Roche

Application Note

WGS on MGI sequencing platforms

Authors

Hui Kang
Associate Science Value Manager

Terry Wang
Science Value Manager

Yi-Wen Yan
Senior Science Value Manager

Roche Diagnostics (Shanghai) Limited

Jacqueline Meyer
Support & Training Senior Specialist, Global Customer Support

Peter Ristow
Senior Scientist, Research and Early Development

Roche Dia Cape Town, South Africa

Andrea Wijeweera
Product Manager, Sequencing Solutions

Kui Qian
Bioinformatics Consultant

Roche Dia Asia Pacific, Singapore

Rui Chen
Principal Scientist, Reagent and Assay Development

Roche Sequencing Solutions, Pleasanton, CA USA

Date of first publication: August 2018

Publication date of this version: August 2022

Implementation of KAPA HyperPrep Kit for whole-genome sequencing on MGI Sequencing Platforms employing DNBSEQ™

MGI sequencing platforms implement DNBSEQ™ technology and provide a cost-effective and flexible alternative to other short-read sequencing technologies. KAPA HyperPrep Kits, combined with KAPA cleanup beads, and KAPA HiFi HotStart ReadyMix Kits provide a sample preparation solution for human whole-genome sequencing on MGI sequencing platforms.

Introduction

MGI sequencing platforms have become strong competitors against other existing short-read sequencing technologies. The key to MGI's growing success is their underlying technology, coined DNBSEQ™, which enables fast turnaround time at significantly reduced per-base cost. This technology combines 1) single-stranded circular (ssCir) library construction, 2) generation and loading of DNA nanoballs (DNBs) onto patterned nanoarrays, and 3) combinatorial probe anchor synthesis (cPAS) sequencing.[1]

Due to the novelty, flexibility and accuracy of the sequencing technology, MGI platforms garnered significant attention since 2016, and an influx of peer-reviewed papers have demonstrated their utility in various applications in key research areas, including whole-genome sequencing (WGS)[2,3], single cell[4] and bulk[5,6] RNA sequencing, metagenomics[7] and targeted applications such as whole exome sequencing.[8,9]

When one examines the technology in more detail, several advantages are evident. Each DNB consists of tandemly connected copies of an original template (ssCir DNA) generated through rolling circle amplification (RCA), therefore does not accumulate clonal replication errors like conventional library amplification methods using polymerase chain reaction (PCR).[1] This advantage enables exceptional base calling accuracy in sequencing data generated on the MGI platforms. The negatively charged DNBs are then adsorbed onto a patterned silicon array with a positively charged surface coating.[1] Due to careful DNB generation and strict size selection, each DNB exclusively occupies one single position on the array through an electrostatic interaction that also serves to repel other DNB. This process achieves a high loading density on the array with a reduced risk of optical duplicates.[1] Moreover, PCR duplicates are reduced and index misassignment is virtually eliminated because DNBs are loaded onto the array in the absence of PCR reagents.[10] The accuracy of the MGI platforms are further improved through the use of their innovative cPAS sequencing chemistry.[1]

A range of low (DNBSEQ E series), medium (DNBSEQ-G50*), high (DNBSEQ-G400/G400 Fast*) and ultra-high throughput sequencers (DNBSEQ-T7) are available. This allows users to select from a series of fast and flexible systems to support individual needs. The core library construction process for the MGI platform is straightforward and similar to that used for other sequencing platforms such as Illumina. Notwithstanding, very few library preparation solutions from third parties are currently available for these MGI platforms.

KAPA HyperPrep Kits offer a very efficient, automation-ready, single-tube library construction protocol. Extensive chemistry and protocol optimization has enabled high conversion rates, thereby expanding the pool of samples that can be successfully processed for a variety of sequencing applications. The flexible protocol can be fine-tuned to optimize performance with specific sample types, or to meet operational requirements. A suite of accessory products, such as KAPA Pure Beads, KAPA HyperPure Beads, and KAPA HiFi HotStart ReadyMix Kits rounds out the sample preparation workflow.

In this Application Note we demonstrate the utility of the KAPA HyperPrep Kit to generate high-quality WGS libraries for sequencing on an MGI platform and provide a step-by-step workflow. Additionally, we present a dataset generated on the DNBSEQ-G400, from libraries prepared using the KAPA HyperPrep Kit and a HapMap DNA sample (NA12891, 500 ng input).

*In parts of Asia-Pacific, e.g., China, the DNBSEQ-G400/G400 Fast is known as MGISEQ-2000/2000 Fast and the DNBSEQ-G50 as the MGISEQ-200. Forthwith this note will refer to DNBSEQ-G400 in reference to MGISEQ-2000.

Materials and methods

Experimental design

Whole-genome human shotgun libraries were prepared from a characterized HapMap genomic DNA sample (NA12891; Coriell Institute of Biomedical Research), using the KAPA HyperPrep Kit (Figure 3, **Appendix**). A detailed, step-by-step protocol is included as an Appendix. DNB preparation and DNBSEQ™ using a DNBSEQ-G400 sequencer was performed by MGI Tech Co., Ltd (Shenzhen).

Library construction and sequencing

Input DNA: Both the quantity and quality of input DNA have a profound impact on the outcome of library construction for next-generation sequencing (NGS).[11] The workflow in this study was executed using a high-quality HapMap sample (NA12891). However, the workflow can also be implemented using both commercial (e.g., FFPE reference standard HD300, Horizon Discovery, Cambridge, UK) and extracted formalin-fixed, paraffin-embedded (FFPE) samples (data not shown; **please contact Technical Support** for guidance when using degraded sample inputs for this application).

DNA fragmentation: NA12891 DNA was diluted to the appropriate concentration, and transferred to a Covaris® MicroTUBE-50. Mechanical fragmentation was performed with a Covaris M220 Focused-Ultrasonicator using the settings indicated in Section 4.1 (**Appendix**). A volume of 50 µL of each sheared DNA sample was recovered for library construction.

