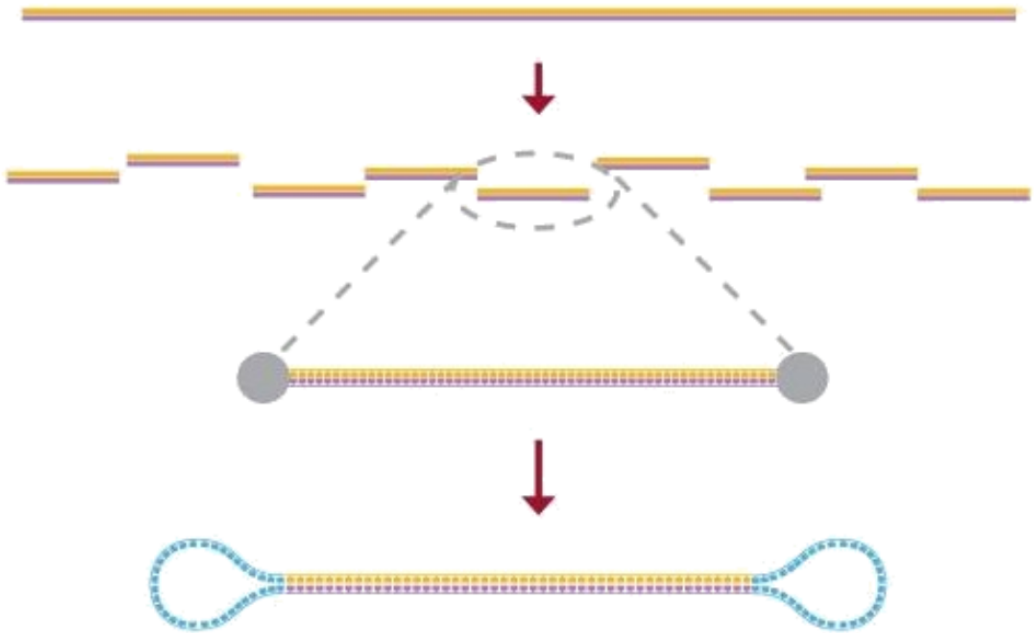


Schéma kruhového konsenzuálního sekvenování - ccs (upraveno podle Travers a kol., 2010)
A) SMRT templát obsahuje dvouřetězcovou oblast (inzert) (žlutá a fialová) na obou koncích uzavřanou jednořetězcovými vlásenkovými smyčkami (zelená). Vlásenkové smyčky představují jednořetězcovou oblast, na kterou se může vázat sekvenační primer (oranžová).
B) Polymeráza (šedá) prodlužující primer z jedné vlásenkové struktury využívá jedno vlákno DNA jako templát a druhé vytěsňuje. Když se polymeráza dostane zpět k 5´konci primeru, začne vytlačovat již nasyntetizovaný řetězec a pokračuje v syntéze DNA (pohybuje se ve směru modré šipky). Výsledná sekvence je odvozená z obou vláken, pozitivního i negativního.

PacBio

Template Preparation





SMRTbell™ Template preparation can be used to create libraries of various insert sizes from 250 bp to 20,000 bp depending on the needs of the application.

http://www.pacb.com/wp-content/uploads/2015/09/PacBioWorkflow.pdf

PaBio

PACBIO® PROJECT SUBMISSION RECOMMENDATIONS

I. Estimating library yield on the PacBio RS II (Table 1)

The PacBio Binding Calculator should be used to estimate yield for all samples. The table below provides estimates of expected yields for various DNA libraries. Please note the assumptions used to generate the table.

Target Library Insert Size	Recommended DNA Quantity for Submission*	Min Input DNA Amount Required (Post-Shearing)	Purified DNA SMRTbell™ Library Concentration (ng/µl)	Est. Total Data Yield (Gb)	
				MIN	MAX
250 bp	600 ng	250 ng	5	95	190
500 bp	600 ng	250 ng	5	18	36
1 kb	1.2 µg	500 ng	10	340	680
2 kb	1.2 µg	500 ng	10	170	340
5 kb	2.4 µg	1 µg	25	170	340
10 kb	2.4 µg	1 µg	25	85	170
10 - 20 kb (AMPure)	15 µg	5 ug	125	35	70
20 kb (BluePippin)	15 µg	5 µg	50	10	20

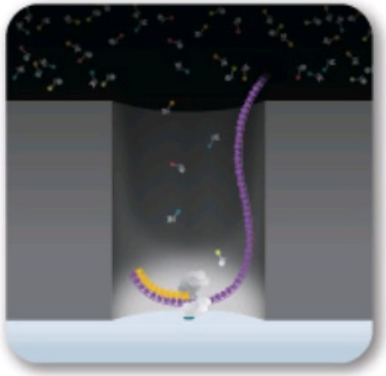
* DNA input amounts recommended for submission represent quantities needed for one (1) SMRTbell library prep and includes extra quantity needed for any additional QC (with conservative excess). Estimated Min and Max total sequencing data yields (Gb) are based on the assumption of a starting SMRTbell DNA template volume of 10 uL, a purified SMRTbell DNA template concentration of 5 - 125 ng/ µl and a throughput of 500 - 1000 MB per SMRT Cell using P6-C4 chemistry. For insert sizes ≥ 1-kb, a Magnetic Bead (one cell per well) loading protocol is used in the SMRT Cell yield calculations. Preparation of 20-kb large-insert libraries can be achieved using either AMPure or BluePippin size selection strategies. Actual data yields may vary depending on sample DNA quality.

