High Data Throughput and Low Cost Ultra Long Nanopore Sequencing

03/11/2021

Dr. Kelvin Liu (kliu@circulomics.com)
CEO

circulomics

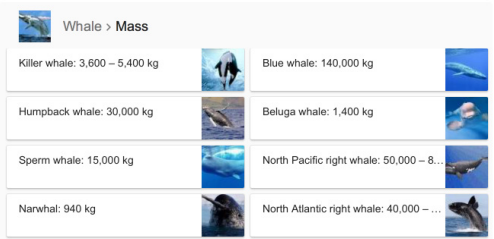
Ultra Long Nanopore Sequencing Pioneers

Nanopore sequencing has no limit to read length. Ultra long=100+ kb. Whales=1+ Mb
Previous groups have pioneered the method and demonstrated utility for spanning long repeats.

Loman Lab

Thar she blows! Ultra long read method for nanopore sequencing

09 Mar 2017



Loose Lab

Matt Loose
@mattloose

Turns out that @nanopore reads are longer than we thought. Longest we've found now is 2,272,580 bases. I reckon this is a small baby blue whale! info here:

bioRxiv
THE PREPRINT SERVER FOR BIOLOGY

Whale watching with BulkVis: A graphical viewer for Oxford ...
Motivation The Oxford Nanopore Technologies (ONT) MinION is used for sequencing a wide variety of sample types with ...
biorxiv.org

UCSC / NHGRI

Article | Open Access | Published: 14 July 2020

Telomere-to-telomere assembly of a complete human X chromosome

Karen H. Miga, Sergey Koren, [...] Adam M. Phillippy

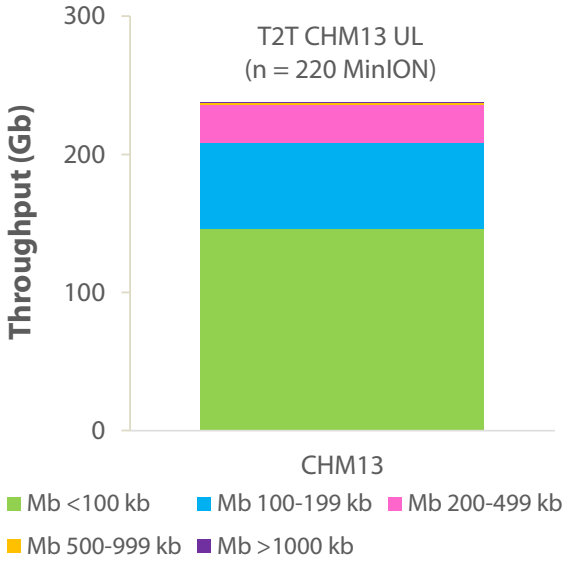
Nature **585**, 79–84(2020) | Cite this article

31k Accesses | **49** Citations | **1339** Altmetric | Metrics

circulomics

Previous UL Sequencing Methods Had Very Low Yield

Previous methods only generated ~1 Gb per flow cell
T2T CHM13 used 220 flow cells to acquire 240 Gb of data

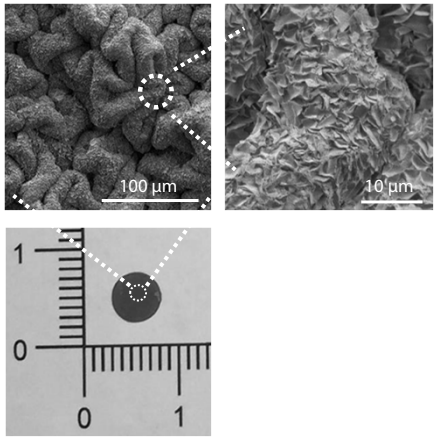


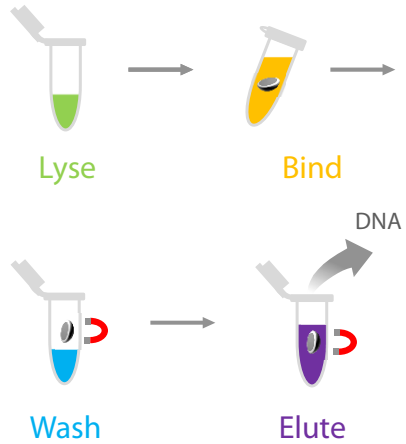
Reasons for Low Yield

- Low pore occupancy
- High pore blocking rates
- Poor response to nuclease flush
- Not compatible with PromethION

