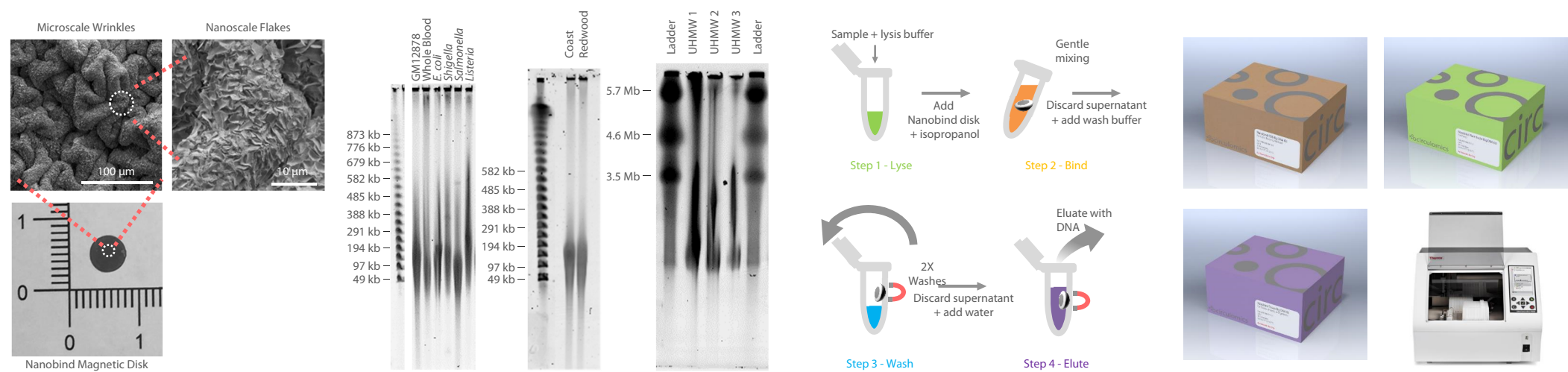


HIGH-THROUGHPUT HIGH MW DNA EXTRACTION AND SIZE SELECTION FOR LONG-READ SEQUENCING

Kelvin J Liu¹, Michelle Kim¹, Renee Fedak¹, Duncan Kilburn¹, Jeffrey M Burke¹, Heather Ferrao³, Brendan Galvin³, Keith Bjornson³, Rachael Workman², Winston Timp²,

¹Circulomics Inc, Baltimore, MD; ²Johns Hopkins University, Baltimore, MD; ³Pacific Biosciences, Menlo Park, CA;

Nanobind Magnetic Disks



Nanobind Magnetic Disks

- High density of micro- and nanostructured silica
- Micro- and nanowrinkles protect DNA from shearing
- High surface area = high binding capacity (nanogram to milligrams)
- Non-porous disk = efficient washing and high purity

HMW and UHMW DNA Extraction

- HMW (50 kb – 300+ kb) DNA extraction protocols
- UHMW (50 kb – 1+ Mb) DNA extraction protocols
- HMW DNA + high purity = long read lengths and high sequencing yields across diverse sample types

