GENERATING LARGE NUMBERS OF ULTRA LONG READS ON PROMETHION

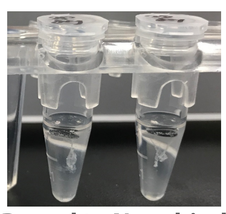
Duncan Kilburn, Jeff Burke, Renee Fedak, Kelvin Liu

Circulomics Inc, Baltimore, MD

circulomics 701 E Pratt St, Baltimore, MD 21202 USA, +1-410-996-4762, www.circulomics.com, info@circulomics.com

UHMW DNA Extraction

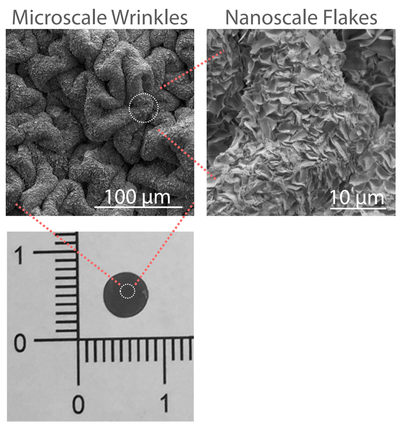
Circulomics Nanobind CBB Big DNA Kit



DNA Bound to Nanobind Disk

Ultra High Molecular Weight (UHMW), megabase DNA was extracted from human cells, bacteria, and human whole blood using the Circulomics Nanobind CBB Big DNA Kit.

Nanobind Magnetic Disks



Nanobind Magnetic Disk

Nanobind Magnetic Disks

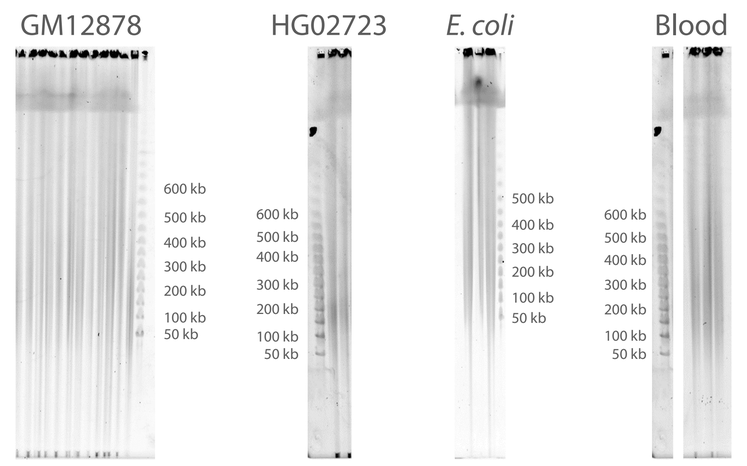
- Micro- and nanowrinkles protect DNA from shearing.
- High binding capacity.
- High purity for single molecule sequencing.

Rapid Magnetic Purification

- Rapid bind, wash, and elute process.
- 1 disk per tube.
- HMW protocols generate 50 – 300+ kb sized DNA.
- UHMW protocols generate 50 kb – 1+ Mb sized DNA.
- Manual processing using magnetic rack.
- Automation compatible.

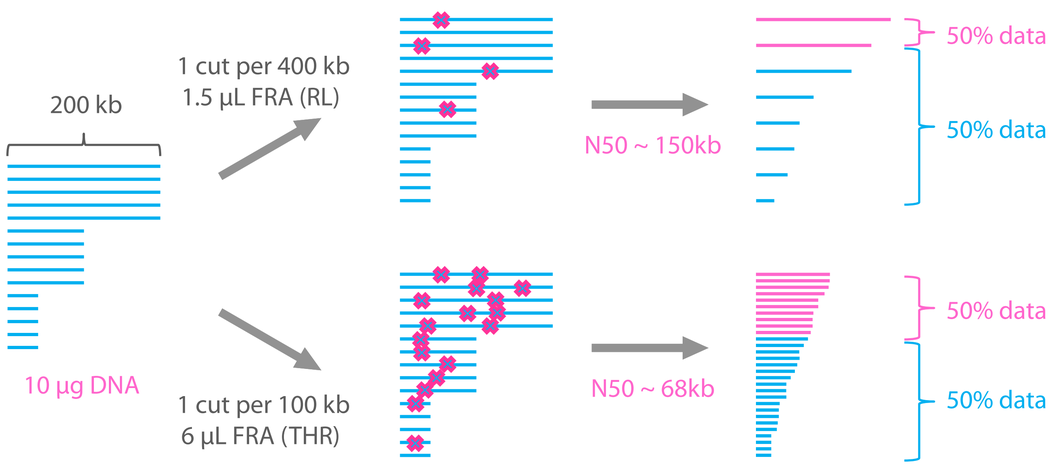
Nanobind extractions can be completed in 1 – 3 hours depending on sample type and is easily scaled to large numbers of samples.