Nanobind Magnetic Disks

Nanostructured magnetic disks for rapid HMW and UHMW DNA extraction





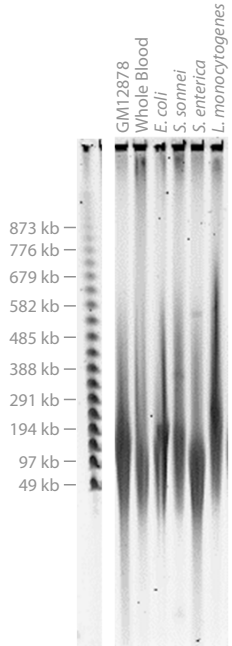
High Surface Area
Low Shear Force

Rapid Bind, Wash, and Elute

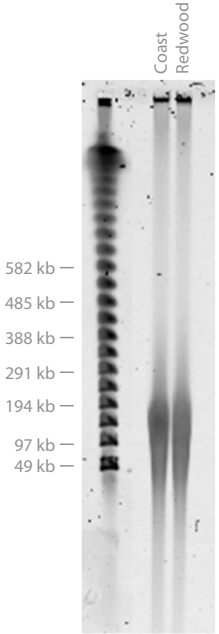
circulomics

HMW DNA Extraction (50 – 300+ kb)

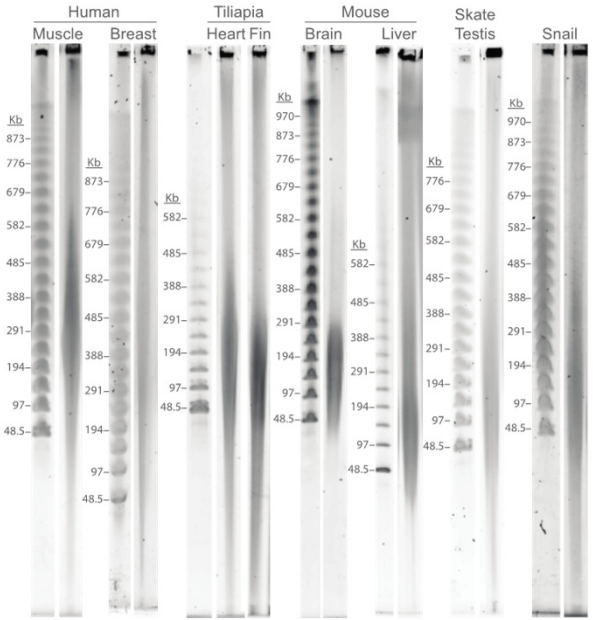
Recommended for: Oxford NANOPORE Technologies PACBIO®