http://mogene.com/userfiles/files/Pacbio%20Guidelines%20-%20DNA%20v2_0.pdf

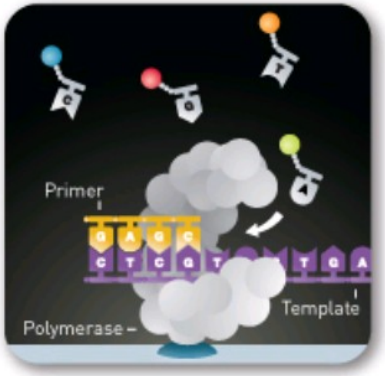
PacBio

SMRT® Cells

Zero-Mode Waveguides

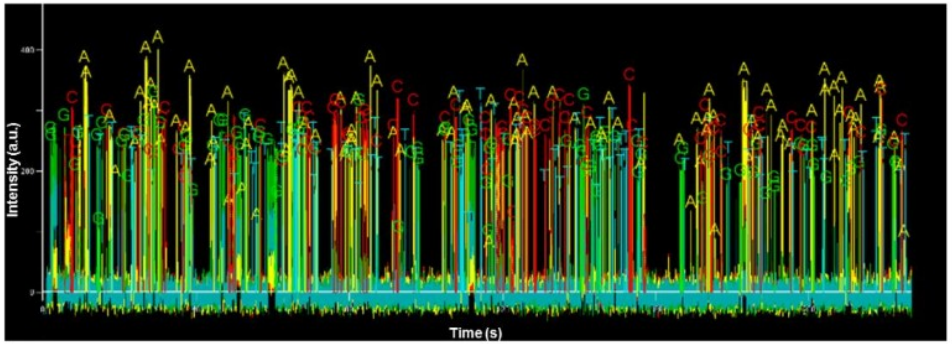


Phospholinked Nucleotides



PacBio® RS II

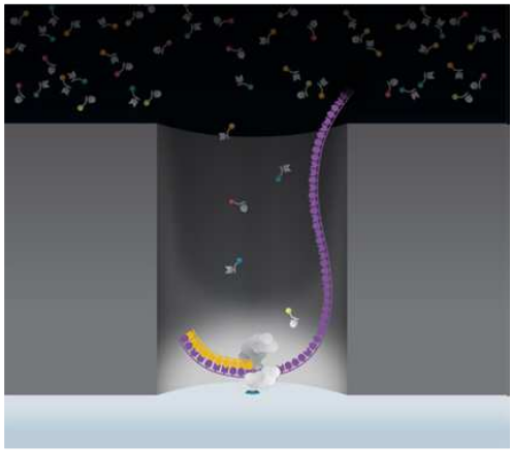
Trace



PacBio – technologie SMRT

Single Molecule, Real-Time Analysis

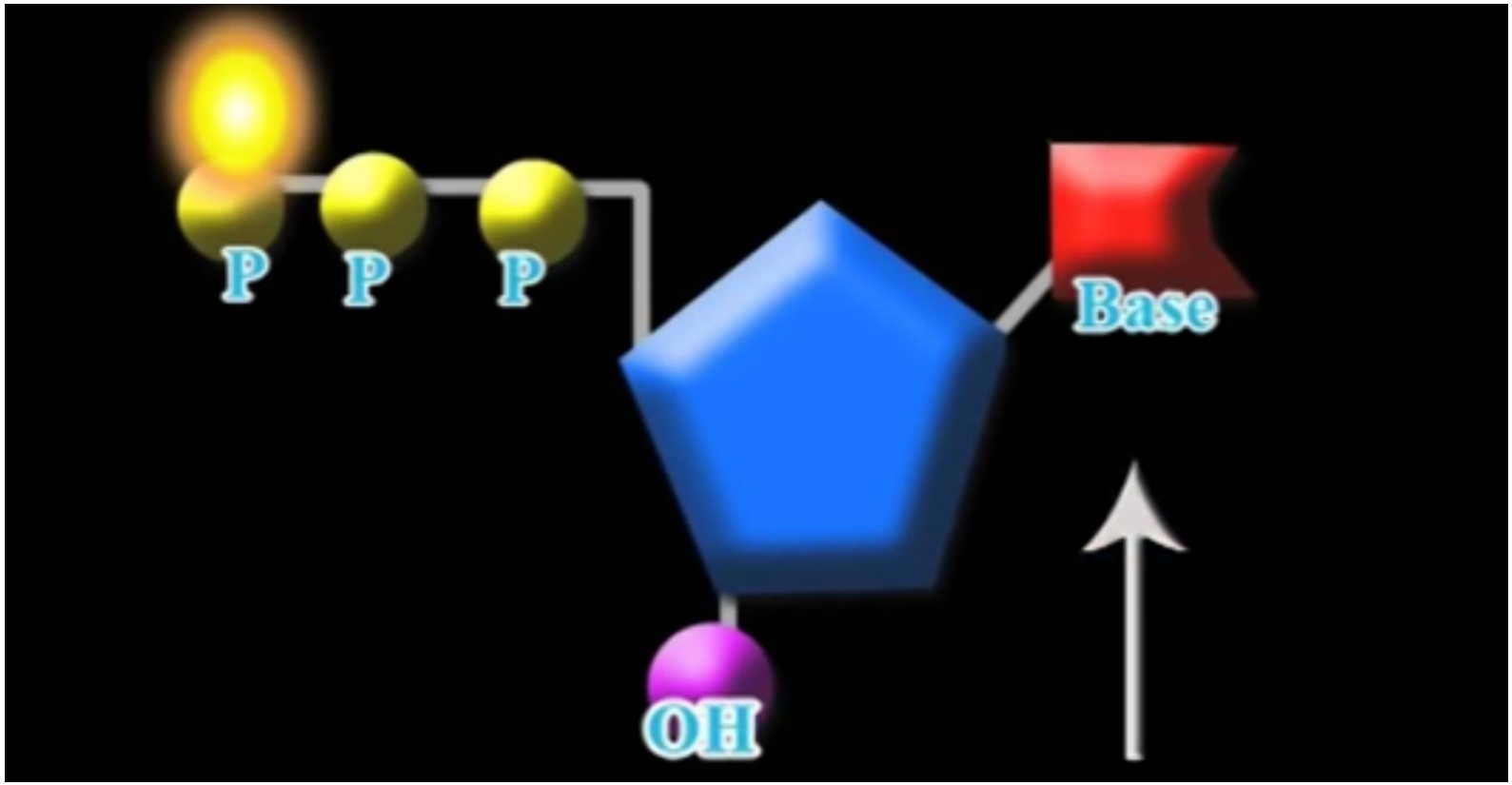
- Single molecule, real-time (SMRT™) detection provides direct measurement of individual molecules, capturing multiple dimensions of data
- Templates can be prepared for the PacBio *RS* without PCR amplification, resulting in more uniform sequence coverage across genomic regions regardless of GC content, facilitating the detection of minor variants in heterogeneous samples
- While observing single molecule sequencing by a highly processive strand-displacing DNA polymerase in real time, the system also records the kinetics of each nucleotide incorporation reaction, identifying base modifications of the native templates, such as DNA methylation



