Introduction

Large-scale genome population studies have historically involved sequencing individual genomes using short-reads and aligning to a consensus reference sequence. Shortcomings of this method include poor performance on structural variation and loss of information relevant to variant calling including haplotype knowledge and genetic phasing information. While short-read sequencing provides sufficient power to call single nucleotide variants (SNVs) across most of the consensus genome, a comprehensive analysis of the genome of each individual sequenced is not possible.

The 10x Genomics® Chromium™ Genome Solution enables researchers to fully reconstruct long range haplotypes, reveal structural variation, and detect variants in previously inaccessible and complex regions of the genome. The 10x Genomics® Chromium™ Genome Solution generates a unique datatype, called Linked-Reads, by tagging short reads derived from the same genomic DNA (gDNA) fragment with a molecular barcode. This molecular barcoding occurs across >1 million partitions of long input DNA, thus allowing the reconstruction of long-range information from short-read data and enabling a more comprehensive analysis of the genome. This unique approach reveals the true diploid nature of the human genome and unlocks the full spectrum of genetic variation, such as SNVs, indels, and large-scale structural rearrangements (including inversions and translocations).

PerkinElmer® and 10x Genomics® have created an efficient, automated workflow to maximize high molecular weight (HMW) gDNA extraction, sample QC, and library preparation, providing Linked-Read sequencing capabilities for a variety of sample types including dried blood spots (DBS).
Dried Blood Spots

DBS sample collection is a powerful tool in screening programs and large population-based surveys, but has been a difficult sample type for long-range sequencing due to the low quantity and poor quality of extracted DNA.

DBS samples have been used for decades to provide samples for a variety of applications. DBS sampling delivers a number of benefits not realized by other blood sampling protocols:

- Minimal blood volume is required (approximately 30 – 100 µL per spot).
- Sample collection is minimally invasive for the donor.
- The samples are easy to handle and transport.
- The samples are stable for days at ambient temperature, and for years under refrigeration.

These characteristics make DBS samples highly desirable for research, pathogen detection, biobanking, and other long-term applications. The use of DBS samples for nucleic acid studies does, however, have challenges:

- Impurities such as cell fragments and proteins carried through the workflow
- Nucleic acid denaturation during traditional alkaline elution and lysis methods, resulting in reduced yield
- Excessive fragmentation during traditional extraction methods, resulting in fragments too small for NGS applications

Most DNA extraction methods using DBS samples result in poor yields and excessive fragmentation, making these archival samples inaccessible to long-range sequencing.

Workflow Overview

Automated workflows provide a number of benefits that enable researchers to gain value by minimizing manual intervention with respect to assay preparation. Not only is assay preparation time dramatically decreased through the automation of a workflow, but also reproducibility, precision, and result quality also typically improve.

10x Chromium™ Genome Library preparation throughout the automated process described here the PerkinElmer suite of products provides advantages such as:

- Extraction of gDNA that preserves the integrity of HMW input DNA
- Optimal manipulation of the novel 10x Genomics® gel bead and partitioning oil reagents
- Reproducible high-throughput automated sample scale-up
- Minimization of process variability and increased efficiency
- Generated libraries are currently compatible with most Illumina® sequencers

Here we describe a sequence-verified, workflow solution for the 10x Genomics® Long Linked-Read Chromium™ Genome v2 DNA Kit using DBS samples.

![Figure 1. PerkinElmer and 10x Genomics® long, Linked-Read workflow.](image-url)
DNA Isolation

The PerkinElmer® chemagen™ technology provides the solution for challenges associated with the isolation of HMW gDNA from DBS samples by eliminating the traditional alkaline elution, lysis, the high temperatures, shaking, and centrifugation steps that result in nucleic acid denaturation and excessive fragmentation. The chemagen™ technology’s automated magnetic separation procedure uses chemagen™ M-PVA magnetic beads to isolate and purify nucleic acid from DBS. The beads have a high affinity for nucleic acids and low protein binding, resulting in very pure DNA/RNA elution. The chemagen™ technology features magnetizable rotating rods, combining the transfer and suspension of magnetic beads, to extract ultra-pure, HMW gDNA fragments.

Figure 2. chemagen™ technology’s automated magnetic separation features magnetizable rotating rods, combining the transfer and suspension of magnetic beads, to isolate ultrapure, HMW gDNA.

The LabChip® GX Touch™ nucleic acid analyzer was used to analyze HMW gDNA extracted using chemagic™ technology. It delivers rapid capillary electrophoresis analysis for DNA sample quality control, with fragment analysis in as few as 30 seconds per sample.

Figure 3. Representative dried blood spot electropherogram using Genomic DNA Assay on the LabChip® GX Touch™ nucleic acid analyzer.

A pulse field gel illustrates the size distribution of the genomic DNA extracted using the chemagic 360 technology. This step is critical in determining the quality of the DNA as HMW gDNA is required for Linked-Read sequencing and can provide visualization and understanding of the quality of DNA for downstream 10X Genomics® automated assay setup.

