Green Mountain DNA Forensics Conference
July 24, 2023

Technologies of the Rapid Changing Landscape of Next Generation DNA Sequencing

Scott Tighe

Technical Director
UVM Genomics Core
Center for Biomedical Research
Firestone 150
Burlington Vermont

Outline

- Background
- UVM Genomics
- Resources
- Current Sequencing Technologies
- New and Emerging Sequencing
- Plug and Play Automation
- Low Input Biomass Challenges
- Controlling Genomic Noise in NGS sequencing
- Other New Technologies

Background

- 20 years in managing a genomics core
- Started Microbiology in 1981 and Genomics in 1995
- ABRF Metagenomics & Microbiome Research Group
- International Microbiome and Multi-Omics Standards Alliance (NIST)
- Genomics Standards Consortium
- Extreme Microbiome Project
- NASA ISS DNA grant and DOD DOA grant
- Extremophiles and difficult to extract samples
- Product development

Genomic Metadata Standards

Checklist	Description
MIGSEukaryote	Minimal Information about a Genome Sequence: eukaryote
MIGSBacteria	Minimal Information about a Genome Sequence: cultured bacteria/archaea
MIGSPlant	Minimal Information about a Genome Sequence: plant
MIGSVirus	Minimal Information about a Genome Sequence: cultured bacteria/archaea
MIGSOrg	Minimal Information about a Genome Sequence: org
MIMS	Metagenome or Environmental
MIMARKSSpecimen	Minimal Information about a Marker Specimen: specimen
MIMARKSSurvey	Minimal Information about a Marker Specimen: survey
MISAG	Minimum Information About a Single Amplified Genome
MIMAG	Minimum Information About a Metagenome-Assembled Genome
MIUVIG	Minimum Information About an Uncultivated Virus Genome

http://w3id.org/mixs

- Agriculture
- QuantityValue
- Food-farmEnvironment
- Food-foodProductionFacility
- Food-animalAndAnimalFeed
- Food-humanFoods
- Symbiont-associated
- Water
- WastewaterSludge
- Soil
- Sediment
- Plant-associated
- MiscellaneousNaturalOrArtificialEnvironment
- MicrobialMatBiofilm
- HydrocarbonResources-fluidsSwabs
- HydrocarbonResources-cores
- Human-vaginal
- Human-skin
- Human-oral
- Human-gut
- Human-associated
- Host-associated
- BuiltEnvironment
- Air

Plus 87 additional

- XMP

ABOUT
TEAM
PROJECTS
METHODS
MEETINGS
PARTNERS
RESOURCES
CONTACT

Extreme Microbiome Project

Profiling Novel and Extreme Environments using Advanced Genomic and Microbiological Technologies




UVM Genomics Core

- Integrated Center –Genomics, Bioinformatics, Proteomics, Imaging, SEM, TEM, Confocal, Flow Cytometry, Spatial
- 21 Faculty and Staff
- 3 Room Genomics Lab- DNA-free sample prep, low amp, high amp

Full service core

DNA, RNA, FFPE extraction
Quantification and QA/QC
Sanger Sequencing
Fragment Analysis-CE
DNA and RNA sequencing
PCR and Ultra-Low amplification
Gene Expression, CNV, HiC, ATAC, Methylation
Microbiome metagenomics
Single cell and Spatial sequencing
Culturing/microscopy
High Volume liquid genomic analysis
Gel doc Imaging
Advanced PCR trouble shooting

Bleach Stations, Nitrogen tanks

Instruments

Singular G4 NGS (660gbases/run)
Illumina NGS
- MiSeq
- MiniSeq
- HiSeq 1500

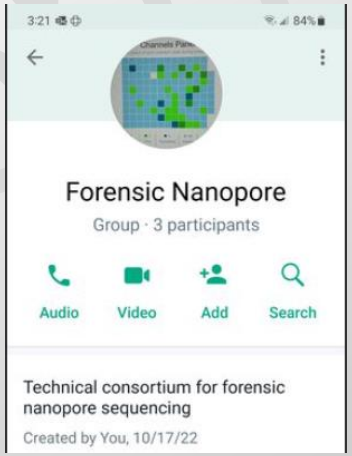
Oxford Nanopore
- 3 MinIONS, 1 Nanopore MK1C
- 1 Nanopore P2 Solo PromethION with RTX4090 GPU
- 1 Nanopore GridION, 4 Flongle adaptors

ddPCR-BioRad QX200
10x Genomics Single Cell system
RTqPCR (QuantStudio 6F, ABI7500 fast)
1 Nanodrop, 4 Qubit fluorometers, 1 Quantus Fluorometer
BioAnalyzer Fragment Analyzer-2100
Covaris S2 Hydroshear
ChemiDoc XRS+ Photodocumentation system
2 Bead Beater systems for extractions
12 thermocyclers, 4 PCR hoods, Speedvac

User Resources and Forums

- WhatsApp for DNA forensics-Open discussion on techniques
- WhatsApp for Nanopore and other sequencing- Nanopore and Long Read Challenges

Please Let me know if you would like to be added

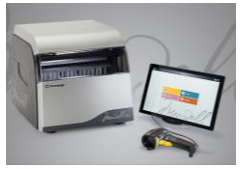


Current Next Gen Sequencing Technologies

Why is NGS important

Sanger vs MPS Workflow Comparison – Whole Mito Analysis

DNA Extraction

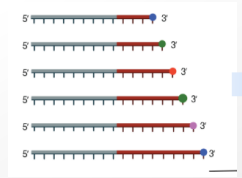


Human DNA Quantification

Sanger Seqeuncing

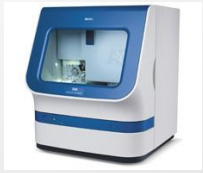
Single Target Amplification

Multiple long-range amplifications per sample



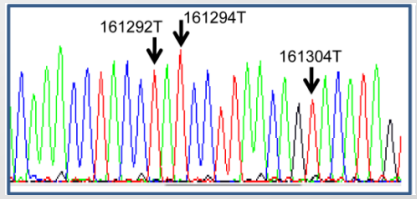
Sanger Sequencing

> 100 reactions for each sample



Separation and detection on CE

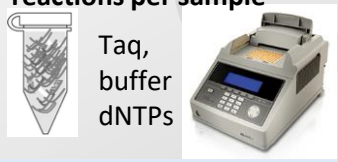
Each reaction sequenced separately



Data analysis

Non-quantitative, difficult deconvolution of mixture/heteroplasmy

Short amplicons in 1-2 reactions per sample



Multiplex Amplification

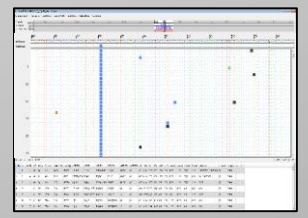
Entire sample library prepared in a single well

Library Preparation

Multiple Samples pooled and sequenced simultanously

MPS Sequencing

Data displayed as individual reads – quantitative heteroplasmy and mixture assessment



Data Analysis

Massively Parallel Sequencing

Nanopore

Current Sequencing Technologies

Sanger Sequencing and CE

- Promega Spectrum (6 and 8 color)
- Life Technologies SeqStudio (4, 8,16 caps)
- ABI 3500 and 3730 (96 caps)
- Standard CE technology
- ProDye, BrightDye, MagaDye, BrilliantDye, SupreDye, BigDye, Gerbera v3.1, QuantumDye, and DYEnamic ET.

Massively Paralleled Sequencing (NGS) Short Read Sequencing Technologies

Illumina MiSeq/Verogen

- Low output short read NGS sequencer
- Single or Paired end sequencing
- Approved for both Clinical and Forensics