Cells, Bacteria, Blood



Plants

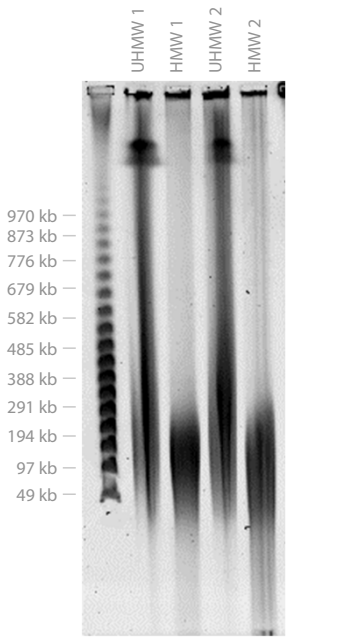


Tissues

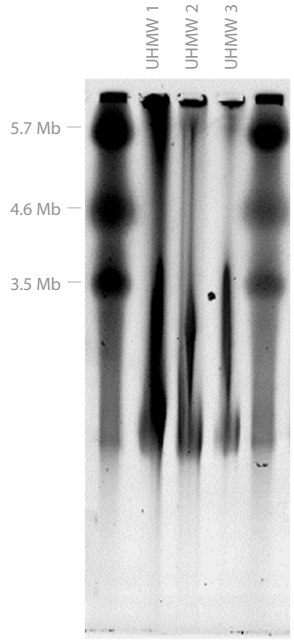
circulomics

UHMW DNA Extraction (50 kb - 1+ Mb)

Recommended for: Oxford NANOPORE Technologies | bionano GENOMICS



22 Hr CHEF Gel



72 Hr CHEF Gel

circulomics

Nanobind Enhanced Ultra Long Sequencing

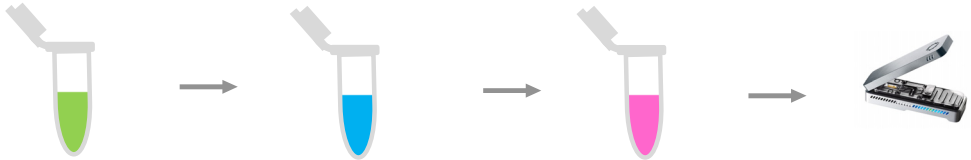
Transposase based preparation generates 6X (Min) or 3X (Prom) scale library
Nanobind disks extract and clean UHMW DNA without fragmentation

Step 1	Step 2	Step 3	Step 4	Step 5
Nanobind UHMW DNA Extraction	Tagmentation	Adaptation	Nanobind Library Cleanup and Concentration	Sequence, Flush, & Reload
circulomics Nanobind Big DNA Kit	NANOPORE Ultra-Long DNA Sequencing Kit		circulomics Nanobind UL Library Prep Kit	NANOPORE R9.4 or R10.3 Flow Cell Flow Cell Wash Kit

Steps to Maximize Pore Occupancy

Target read length N50 of 100 kb balances pore occupancy and read length.
10-100X increase in total throughput.