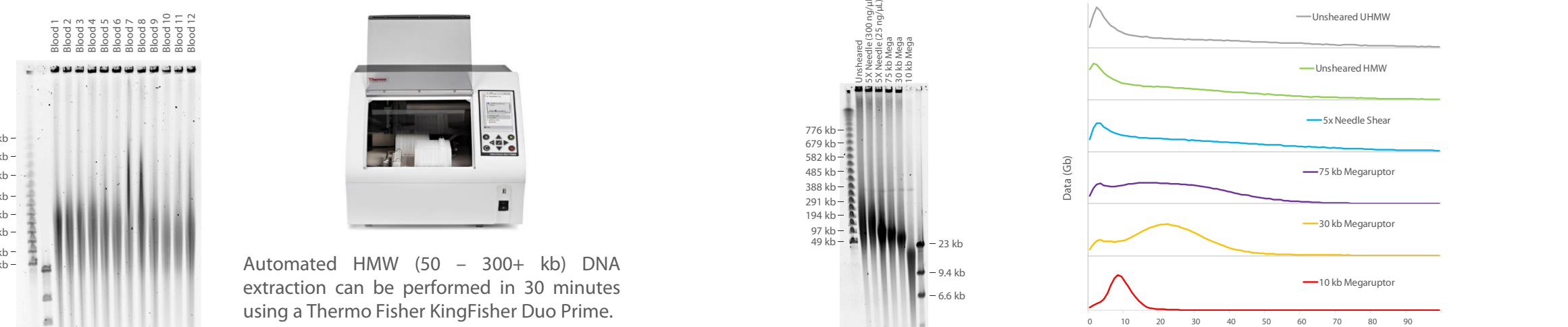
Rapid Magnetic Purification

- Simple bind, wash, and elute process
- 1 – 4 hour process time depending on sample type
- 1 mm – 6 mm diameter disks for various sample volumes
- Convenient kit format
- Manual or automated processing

Broad Applications

- Nanobind CBB Big DNA Kit for cells, bacteria, and whole blood
- Nanobind Plant Nuclei Big DNA Kit for plant samples
- Nanobind Tissue Big DNA Kit for tissue, insects, small organisms
- Nanobind Size Select Kit for reaction and size selection purification
- Plasmid purification
- Microvolume samples

Cells / Bacteria / Blood + Automation



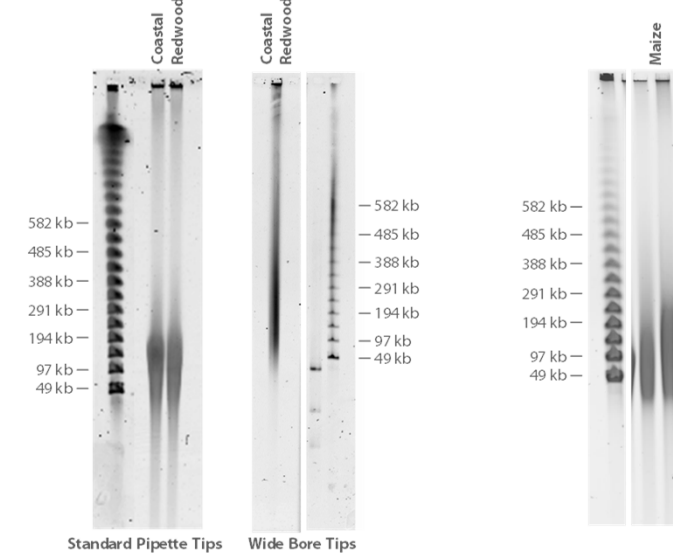
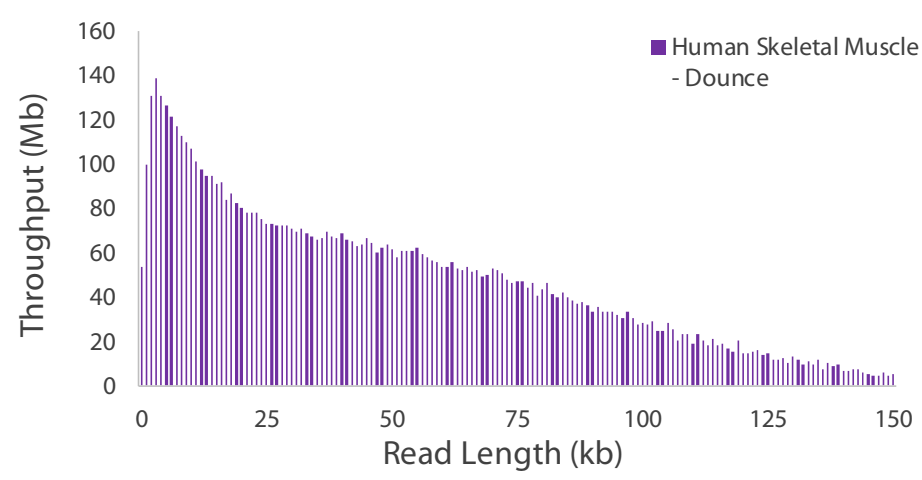
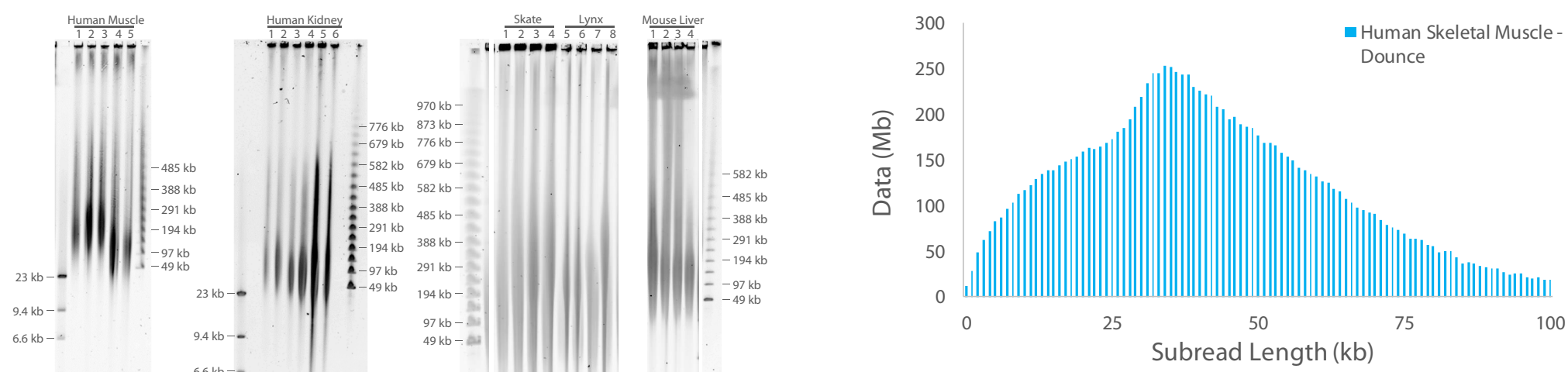
Nanobind CBB Big DNA Kit				
Sample	Input	260/280	260/230	Yield (µg)
MCF-7 Cells (tetraploid)	1x10 ⁶ cells	2.0	1.9	18.3
GM12878 Cells (diploid)	2.5x10 ⁶ cells	1.9	2.1	12.7
<i>Escherichia coli</i>	2x10 ⁹ cells	1.8	1.4	21.0
<i>Shigella sonnei</i>	1x10 ⁹ cells	1.8	1.4	27.7
<i>Salmonella enterica</i>	1x10 ⁹ cells	1.8	1.5	23.4
<i>Listeria monocytogenes</i>	1x10 ⁹ cells	1.8	1.9	21.7
Human Whole Blood	200 µL	1.9	1.9	5.6