Size selection: Post-fragmentation size selection of sheared DNA (500 ng per library in 50 µL volumes) was performed with either KAPA Pure Beads (Roche) (bead-to-sample volume ratio of 0.6X for the first size cut i.e., the upper boundary, and 0.8X for the second size cut i.e., the lower boundary) or KAPA HyperPure Beads (Roche) (bead-to-sample volume ratio of 0.55X for the first size cut and 0.75X for the second size cut). To accommodate ease of pipetting, the volume of the Covaris-sheared DNA input was adjusted to 100 µL with 10 mM Tris-HCl (pH 8.5) prior to size selection. These ratios were previously optimized to yield fragmented DNA with a mode size of ~400 bp.

Library construction: Library construction was modelled on the workflow of the MGIEasy Universal DNA Library Prep Kit. The standard KAPA HyperPrep protocol was followed as per **KAPA HyperPrep Kit Instructions for Use** (KR0961 v7.20), up to the adapter ligation step. For adapter ligation, 10 µL of MGIEasy DNA Adapters (10 µM) was used per library, with a 30 minutes ligation reaction. To ensure a color-balanced base composition, adapters were mixed in equal volume prior to addition to each sample as per manufacturer's instructions. Library amplification was performed using the KAPA HiFi HotStart ReadyMix, with 5 µL MGI Primer Mix (provided by MGI Tech Co., Ltd in Shenzhen) for seven amplification cycles.

Library quantification and fragment size assessment: Library size distributions were analyzed with an Agilent® 2100 Bioanalyzer instrument and Agilent High Sensitivity DNA Kit (Agilent Technologies; Santa Clara, CA). Each amplified library was quantified using the Invitrogen Qubit™ Fluorometer and Qubit dsDNA HS Assay Kit (ThermoFisher Scientific).

DNA Nanoball (DNB) Preparation: Library denaturation, single-stranded circularization (ssCir), Exo digestion and cleanup, and DNB generation was performed at MGI Tech Co., Ltd (Shenzhen), as per manufacturer's instructions. The recommended input amount (1 pmol amplified library) was used for ssCir using the MGIEasy Circularization Kit,[12] and subsequently 40 fmol of purified ssCir library was converted to DNBs using the DNBSEQ-G400RS High-throughput Sequencing Set.[13]

Sequencing: WGS was performed on a DNBSEQ™ platform DNBSEQ-G400 instrument (2 x 150 bp paired-end sequencing) using Standard MPS technology. Three replicate libraries (one sample per lane) were sequenced.

Alignment: Read quality was assessed using FASTQC (v0.10.1). Adapter and quality trimmed sequencing reads (SOAPnuke, v2.1.2) were aligned to the human reference genome (build GRCh37/hg19) using BWA-MEM (v0.7.12) with default parameters. Duplicates were marked with Picard (v1.134) and the output SAM file was converted to BAM format as input for GATK (v4.17.0) analysis and variant calling tools.

Data on file.
For Research Use Only. Not for use in diagnostic procedures.

Quality control analysis: The quality control metrics for this workflow was calculated with Picard and included library insert size, alignment, GC bias and genome coverage metrics. Picard *AlignmentSummaryMetrics* was used to summarize the alignment metrics.

Variant calling: Germline variants (SNP and InDel) were called using GATK (*HaplotypeCaller*), and compared to a NA12891 truth variant list from GATK resource bundle (Broad Institute)[14] using the GATK *GenotypeConcordance* tool.

Results and discussion

Library construction metrics

The aforementioned Covaris®-shearing parameters yielded fragmented DNA with an average size of approximately 300 – 400 bp (Figure 4, **Appendix**). Once combined with post-fragmentation size selection [using either KAPA Pure Beads or KAPA HyperPure Beads] DNA fragments with an average size of 400 – 450 bp were obtained (Figure 5, **Appendix**). The average yield after fragmentation and size selection was ~130 ng (~2.7 ng/µL) equating to an approximate loss of 70% of the input material. The loss is largely inherent to the nature of size selection, which removes fragments whose size falls out of the selected band. To account for this anticipated loss, an input of 500 ng was used to ensure that at least 50 ng of properly fragmented, size-selected DNA would be available for end repair and A-tailing.

After ligation, the average yield varied between 75 and 120 ng (12 – 20 nM), indicating high DNA retainment and efficiency in library conversion using the KAPA HyperPrep Kit. To continue with ssCir and DNB preparation, the protocol recommends at least 1 pmol of material with a minimum of no less than 0.5 pmol.

With an average library size of ~500 bp, 1 pmol of material is the equivalent of 330 ng of library. In this study, libraries were amplified to generate sufficient material for downstream ssCir processing. On average, 600 – 1000 ng of amplified library was obtained when using seven amplification cycles, which was sufficient for two ssCir events. An example electropherogram of an amplified library generated with this workflow is provided in Figure 6 (**Appendix**). Users may choose to reduce the number of cycles depending on the actual yield after ligation. To increase pre-ssCir library yield, another possible workflow optimization is to improve ligation efficiency, through fine-tuning DNA input amounts, size selection parameters (single- versus double-sided), ligation time, and adapter concentration.

Sequencing metrics

A summary of the sequence data generated in this study is given in Table 1, whereas alignment and coverage statistics for the replicate libraries are summarized in Table 2. As each replicate was sequenced on a separate lane, the amount of data obtained per sample was ~140 Gb resulting in a mean coverage of ~50X. Figure 1 illustrates the coverage uniformity plots.

Duplication rate: Despite performing library amplification low duplication rates were observed (<4%) demonstrating one key advantage of the technology (Table 2). Optical duplicates were not observed.

GC-bias: GC-bias plots are shown in Figure 2. The genome was well represented in the data with extremely GC-rich areas moderately under-represented.

Variant Calling: Variant calling results and concordance to the NA12891 truth variant set from the GATK resource bundle (Broad Institute) are given in Table 3. SNPs and InDels were detected with high sensitivity and specificity in all three replicates.