http://cgs.hku.hk/portal/files/GRC/Events/Seminars/2011/20110513/pacbiors_overview.pdf

Video – SMRT technologie: http://www.pacb.com/smrt-science/smrt-sequencing/

PacBio



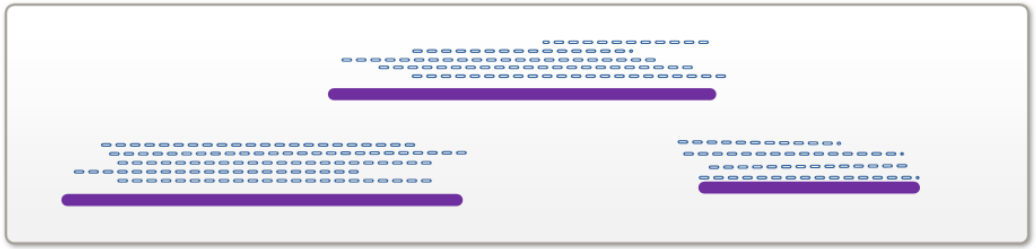
https://www.youtube.com/watch?v=v8p4ph2MAvI

PacBio

What could you do with a complete genome?

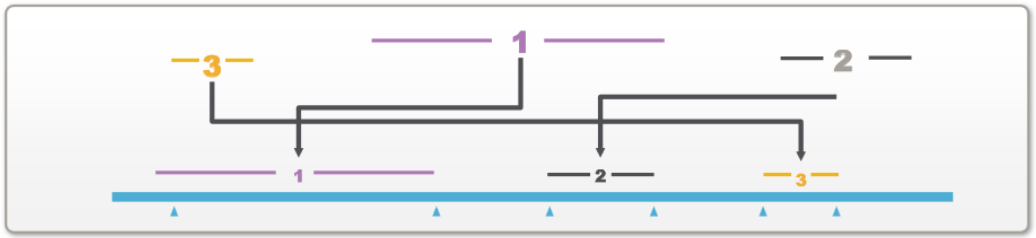
De novo Assembly

Complete genomes using only PacBio® reads or combine technologies



Scaffold

Establish framework for genome and resolve ambiguities