Quick et al Ultra Long Sequencing



Step 1
DNA Spooling

Step 2
Tagmentation
1 µg/µL DNA

Step 3
Adaptation

Step 4
Sequence

Nanobind Enhanced Ultra Long Sequencing

Step 1
Nanobind UHMW DNA Extraction

Step 2
Tagmentation
0.1 µg/µL DNA

Step 3
Adaptation

Step 4
Nanobind Cleanup

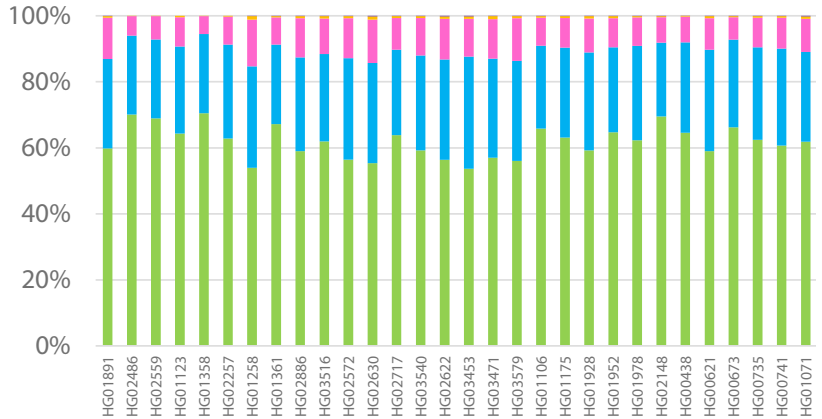
Step 5
Sequence, Flush, & Reload

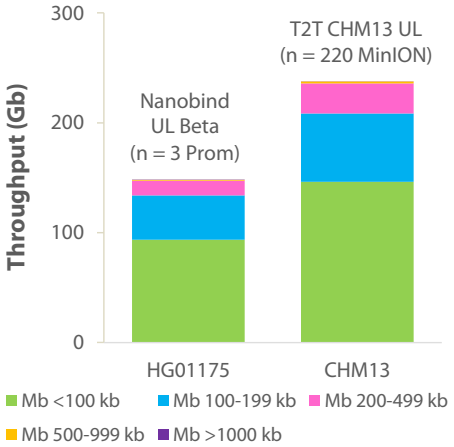
Reduced Viscosity | High Adaptation Efficiency | Clean + Concentrate | Reload

circulomics

Human Pangenome Beta Test

Beta tested an early version of the protocol on 1st 30 pan genome samples
High coverage generated using only 3-5 flow cells



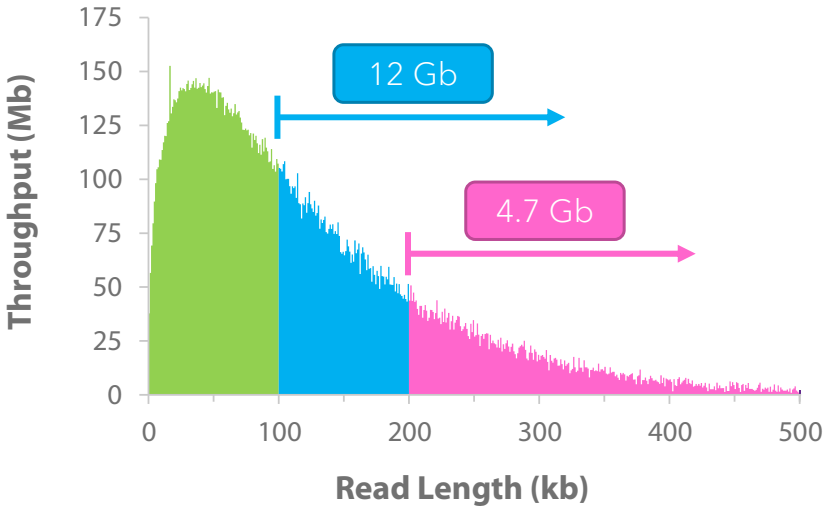


	N50	Throughput	Gb >100 kb	Gb >200 kb
Average By Flow Cell (n = 91)	78 ± 12 kb	31 ± 15 Gb	12 ± 1.7 Gb	3.2 ± 0.5 Gb
Average By Genome (n = 30)	77 ± 9 kb	95 ± 33 Gb	36 ± 12 Gb	9.7 ± 3.6 Gb

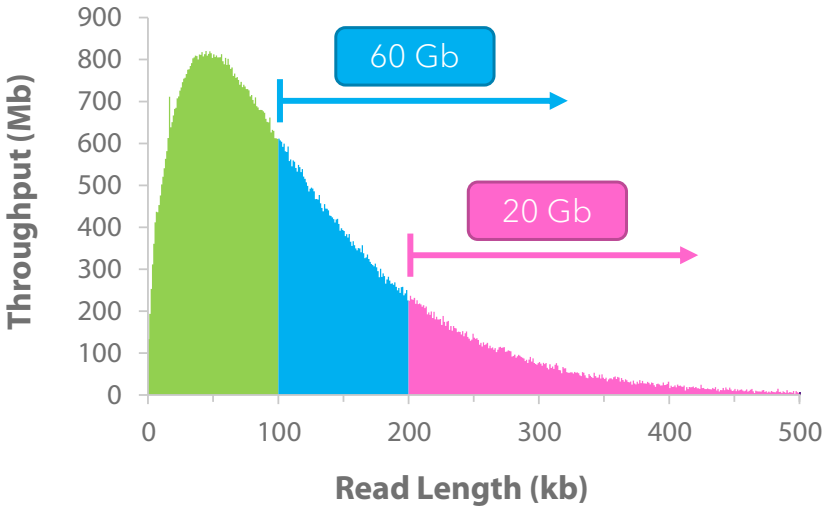
Increased Yields by Additional 2-3X Since Beta Test

100 kb N50. Up 15-25 Gb on MinION/GridION and 50-130 Gb on PromethION.