Oxford Nanopore GridION				
Sample	Sequencing	Read Length N50 (bp)	Longest Read (bp)	Total Data (Gb)
GM12878 Cells	5X NS, No SS, LSK109, 106C	34,417	253,002	5.8
Jurkat Cells	No shear, No SS, LSK109, 106D	24,977	306,313	7.2
<i>E. coli</i>	5X NS, 30 kb SS, LSK108	46,156	191,836	6.8
<i>L. monocytogenes</i>	5X NS, 30 kb SS, LSK108	48,391	254,985	8.0
Whole Blood	No shear, No SS, LSK109, 106D	30,935	301,055	9.1

PacBio Sequel				
Sample	Sequencing	Subread N50 (bp)	Max Subread Length (bp)	Total Data (Gb)
Manual GM12878	5X NS, 30 kb SS, v2.0	32,813	95,946	10.4
Automated GM12878	5X NS, 30 kb SS, v2.0	31,110	97,357	7.5
Manual <i>E. coli</i>	5X NS, 30 kb SS, v2.0	29,611	80,713	6.8
Automated <i>E. coli</i>	5X NS, 30 kb SS, v2.0	32,840	78,470	4.2
Manual Whole Blood	5X NS, 30 kb SS, v2.0	28,443	95,679	6.7
Automated Whole Blood	5X NS, 30 kb SS, v2.0	28,154	104,768	6.1

