Solution-based isolation of ultra-high molecular weight (UHMW) genomic DNA from fresh/frozen human blood and cultured cells in less than 4 hours

H. B. Sadowski, C. Proskow, A. Files, K. Pham, Y. Zhang, G. Pljevaljcic, A. Hastie and M. Borodkin

Bionano Genomics, San Diego, California, United States of America

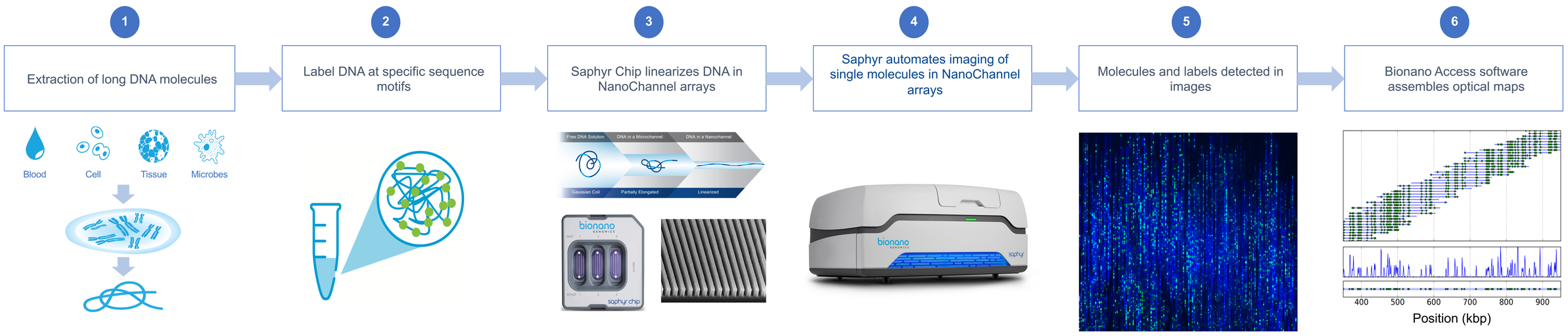
Abstract

Optical mapping of genomic DNA on the Bionano Genomics Saphyr® system for genome assembly or structural variation detection relies on starting with UHMW DNA. To achieve this, we developed methods for the isolation of genomic DNA that involves the embedding of material in agarose plugs, and overnight lysis *in situ* with detergent and proteinase K. After extensive washing on day 2, the agarose plug is melted, treated with beta agarase, and the liberated DNA drop dialyzed. The UHMW DNA is ready for quantification and labeling on day 3. This "plug lysis" method is extremely robust, but it is also labor intensive, difficult to automate, lengthy and expensive. To address these shortcomings, we coupled solution-based lysis with a purification step that leverages a novel process to bind, wash and elute UHMW genomic DNA. This entire protocol can be conducted in less than 4 hours on a batch of 6 samples, allowing 12 samples to be processed in one day. The eluted material is ready to use by day 2 and contains high quality DNA that is clean enough for the direct label and stain (DLS) protocol. The resulting labeling metrics of this labeled DNA on a Saphyr Chip® are comparable to labeled DNA isolated by the traditional "plug lysis" protocol. We have validated protocols for fresh/frozen human blood and cells, and are developing protocols for plant and animal tissue. These protocols are automatable, providing the needed solution for researchers needing to purify DNA from hundreds to thousands of individuals per year.

Background

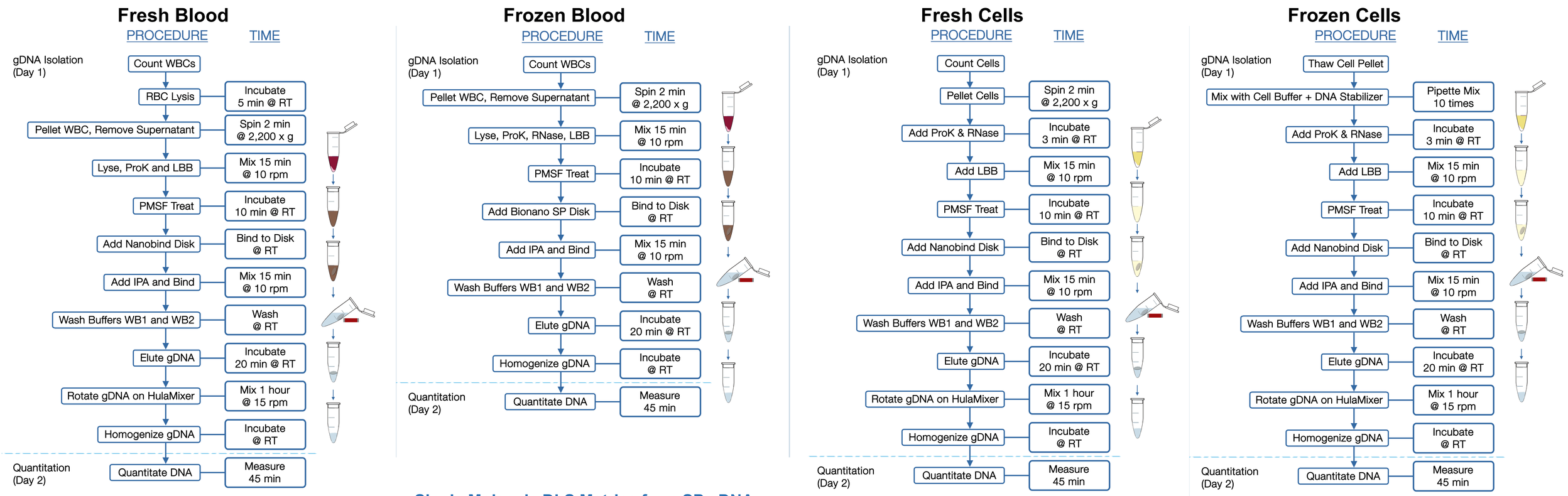
Generating high-quality finished genomes replete with accurate identification of structural variation and high completion (minimal gaps) remains challenging using short read sequencing technologies alone. The Saphyr™ system provides direct visualization of long DNA molecules in their native state, bypassing the statistical inference needed to align paired-end reads with an uncertain insert size distribution. These long labeled molecules are *de novo* assembled into physical maps spanning the entire diploid genome. The resulting provides the ability to correctly position and orient sequence contigs into chromosome-scale scaffolds and detect a large range of homozygous and heterozygous structural variation with very high efficiency.