Table 1. Sequencing statistics

KAPA HyperPrep Technical Replicate	Lane	PE Reads PF (M)	Total Gb PF	% >Q30 (average)	% >Q30 (Read 1)	% >Q30 (Read 2)	% ESR
1	L01	461.11	144.3	84.94	85.21	84.67	79.96
2	L02	463.69	145	86.25	86.27	86.22	80.63
3	L03	470.41	147.8	86.33	86.03	86.64	81.95

PE=paired-end; PF=passed filter; Gb=gigabases; M=million
ESR=Effective spot rate; number of successfully fixed DNBs on patterned array that has the ability to produce sequencing reads

Table 2. Alignment and coverage statistics

KAPA HyperPrep Technical Replicate	% PF reads aligned	Duplication rate	Optical duplicates	Median insert size	Mean coverage
1	99	3.77	0	384	47.8
2	99	3.93	0	366	48.1
3	99	3.87	0	369	48.9

PF=passed filter

Table 3. Variant analysis

KAPA HyperPrep Technical Replicate	Mismatch rate	InDel rate	% Chimeras	Sensitivity		Specificity	
				SNPs	InDels	SNPs	InDels
1	0.0103	0.0003	0.38	99.8	95.8	98.8	93.7
2	0.0092	0.0003	0.44	99.8	96.0	98.7	93.8
3	0.0094	0.0003	0.42	99.8	96.1	98.7	93.8

Data on file.
For Research Use Only. Not for use in diagnostic procedures.

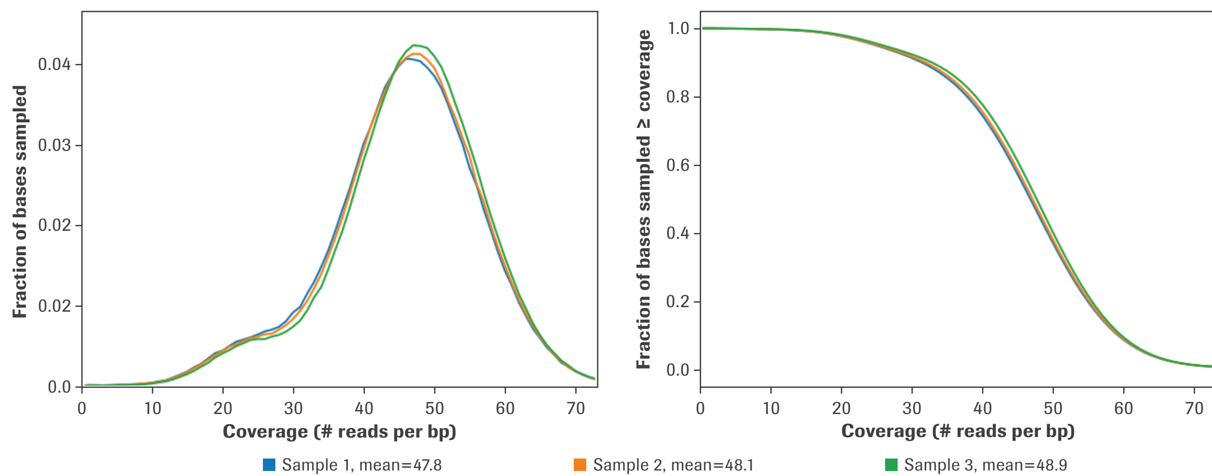

Figure 1. Coverage uniformity plots for human WGS libraries prepared with the KAPA HyperPrep Kit and sequenced on the DNBSEQ-G400. Replicate libraries yielded identical coverage profiles, with a sharp peak indicating coverage uniformity. The mean coverage for each sample was ~48X with a high proportion of bases covered >30X.

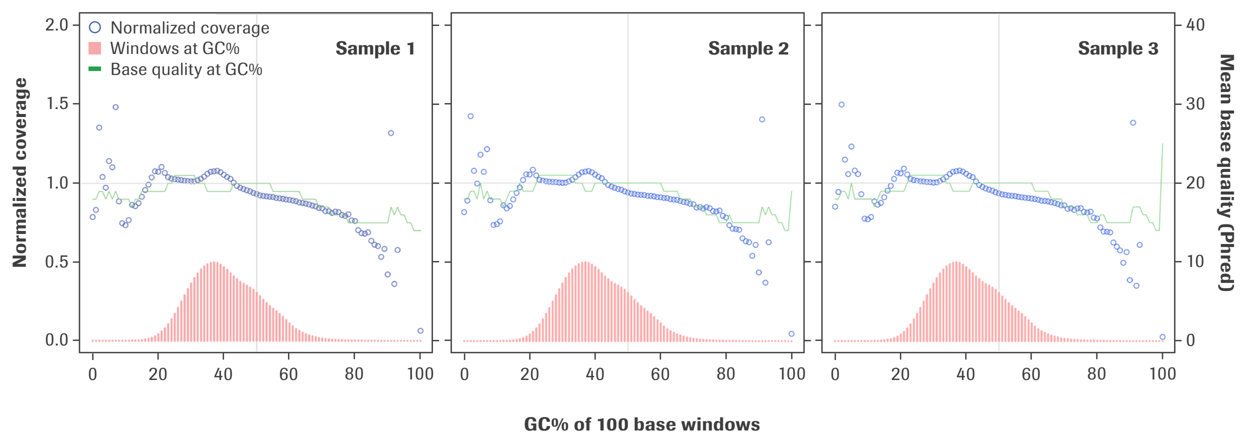

Figure 2. GC-bias plots for human WGS libraries prepared with the KAPA HyperPrep Kit and sequenced on the DNBSEQ-G400. Bins with a balanced GC content (30% – 50%) were well represented in the data. Extremely GC-rich bins were moderately under-represented. Plots were generated with Picard *CollectGCBiasMetrics*. Red histograms represent the distribution of genomic GC content for NA12891, calculated for the reference sequence in 100 bp bins. GC bias was assessed by plotting normalized coverage for each bin. If all sample-to-data processes were completely unbiased, all bins would be equally represented, i.e., the plot for each workflow would be a horizontal distribution centered on a normalized coverage of 1.

Conclusions

The aim of this study is to demonstrate that the KAPA HyperPrep Kit and accessory reagents from Roche can be used to construct MGI sequencing platform-compatible libraries. Specific recommendations and a step-by-step protocol for the construction of human WGS libraries are provided in this Application Note. We show that libraries prepared from 500 ng of a commercial preparation of NA12891 human gDNA yielded high-quality sequencing data as evidenced by alignment, coverage, and variant calling statistics.