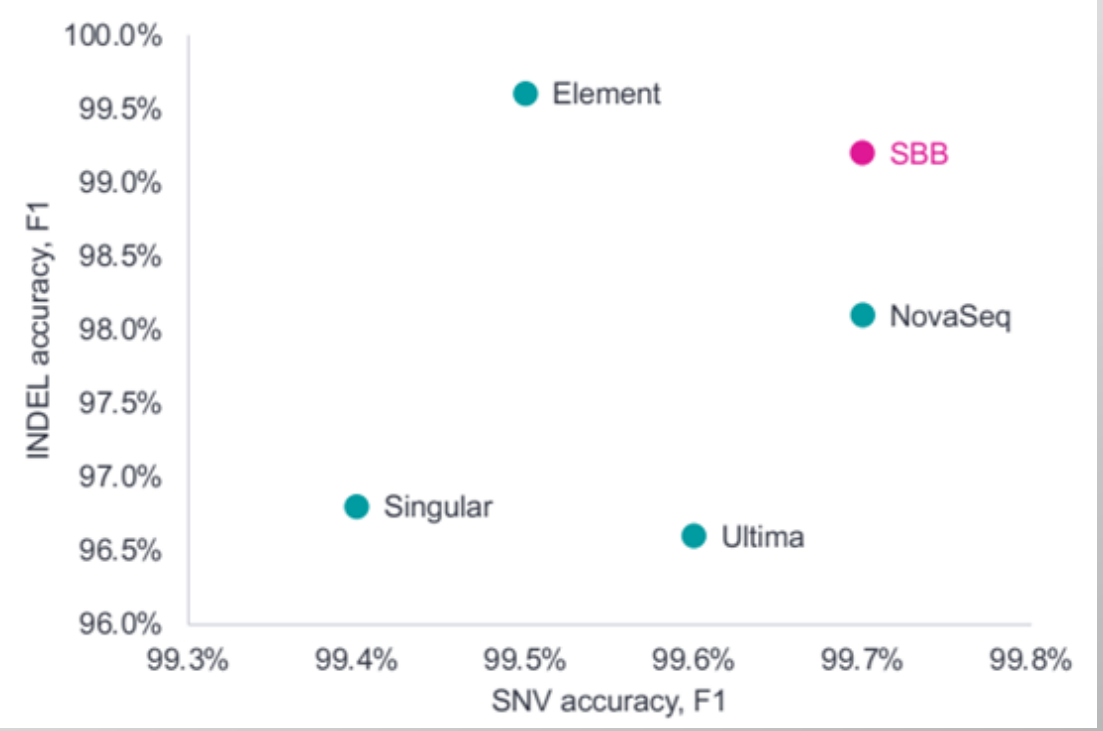
ThermoFisher Scientific

- Ion Torrent PGM series continues
- S5
- GeneStudio and Genexus sequencers

New and Emerging Sequencing

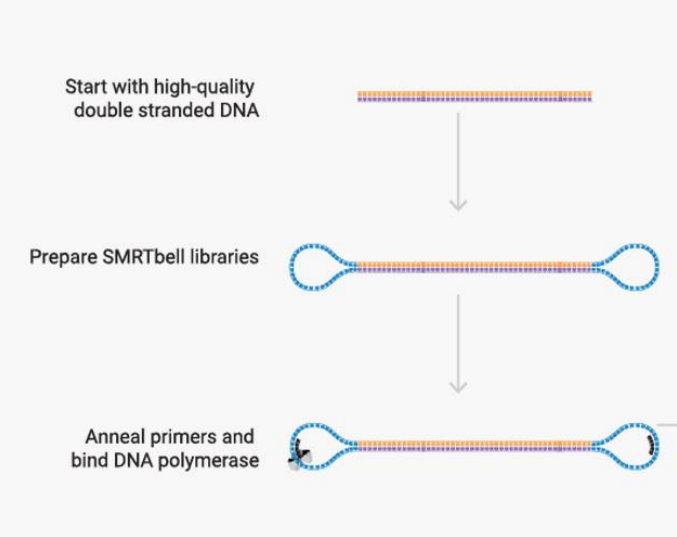
New NGS systems are Rapidly Coming to Market

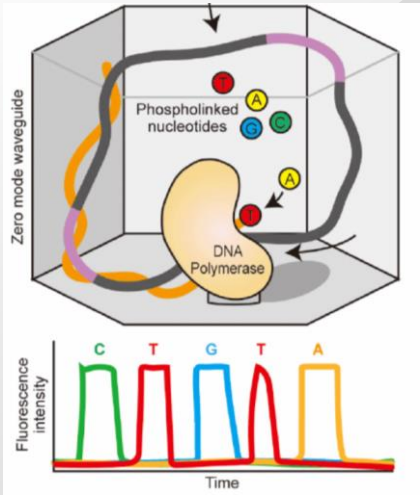
- PacBio Revio
- PacBio Onsu
- Singular G4
- Element Biosciences Aviti
- Ultima
- Illumina NovaSeq X

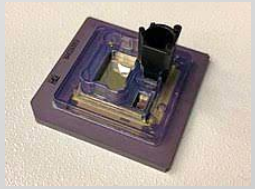


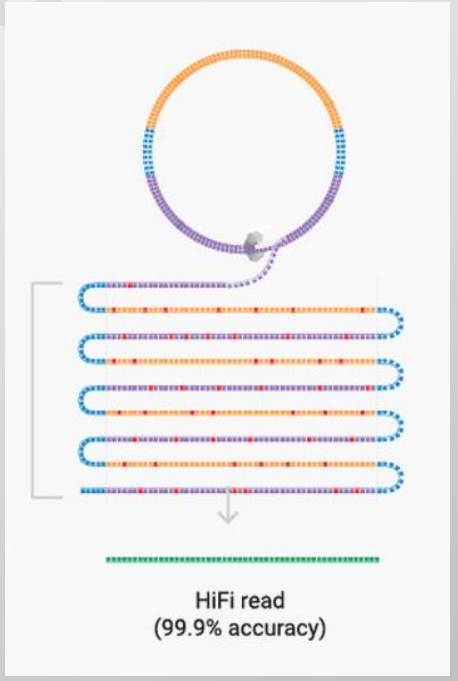
Pacific Biosciences Revio System

- Nanofabricated Revio SMRT Cells 25 million zero-mode waveguide wells
- 4 stages run 4 SMRT flow cells at once
- HiFi yield 15–20 kb fragments 90 Gb/SMRT (5mC at CpG sites)
- Q Score approaching 40
- ~$800K



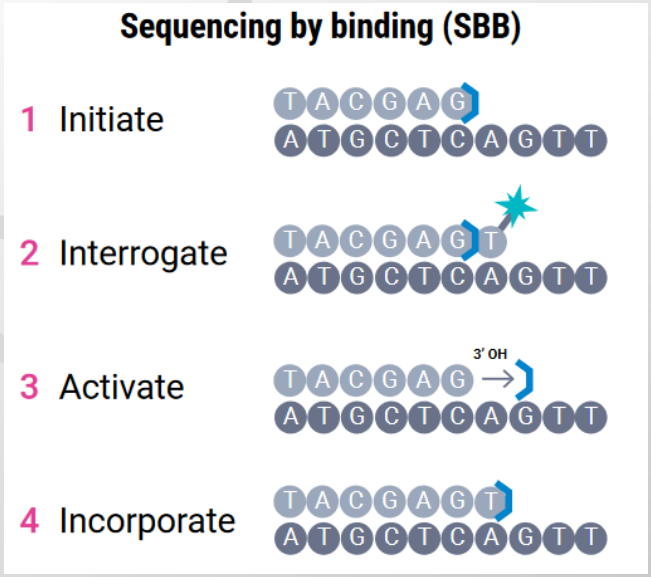


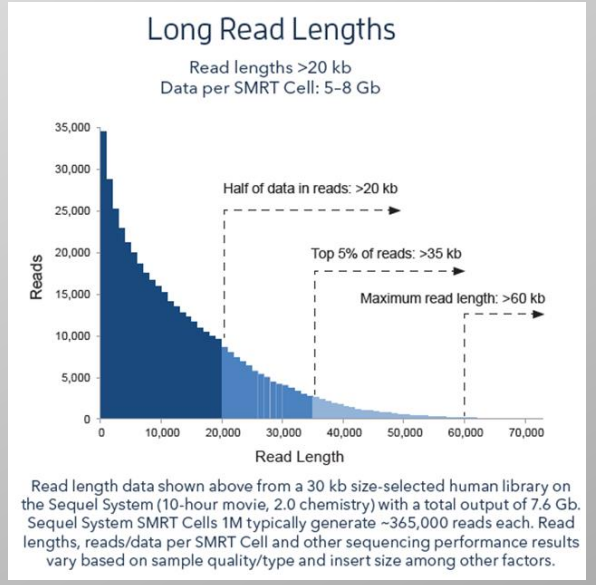




Pacific BioSciences Onsu

- Sequencing by Binding
- Short Read Sequencer
- Formerly Omniome
- 2 x 150 bp (400 mr (PE)120 Gb)
- Q-Scores 35-44
- No Index hopping
- No incorporated scared bases
- 90%+ Q40
- Amplicon linearity to 0.001%
- Beta currently but Q3 2023
- $400K

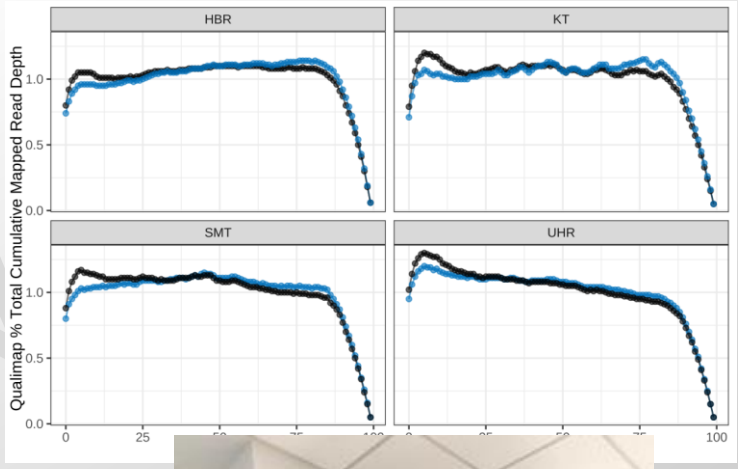




Updates on the Singular Genomics G4

- Short read sequencer similar to Illumina NextSeq and Novaseq
- 1/3 running cost as Illumina
- 15—400 Gb output
- 75-90% Bases ≥ Q30
- Accuracy 99.6 – 99.9%
- Index hopping at 0.07% Very good

- 1 lane (50mr or 100mr) $150/lane
- 1 to 16 lanes at a time-24hr
- 2.2 billion clusters-Similar to NovaSeq6000
- $250K
- Could Replace the MiSeq