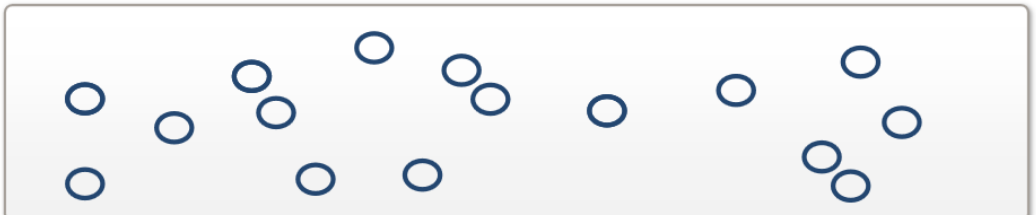


Span Gaps

Polish genomic regions with up to 10x improvement

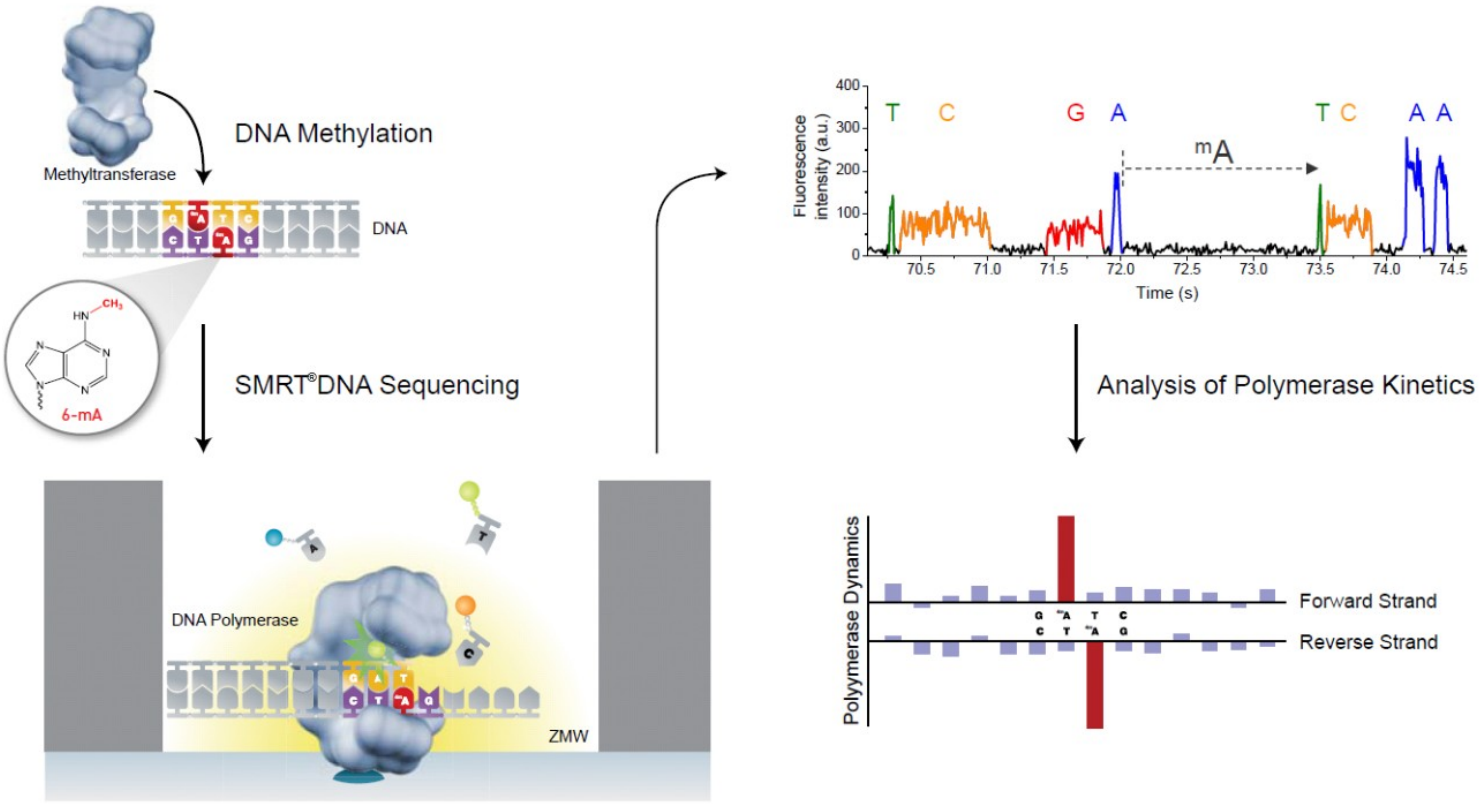
BAC Sequencing

Sequence complex and difficult BACs. Single or pooled.



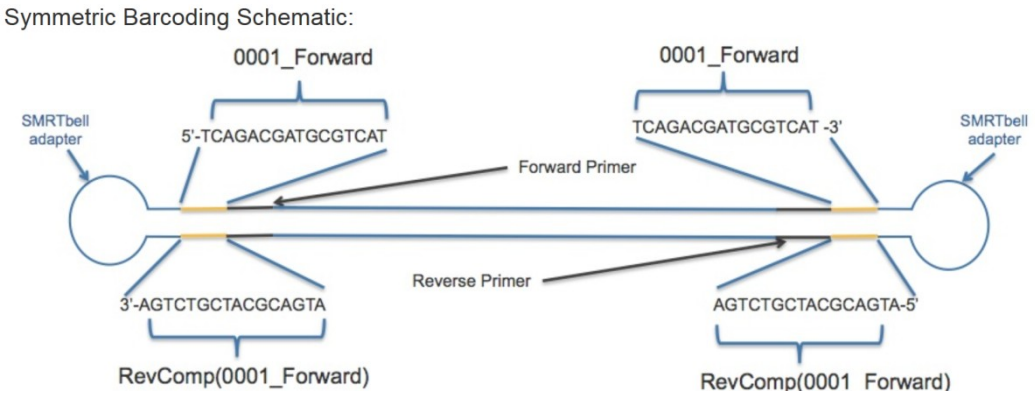
PacBio

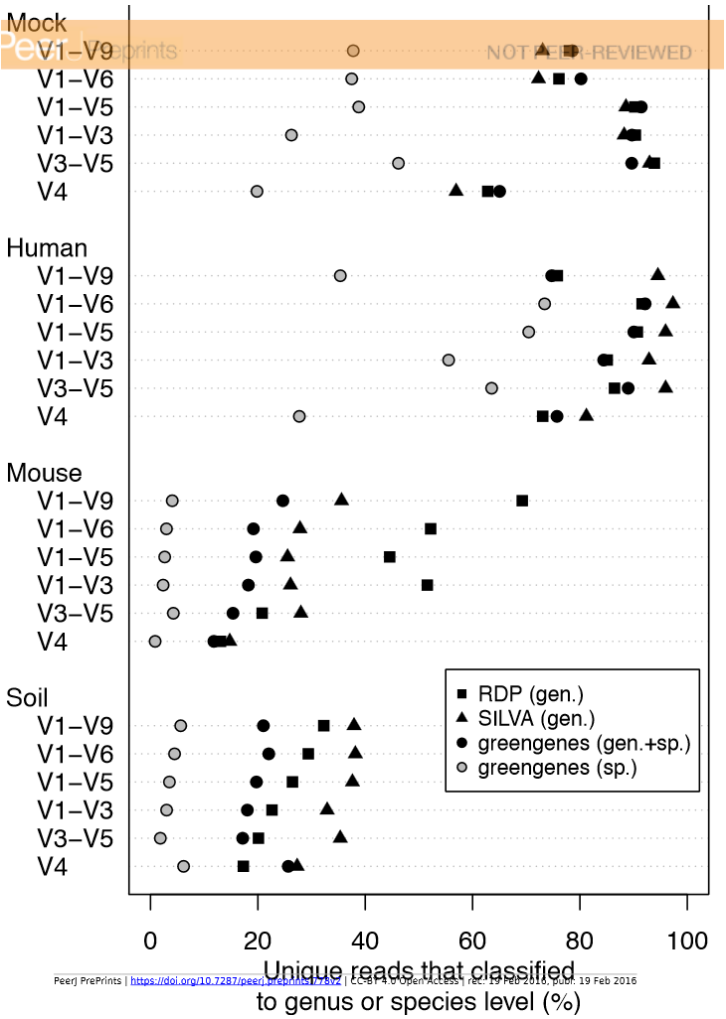
Base Modification: Discover the Epigenome



Detect base modifications using the kinetics of the polymerization reaction during normal sequencing