Extraction Results

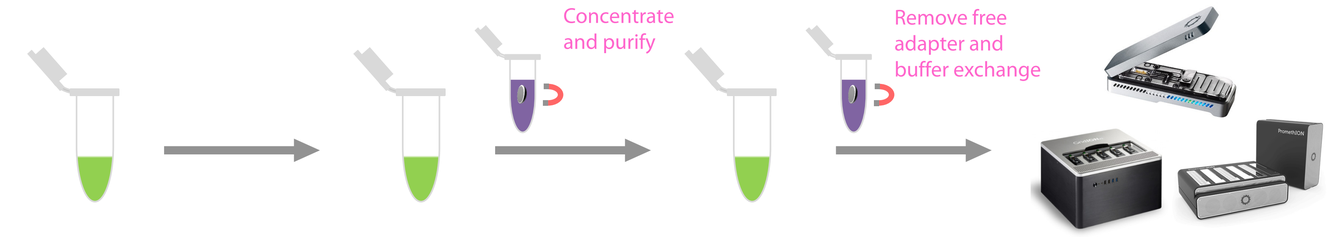


- A new UHMW extraction protocol was developed with gentler handling and modified chemistry to further increase DNA size and improve handling.
- UHMW, megabase-sized DNA can be identified on PFGE images by 1) streaking up to 1 Mb and 2) compression zone banding.

Nanobind-Enhanced Ultra Long Library Preparation (SQK-RAD004)



FRA : DNA ratio controls the number of cuts per strand made by transposase, enabling users to bias sequencing for read length vs. throughput .

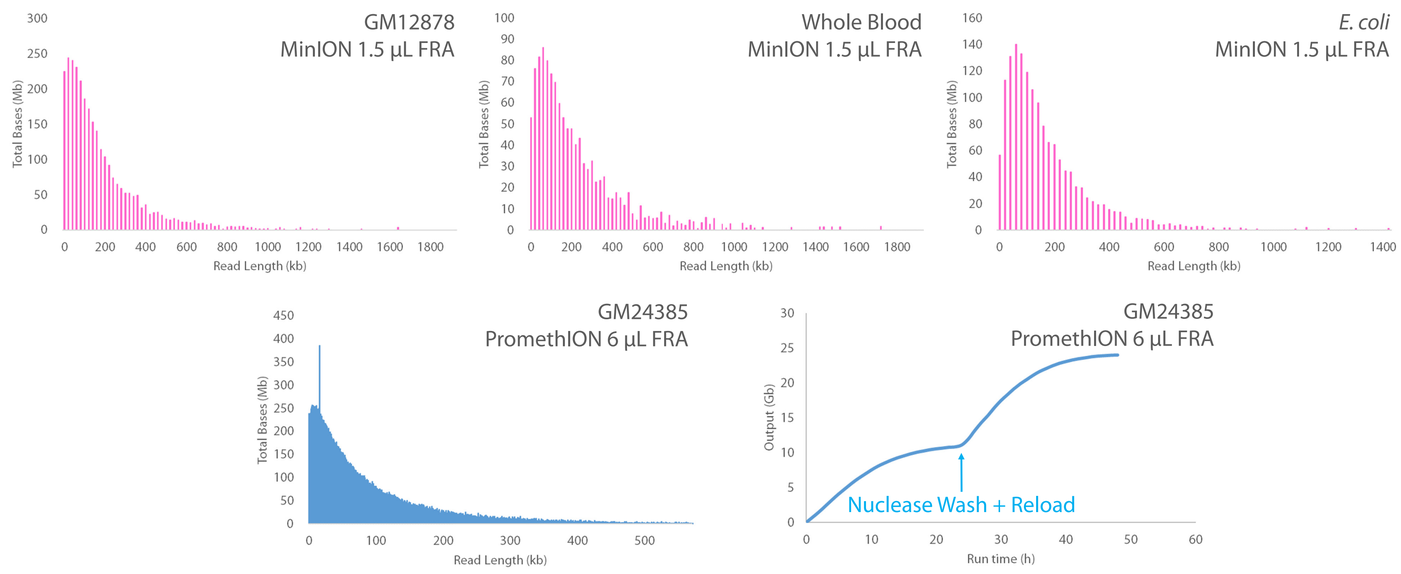


- Compared to standard Rapid Library Prep, FRA : DNA ratio is decreased to reduce fragmentation.
  - 1.5 µL FRA offers enhanced read length vs. 6 µL FRA offers enhanced throughput
- Increased reaction volume facilitates handling and mixing of viscous UHMW DNA.
- First, Nanobind reaction cleanup concentrates and purifies without shearing megabase-sized DNA.
- Second, alcohol-free, Nanobind reaction cleanup removes free adapters and performs buffer exchange.
- Optional nuclease wash and reload step can be used to increase throughput.