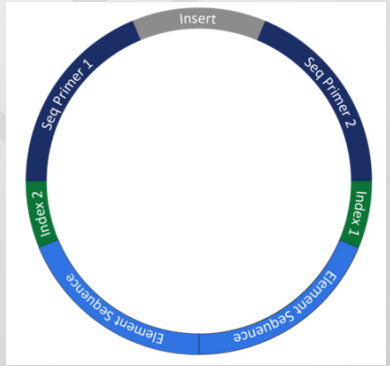
Element Biosciences-AVITI

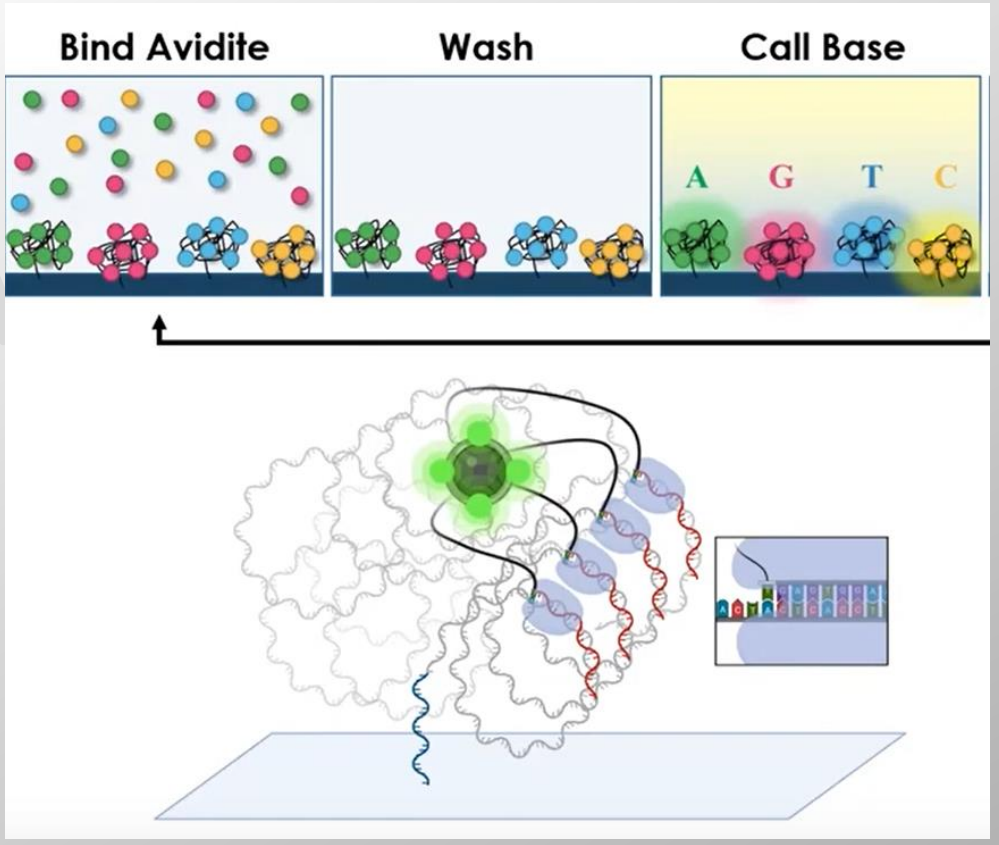
Short Read Paired End MPS sequencing
Error rate 0.005
80-280 gbases
Index hopping 0.01%
Solid State Rolling Circle Amplification (similar to nanoball)
PE 500 maybe PE1000 eventually

Typical run in Chris Lab:

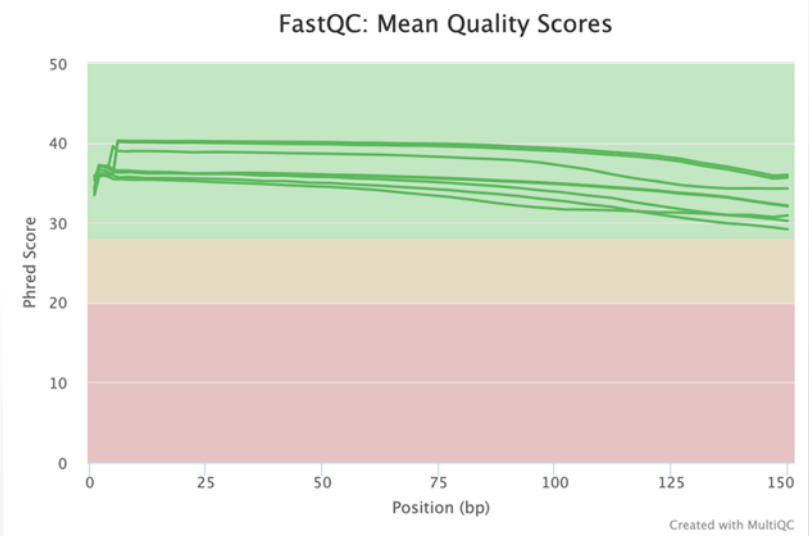
- Reads Generated (million) 1085
- Assigned Reads (million) 1053
- Q30 93.04%
- Mapping 97.03%

- $250K

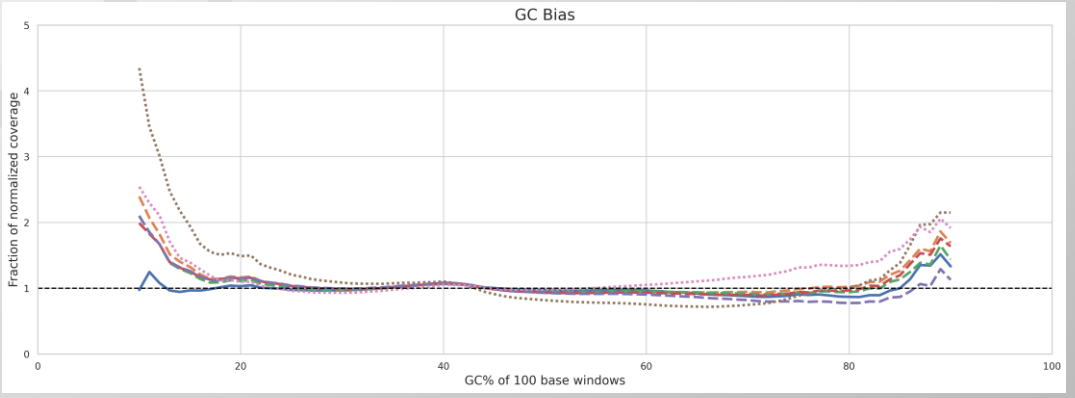




Q-Score vs read length

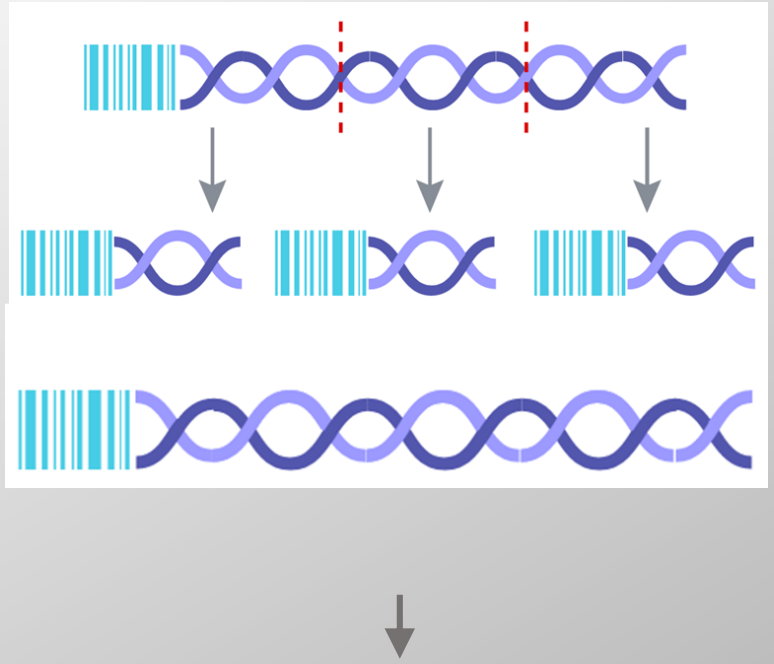


Coverage as a function of GC content



Loop-Seq (Element BioSciences)

- Long Read sequencing with Short read platform
- Amplify a large fragment with one Barcode
- Digest and redistribute same barcode across the digested fragment
- Reassemble computationally
- Simple and clever



Illumina MiSeq, NextSeq2000, NovaSeqX

- MiSeq
  - Still great
  - No upgrades planned
  - Very expensive to run
  - PE300
  - Very flexible running formats
  - User Friendly

	MiSeq Reagent Kit v2	MiSeq Reagent Kit v3	MiSeq Reagent Kit v2 Micro
Single Reads	12–15 million	22–25 million	4 million
Paired-End Reads	24–30 million	44–50 million	8 million

- NextSeq 2000
  - P1, 2, 3 flow cells
  - Sequencing by Synthesis
  - Barcode hopping require UDI and UMI
  - 3 times more expensive than Element and Singular to operate
  - $300K

Configuration	Flow Cell		
	P1	P2	P3
PE150	30gb	120gb	360gb
PE300	60 gb	180 gb	

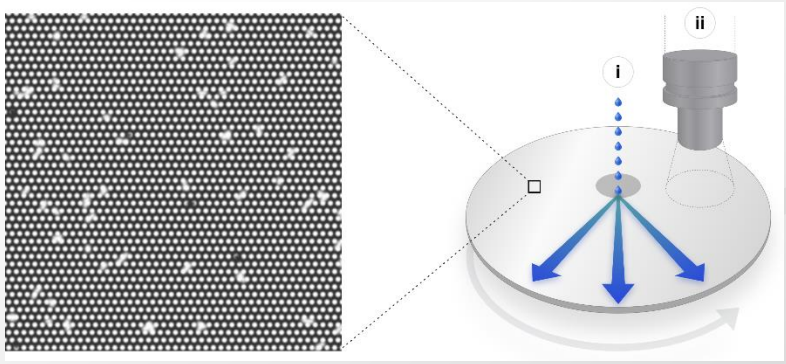
- NovaSeqX (Genome Sequencer)
  - million dollars
  - $10K/Month Service Contract
  - terabase sequencing
  - 150gb-8 tb
  - Not useful for amplicon or panels
  - Max PE150

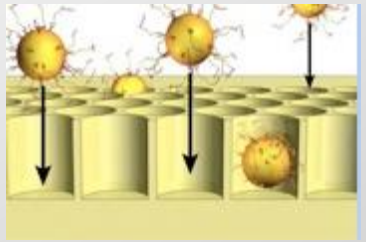
Configuration	Flow Cell		
	1.5	10	25
PE150	500 gb	3000 gb	8000 gb