PacBio a 16S rRNA





PeerJ PrePrints NOT PEER-REVIEWED

https://peerj.com/preprints/778.pdf

PeerJ PrePrints | https://doi.org/10.7287/peerj.preprints.778v2 | CC-BY 4.0 Open Access | rec: 19 Feb 2016, publ: 19 Feb 2016

NanoPore

	MinION	PromethION	
			
	Mk 1 MinION	Single PromethION Flow Cell	PromethION (48 Flow Cells)
			
Availability	Commercially available since May 2015. Start using MinION	Register for PEAP	Register for PEAP
Number of channels available for sequencing	Up to 512	Up to 3,000	Up to 144,000

NanoPore

Your Sample

Sample input Requirement PCR Free[1]	10pg - 1µg	10pg - 1µg	10pg - 1µg
Flow cell input volume[2]	50-160µl	35µl per sample well (4 wells in a flow cell)	35µl per sample well (192 wells in a PromethION)
Sample preparation time 1D[3]	15 minutes	15 minutes	15 minutes
Sample preparation time 2D[3]	90 minutes	90 minutes	90 minutes

Sample preparation can quickly prepare linear reads (1D) by ligating on a single adaptor or generate 2D reads with a leader adaptor and Hairpin (joins both the template and complement strands together) giving the system two looks at each region of the DNA

NanoPore

System Operation

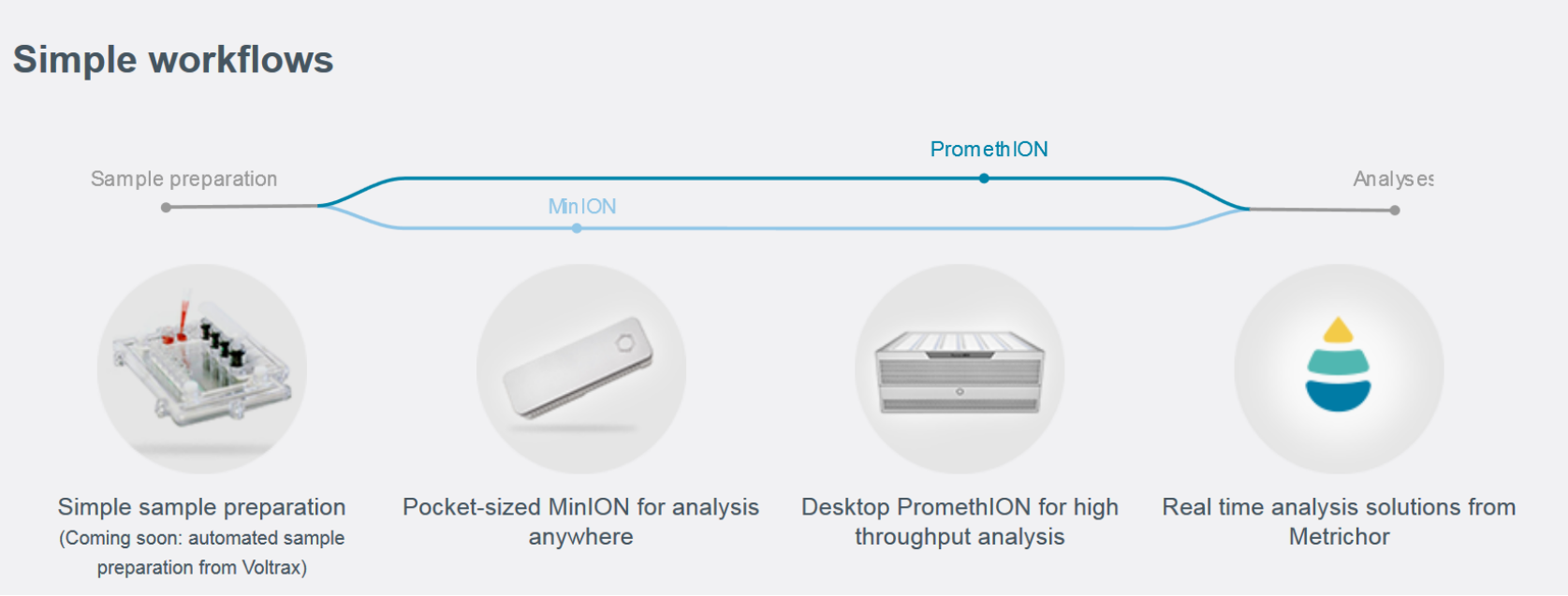
Run time[4]	1 minute - 48 hours	1 minute - 48 hours	1 minute - 48 hours
Flow cell lifetime[4]	~72hrs	>= 72hrs	>= 72hrs
Time to first usable read (data available in real time)	2 minutes	2 minutes	2 minutes
Number of reads at 10Kb at standard speed (70bps)[4]	Up to 600,000	N/A	N/A
Number of reads at 10kb in Fast Mode (500bps)[4]	Up to 4.4M	Up to 26M	Up to 1250M
Read Length	Read length = fragment length Longest reported between 230-300 Kilobases (1D)	Read length = fragment length Longest reported between 230-300 Kilobases (1D)	Read length = fragment length Longest reported between 230-300 Kilobases (1D)
1D Yield[5] at 70 bps in 48 hours	Up to 6 Gb	N/A	N/A
1D Yield[5] at 500 bps in 48 hours	up to 42 Gb	up to 256 Gb	up to 12 Tb
Base calling accuracy[6]	up to 96%	up to 96%	up to 96%
Raw data available	Yes	Coming soon	Coming soon
Modified Base Detection	Yes	Yes	Yes
Data Analysis	Cloud Based	On Unit / On line	On Unit / On line