We did not benchmark the performance of MGI sequencers against Illumina sequencers. Numerous comparative studies have been performed and illustrate that both platforms produce equivalent, high-quality data in a range of applications such as whole-genome sequencing,[15] scRNA-seq,[16] and bulk RNA-seq,[17] just to name a few.

There are several key considerations when following this workflow:

- High DNA inputs (≥500 ng) are essential given the losses associated with mechanical shearing and the requirement for size selection to maintain a uniform size distribution for DNB preparation.
- Library construction was modelled on the workflow of the MGIEasy Universal DNA Library Prep Kit, which includes amplification to achieve the input requirements for DNB preparation. Users may consider a PCR-free workflow, however the workflow and required reagents will be significantly different.

In summary, Roche provides support for solutions that best meet end-user needs based on extensive experience and a deep understanding of the parameters that impact the efficiency of sample preparation. Here, we confirmed the sample preparation solution from Roche to be highly suitable for routine human WGS on DNBSEQ™ MGI platforms.

Data on file.
For Research Use Only. Not for use in diagnostic procedures.

Appendix

Key considerations during experimental planning

Selection of an appropriate KAPA Library Preparation Kit: MGI manufactures library preparation kits that are suitable for use with fragmented DNA (MGIEasy Universal DNA Library Prep Kit) as well as kits with an enzymatic fragmentation module (MGIEasy FS DNA Library Prep Set). Depending on user preference, the KAPA portfolio presents similar solutions—the KAPA HyperPrep Kit, a single tube chemistry and streamlined workflow for mechanically sheared DNA, or the KAPA HyperPlus Kit with integrated, low-bias enzymatic fragmentation. The core library construction steps are similar (despite subtle differences in incubation times and temperatures): both MGI and KAPA kits employ end repair and A-tailing and adapter ligation to generate libraries for downstream processing.

In this study, we evaluated the use of the KAPA HyperPrep Kit for the preparation of adapter-ligated libraries suitable for single strand circularization, Exo digestion, DNB generation and sequencing on the MGI platforms.

Size selection for whole-genome sequencing applications: Mechanical shearing, such as DNA fragmentation using the Covaris® Focused-ultrasonicators, routinely yields a relatively broad size distribution around the targeted fragment length. A uniform size distribution is crucial for DNBSEQ™ technology as the size of the DNBs after RCA are a direct function of polymerization time. Therefore, after a set amount of time all ssCir molecules of equivalent length will have similar DNB dimensions. Size selection is therefore inevitable to ensure a uniform size distribution when preparing libraries for human WGS on an MGI platform. While size selection results in a much narrower final library size distribution, it comes at the cost of a significant amount of input DNA. This can have a profound impact on library yield and complexity, particularly if the mean size of sheared DNA does not correspond well to the desired insert size of the final library.

The KAPA portfolio consists of two types of paramagnetic cleanup beads for tunable size selection; KAPA Pure Beads and KAPA HyperPure Beads. The KAPA HyperPrep Kit further supports post-fragmentation, post-ligation or post-amplification size selection. Each strategy has potential advantages and disadvantages. Prior experience with KAPA Pure Beads showed that recovery of fragmented DNA was less efficient when performing post-fragmentation size selection compared with size selection later in the library construction workflow.[18] KAPA HyperPure Beads is a product upgrade that enables improved recovery of nucleic acids. It is important to note that the bead solids and buffer formulation of KAPA HyperPure Beads is different from that of KAPA Pure Beads. Thus size selection parameters will vary depending on the bead product utilized. In this study, both bead types were evaluated for post-fragmentation size selection and reaction cleanups.

Detailed library construction workflow

1. Required reagents

The following reagents are supplied by Roche:

- KAPA HyperPrep Kit
  - Roche PN: 07962312001 (8 libraries)
  - Roche PN: 07962347001 (24 libraries)
  - Roche PN: 07962363001 (96 libraries)
- KAPA Pure Beads
  - Roche PN: 07983271001 (5 mL)
  - Roche PN: 07983280001 (30 mL)
  - Roche PN: 07983298001 (60 mL)
- KAPA HyperPure Beads
  - Roche PN: 08963835001 (5 mL)
  - Roche PN: 08963843001 (30 mL)
  - Roche PN: 08963851001 (60 mL)
  - Roche PN: 08963860001 (450 mL)
  - Roche PN: 08963878001 (4 x 60 mL)

The following reagents are not supplied by Roche, and must be sourced from general laboratory stocks or third-party suppliers:

- MGIEasy DNA Adapters Kit V1.0
  - MGI PN: 1000005284 (16 rxn, 16 barcodes)
  - MGI PN: 1000005282 (96 rxn, 96 barcodes)
- MGIEasy Primer Mix*
- MGIEasy Circularization Kit
  - MGI PN: 1000005259 (16 rxn)
- Freshly prepared 80% ethanol
- Elution buffer (10 mM Tris-HCl, pH 8.0 at 25°C)
- Nuclease-free water

**Amplification Primer Ordering Info:* To perform library amplification, user-supplied primer mixes are required. Sequences can be obtained from https://en.mgitech.cn/download/files/ by filtering for “Oligos and primers for BGISEQ&DNBSEQ NGS system”.[19]

The following equipment and plasticware are required for this workflow:

- Covaris M220 instrument
- Covaris microTUBE-50 AFA Fiber Screw-Cap
- Themocyclers (for standard PCR)
- Magnetic block e.g., DynaMag 96 Side Magnet (ThermoFisher Scientific Cat. No. 12331D)
- PCR tubes/96-well plates (0.2 mL)
- Pipettes
- Low DNA-binding microtubes (1.5 mL)
- Agilent® Bioanalyzer 2100 or similar instrument
- Qubit™ Fluorometer and appropriate assays