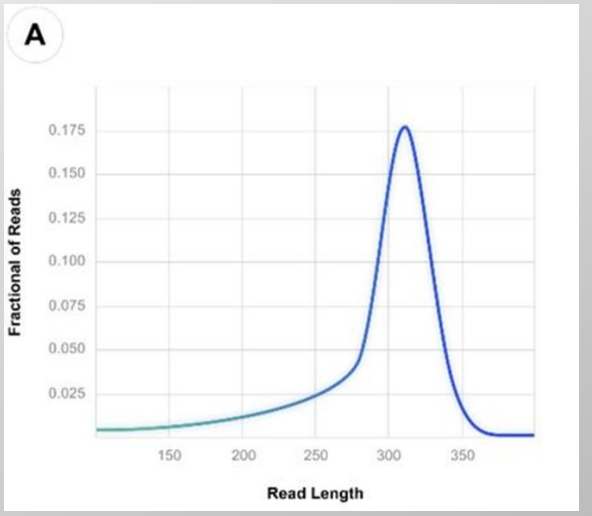
**Illumina has big problems right now

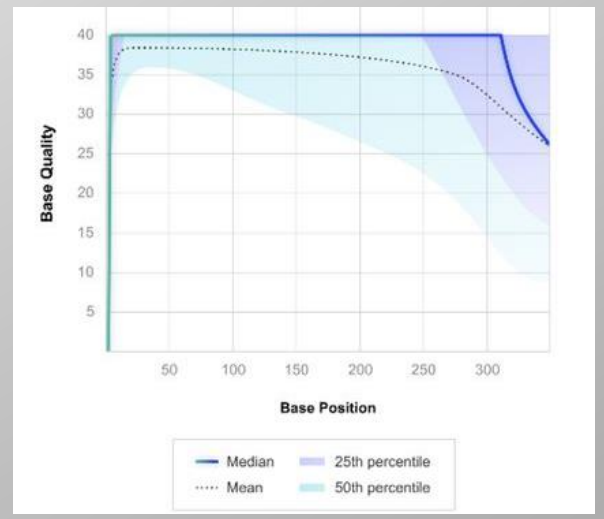
Ultima Genomics UG100 Sequencer

Intended for large Genome Center such as the Broad and NYGC
SE300 bp 3,000 Gb (>3 tbases)
NEBNext Ultra II FS DNA Library
No Flow cells- Uses Spinning Silcon wafer
Library Bead-based sequencing (similar to 454 and Genapsys)
Sequencing-By-Synthesis



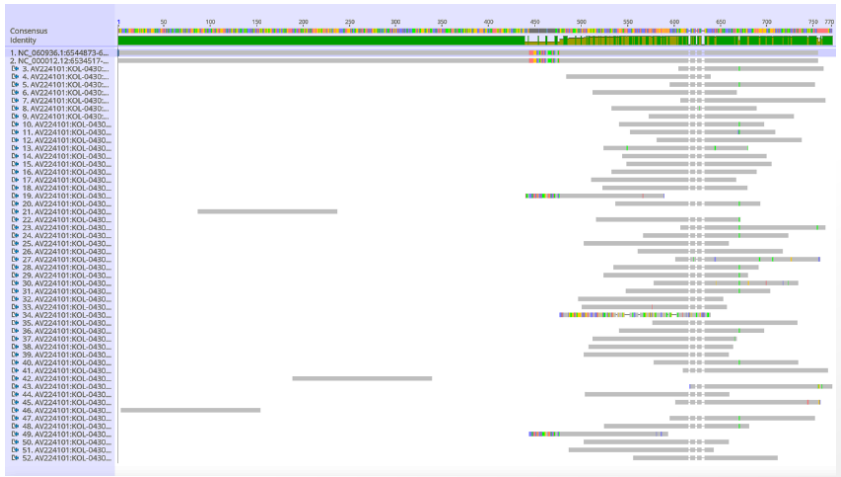






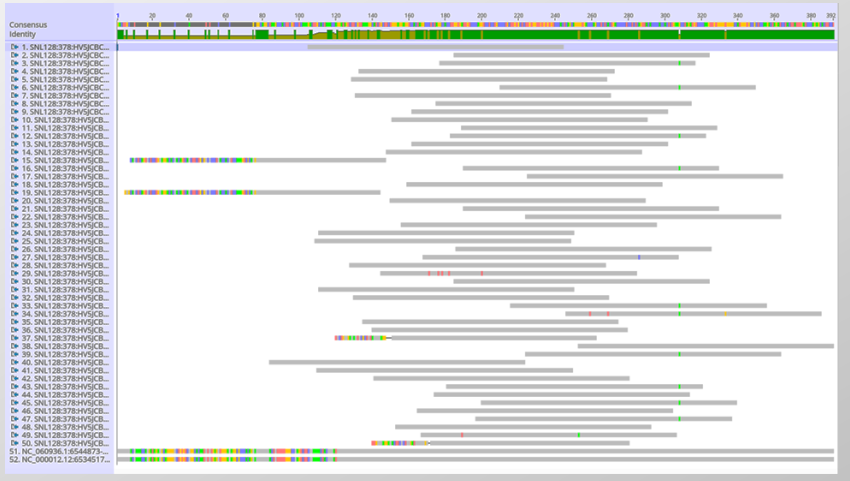
Sequencing Platform Comparisons Using HB RNA

Element Bio Aviti



Singular G4

Illumina HiSeq

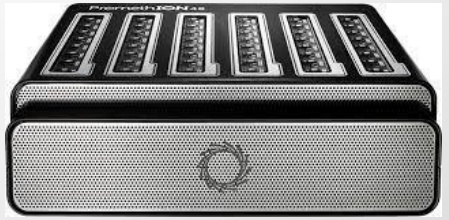


- GAPDH was compared at the same region
- Same sample / Same library
- Images of alignments of 50 Bui gapDH reads

Nanopore Sequencing

- Read Lengths 400bp to 1 x10^6
- Benchtop and Hand held sequencing
- Q scores up to 22 or 26+ with duplexing
- Many library approaches
  - Amplicon
  - Artic
  - Ligation
  - Rapid DNA sequencing
  - DNA base modification detection
  - Direct RNA sequencing
  - RNA base modifications
  - Low input PCR barcoding
  - Native Ligation Barcoding
  - Short read sequencing

GridION
8 gbase/flow cell
10.4.1



PromethION/P2
50 gbase/flow cell
V10 Flow cell

MinION MK1C
8 gbase/flow cell
10.4.1

Rapid Nanopore DNA Screening

ONT

100ul PBS pH7.5 with diamond Beads

Swab surface

Add 4ul MetaPolyzyme ProK 35-55C 20 min

Add 10-100ul of Prepman Ultra™

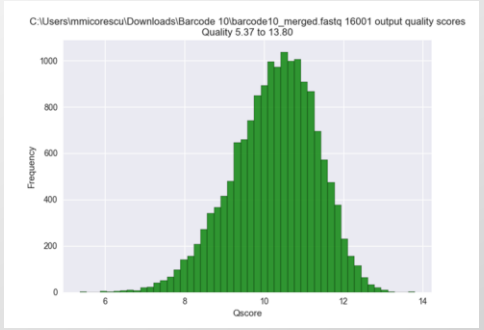
100C 10 min -Vortex

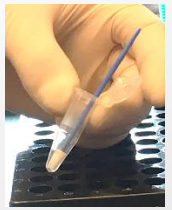
1-4ul input

Use Clontech/Takara Titanium Taq

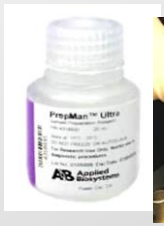
25 cycles

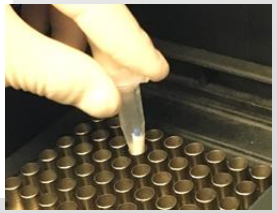
150 ng input to SQK-RAB201





35C 20m





100C 10m

4ul

50ul Titanium Taq PCR

95/1:00
95/30
54/30 25cyc
68/1:30
68/3:00

Ampure Beads

Oxford Nanopore Improvements

- Release of 10.4.1 Flow cell (PromethION and MinION)
  - Increased accuracy to mode Q19 singleplex Ligation LSK114
  - Reduced output in our hands (9.4 to 10.4) (16 gb to 8 gb)
  - Duplex analysis increases accuracy to greater than Q25
  - New Motor

- New version 10 Flongles for low output applications (700mb)

- Adaptive Sampling (sequencing)
  - Ability to selectivity "reject" pre-defined sequences for the Nanopore
  - Great for un-want high background sequencing
    - Human DNA in a microbiome sample
    - Tick DNA in microbiome sample
  - Need a fast computer to keep up with basecalling GPU (280-400 bases/ second)

Gas Sulfur cave

Fast Computer for Rapid Base calling and Adaptive Sampling

At 400 bases/sec,
Must reject sample from pore in 20 bases or
50 milliseconds

ORIGIN
HIGH-PERFORMANCE PCS

Origin PC Quote
Date: Thursday, July 6, 2023
Quote number: 1102054
Valid until: 7/11/2023

Corsair, 115 N. McCarthy Blvd
Milpitas, CA 95035
1-877-674-4460
originpc.com
sales@originpc.com

Customizations/Descriptions:

	QTY	SKU	Unit	Total
	1	L-CLASS	$7,230.20	$7,230.20

Case: Corsair 7000D Airflow
Exterior Color: Black
Processors: Intel Core i9-13900KS 24-Cores 3.2GHz (6.0GHz TurboBoost)
  CPU: Thermal Compound
Motherboard: ASUS ROG MAXIMUS Z790 HERO DDR5
Memory: 96GB CORSAIR VENGEANCE DDR5 (2x48GB) 5600MHz
System Cooling: iCUE H150i LCD ELITE CAPELLIX XT Liquid CPU Cooler
System Fans: ORIGIN PC Standard Fan Kit
Graphics Cards: NVIDIA 24GB GeForce RTX 4090
Operating System: No OS (Installation not supported by Origin PC)
Operating System Drive: Corsair 8TB MP600 PRO XT GEN 4
  RAID: 1
Hard Drive: Corsair 8TB MP600 PRO XT GEN 4
  Hard Drive: 1
Power Supply: CORSAIR AX1600i ATX Titanium
  Power Supply Sleeved Cable: Black Sleeved Cables +90
Audio: Integrated High-Definition Audio
Networking: 10GbE Dual-Port PCIe Network Interface Adapter
The ORIGIN Difference: Unrivaled Quality &Performance - L-Class 7000A