NanoPore

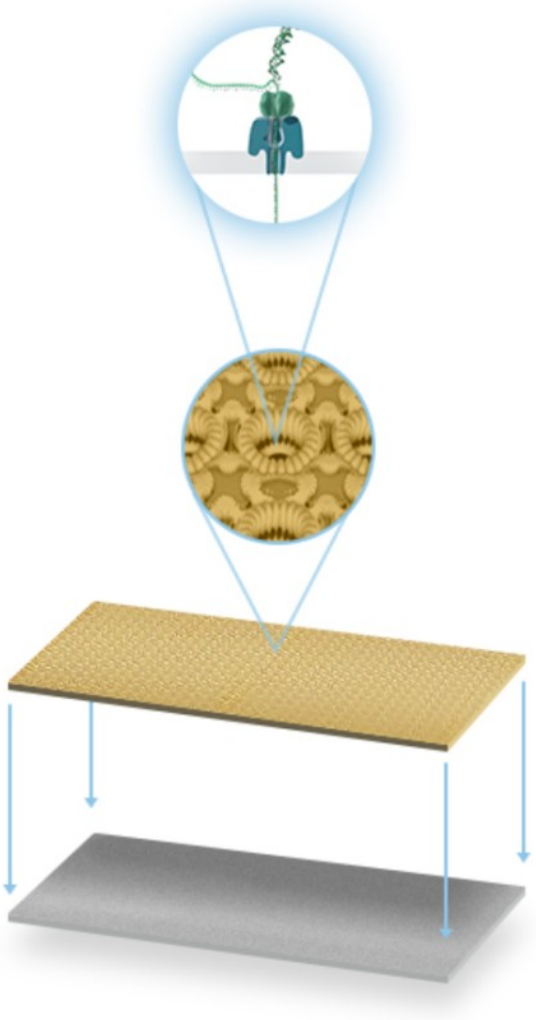
Price on Application

Annual service contract[7]	None	None	None
Reagent cost per run[8]	$ 99	POA	POA
Flow Cell Cost (depending on order type and volume)[8]	$270 - $900	POA	POA
Instrument Access Fee	Starter kit (includes MinION and all materials for 2 runs) $1000	No instrument cost - $75K deposit to be called off against consumable purchases	No instrument cost - $75K deposit to be called off against consumable purchases

NanoPore


Simple workflows
Sample preparation
PromethION
MinION
Analyses
Simple sample preparation
(Coming soon: automated sample preparation from Voltrax)
Pocket-sized MinION for analysis anywhere
Desktop PromethION for high throughput analysis
Real time analysis solutions from Metrichor


Flowcell
MinION Mkl device
USB port



Nanopore

A protein nanopore is set in an electrically resistant polymer membrane. An ionic current is passed through the nanopore by setting a voltage across this membrane. If an analyte passes through the pore or near its aperture, this event creates a characteristic disruption in current. Measurement of that current makes it possible to identify the molecule in question. For example, this system can be used to distinguish between the four standard DNA bases G, A, T and C, and also modified bases.

Array of Microscaffolds

An array of microscaffolds holds the membrane in which the nanopore embedded. This keeps the membrane stable during shipping and usage.

Array Chip

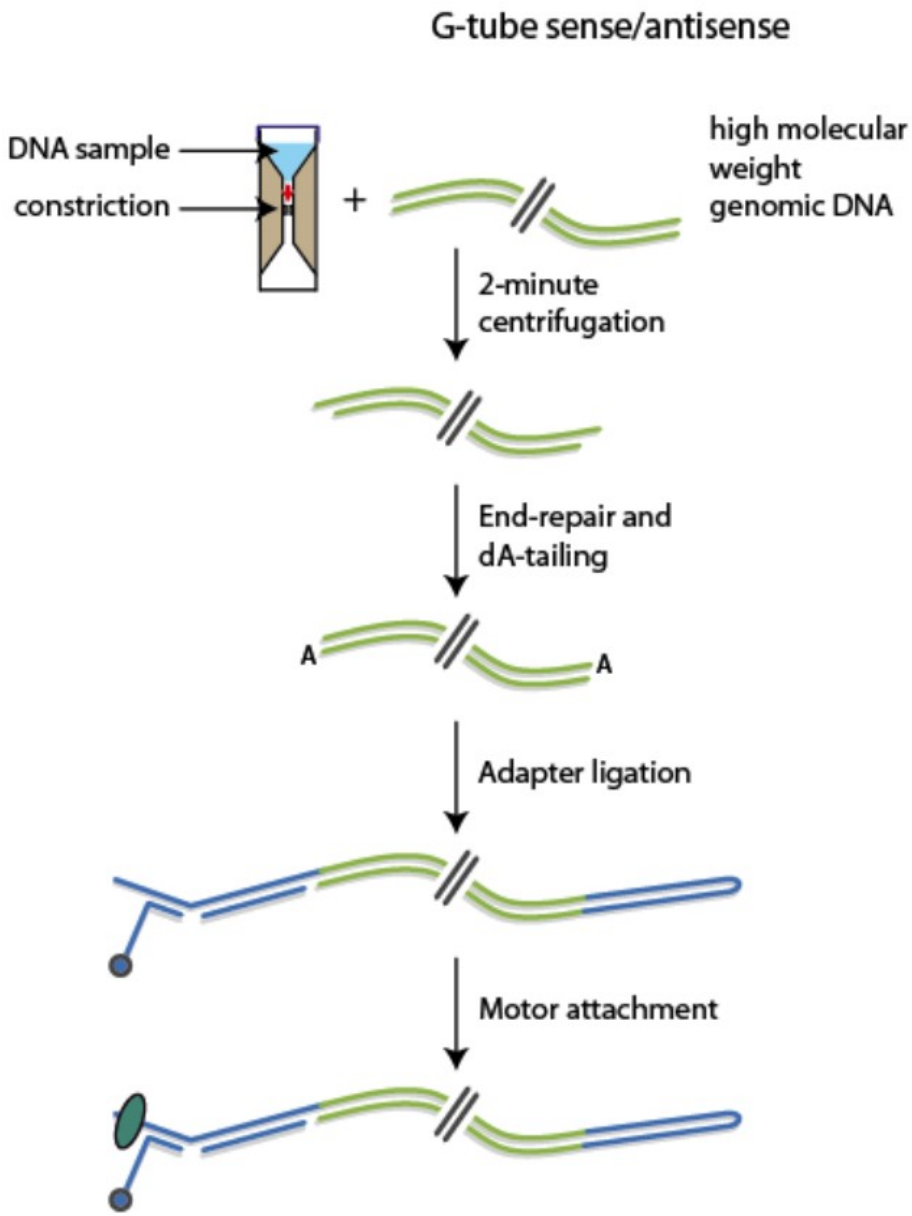
Each microscaffold on the sensor array chip contains an individual electrode, allowing for multiple nanopore experiments to be performed in parallel. Sensor arrays may be manufactured with any number of channels.