2. Workflow overview

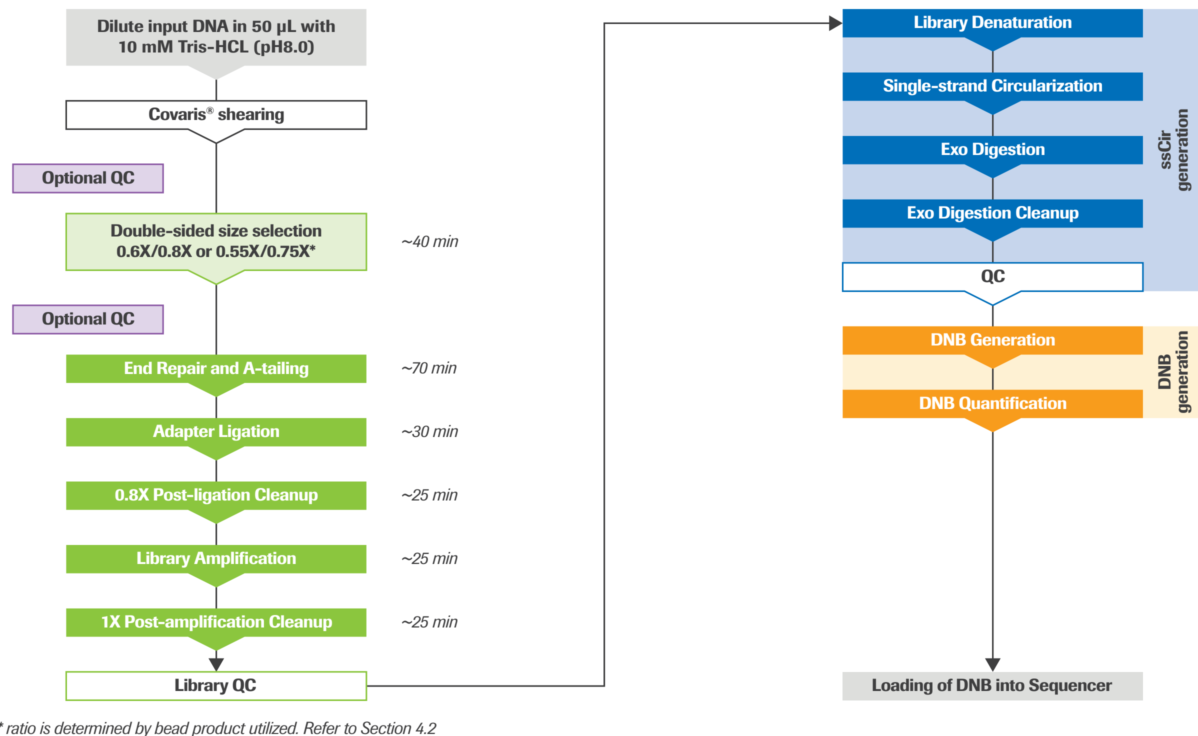

Figure 3. KAPA HyperPrep workflow adapted for DNBSEQ-G400. Replicate aliquots of NA12891 human genomic DNA (500 ng) were sheared using a Covaris M220 (as described in **Section 4.1**). A post-fragmentation size selection step was performed after Covaris shearing followed by KAPA HyperPrep library construction as described in **Section 4.3** and following. After library QC, ssCir and DNB generation utilizes MGI reagents. The total workflow time is approximately 5 hours, excluding the ssCir generation and DNB generation workflow.

3. Quality control of input DNA

Quality control of input DNA is strongly recommended. In this workflow, DNA was quantified using an Invitrogen Qubit™ Fluorometer and Qubit dsDNA HS (High Sensitivity) Assay Kit (ThermoFisher Scientific). These assays are designed to provide accurate quantification of double-stranded DNA molecules present in the sample.

4. Library construction protocol

This protocol describes library construction using the KAPA HyperPrep Kit with size selection after Covaris shearing (post-fragmentation) and is based on the protocol outlined in the **KAPA HyperPrep Kit Instructions for Use** (KR0961 v7.20 or later). Optimal shearing parameters for Covaris instruments other than the M220 will have to be empirically determined.

4.1 Covaris shearing

4.1.1 Prepare a Covaris M220 instrument as per the manufacturer's instructions.

4.1.2 Transfer 550 ng of DNA in a total volume of 55 µL into a Covaris microTUBE-50 AFA Fiber Screw-Cap.

4.1.3 Shear using the following settings:
- Peak Incident Power (Watt): 75
- Duty Factor (percent): 10
- Cycles/Burst (count): 200
- Duration (seconds): 90

4.1.4 Transfer 50 µL of sheared DNA into a new plate/tube and proceed with **Post-fragmentation double-sided size selection**.

NOTE: The Covaris shearing settings provided are expected to yield DNA with the fragment size distribution depicted in Figure 4.

4.2 Post-fragmentation double-sided size selection

4.2.1 Adjust the volume of the Covaris-sheared DNA to 100 µL with 10 mM Tris-HCl (pH 8.0 – 8.5)

4.2.2 Perform the first size cut by adding 0.6X volume of KAPA Pure Beads or 0.55X volume of KAPA HyperPure Beads to the sample:

Component	Volume for size selection: with KAPA Pure Beads*	with KAPA HyperPure Beads*
DNA sample	100 µL	100 µL
KAPA cleanup beads	60 µL	55 µL
Total volume	**160 µL**	**155 µL**
Volume of supernatant to transfer	155 µL	150 µL

*The volume of beads required will differ depending on the bead product utilized. The bead solids and buffer formulation of KAPA HyperPure Beads is different from KAPA Pure Beads, so please pay special attention.

4.2.3 Mix thoroughly by vortexing and/or pipetting up and down multiple times.

4.2.4 Incubate the plate/tube(s) at room temperature for 5 min to bind DNA to the beads.

4.2.5 Place the plate/tube(s) on the magnetic stand to capture the beads. Incubate until the liquid is clear.

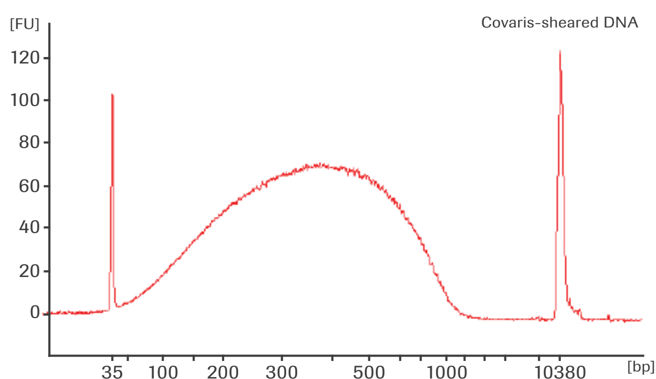

Figure 4. Expected size distribution profile of fragmented DNA sheared using a Covaris M220 instrument. Input gDNA was fragmented with the Covaris M220 instrument using conditions as described in section 4.1. The average size of the fragments were approximately 300 – 400 bp. For size determination, each sample was diluted to within the detection limit of the assay and analyzed using an Agilent® 2100 Bioanalyzer instrument and High Sensitivity DNA Kit.