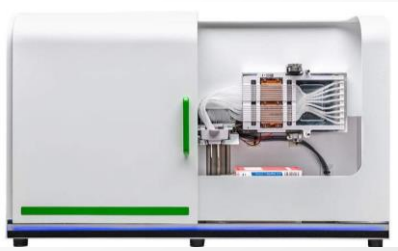
NGS Plug and Play Automation

Extraction → Short Read Library Synthesis → Long read Library Synthesis

MagBio NP16

Magnetic Bead DNA extraction

Revvity Bioqule



NextFlex DNASeq
NextFlex XL
Illumina DNA Prep protocol
Roche KAPA® HyperPrep™ protocol

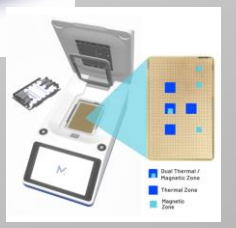
Tecan Magic Prep

Revelo mRNA-Seq
Revelo DNA-Seq Enz

Oxford Nanopore VolTRAX

Works on flat surface for RAD004
Waiting for updates

Miroculus MiroCanvas



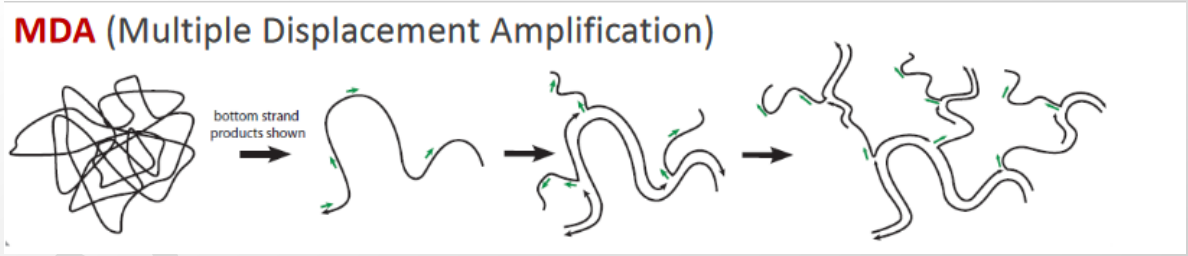
Digital microfluidics platform
Thermal and magnetic zones
Single use cartridge
12 ports for reagents
Any library prep- flexible

Low Input Biomass Library Amplification

Low Input DNA Amplification

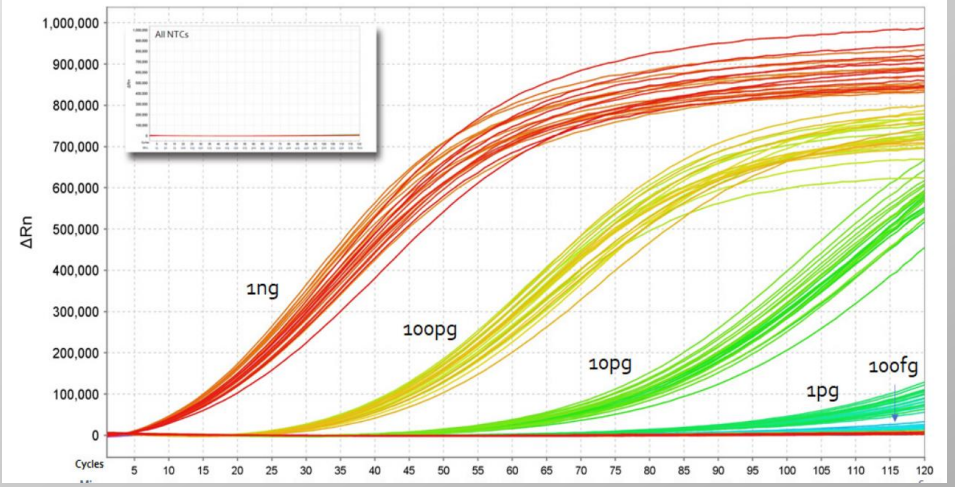
Multiple Displacement Amplification (MDA-Phi29)

- Isothermal down to 10pg
- Phi29 RepliQ polymerase
- Bias are well known
- Results in extensively branched product
- Poor for Short fragment DNA
- T7 Endo I



Primary Template-directed Amplification (PTA)

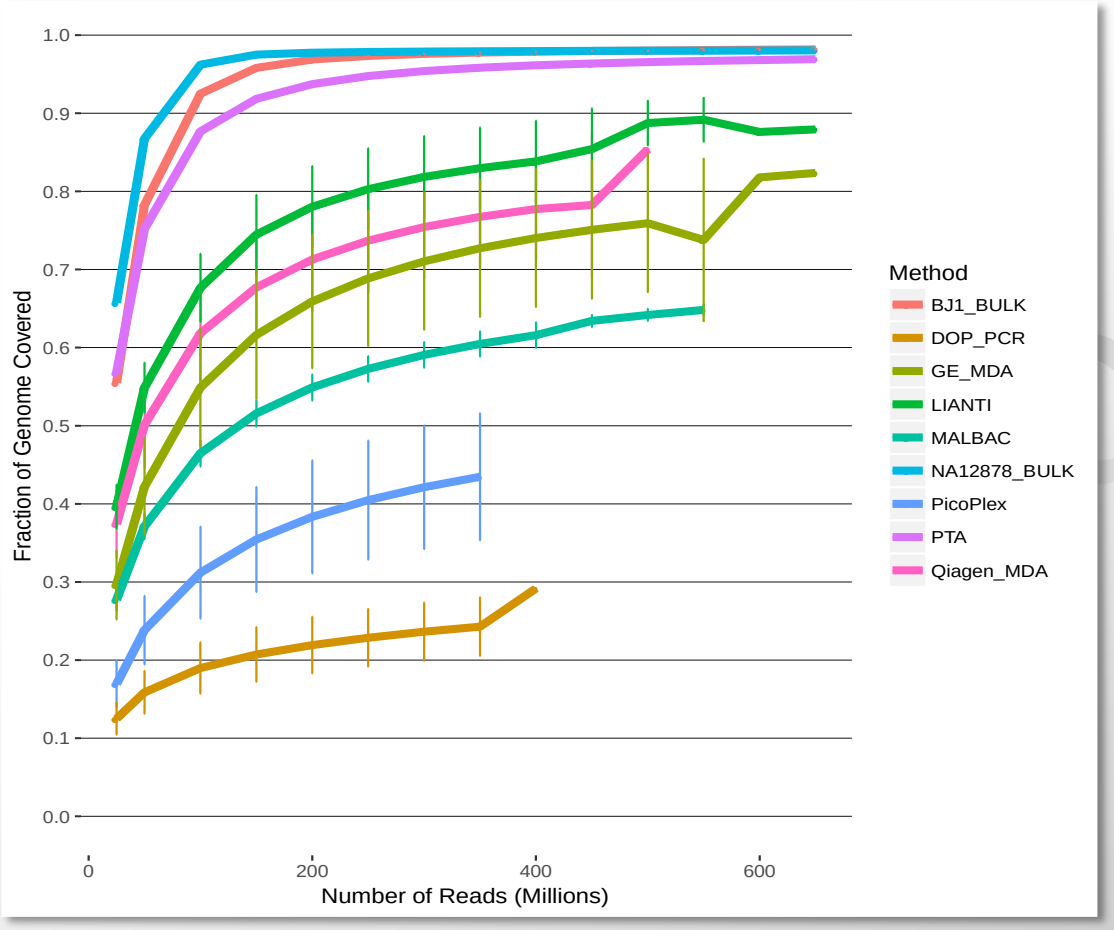
- Isothermal
- Low bias as reported by literature
- 1pg
- 1kb limit