Oxford Nanopore GridION Shearing Study					
Sample	Shearing	Sequencing	Read Length N50 (bp)	Longest Read (bp)	Reads >100 kb
UHMW GM12878	None	No SS, LSK109, 106D	23,896	456,283	2,866
HMW GM12878	None	No SS, LSK109, 106D	21,380	292,236	1,538
HMW GM12878	5X Needle Shear	No SS, LSK109, 106D	25,530	285,826	1,392
HMW GM12878	75 kb Megaruptor	No SS, LSK109, 106D	22,487	149,638	114
HMW GM12878	30 kb Megaruptor	No SS, LSK109, 106D	22,111	128,601	13
HMW GM12878	10 kb Megaruptor	No SS, LSK109, 106D	9,026	102,732	1

Tissue



Nanobind Tissue Big DNA Kit				
Sample	260/280	260/230	DNA (ng) / Tissue (mg)	
Human Muscle – TissueRuptor (L3)	1.86	2.08	246	
Human Muscle – Dounce (L2)	1.88	1.95	363	
Human Kidney – TissueRuptor (L3, L4)	1.81	1.97	294	
Skate Testis – TissueRuptor (L3, L4)	1.85	2.29	1359	
Lynx Muscle – TissueRuptor (L7, L8)	1.90	2.16	471	
Mouse Liver – TissueRuptor (L3, L4)	1.86	2.09	925	

Skate and lynx samples courtesy of Vertebrate Genome Lab at Rockefeller University. Mouse samples courtesy of DNA Sequencing Core at Max Planck Institute. Human samples courtesy of W. Mojica at SUNY Buffalo.

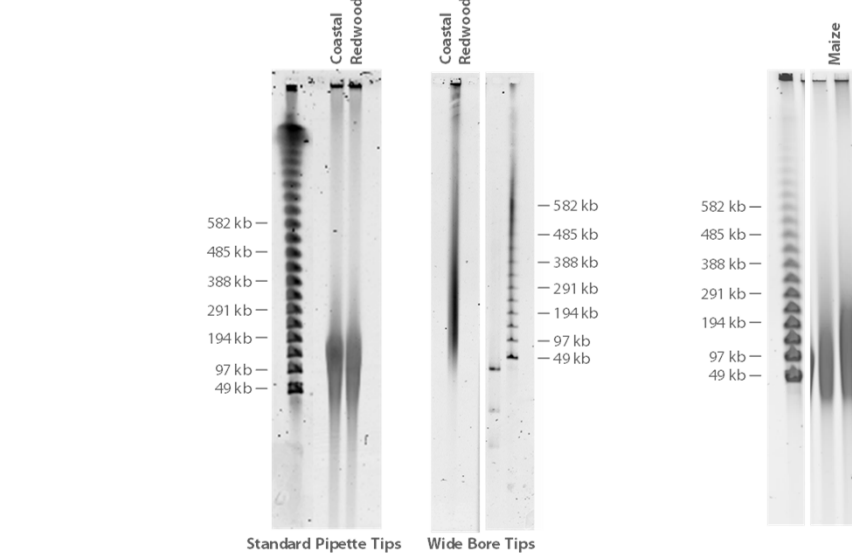
PacBio Sequel – Tissue				
Nanobind Protocol	Sequencing	Subread N50 (bp)	Max Subread (bp)	Total Data (Gb)
Human Muscle – TissueRuptor	5X NS, 30 kb BP	36,369	218,345	14.8
Human Muscle – Dounce	5X NS, 30 kb BP	40,229	240,443	12.7
Mouse Liver	5X NS, 30 kb BP	29,038	182,344	9.8
Skate Testis	5X NS, 30 kb BP	28,910	211,315	6.2
Human Kidney	5X NS, 30 kb BP	29,443	186,687	11.8
Human Colon	5X NS, 30 kb BP	30,985	201,706	4.2

Tissue samples extracted using the Nanobind Tissue Big DNA Kit result in high subread N50 and yields on PacBio Sequel (Express Template Kit 1.0, Sequel Chemistry 3.0., 20 hour movie). Lower sequencing yields on some samples due to underloading of flow cells. BP = BluePippin

Oxford Nanopore GridION – Tissue				
Nanobind Protocol	Sequencing	Read Length N50 (bp)	Longest Read (bp)	Total Data (Gb)
Human Muscle	5X NS, No SS, LSK109, 106D	42,574	474,031	7.4
Human Breast	5X NS, No SS, LSK109, 106D	36,420	319,125	8.2
Human Kidney	5X NS, No SS, LSK109, 106D	30,520	200,567	8.3
Human Colon	5X NS, No SS, LSK109, 106D	20,130	192,816	15.8
Zebrafinch	5X NS, No SS, LSK109, 106D	24,071	170,941	4.6
Mouse Liver	5X NS, No SS, LSK109, 106D	49,011	241,282	3.3