4.2.6 Carefully transfer the appropriate amount of the supernatant as shown in the table above (containing DNA fragments smaller than those that were intentionally excluded) to a new plate/tube(s). It is critical that no beads are transferred with the supernatant. Discard the plate/tubes(s) with beads to which the unwanted, large DNA fragments are bound.

4.2.7 Perform the second size cut by adding 0.2X volume of KAPA cleanup beads to the supernatant from the first cut:

Component	Volume for size selection	
	with KAPA Pure Beads*	with KAPA HyperPure Beads*
Supernatant from first size cut	155 μL	150 μL
KAPA cleanup beads	20 μL	20 μL
Total volume	**175 μL**	**170 μL**

4.2.8 Mix thoroughly by vortexing and/or pipetting up and down multiple times.

4.2.9 Incubate the plate/tube(s) at room temperature for 5 min to bind DNA to the beads.

4.2.10 Place the plate/tube(s) on a magnetic stand to capture the beads. Incubate until the liquid is clear.

4.2.11 Carefully remove and discard the supernatant.

4.2.12 Keeping the plate/tube(s) on the magnetic stand, add 200 μL of 80% ethanol (freshly prepared).

4.2.13 Incubate the plate/tube(s) on the magnetic stand at room temperature for at least 30 sec.

4.2.14 Carefully remove and discard the ethanol.

4.2.15 Keeping the plate/tube(s) on the magnetic stand, add 200 μL of 80% ethanol.

4.2.16 Incubate the plate/tube(s) on the magnetic stand at room temperature for at least 30 sec.

4.2.17 Carefully remove and discard the ethanol. Try to remove any residual ethanol without disrupting the beads.

4.2.18 Dry the beads at room temperature for 3 – 5 min, or until all of the ethanol has evaporated. Be careful not to over dry the beads as this may result in reduced yield.

4.2.19 Remove the plate/tube(s) from the magnet.

4.2.20 Thoroughly resuspend the beads in 55 μL of the elution buffer (10 mM Tris-HCl, pH 8.0 – 8.5).

4.2.21 Incubate the plate/tube(s) at room temperature for 2 min to elute DNA off the beads.

4.2.22 Place the plate/tube(s) on the magnetic stand to capture the beads. Incubate until the liquid is clear.

4.2.23 Transfer 50 μL of the clear supernatant to a new plate/tube(s) and proceed to **End Repair and A-tailing**.

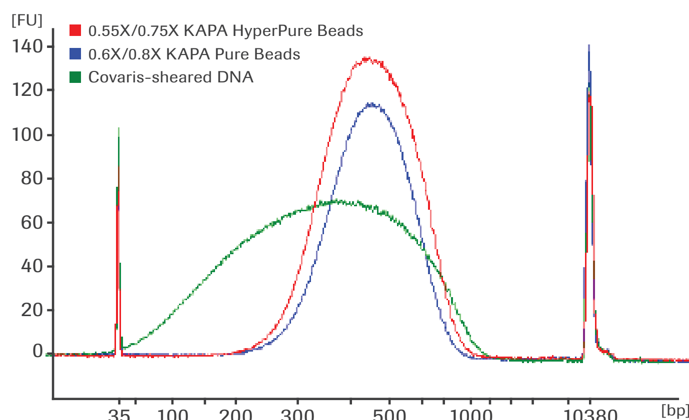

Figure 5. Expected size distribution post size selection. Input DNA was sheared using Covaris® M220 instrument (green curve) as described above. Fragmented DNA was subjected to 0.6X – 0.8X (KAPA Pure Beads, blue curve) or 0.55X – 0.75X (KAPA HyperPure Beads, red curve) double-sided size selection prior to end repair/A-tailing. For size determination, samples were analyzed without dilution using an Agilent® 2100 Bioanalyzer instrument and High Sensitivity DNA Kit.

4.3 End Repair and A-tailing

4.3.1 Assemble each End Repair and A-tailing reaction in a tube or well of a PCR plate as follows:

Component	Volume
Fragmented, size-selected DNA	50 μL
End Repair & A-Tailing Buffer*	7 μL
End Repair & A-Tailing Enzyme Mix*	3 μL
Total volume	**60 μL**

*It is recommended to pre-mix the buffer and enzyme mix, before adding it to the DNA sample in a single pipetting step. Premixes are stable for <24 hr at room temperature, for <3 days at 4°C, and for <4 weeks at -20°C.

4.3.2 Vortex gently and spin down briefly. Return the plate/tube(s) to ice. Proceed immediately to the next step.

4.3.3 Incubate in a thermocycler programmed as outlined below:

Step	Temp	Time
End Repair and A-tailing	20°C	30 min
	65°C	30 min
HOLD	4°C	∞

*A heated lid is required for this incubation. If possible, set the lid temperature to 85°C, instead of the default ~105°C.

4.3.4 Centrifuge the plate/tubes(s) briefly and proceed immediately to **Adapter ligation**.

4.4 Adapter ligation

4.4.1 In the same plate/tubes(s) in which end repair and A-tailing was performed, assemble each adapter ligation reaction as follows:

Component	Volume
End repair and A-tailing reaction product	60 µL
MGIEasy DNA Adapter (10 µM)	10 µL
Ligation Buffer*	30 µL
DNA Ligase*	10 µL
Total volume	**110 µL**

*The buffer and ligase enzyme should preferably be premixed and added in a single pipetting step. Premixes are stable for <24 hr at room temperature, for <3 days at 4ºC

4.4.2 Mix thoroughly and centrifuge briefly.

4.4.3 Incubate in a thermocycler programmed as outlined below:

Step	Temp	Time
Set heated lid	30ºC	On
Adapter ligation	20ºC	30 min
HOLD	4ºC	∞

4.4.4 Centrifuge plate/tube(s) briefly and proceed immediately to **Post-ligation cleanup**.

4.5 Post-ligation cleanup

4.5.1 In the same plate/tube(s), perform a 0.8X bead-based cleanup by combining the following:

Component	Volume
Adapter ligation reaction product	110 µL
KAPA cleanup beads*	88 µL
Total volume	**198 µL**

*Users may employ KAPA Pure Beads or KAPA HyperPure Beads for this step.