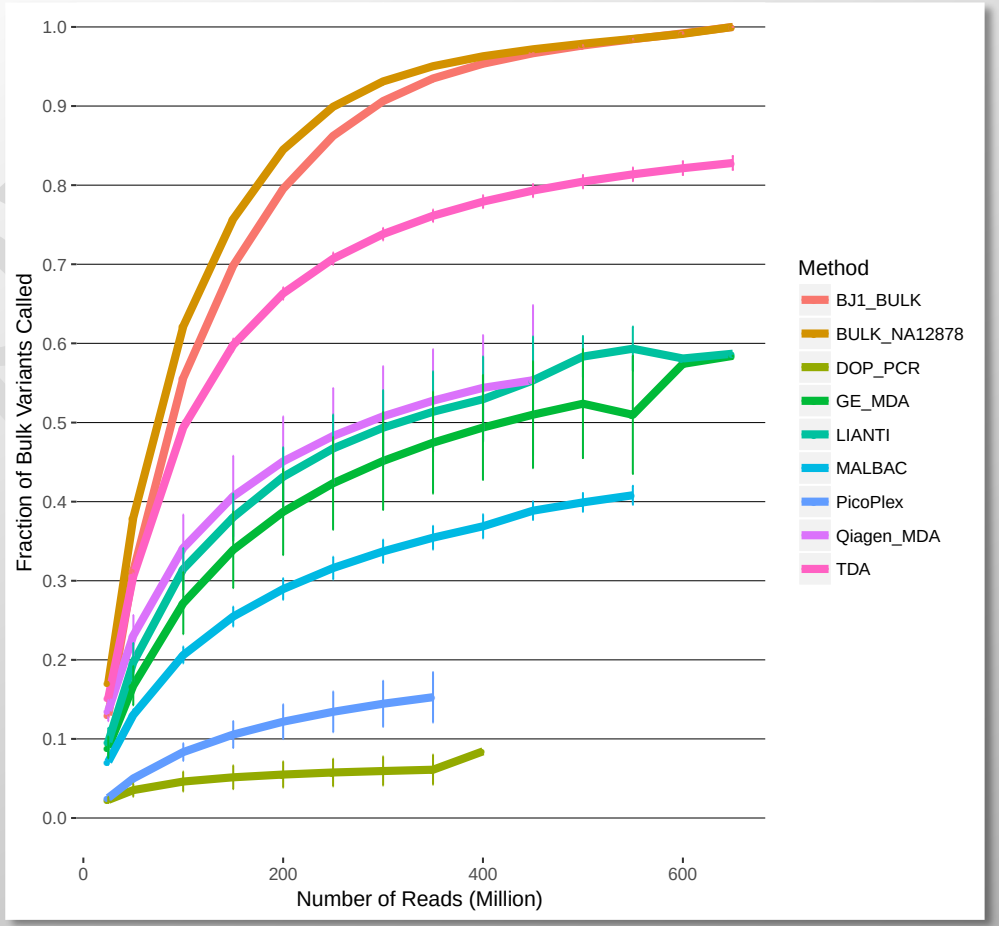
Comparing Low Input Amplifications

Coverage and SNPs

Fraction of Genome covered



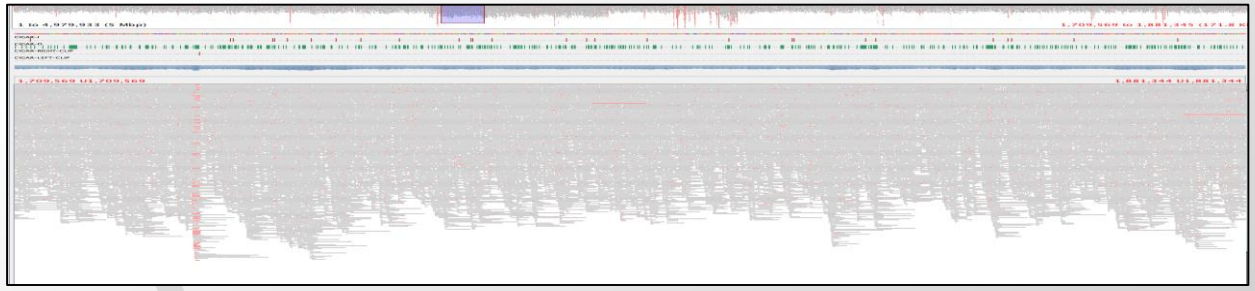
Fraction of SNVs called



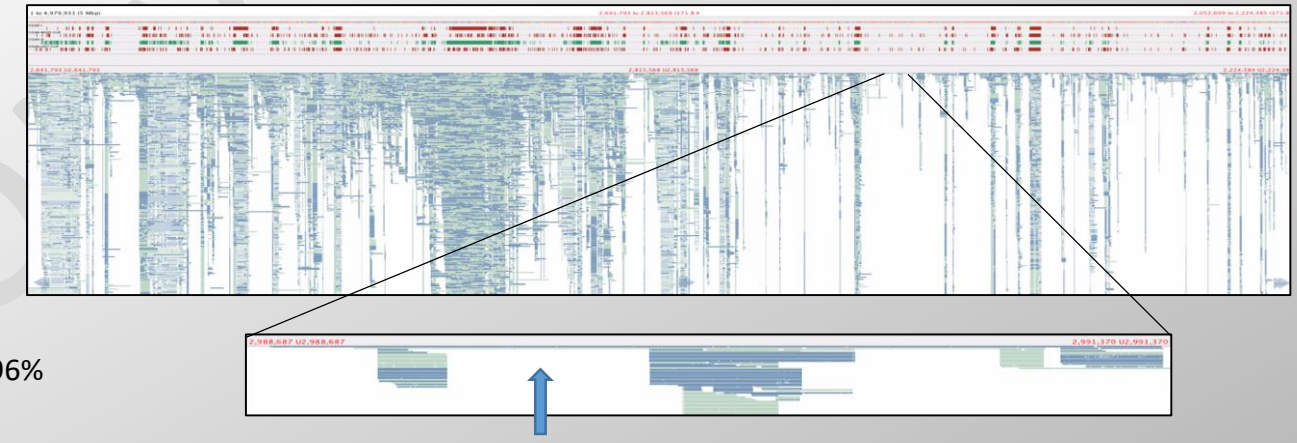
Ultra Low Input Biomass for Nanopore

- Multiple Displacement Amplification
- Ultra low Sample = 200pg
- 5hr Qiagen RepliG
- Result 5ug product
- Debranched 2.5ug T7 Endo I Ratio 8.3uL T7EI: 2.5ug MDA product
- Sequence Nanopore LSK114
- /10.4 flow cell 24hr
- Result 5gb/24hrs

500ng No MDA



200 pg MDA



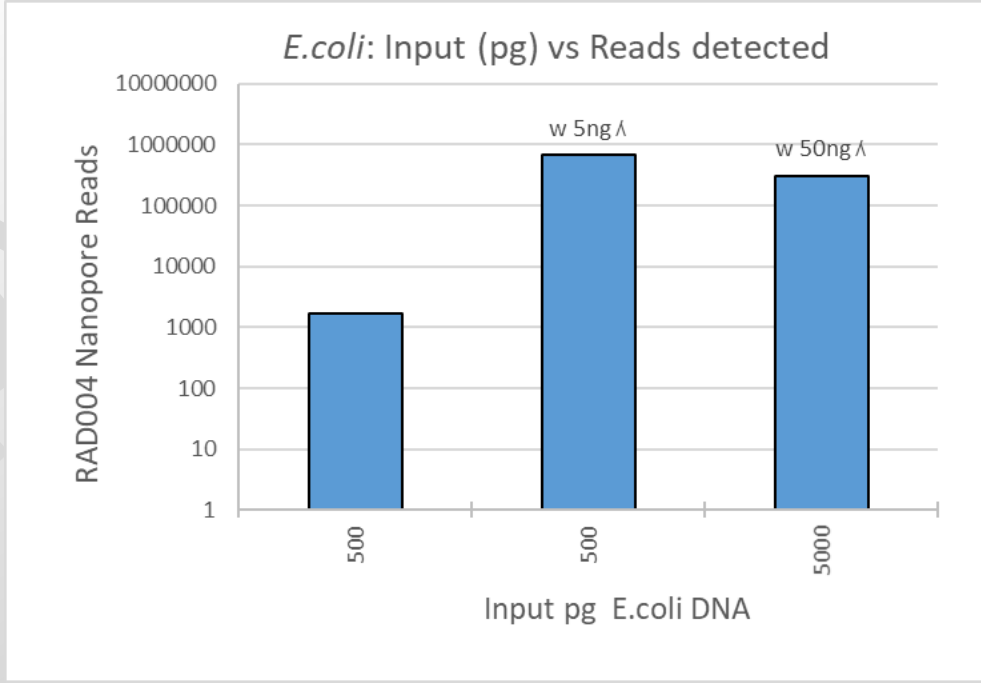
Mycobacterium spp. GC 67.96%

3x coverage

Sample	Reads	Average length	Ave Depth	Max Depth	% Coverage
MDA 16	165501	3600	64	250	71
MDA 17	258790	3600	97	410	80.5
MDA 18	532892	3600	212	570	93
MDA 19	1,169,540	3600	398	1200	98.86

Oxford Nanopore-How long can you go?

- Native DNA detection
  - RAD004-Rapid singe tube
  - Molecular counting but requires carrier DNA
  - Methylation calling
- Rapid PCR Barcoding
  - 200 pg
  - MW limit
- Avoid PCR without dUTP/UNG chemistry
- DNA sequencing



NASA Study- how low can you go
E. coli DNA titration with and without lambda

Controlling Low Level Genomic Noise

Low input Noise and the Kitome are Well Recognized

SUBSCRIBE

SCIENTIFIC AMERICAN

THE SCIENCES MIND HEALTH TECH SUSTAINABILITY EDUCATION VIDEO PODCASTS BLOGS STORE

THE SCIENCES 60-SECOND SCIENCE

Microbiome Studies Contaminated by Sequencing Supplies

By Christopher Intagliata on November 11, 2014

Trends in Microbiology

CellPress REVIEWS

Opinion

Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations

Raphael Eisenhofer[1,2,*], Jeremiah J. Minich[3], Clarisse Marotz[4], Alan Cooper[1,2], Rob Knight[4,5,6] and Laura S. Weyrich[1,2]

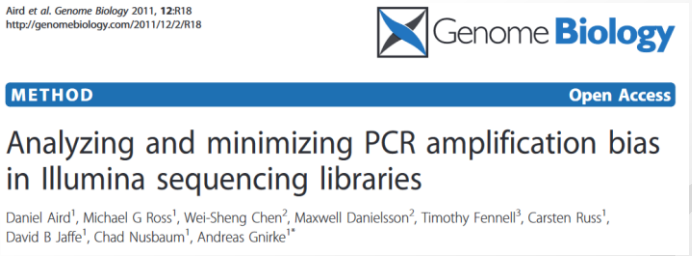
Kim *et al. Microbiome* (2017) 5:52
DOI 10.1186/s40168-017-0267-5

Microbiome

REVIEW Open Access

Optimizing methods and dodging pitfalls in microbiome research

Dorothy Kim[1†], Casey E. Hofstaedter[1†], Chunyu Zhao[1], Lisa Mattei[1], Ceylan Tanes[1], Erik Clarke[2], Abigail Lauder[2], Scott Sherrill-Mix[2], Christel Chehoud[2], Judith Kelsen[1], Máire Conrad[1], Ronald G. Collman[3], Robert Baldassano[1], Frederic D. Bushman[2] and Kyle Bittinger[1*]



Aird *et al. Genome Biology* 2011, **12**:R18
http://genomebiology.com/2011/12/2/R18

Genome Biology

METHOD Open Access

Analyzing and minimizing PCR amplification bias in Illumina sequencing libraries

Daniel Aird[1], Michael G Ross[1], Wei-Sheng Chen[2], Maxwell Danielsson[2], Timothy Fennell[3], Carsten Russ[1], David B Jaffe[1], Chad Nusbaum[1], Andreas Gnirke[1*]

Drengenes *et al. BMC Microbiology* (2019) 19:187
https://doi.org/10.1186/s12866-019-1560-1

BMC Microbiology

RESEARCH ARTICLE Open Access

Laboratory contamination in airway microbiome studies

Christine Drengenes[1,2*], Harald G. Wiker[2,3], Tharmini Kalananthan[3], Eli Nordeide[1], Tomas M. L. Eagan[1,2] and Rune Nielsen[1,2]

› J Microbiol Methods. 2006 Jul;66(1):21-31. doi: 10.1016/j.mimet.2005.10.005. Epub 2005 Nov 21.