HG02723 on MinION
24 Gb, 97 kb N50, 1.2 Mb max



HG02723 on PromethION
130 Gb, 93 kb N50, 1.2 Mb max

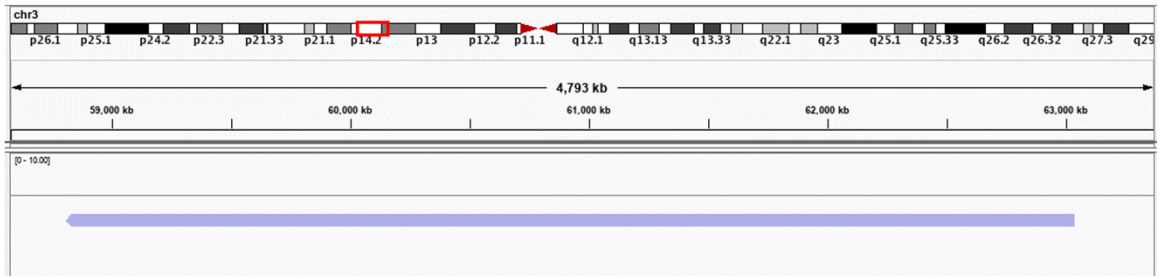


circulomics

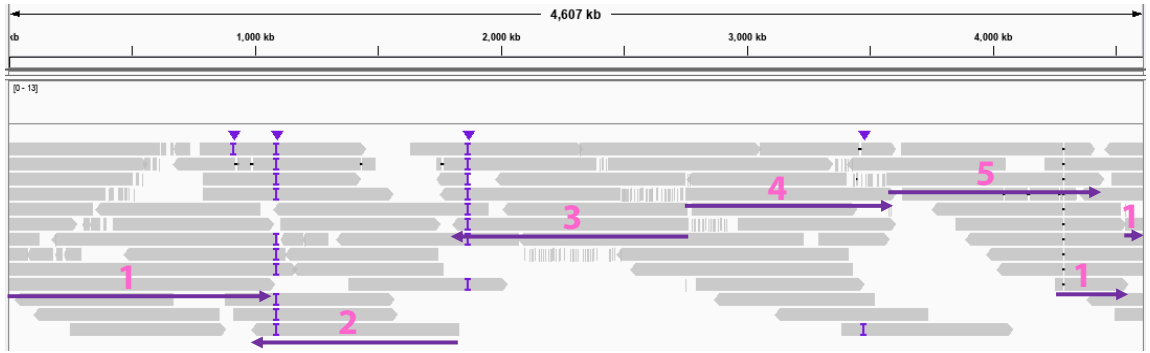
The Longest Sequencing Reads Ever Generated

Continuously increased the record for longest sequencing read. Currently at 4.2 Mb.
Assembled *E. coli* genome using only 5 reads.

4.2 Mb Read Aligning to Human Chr. 3



Complete *E. coli* Assembly Using Only 5 UL Reads

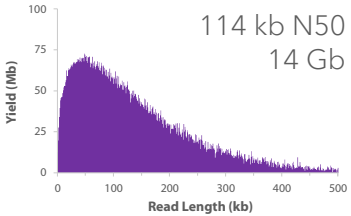


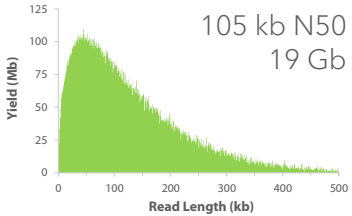
circulomics

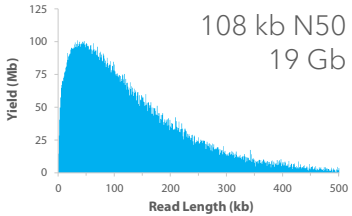
Robust Read Length and Throughput

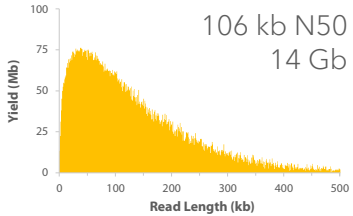
Consistent read length and throughputs across independent samples and replicates.