4.5.2 Mix thoroughly by vortexing and/or pipetting up and down multiple times.

4.5.3 Incubate the plate/tube(s) at room temperature for 5 min to bind DNA to the beads.

4.5.4 Plate the plate/tube(s) on a magnetic stand to capture the beads. Incubate until the liquid is clear.

4.5.5 Carefully remove and discard the supernatant.

4.5.6 Keeping the plate/tube(s) on the magnetic stand, add 200 µL of 80% ethanol.

4.5.7 Incubate the plate/tube(s) on the magnetic stand at room temperature for >30 sec.

4.5.8 Carefully remove and discard the ethanol.

4.5.9 Keeping the plate/tube(s) on the magnetic stand, add 200 µL of 80% ethanol.

4.5.10 Incubate the plate/tube(s) on the magnetic stand at room temperature for >30 sec.

4.5.11 Carefully remove and discard the ethanol. Try to remove all residual ethanol without disturbing the beads.

4.5.12 Dry the beads at room temperature for 3 – 5 min, or until all of the ethanol has evaporated. Be careful not to over dry the beads as this may result in reduced yield.

4.5.13 Remove the plate/tube(s) from the magnetic stand.

4.5.14 Thoroughly resuspend the beads in 22.5 µL of the elution buffer (10 mM Tris-HCl, pH 8.0 – 8.5).

4.5.15 Incubate the plate/tube(s) at room temperature for 2 min to elute DNA off the beads.

4.5.16 Place the plate/tube(s) on a magnetic stand to capture the beads. Incubate until the liquid is clear.

4.5.17 Transfer the clear supernatant to a new plate/tube(s) and proceed with **Library amplification**.

4.6 Library amplification

4.6.1 Assemble the library amplification reaction as follows and add 25 µL of the library amplification mastermix to 20 µL of purified, adapter-ligated DNA:

Component	Volume
Purified, adapter-ligated DNA	20 µL
KAPA HiFi HotStart ReadyMix (2X)*	25 µL
MGI Library Amplification Primers*,†	5 µL
Total volume	**50 µL**

*ReadyMix and primers should be premixed and added in a single pipetting step.

†Primer sequences can be found on the MGI website, under the document name "Oligos and primers for BGISEQ&DNBSEQ NGS system"[19]

4.6.2 Mix well by pipetting up and down several times.

4.6.3 Amplify the library using the following thermocycling profile:

Step	Temp	Time	Cycles
Initial denaturation	98ºC	45 sec	1
Denaturation	98ºC	15 sec	7
Annealing	60ºC	30 sec	
Extension	72ºC	30 sec	
Final Extension	72ºC	1 min	1
HOLD	4ºC	∞	1

*The number of PCR cycles may need to be optimized depending on the sample type and library yield needed for DNB generation.

4.6.4 Proceed immediately to **Library amplification cleanup**.

Data on file.
For Research Use Only. Not for use in diagnostic procedures.

4.7 Library amplification cleanup

4.7.1 In the same plate/tube(s), perform 1X bead-based cleanup by combining the following:

Component	Volume
Amplified library DNA	50 µL
KAPA cleanup beads*	50 µL
Total volume	**100 µL**

*Users may employ KAPA Pure Beads or KAPA HyperPure Beads for this step.

4.7.2 Thoroughly resuspend the beads by pipetting up and down multiple times.

4.7.3 Incubate the plate/tube(s) at room temperature for 5 min to bind DNA to the beads.

4.7.4 Place the plate/tube(s) on a magnetic stand to capture the beads. Incubate until the liquid is clear.

4.7.5 Carefully remove and discard the supernatant.

4.7.6 Keeping the plate/tube(s) on the magnetic stand, add 200 µL of 80% ethanol.

4.7.7 Incubate the plate/tube(s) on the magnetic stand at room temperature for >30 sec.

4.7.8 Carefully remove and discard the ethanol.

4.7.9 Keeping the plate/tube(s) on the magnetic stand, add 200 µL of 80% ethanol.

4.7.10 Incubate the plate/tube(s) on the magnetic stand at room temperature for >30 sec.

4.7.11 Carefully remove and discard the ethanol. Try to remove all residual ethanol without disturbing the beads.

4.7.12 Dry the beads at room temperature for 3 – 5 min, or until all of the ethanol has evaporated. Be careful not to over dry the beads as this may result in reduced yield.

4.7.13 Remove the plate/tube(s) from the magnetic stand.

4.7.14 Thoroughly resuspend the beads in 30 µL of the elution buffer (10 mM Tris-HCl, pH 8.0 – 8.5).

4.7.15 Incubate the plate/tube(s) at room temperature for 2 min to elute DNA off the beads.

4.7.16 Place the plate/tube(s) on a magnetic stand to capture the beads. Incubate until the liquid is clear.

4.7.17 Transfer the clear supernatant to a new plate/tube(s) and proceed with **Library QC**.

4.8 Library QC

Evaluate the success of library construction by determining the library size distribution via an electrophoretic method. In this workflow, the Agilent® Bioanalyzer 2100 instrument was used together with the High Sensitivity DNA Kit. Quantify the amplified library using the Invitrogen Qubit Fluorometer and Qubit™ dsDNA HS Assay Kit (ThermoFisher Scientific).