DNA extraction from low-biomass carbonate rock: an improved method with reduced contamination and the low-biomass contaminant database

H A Barton [1], N M Taylor, B R Lubbers, A C Pemberton

OPEN ACCESS | Research Article | 25 June 2019

Quantifying and Understanding Well-to-Well Contamination in Microbiome Research

Authors: Jeremiah J. Minich, Jon G. Sanders, Amnon Amir, Greg Humphrey, Jack A. Gilbert, Rob Knight | AUTHORS INFO & AFFILIATIONS

DOI: https://doi.org/10.1128/mSystems.00186-19

Letters in Applied Microbiology

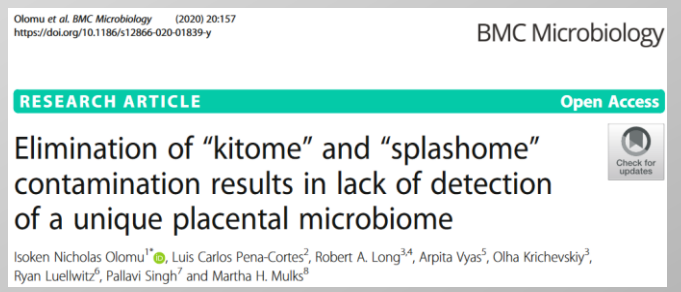
Letters in Applied Microbiology ISSN 0266-8254

ORIGINAL ARTICLE

Identification and removal of contaminating microbial DNA from PCR reagents: impact on low-biomass microbiome analyses

L.F. Stinson, J.A. Keelan and M.S. Payne

Division of Obstetrics and Gynaecology, Faculty of Health & Medical Sciences, The University of Western Australia, Perth, WA, Australia



Olomu *et al. BMC Microbiology* (2020) 20:157
https://doi.org/10.1186/s12866-020-01839-y

BMC Microbiology

RESEARCH ARTICLE Open Access

Elimination of "kitome" and "splashome" contamination results in lack of detection of a unique placental microbiome

Isoken Nicholas Olomu[1*], Luis Carlos Pena-Cortes[2], Robert A. Long[3,4], Arpita Vyas[5], Olha Krichevskiy[3], Ryan Luellwitz[6], Pallavi Singh[7] and Martha H. Mulks[8]

DNA-Free Reagents Are Important to Avoid False Results in NGS

Bioflms and Cells walls

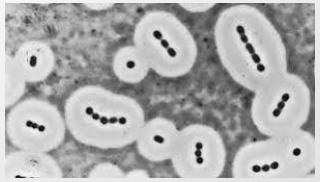
MetaPolyzyme (cell wall and membrane lysis)

- Achromopeptidase
- Chitinase
- Lyticase
- Lysostaphin
- Lysozyme
- Mutanolysin

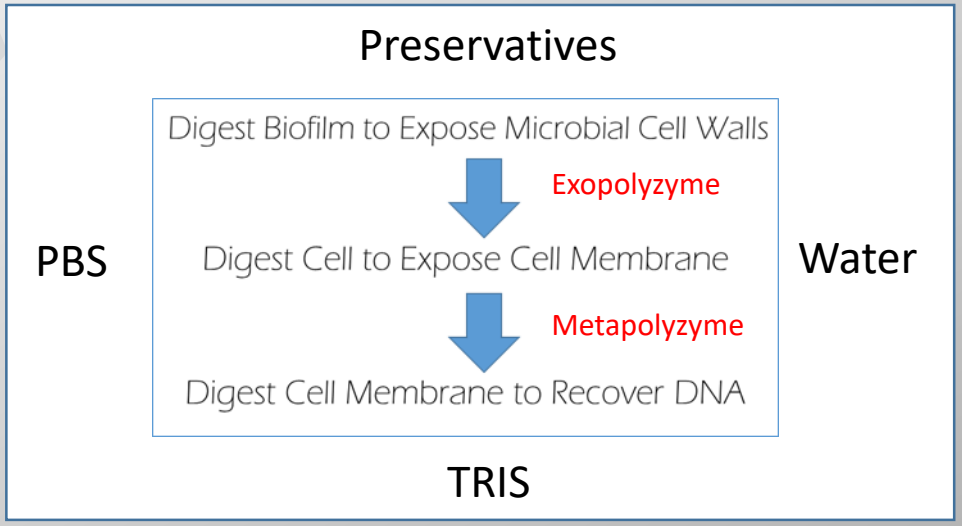
Exopolyzyme Multi-enzyme (under development)

- Slime, Capsules, ECP, ECM, polyglutamic acid, Gylcocalyx
- Protects the cell
- Allows nutrient trapping
- Common in extreme environments
  - α-Amylase
  - Cellulase
  - β-Glucosidase
  - Lyticase
  - β-N-Acetylglucosaminidase
  - Alginate lyase
  - Lipase

Nucleic Acid Preservative Buffer
TRIS Solution
Water
Mycopolyzyme
PBS



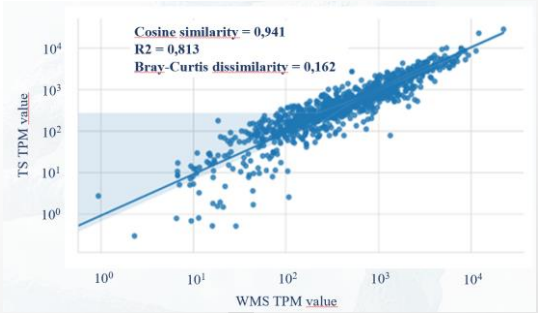
Bacterial capsules



Other Novel Technologies in NGS

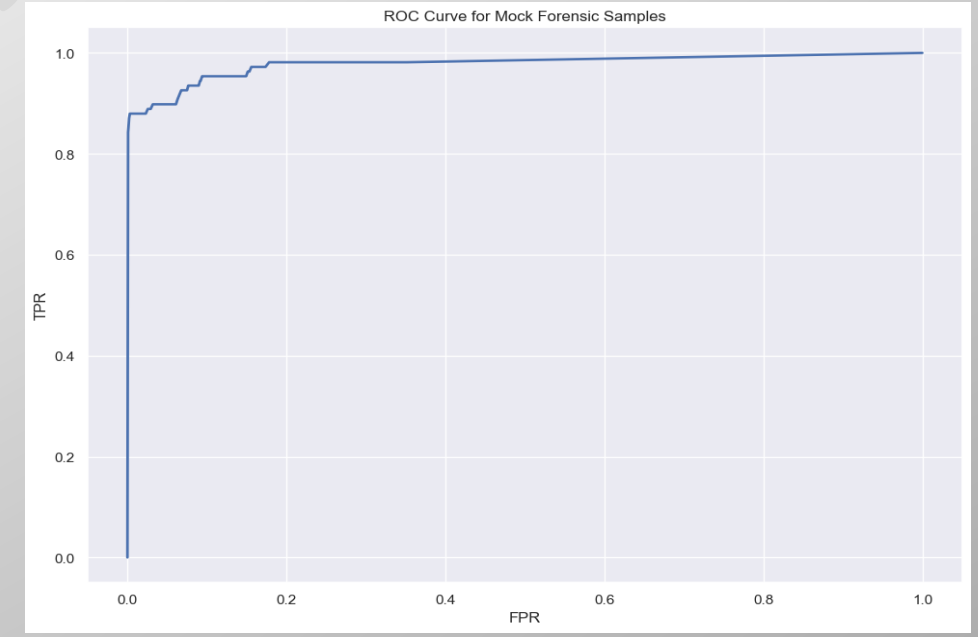
Soil Microbiome Analysis Forensic Tool (SMAFT)

- Tracking microbiomes
- Good application of Loop-Seq
- Reference database free prediction of the origin of soil samples
- Developed on basis of information theory, graph theory (Metagraph), similarity learning and feature learning
- Based on findings of the Mason et al
- Universal methodology developed by MetaSub and Biotia



Cost efficient optimization of Targeted Sequencing technology, to resemble Whole-Metagenome-Sequencing levels of signal

Calibrated system for real world cases, on petabase scale dataset:
~1000 training samples
~600 test & mock samples



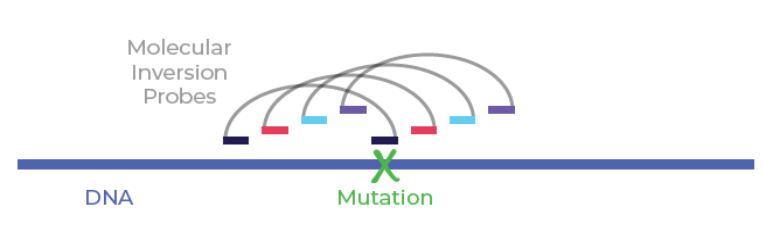
Project Funded by National Centre for Research and Development of Poland
Michał Kowalski, Kamila Marszałek, Alina Frolova, Agata Jagiełło, Anna Woźniak, Łukasz Nowak, Andrzej Ossowski, Rafał Płoski, Renata Zbieć-Piekarska, Paweł P. Łabaj, Wojciech Branicki