ASIC

Each nanopore channel is controlled and measured by an individual channel on a corresponding, bespoke Application Specific Integrated Circuit (ASIC). More than one ASIC may be included in a device and Oxford Nanopore is building ASICs of different sizes for different purposes.

https://nanoporetech.com/applications/dna-nanopore-sequencing

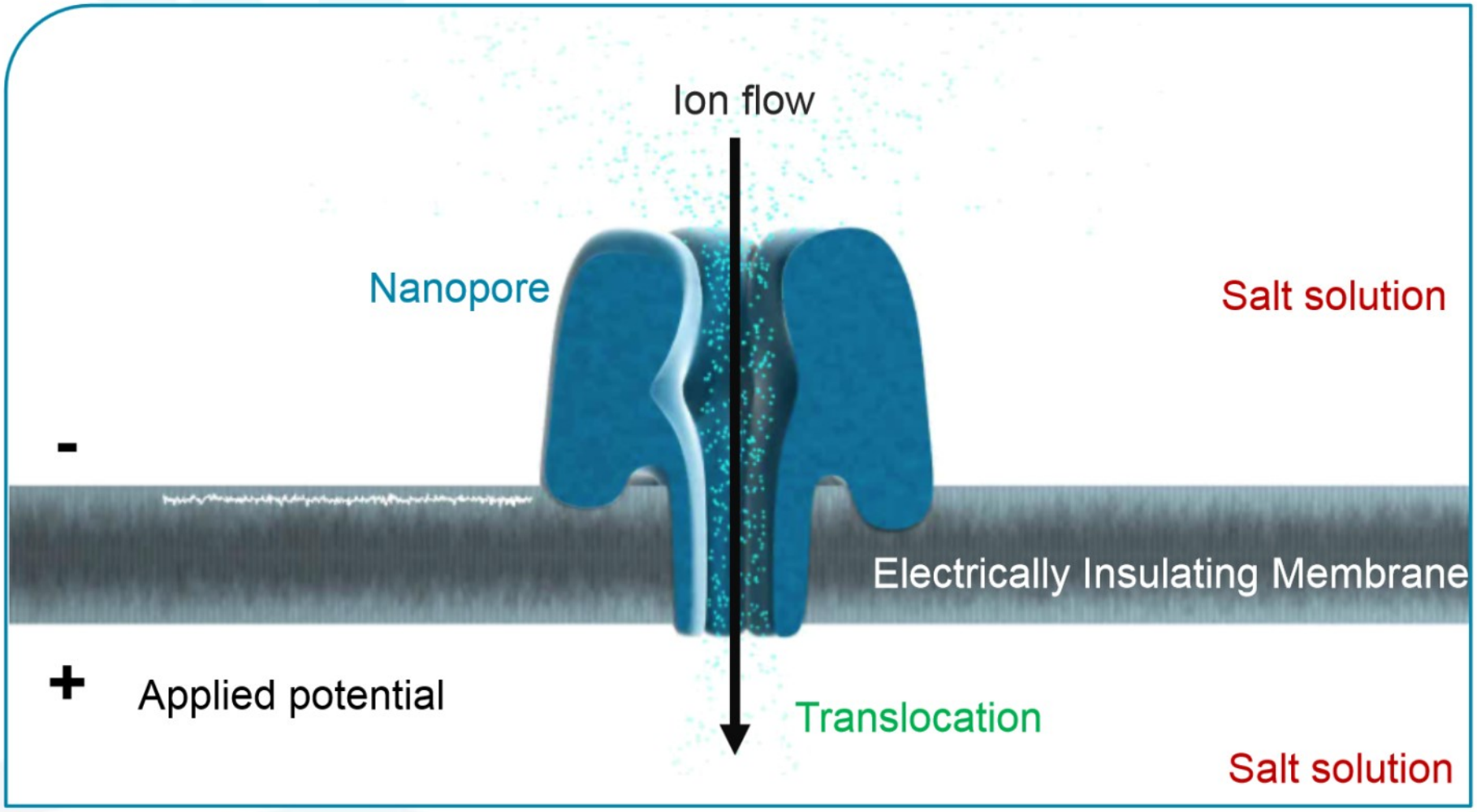
https://www.youtube.com/watch?v=AR6A-zZgX6E