Oxford Nanopore Sequencing Metrics

Sample	FRA	Reloads	Sequencer	Total Data (GB)	Read Length N50 (kb)	Read Length Max (Mb)	Number Reads >1Mb
GM12878	1.5	-	MinION	0.7	138 (179)	2.1 (2.1)	7 (13)
GM12878	1.5	1	MinION	2.9	134 (150)	1.7 (1.9)	16 (36)
Whole Blood	1.5	-	MinION	1.2	168 (193)	1.8 (2.44)	16 (28)
E. coli	1.5	-	MinION	1.5	134 (175)	1.4 (1.6)	6 (15)
GM24385	6	1	PromethION	24.2	67	1.7	10

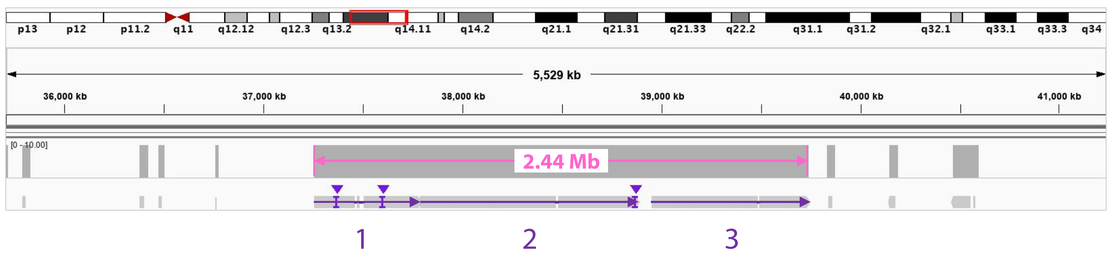
- Libraries were prepared using the 1.5 µL FRA or 6 µL FRA protocol.
- Sequencing was performed on MinION or PromethION for 48 – 72 hours.
- A single nuclease flush and reload was tested.
- Raw reads were mapped against hg38 or *E. coli* K-12 reference using Minimap2.
- Incorrectly split reads were identified and fused together using whale_watch.py (Payne *et al*, Bioinformatics 2019). Fused metrics are provided in parenthesis.
- A 2.1 Mb read mapping to chromosome 3 was obtained from GM12878 cells.
- A 2.44 Mb fused read mapping to chromosome 13 was obtained from whole blood.



- As expected, 1.5 µL FRA and 6 µL FRA protocols generated enhanced read length and throughput, respectively.
- Consistently achieved 100+ kb N50 and large numbers of 1 – 2+ Mb reads.
- Final buffer exchange enables ultra long sequencing on PromethION.
- Nuclease wash and reload increases throughput by ~50 – 100%.

2.44 Mb Read from Human Blood

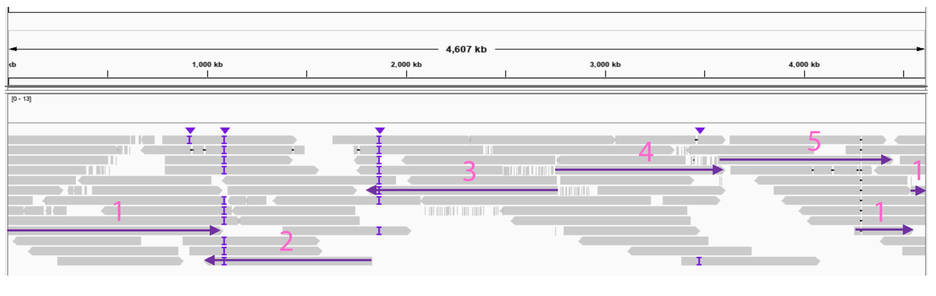
Human Chromosome 13: 35,705,735 – 41,275,889 (hg38)



Read	1	2	3
Start Time	10876.7 s	12210.8 s	15028.2 s
End Time	12210.7 s	15027.8 s	17246.7 s
Size	0.51 Mb	1.08 Mb	0.85 Mb

- Whale_watch.py identified 3 reads that were inadvertently split by the sequencing software and fused back together into a single 2.44 Mb read.
- They sequenced consecutively through pore 376 and mapped to contiguous regions of chromosome 13.

E. coli assembly in 5 reads



Read	Read Length (Mb)	Start Position	End Position
1	1.44	4255205	1070891
2	0.83	1010582	1842841
3	0.95	1829159	2776627
4	0.83	2754343	3595920
5	1.13	3293599	4444399

- Mapped all reads > 600 kb, 9.2X total coverage.
- Minimum number of reads needed to generate Miniasm assembly = 5
  - Contig N50 = 4,597,548 bp. Median depth = 1.00X.

Acknowledgements

Simon Mayes, Vania Costa, David Stoddart, and Michelle Hiscutt at Oxford Nanopore Technologies helped with incorporating the Nanobind reaction cleanup protocol into the RAD004 library preparation. Simon Mayes generated the GM24385 PromethION sequencing data.