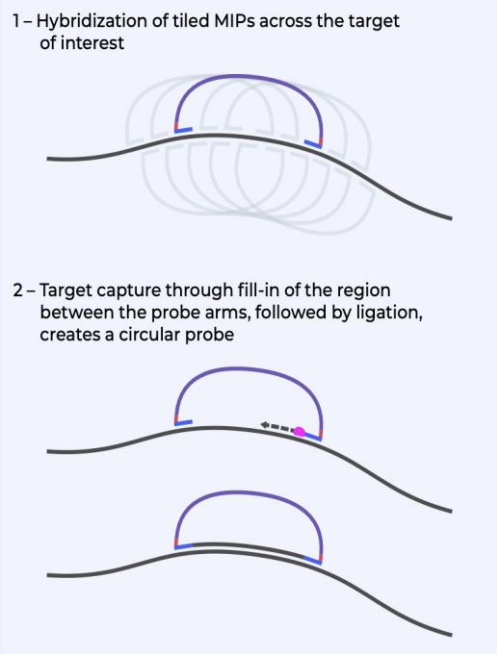
www.smaft.eu

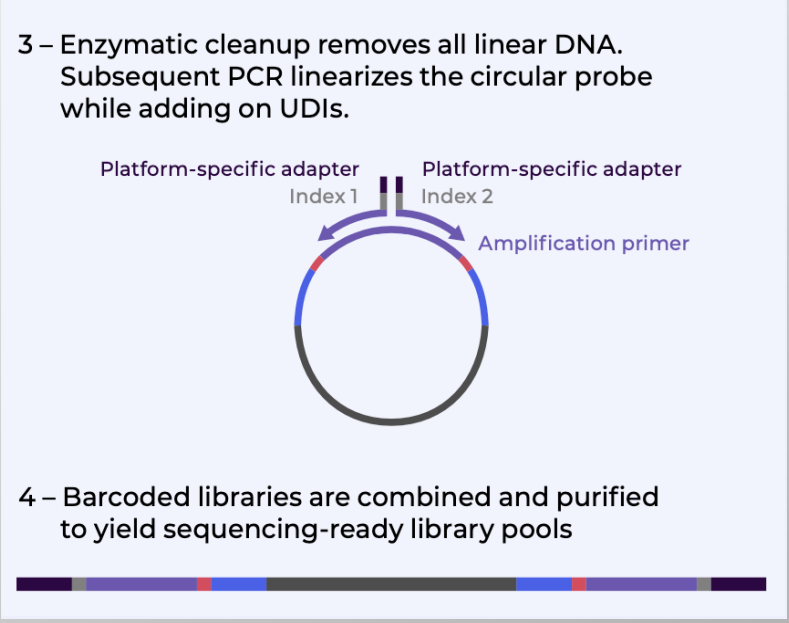
DOB-BIO10/03/01/2019

Molecular Loop Inversion Probes

- Rapid targeted library probes for NGS sequecing
- Targeted
- Substitute for PCR Ampliseq
- Genome Tiling
    - Carrier screening
    - Hereditary and somatic oncology research
    - Molecular characterization of rare and
    - complex diseases
    - Pathogen characterization and surveillance
    - Genotyping-by-sequencing



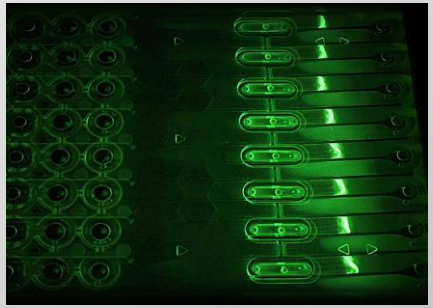




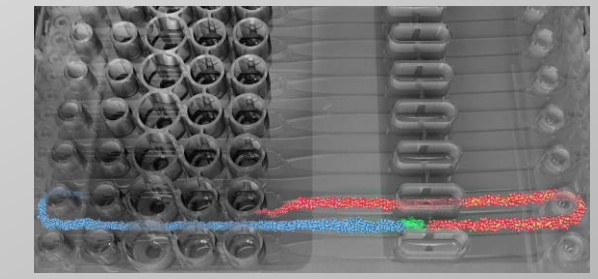
BioNano Ionic DNA Purification

Uses isotachophoresis (ITP) to isolate, purify, and concentrate genomic DNA and RNA from cells, tissue, and FFPE samples

Applied electric field across the length of the microchannels, tseparates and concentrates nucleic acid between buffers with higher and lower mobilities

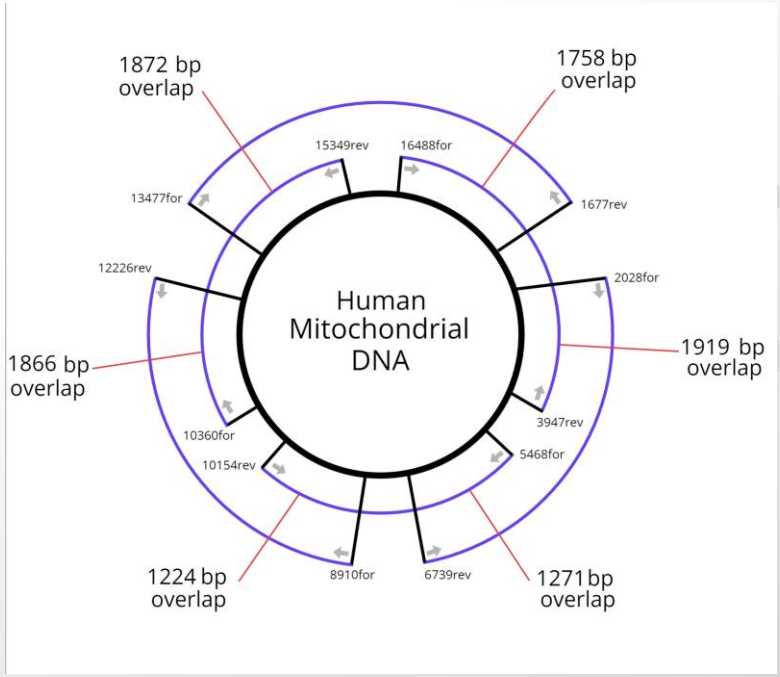


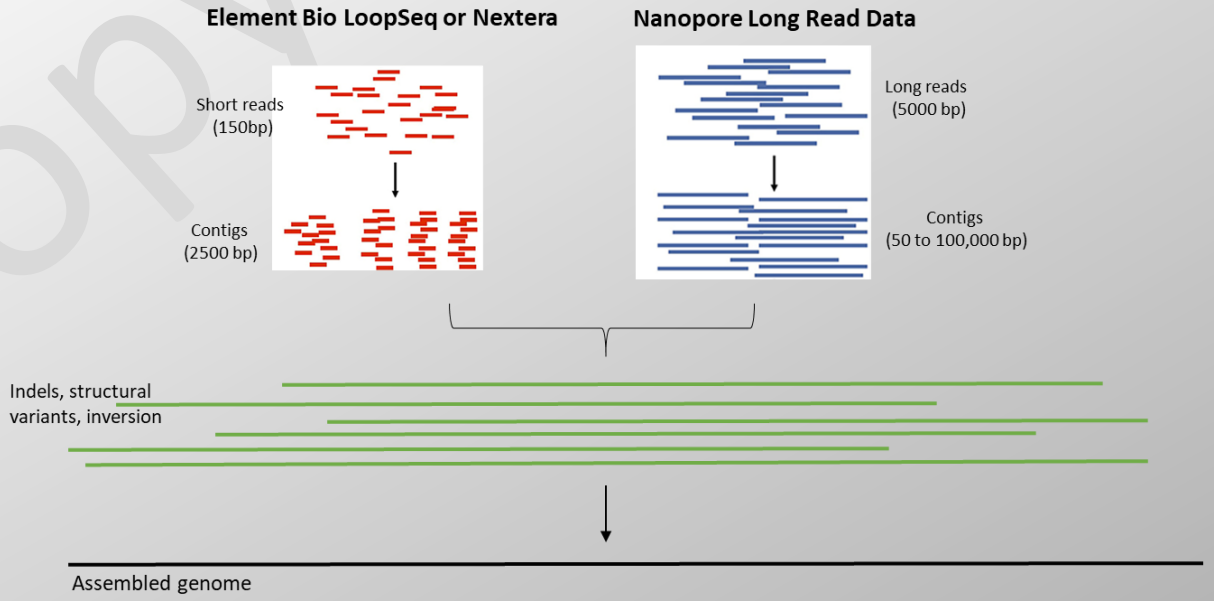
No organic solvents or high-salt buffers



Sequencing Degraded and Mixed Mitochondria Species Using Nanopore

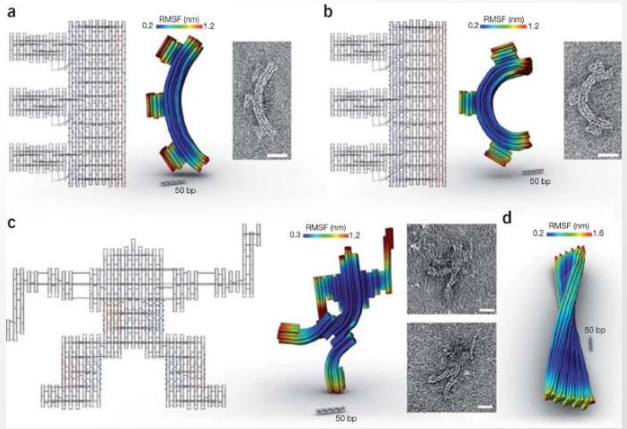
- Small POP Project with Bruce, George, and my team.
- Possible application for augmenting LoopSeq with long read Nanopore may prove useful for mild degradation values and mixed samples



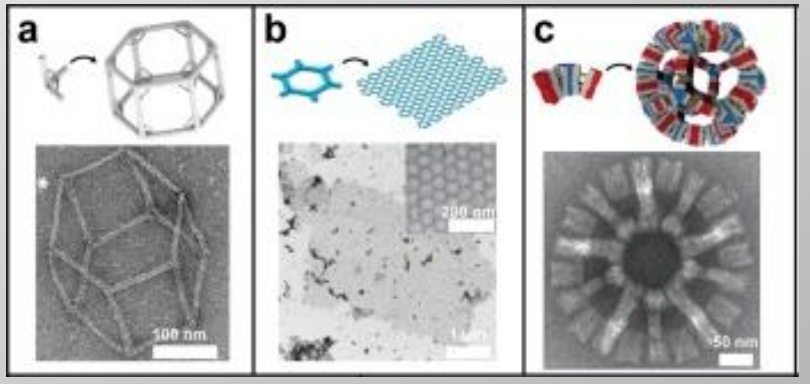


Closing Comments

- The landscape of Next Generation sequencing is changing so fast, it is hard to follow.
- The number of new developments and instrument are happening at the monthly level, not by the year or 5 year cycle
- Most are occurring for Single cell and Spatial sequencing



Is DNA Origami the next disruptive technology?



Thank You for listening