Mouse samples courtesy of DNA Sequencing Core at Max Planck Institute. Zebrafinch courtesy of Rockefeller University. Human samples courtesy of W. Mojica at SUNY Buffalo. NS = needle shear. SS = size selection

Plant



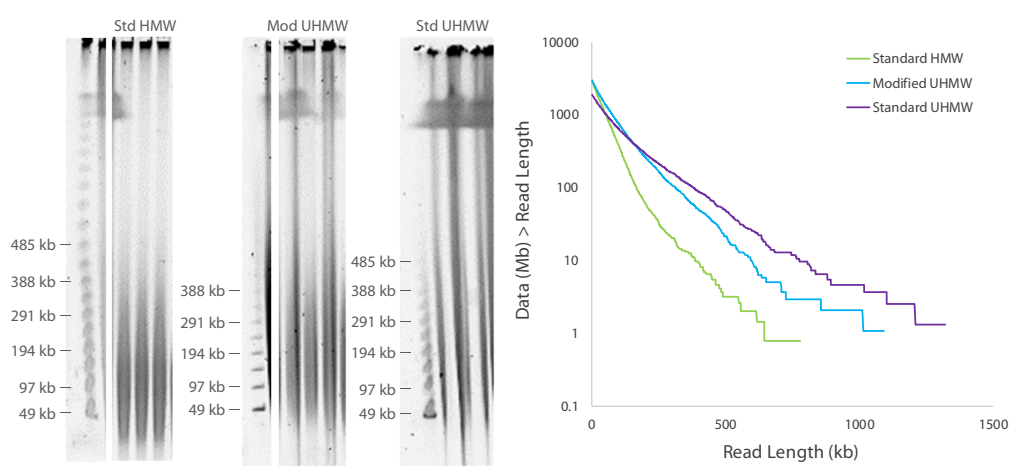
Nanobind Plant Nuclei Big DNA Kit				
Sample	Plant Input	260/280	260/230	Yield (µg)
Coast Redwood	1g	1.8	1.4	13.4
Coast Redwood	1g	1.8	1.4	11.5
Maize	1g	1.9	1.9	5.8
Rice	1g	1.9	1.7	12.0
Brazilian Hyacinth	1g	1.8	2.0	26.0

PacBio Sequel - Plant Samples				
Sample	Sequencing	Polymerase N50 (bp)	Subread N50 (bp)	Total Data (Gb)
Coast Redwood	5X NS, 30 kb SS, v2.0	35,250	33,250	5.6
Rice	No shear, 30 kb SS, v2.0	31,750	29,750	8.5
Brazilian Hyacinth	No shear, 30 kb SS, v2.0	31,750	29,750	11.0

Oxford Nanopore - Plant Samples				
Sample	Sequencing	Read Length N50 (bp)	Longest Read (bp)	Total Data (Gb)
Giant Sequoia	8 kb Mega, No SS, LSK108, 106C	7,039	121,596	7.0
Coast Redwood	8 kb G-tube, No SS, LSK108, 106C	6,567	78,241	10.1
Coast Redwood	5X NS, No SS, LSK108, 106C	29,196	227,249	3.3
Raspberry	LSK109, PromethION	24,616	821,689	47.9

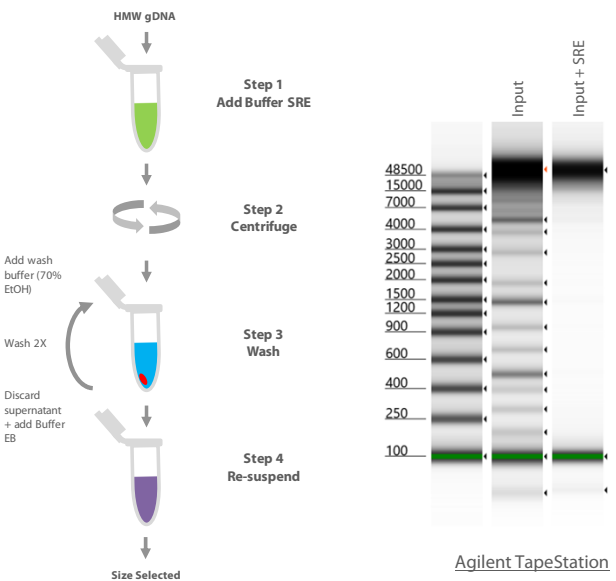
Maize data generated in collaboration with Buell Lab at Michigan State University. Rice and Brazilian hyacinth data generated in collaboration with Arizona Genomics Institute.