GridION (R9.4)
Individual Libraries
Blood

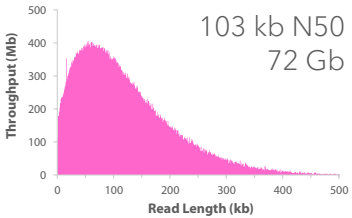


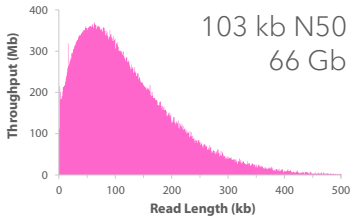


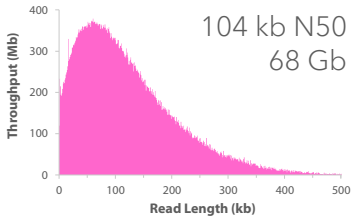




PromethION (R10.3)
Split Library
Cells







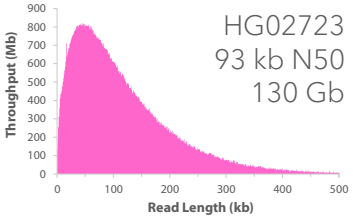
Generated Using
a Single 9X
Library Prep!

circulomics

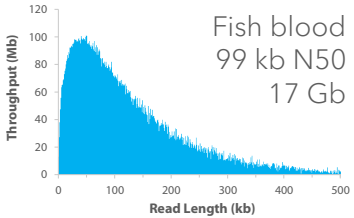
Diverse Sample Types

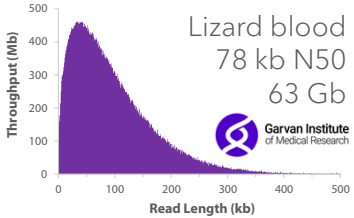
Launching with cultured cell, blood and bacteria samples.
Expansion to plant and tissue samples is underway.

Cells

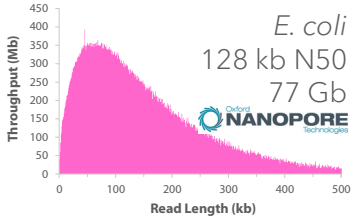


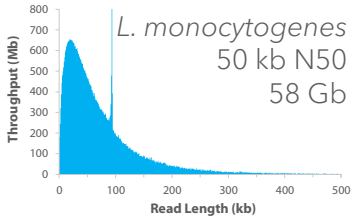
Blood



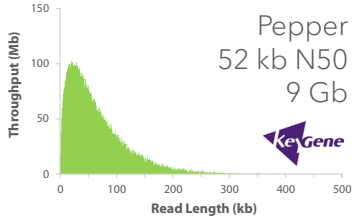


Bacteria

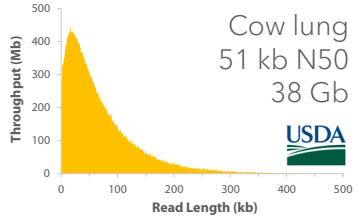


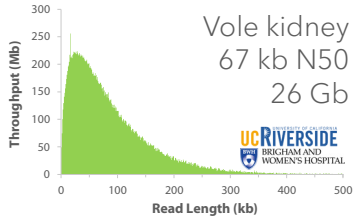


Plants (In Dev)



Tissues (In Dev)