Quantitation and normalization of gDNA is required prior to 10x Genomics® chip setup. Maintaining the integrity of the HMW gDNA for long-range genomic performance and phasing is critical to this process. Sample normalization was performed on the Sciclone® NGSx workstation utilizing intuitive normalization protocol optimized to eliminate gDNA degradation.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Date of Sample Collection</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2016</td>
</tr>
<tr>
<td>2</td>
<td>2016</td>
</tr>
<tr>
<td>3</td>
<td>2010</td>
</tr>
<tr>
<td>4</td>
<td>2010</td>
</tr>
<tr>
<td>5</td>
<td>2016</td>
</tr>
<tr>
<td>6</td>
<td>2016</td>
</tr>
<tr>
<td>7</td>
<td>2012</td>
</tr>
<tr>
<td>8</td>
<td>2012</td>
</tr>
<tr>
<td>9</td>
<td>2012</td>
</tr>
<tr>
<td>10</td>
<td>2012</td>
</tr>
<tr>
<td>11</td>
<td>2016</td>
</tr>
<tr>
<td>12</td>
<td>2016</td>
</tr>
</tbody>
</table>

Figure 4. Results – Pulsed Field Electrophoresis
**Generation of Barcoded Libraries**

The 10x Genomics® Chip Holder was placed within a custom automation fixture on the Sciclone G3 NGSx workstation. Two chips can be placed within the holder for increased throughput workflows. Rows 1, 2, and 3 were filled with genomic DNA solution (see image above), gel-beads, and partitioning oil, respectively. Gel-beads were thoroughly mixed before use with specific speeds and careful meniscus tracking.

Once the chip was loaded, it was placed within the 10x Genomics® Chromium™ Controller. The onboard microfluidic process partitions ~1 ng of DNA (~375-450 haploid genome equivalents) into >1 million nanoliter droplets, each containing a unique barcode introduced by a co-partitioned gel bead, resulting in gel-beads-in-emulsion (GEMs).

While not part of the protocol, the GEMS were measured by high-speed video to demonstrate proper single gel-bead loading and gel-bead:reagent ratios using the Sciclone G3 NGSx workstation.

GEMs were recovered from the chip and pipetted into a 96-well plate for subsequent isothermal reaction to produce barcoded fragments.

Following an isothermal amplification reaction, the recovery reagent was mixed with the GEMs to disrupt the emulsion and recover the reaction products. Extraction of phase separated GEM products during the recovery requires carefully choreographed pipetting speeds and head positioning motions.

Sample clean-up was performed post-recovery through successive Dynabead® Beads (Thermo Fisher Scientific®) and SPRI clean-up steps. The bead clean-up requires careful handling as pipetting beads can be difficult to disperse. Stringent control over pipetting was required to ensure robust bead handling, thus, eliminating the need for the visual feedback that an operator would normally provide during a manual process.

The Sciclone G3 NGSx automated 10x Genomics® protocol has been designed to take careful consideration of the bead clean-up to ensure efficient nucleic acid recovery for library preparation.
Automated Library Preparation and Quality Control

An automated library preparation protocol was developed on the Sciclone G3 NGSx workstation to complete the 10x Genomics® Chromium™ Genome Workflow. The protocol was verified through library quality control on the LabChip® GX Touch™ nucleic acid analyzer (shown below), and subsequent sequencing. Final libraries are compatible for sequencing on most Illumina® systems, including the NovaSeq® sequencer.

Completed libraries are quantified using qPCR and Illumina® sequencing protocols.

RESULTS

DNA Quality by Linked-Read Sequencing Data

The performance of the Chromium™ Genome v2 DNA Kit is a function of both DNA quality and sequencing depth. Starting the process with HMW gDNA will typically result in better application performance, including increased haplotype phase block length and ability to call structural variants (see Technical note CG00045 on the 10x Genomics® website for more information). Libraries generated from five DBS samples have an average molecule length ranging from ~5 kb to ~11 kb with up to 20% of DNA molecules measuring longer than 20 kb (Figure 10).

Long Range Haplotype Reconstruction

Some application metrics such as haplotype block size are dependent on DNA quality (Figure 11), while others, like single nucleotide polymorphism (SNP) sensitivity, are minimally affected by DNA quality.

Figure 11. Examples of Chromium™ Genome Technology Performance: Left panel. Phase block distribution of two DBS samples with different molecule lengths as measured by the Chromium™ system. DBS_3 (upper left panel) has a calculated molecule length of ~5.7 kb, where DBS_2 (lower left panel) has a calculated molecule length of ~10.7 kb. Phase block distribution is shifted toward longer phase blocks as the DNA quality improves. Right panel. The magnitude of difference in the N50 Phase block length (upper right panel) is consistent with the difference in molecule length of the respective samples (lower right panel).
Variant Calling

The Chromium™ Software Suite provides a variety of algorithms and visualization tools that make use of the molecular barcoding to call a wide range of variant types. Figure 12 shows examples of detected structural variations, including copy neutral and copy number variants, in sample DBS_5.

![Sample Matrix View](image)

Figure 12. Example of matrix view of various structural variants in DBS. A 90 kb inversion, a homozygous and a heterozygous deletion were successfully identified in DBS_5 which average molecule length was ~5.7 kb

SUMMARY

An automation workflow for sample extraction and preparation in conjunction with 10x Genomics® chemistry is used to generate long range genomic context using today’s short read sequencing technology. The specific 10x Genomics® material handling can currently only be accomplished using flexible control features such as those found on the Sciclone G3 NGSx workstation.

The combination of the chemagen™ technology, a LabChip® GX Touch™ nucleic acid analyzer, the Chromium™ Genome technology, and a Sciclone G3 NGSx workstation employed in the generation of sequenceable library produced linked read sequencing results of up to 50 kb allowing for analysis of known and unknown genetic variants.

We would like to acknowledge the Center for Applied Genomics, Children’s Hospital of Philadelphia, for their assistance in developing the Sciclone script with PerkinElmer.