Ultra-Long Reads



Oxford Nanopore GridION - Whale Hunting				
Nanobind Protocol	Read Length N50 (bp)	Longest Read (Mb)	Reads >1 Mb	Data >100 kb (Gb)
Standard HMW	33,669	.78	0	0.39
Modified UHMW	44,849	1.1	2	0.77
Standard UHMW	60,734	1.3	4	0.65

DNA was extracted from GM12878 cells using 3 different Nanobind protocols and sequenced on Oxford Nanopore GridION (FLO-MIN106C) using the Rapid Sequencing Kit (SQK-RAD004). Ultra long reads were identified using whale_watch.py (https://doi.org/10.1101/312256). The extraction protocol and DNA size differences lead to tradeoffs in sequencing yield, read length N50, and the total number of long and ultra-long reads.

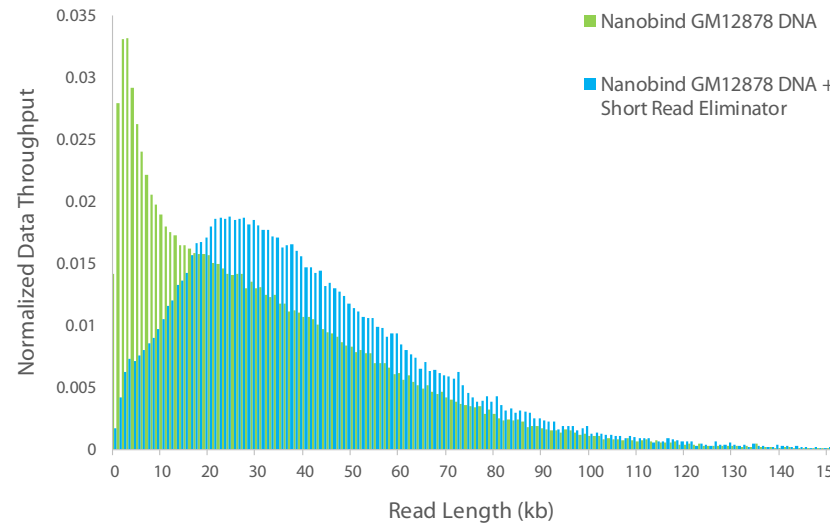


Circulomics Short Read Eliminator Kits

- Variable cutoffs from 10 kb to 40 kb
- High recovery (~40-80%)
- Fast processing (1-2 hours)
- High throughput (10-20 samples in parallel)
- Compatible with Oxford Nanopore MinION / GridION/ PromethION and PacBio Sequel
- Nanobind Reaction Cleanup (RXN) coming soon

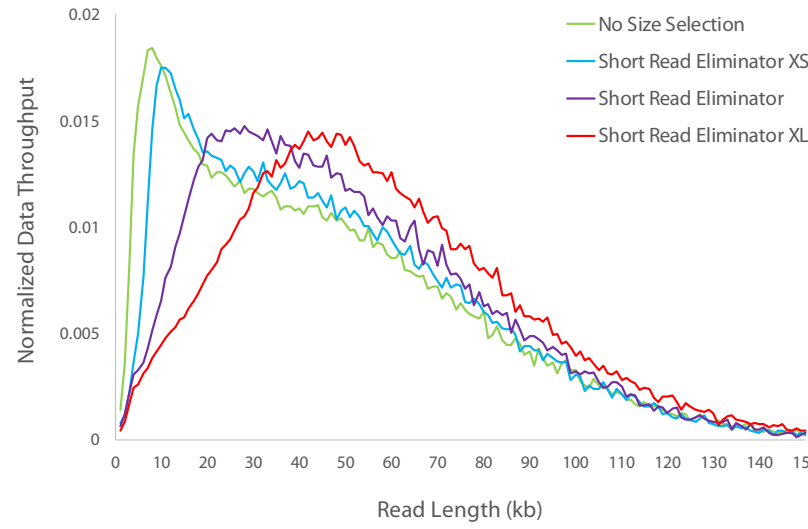
Short Read Eliminator Family				
Version	Size Selection	DNA Input	Recovery Efficiency	Notes
Short Read Eliminator XS	Progressive depletion up to 10 kb	50 – 150 ng/µL	80%	Suitable for fragmented DNA
Short Read Eliminator	Progressive depletion up to 25 kb	50 – 150 ng/µL	60%	Requires high quality DNA (>48 kb)
Short Read Eliminator XL	Progressive depletion up to 40 kb	50 – 150 ng/µL	40%	Requires very high quality DNA (>>48 kb)