Methods



(1) Long molecules of DNA are labeled with Bionano reagents by (2) incorporation of fluorophores at a specific sequence motif throughout the genome. (3) The labeled genomic DNA is then linearized in the Saphyr Chip using NanoChannel arrays (4) Single molecules are imaged by Saphyr and then digitized. (5) Molecules are uniquely identifiable by distinct distribution of sequence motif labels (6) and then assembled by pairwise alignment into *de novo* genome maps.

SP Workflows and Results



Bloods (EDTA) from the San Diego Blood Bank were mixed well & WBCs counted by Hemocue. For fresh bloods, the volume required for 1.5 million cells is placed in a 1.5 ml microfuge tube. After RBC lysis, the resulting WBC pellet is resuspended in cell buffer, transferred to a protein lobind tube & subjected to SP protocol. For frozen bloods, 1ml aliquots are frozen and stored at -80C. After thawing, mixing and Hemocue counting, the volume required for 1.5 million cells is transferred to a protein lobind tube. After centrifugation, most of the sup is removed and remaining WBCs subjected to SP protocol. GM12878 cells were grown & cells counted by hemocytometer. For fresh cells, 1.5 million cells in growth medium were pelleted, resuspended in cell buffer, transferred to a protein lobind tube and subjected to SP protocol. For frozen cells, after pelleting 1.5 million cells, the cells were resuspended in cell buffer with DNA stabilizer, repelleted and frozen at -80C until use. After quick thawing, the cell pellets were resuspended with cell buffer with DNA stabilizer, transferred to protein lobind tubes & subjected to SP protocol. After elution from Nanobinds, the gDNAS are mixed end over end for one hour and allowed to rest overnight at room temperature. gDNA concentrations of eluted DNAs were measured on Day 2 and 750 ng of input DNA used for DLS labeling.

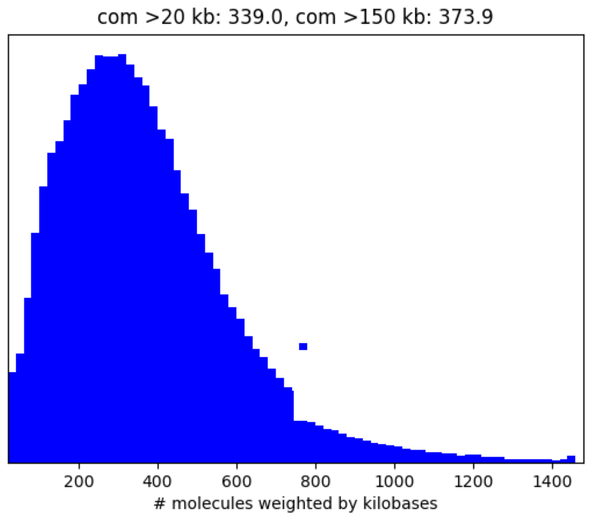
Single Molecule DLS Metrics from SP gDNA

Sample Type	gDNA [ ] (ng/uL)	Gbp / Scan	N50		Label Density	Map Rate
			> 20kbp	> 150kbp		
Fresh Blood (4 users, 8 bloods, 24 samples)						
Mean	59.0	36.1	233.4	300.5	15.4	89.6%
SD	12.1	9.3	58.4	42.0	0.4	3.6%
Frozen Blood (3 users, 17 bloods, 34 samples)						
Mean	57.4	27.8	203.0	264.6	15.2	90.2%
SD	12.0	8.5	28.5	22.1	0.3	2.5%
Fresh Cells (4 users, 2 cultures, 10 samples)						
Mean	113.4	31.8	283.2	317.7	14.9	93.1%
SD	16.2	14.0	39.0	37.4	0.3	0.8%
Frozen Cells (4 users, 2 cultures, 12 samples)						
Mean	106.9	21.4	264.4	319.4	15.2	89.6%
SD	29.9	11.7	40.2	32.1	0.4	5.2%

DLS labeled SP DNAs from fresh/frozen blood and cullured GM12878 cells were run on the Saphry platform and single molecule statistics summarized

DLS labeled SP gDNA from a a fresh pellet of GM 12878 cells was analyzed on the Bionano Saphyr system and the the number of accumulated DNA molecules was plotted as a function of size (x kbp bins).

Size distribution of DLS labeled SP gDNA molecules isolated from fresh GM 12878 cells



Conclusions

- Bionano Prep SP provides a rapid gDNA isolation protocol from human blood and cells (~ 4 hours total time)
- The single molecule DLS labeling metrics of SP gDNA are comparable or exceed those from plug lysis DNA
- Bionano Prep SP protocols have been extended to include heparin bloods and bone marrows
- They are automatable

Reference

Mak AC et al. Genome-wide structural variation detection by genome mapping on nanochannel arrays. Genetics (2016); 202: 351-62.

Cao H et al. Rapid detection of structural variation in a human genome using nanochannel-based genome mapping technology. GigaScience (2014); 3(1):34