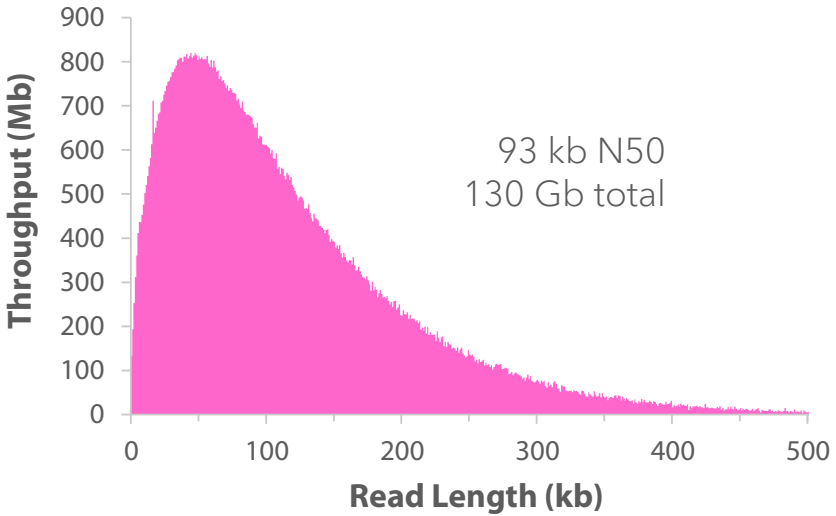


circulomics

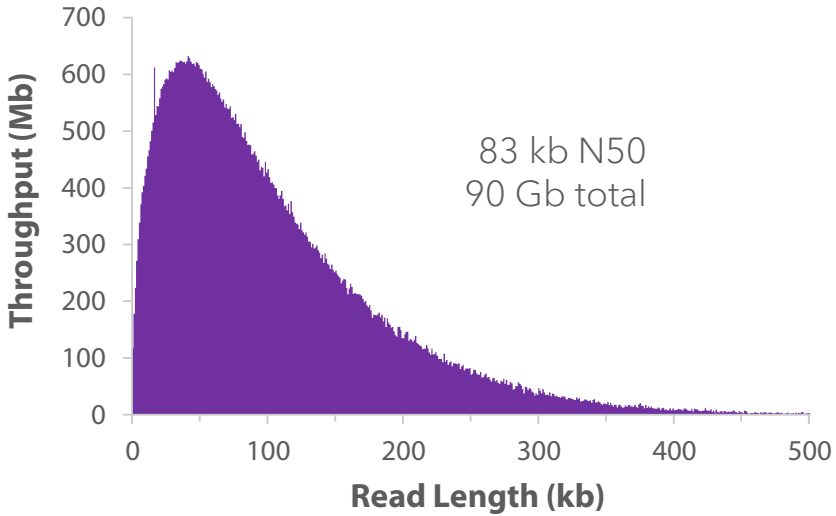
High Accuracy Ultra Long Reads

Nanobind Enhanced Ultra Long Sequencing generates high yields on both R9.4 and R10.3 flow cells. R10.3 pores have dual reader head for increased accuracy.