Short Read Eliminator (SRE) Kit



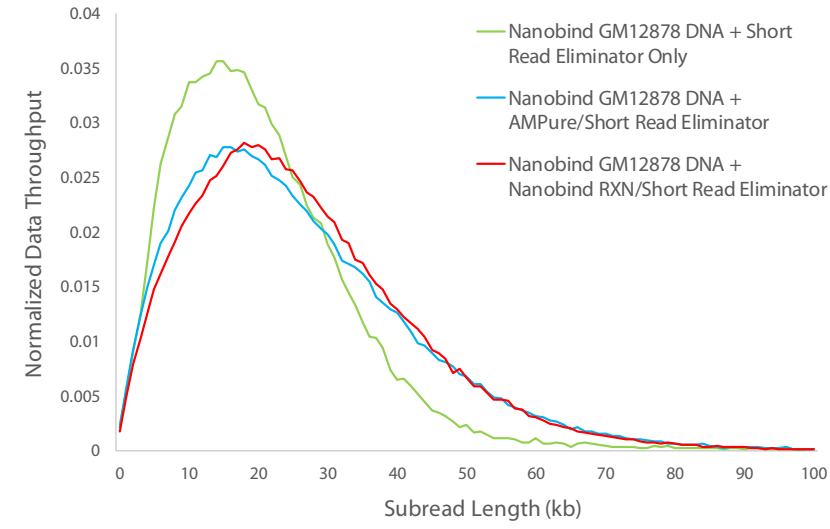
Oxford Nanopore MinION/GridION				
Sample	Size Selection	Shear	Read Length N50 (bp)	Total Data (Gb)
Nanobind GM12878 DNA	None	5X NS	25,530	7.6
Nanobind GM12878 DNA	SRE	5X NS	36,022	4.9

DNA was extracted from GM12878 cells using the Nanobind CBB Big DNA Kit HMW protocol. The Short Read Eliminator Kit was then used to remove short DNA. The samples were sequenced on Oxford Nanopore GridION (FLO-MIN106D, SQK-LSK109). Short reads <25 kb were progressively removed, increasing read length N50 from 25.5 kb to 36 kb.



Oxford Nanopore MinION/GridION				
Sample	Size Selection	Shear	Read Length N50 (bp)	
Nanobind Blood DNA	None	5X NS	38,985	
Nanobind Blood DNA	SRE XS	5X NS	41,326	
Nanobind Blood DNA	SRE	5X NS	45,513	
Nanobind Blood DNA	SRE XL	5X NS	53,159	

The Short Read Eliminator Kits feature different cutoffs for different sample types and applications.



PacBio Sequel				
Sample	Final Purification / Size Selection	Shear	Subread N50 (bp)	Total Data (Gb)
Nanobind GM12878 DNA	SRE Only	5X NS	19,061	8.0
Nanobind GM12878 DNA	AMPure + SRE	5X NS	23,379	16.4
Nanobind GM12878 DNA	Nanobind RXN + SRE	5X NS	24,408	13.8

The Short Read Eliminator can be used to replace the final AMPure/BluePippin/AMPure size selection with a single purification, resulting in dramatically higher recovery and faster processing time. The Short Read Eliminator can also be combined with Nanobind Reaction Cleanup (Nanobind RXN) and/or AMPure to further improve read length N50.