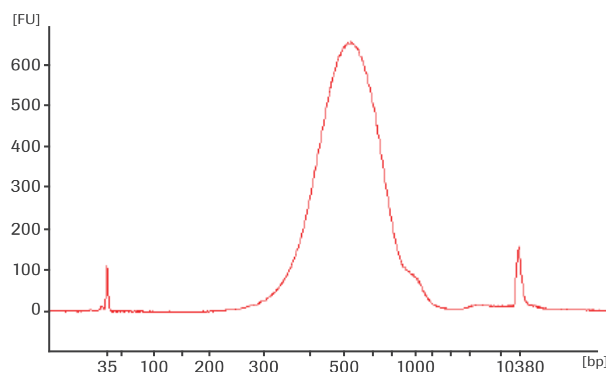

Figure 6. Expected size distribution of amplified libraries. Adapter-ligated libraries were amplified for seven cycles using KAPA HiFi HotStart ReadyMix and the MGI Library Amplification Primer Mix. The size distribution was assessed using an Agilent 2100 Bioanalyzer instrument and High Sensitivity DNA Kit.

Acknowledgements

The authors wish to thank staff from the MGI Tech Co., Ltd in Shenzhen for generously providing us with the MGIEasy DNA Adapters Kit, MGIEasy Primer Mix and sequencing services (DNBSEQ-G400) used in this project.

Data on file.
For Research Use Only. Not for use in diagnostic procedures.

References

1. Drmanac R, Sparks AB, Callow MJ, et al. Human Genome Sequencing Using Unchained Base Reads on Self-Assembling DNA Nanoarrays. *Science.* 2009;327(78). doi:10.1126/science.1181498
2. Zhao YP, Fan G, Yin PP, et al. Resequencing 545 ginkgo genomes across the world reveals the evolutionary history of the living fossil. *Nat Commun.* 2019;10(1):4201. doi:10.1038/s41467-019-12133-5
3. Fan, G, Zhang, Y, Liu, X, et al. The first chromosome-level genome for a marine mammal as a resource to study ecology and evolution. *Mol Ecol Resour.* 2019;19:944 – 956. https://doi.org/10.1111/1755-0998.13003
4. Sjöstedt E, Zhong W, Fagerberg L, et al. An atlas of the protein-coding genes in the human, pig, and mouse brain. *Science.* 2020; 367(6482):eaay5947. doi:10.1126/science.aay5947
5. Xiong Y, Liu Y, Cao L, et al. Transcriptomic characteristics of bronchoalveolar lavage fluid and peripheral blood mononuclear cells in COVID-19 patients. *Emerg Microbes Infect.* 2020;9(1):761 – 770. doi:10.1080/22221751.2020.1747363
6. Wang R, Li H, Wu J, et al. Gut stem cell necroptosis by genome instability triggers bowel inflammation. *Nature.* 2020;580(7803):386 – 390. doi:10.1038/s41586-020-2127-x
7. Zhu F, Ju Y, Wang W, et al. Metagenome-wide association of gut microbiome features for schizophrenia. *Nat Commun.* 2020;11(1):1612. doi:10.1038/s41467-020-15457-9
8. Huang Y, Situ B, Huang L, et al. Nondestructive Identification of Rare Trophoblastic Cells by Endoplasmic Reticulum Staining for Noninvasive Prenatal Testing of Monogenic Diseases. *Adv Sci.* 2020;7(7):1903354. doi:10.1002/advs.201903354
9. Cai L, Bai H, Duan J, et al. Epigenetic alterations are associated with tumor mutation burden in non-small cell lung cancer. *J Immunother Cancer.* 2019;7(1):198. doi:10.1186/s40425-019-0660-7
10. Li Q, Zhao X, Zhang W, et al. Reliable multiplex sequencing with rare index mis-assignment on DNB-based NGS platform. *BMC Genomics.* 2019;20(1):215. doi:10.1186/s12864-019-5569-5
11. Appel M, Becker M, Berger M, et al. qPCR-based Quantification and QC Assays Provide Actionable Data for NGS Library Construction from FFPE Samples Of Variable Quality. Poster presented at AGBT 2014.
12. MGI Tech Co., Ltd. en.mgitech.cn/Uploads/Temp/picture/20191016/5da6dbb28f8ed.pdf. MGIEasy Circularization Kit User Manual.
13. MGI Tech Co., Ltd. en.mgitech.cn/Uploads/Temp/file/20191223/5e007f7f6da8c.pdf. DNBSEQ-G400RS High-throughput Sequencing Set User Manual.
14. Broad Institute. https://gatk.broadinstitute.org/hc/en-us/articles/360035890811. Truth variant dataset for NA12891: CEUTrio.HiSeq.WGS.b37.bestPractices.hg19.vcf.gz
15. Korostin D, Kulemin N, Naumov V, et al. (2020) Comparative analysis of novel MGISEQ-2000 sequencing platform vs Illumina HiSeq 2500 for whole-genome sequencing. PLOS ONE. 15(3): e0230301. https://doi.org/10.1371/journal.pone.0230301
16. Senabouth A, Andersen S, Shi Q, et al. Comparative performance of the BGI and Illumina sequencing technology for single-cell RNA-sequencing. *NAR Genomics and Bioinformatics.* 2020;2(2): lqaa034. https://doi.org/10.1093/nargab/lqaa034
17. Jeon SA, Park JL, Kim JH, et al. Comparison of the MGISEQ-2000 and Illumina HiSeq 4000 sequencing platforms for RNA sequencing. *Genomics Inform.* 2019;17(3):e32. doi:10.5808/GI.2019.17.3.e32
18. Appel M et al. (2018). KAPA HyperPrep Kits offer a flexible, high efficiency library preparation solution for PCR-free human whole-genome sequencing. Roche Application Note, MC--01908
19. Oligos and primers for BGISEQ&MGISEQ NGS system [Internet]. Available from: https://en.mgitech.cn/Download/download_file/id/71

Published by:

Roche Sequencing Solutions, Inc.
4300 Hacienda Drive
Pleasanton, CA 94588

sequencing.roche.com

Data on file.
For Research Use Only. Not for use in diagnostic procedures.
KAPA, KAPA HYPERPLUS, KAPA HYPERPURE and KAPA HYPERPREP are trademarks of Roche. All other product names and trademarks are the property of their respective owners.
© 2023 Roche Sequencing Solutions, Inc. All rights reserved. MC--05005 06/23