Příprava knihovny MinION

NanoPore

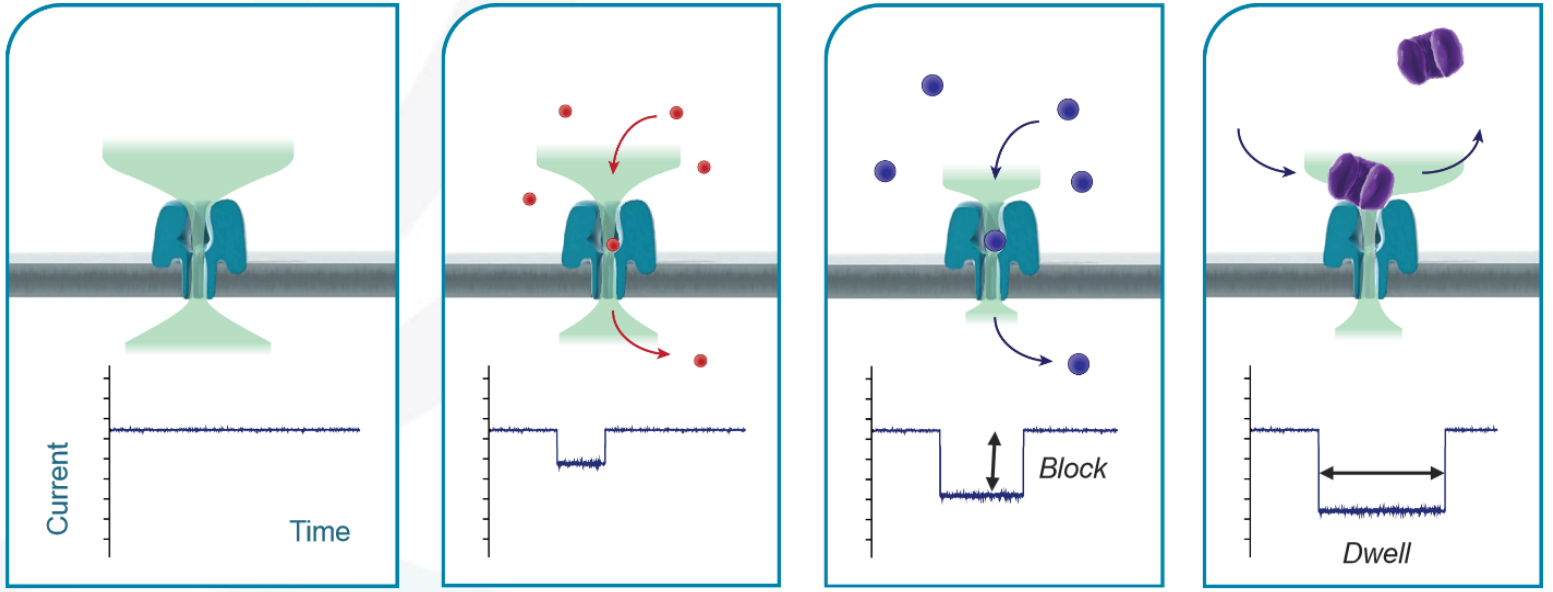
- Nanopore = ‘very small hole’



NanoPore

Nanopore Sensing Summary

- Nanopore = ‘very small hole’
- Ionic current flows through the pore
- Introduce analyte of interest into the pore
  - Identify target analyte by the characteristic disruption or block to the electrical current
  - Block or ‘State’, Dwell, Noise



MinION a metagenomika

Greninger *et al. Genome Medicine* (2015) 7:99
DOI 10.1186/s13073-015-0220-9

Genome Medicine

METHOD Open Access

Rapid metagenomic identification of viral pathogens in clinical samples by real-time nanopore sequencing analysis

Alexander L. Greninger[1,2], Samia N. Naccache[1,2†], Scot Federman[1,2†], Guixia Yu[1,2], Placide Mbala[3,6], Vanessa Bres[4], Doug Stryke[1,2], Jerome Bouquet[1,2], Sneha Somasekar[1,2], Jeffrey M. Linnen[4], Roger Dodd[5], Prime Mulembakani[6], Bradley S. Schneider[6], Jean-Jacques Muyembe-Tamfum[3], Susan L. Stramer[5] and Charles Y. Chiu[1,2,7*]

Abstract

We report unbiased metagenomic detection of chikungunya virus (CHIKV), Ebola virus (EBOV), and hepatitis C virus (HCV) from four human blood samples by MinION nanopore sequencing coupled to a newly developed, web-based pipeline for real-time bioinformatics analysis on a computational server or laptop (MetaPORE). At titers ranging from 10^7–10^8 copies per milliliter, reads to EBOV from two patients with acute hemorrhagic fever and CHIKV from an asymptomatic blood donor were detected within 4 to 10 min of data acquisition, while lower titer HCV virus (1×10^5 copies per milliliter) was detected within 40 min. Analysis of mapped nanopore reads alone, despite an average individual error rate of 24 % (range 8–49 %), permitted identification of the correct viral strain in all four isolates, and 90 % of the genome of CHIKV was recovered with 97–99 % accuracy. Using nanopore sequencing, metagenomic detection of viral pathogens directly from clinical samples was performed within an unprecedented <6 hr sample-to-answer turnaround time, and in a timeframe amenable to actionable clinical and public health diagnostics.

Background

Acute febrile illness has a broad differential diagnosis and can be caused by a variety of pathogens. Metagenomic next-generation sequencing (NGS) is particularly attractive for diagnosis and public health surveillance of febrile illness because the approach can broadly detect viruses, bacteria, and parasites in clinical samples by uniquely identifying sequence data [1, 2]. Although currently limited by sample-to-answer turnaround times typically exceeding 20 hr (Fig. 1a), we and others have reported that unbiased pathogen detection using metagenomic NGS can generate actionable results in timeframes relevant to clinical diagnostics [3–6] and public health [7, 8]. However, timely analysis using second-generation platforms such as Illumina and Ion Torrent has been hampered by the need to wait until a sufficient read length has been achieved for diagnostic pathogen identification, as sequence reads for these platforms are generated in parallel and not in series.

Nanopore sequencing is a third-generation sequencing technology that has two key advantages over second-generation technologies – longer reads and the ability to perform real-time sequence analysis. To date, the longer nanopore reads have enabled scaffolding of prokaryotic and eukaryotic genomes and sequencing of bacterial and viral cultured isolates [9–13], but the platform's capacity for real-time metagenomic analysis of primary clinical samples has not yet been leveraged. As of mid-2015, the MinION nanopore se-

https://genomemedicine.biomedcentral.com/articles/10.1186/s13073-015-0220-9

Shrnutí rozdílů 1, 2 a 3 generace sekvenování

https://www.youtube.com/watch?v=_ApDinCBt8g