HG02723 on PromethION R9.4



HG02723 on PromethION R10.3

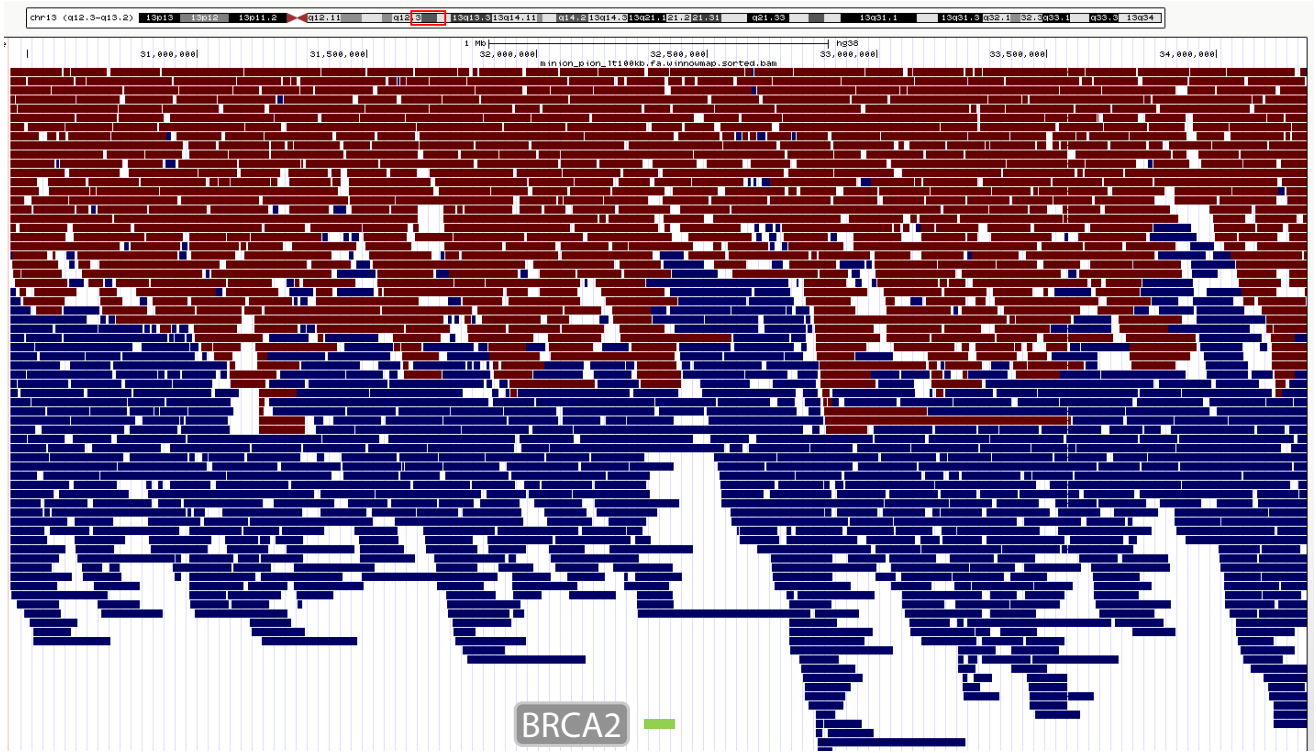


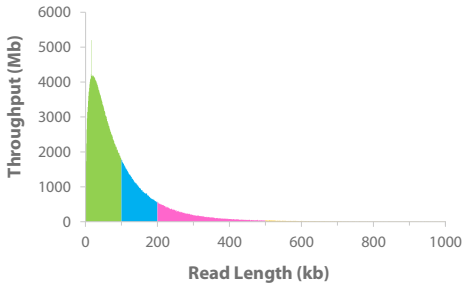
circulomics

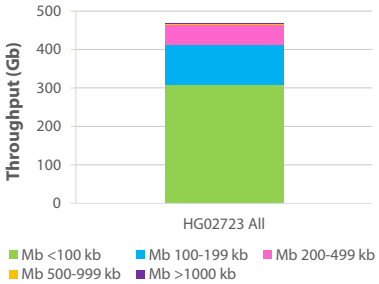
High Coverage UL Reads for Human Assembly

Generated 468 Gb of ultra long data on HG02723
50X coverage 100+ kb reads

UNIVERSITY OF CALIFORNIA SANTA CRUZ







circulomics

Acknowledgements

Questions?

circulomics

Duncan Kilburn
Jeffrey Burke
Michelle Kim
Kelvin Liu

Oxford NANOPORE Technologies

Simon Mayes
David Stoddart
Rosemary Dokos
Paul Kotturi

University of California Santa Cruz

Karen Miga
Miten Jain
Hugh Olsen

NIH National Institute of General Medical Sciences

NIH National Human Genome Research Institute

UC Riverside University of California

BWH Brigham and Women's Hospital

Polly Campbell
Brian Couger

Garvan Institute of Medical Research

Jillian Hammond
Ira Deveson

USDA

Kristen Kuhn
Tim Smith

KeyGene

Raymond Hulzink
Dick Roelofs
Alexander Wittenberg