

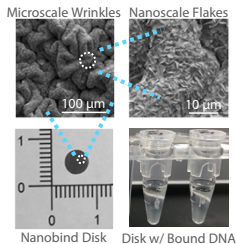
HIGH DATA THROUGHPUT AND LOW COST ULTRA LONG NANOPORE SEQUENCING FOR HUMAN PANGENOME ASSEMBLY

Duncan Kilburn¹, Jeff Burke¹, Renee Fedak¹, Hugh Olsen², Miten Jain², Karen Miga², Simon Mayes³, Kelvin Liu¹

¹Circulomics Inc, Baltimore, MD; ²UC Santa Cruz Genomics Institute, Santa Cruz, CA; ³Oxford Nanopore Technologies, Oxford, UK

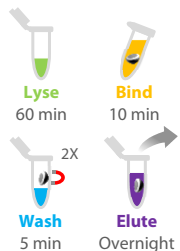
 **circulomics** www.circulomics.com
info@circulomics.com

Nanobind Magnetic Disks



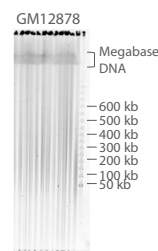
- Ultra high molecular weight (UHMW) DNA binds and releases from disk without shearing
- Micro- and nanowrinkles protect DNA from fragmentation
- 1 disk per tube
- Rapid bind, wash, and elute process
- Contaminants cleanly washed away
- Extraction protocols for nearly all sample types
- Manual or automated processing

Nanobind UHMW DNA Extraction



Nanobind CBB Big DNA Kit

- New UHMW DNA extraction protocols and chemistries were developed to increase size and reduce sample viscosity, resulting in longer reads and higher throughputs
- UHMW, megabase-sized DNA (50 kb – 1+ Mb) can be identified on PFGE images by 1) streaking up to 1 Mb and 2) compression zone banding



Nanobind UL Library Cleanup



Nanobind UL Library Prep Kit

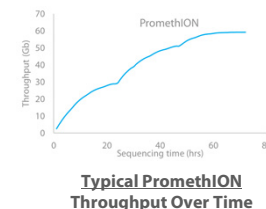
- Nanobind cleanup is used to remove free adapters, perform buffer exchange, and concentrate the library before sequencing, leading to higher adaptation efficiency, pore occupancy, and sequencing throughput
- Alcohol-free chemistries don't damage proteins and are compatible with all nanopore sequencing methods
- Nanobind cleanup does not reduce read length or shear megabase DNA

Nanobind-Enhanced Ultra Long (UL) Sequencing



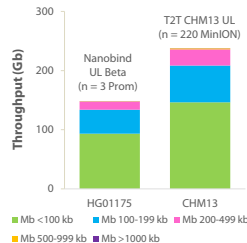
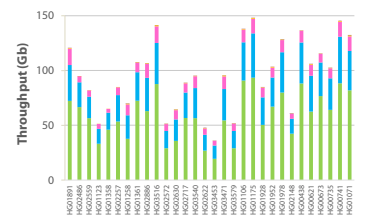
- Optimized UHMW DNA extraction and transposase-based library preparation maximizes adaptation of ultra long DNA and pore occupancy, leading to 10-100X greater throughput than previous ultra long sequencing methods

- UHMW DNA is extracted using Nanobind CBB Big DNA Kit
- Library is prepared using Oxford Nanopore Ultra-Long DNA Sequencing Kit
- Library cleanup is performed using Nanobind UL Library Prep Kit

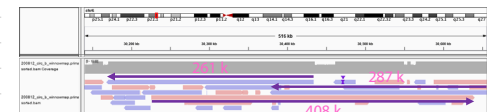
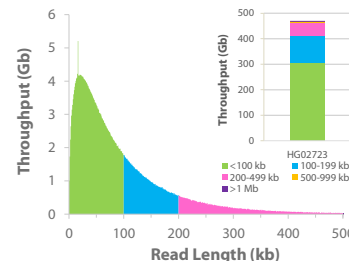


- Standard protocol creates a 3X library that is split into thirds. Each third is sequenced for 24 hours, using the ONT Flow Cell Wash Kit to flush and reload. This maximizes flow cell throughput
- Library is full compatible with R9.4 and R10.3 MiniON, GridION, and PromethION flow cells
- Method typically generates read length N50 of 50-100+ kb and data throughputs of 10-25 Gb on MiniON/GridION and 40-100 Gb on PromethION with longest reads of 2-4+ Mb

Human Pangenome Reference Consortium Pilot Study



- Inexpensive, high coverage UL reads (>100 kb in length) can unambiguously span long repetitive and difficult to assemble regions
- 30 HG cell lines were sequenced on three R9.4 PromethION flow cells each using a beta version of the method
- Read length N50 ranged from 61 kb (HG02148) to 93 kb (HG03453). The longest read was 3.5 Mb.
- T2T CHM13 assembly required 220 MiniON flow cells to generate sufficient ultra long coverage.
- New method can generate similar amounts of ultra long data using only 3-5 PromethION flow cells



UL Reads From a MinION Run Covering Portion of MHC Region

- An additional 468 Gb of R9.4 UL data was generated on HG02723 (left). Data set contains 50X coverage in reads >100 kb and 20X coverage in reads >200 kb.
- High coverage R10.3 UL data is also being generated
- Phased T2T assembly, SNP, and SV analysis is underway

Total Throughput By Genome + Read Length