Technical note

HIGH-THROUGHPUT DNA SHEARING FOR LONG-READ MICROBIAL WGS

Introduction

High-throughput sequencing applications, such as microbial whole genome sequencing (WGS), require fast and cost-effective methods for shearing DNA to the size ranges suitable for PacBio® long-read sequencing. Plate-based shearing using the SPEX SamplePrep *1600 MiniG* homogenizer is a high-throughput and cost-effective alternative to the standard, lower-throughput options recommended for HiFi sequencing on PacBio long-read systems.[1]

This method is robust across a wide range of genomic DNA (gDNA) input amounts and is insensitive to spatial position on the plate. This provides users with greater flexibility in their batch sizes and experimental designs. For example, users can shear anywhere from 1 to 96 samples per *MiniG* run. Furthermore, changes to sample volume, time, and speed can be used to tune the mean size of the sheared DNA for microbial, or other shotgun WGS applications. For example, larger inserts can be used if sequencing 30-hour movies on Sequel II or IIe systems, or 24-hour movies on the Revio™ system.

Plate-based shearing is also compatible with other tissue homogenizers. As stroke length, speed, maximum volume, time intervals, and other parameters can vary between homogenizers, users will need to optimize run conditions. This technical note can serve as a starting point for adjusting conditions to suit individual needs.

1600 MiniG specifications

SPEX SamplePrep *1600 MiniG*	
Sheared fragment size	7–10 kb
Input volume	Up to 300 µL
Recommended input	300 ng–3 µg
Run time	3 minutes
Speed	1,500 rpm
Throughput	Up to 96 samples
Cost	<$1.00/sample

Required materials and equipment

Material	Part number
1600 MiniG	SPEX SamplePrep 1600
96 deep-well plate	Thermo Fisher Scientific 95040450
Sequel® sample plate foil	PacBio 100-667-400
Low TE buffer	PacBio 102-178-400
Centrifuge	Eppendorf *5804 R* or similar
ALPS 50 V manual heat sealer	Thermo Fisher Scientific AB-1443A

Protocol

1. For each sample, dilute 300 ng–3 µg of gDNA into a total volume of 300 µL with low TE buffer.
2. Transfer samples to a 96 deep-well plate and seal the plate with the Sequel sample plate foil for 2.5 seconds at 172°C using the *ALPS 50 V* manual heat sealer.
3. Open the lid to the SPEX SamplePrep *1600 MiniG* instrument, loosen the locking nut, and lift the clamps to place the sealed plate within the clamping mechanism. Secure the plate by lowering the clamps and tightening the locking nut. *Note: the plate is at risk of cracking if not properly secured.*
4. Shear the gDNA by running the *MiniG* instrument at 1,500 rpm for 3 minutes.
5. Centrifuge the sample plate for 1 minute at 150 rcf to collect liquid at the bottom of the plate.
6. Transfer the sheared gDNA to a new plate. Volume loss of up to 10% and foaming are typical.
7. Proceed to the post-shearing SMRTbell® cleanup step in Procedure & checklist — Preparing whole genome and metagenome libraries using SMRTbell prep kit 3.0.[1] Adjust the volume of SMRTbell cleanup beads as appropriate for a 1× (v/v) bead concentration.

PacBio

Conditions

Several conditions including the gDNA input mass, speed, volume, and time can be varied to achieve the desired target insert size. For microbial inserts (7–10 kb) on the SPEX *MiniG* homogenizer, we recommend starting with 300 ng–3 µg of gDNA in 300 µL of low TE buffer and shearing at 1,500 rpm for 3 minutes. For larger insert sizes or when using other plate-based homogenizers (e.g., SPEX *Geno/Grinder* or MP Biomedicals *FastPrep 96*), these conditions will need to be optimized.

DNA input

Plate-based shearing can accommodate a wide range of gDNA inputs. Figure 1 shows 4 bacterial species (*E. coli, S. enterica, L. monocytogenes,* and *B. cepacia*) sheared using the recommended protocol (300 µL sheared at 1,500 rpm for 3 minutes) with inputs ranging from 100 ng to 3 µg. All 4 bacteria consistently yielded sheared DNA within the target insert size of 7–10 kb, illustrating that for this insert size, plate-based shearing is relatively insensitive to input mass across the input range tested.

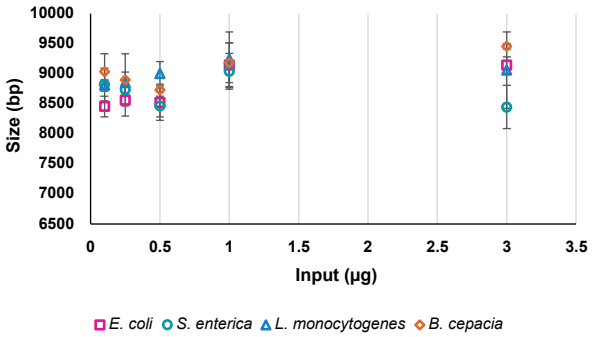

Figure 1. Size of sheared gDNA samples of 4 bacteria with increasing input. Each input mass was done in triplicate. Error bars show standard deviation.

Time

DNA size decreases with increasing shearing time. For microbial applications, the risk of over-shearing is low as the sheared size levels off to ~7 kb when sheared for greater than 3 minutes. We recommend 3 minutes to reduce processing time, but shearing times of 3–10 minutes will yield sheared DNA that is consistently sized 7–10 kb.

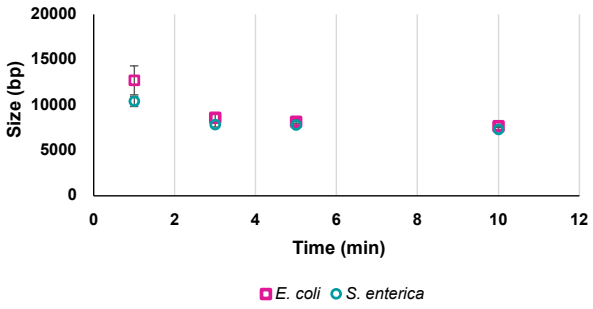

Figure 2. Size of sheared DNA samples of 2 bacteria with increasing time. Three micrograms of gDNA were diluted with low TE buffer to a final volume of 300 µL and sheared at 1,500 rpm for the indicated times. Each time point was performed in triplicate. Error bars show standard deviation. Sizes are modal sizes calculated from DNA size distributions on the *Femto Pulse* system (Agilent Technologies M5330AA).

Speed

Shear speed directly affects the size of the sheared gDNA and the shape of the distribution. Three micrograms of *E. coli* gDNA were sheared in 300 µL of low TE buffer for 5 minutes in the *MiniG* homogenizer with speeds ranging from 500–1,500 rpm. Increased shearing speed produces smaller inserts with narrower distributions. For microbial inserts (7–10 kb), we recommend 1,500 rpm max setting on the *MiniG* instrument.

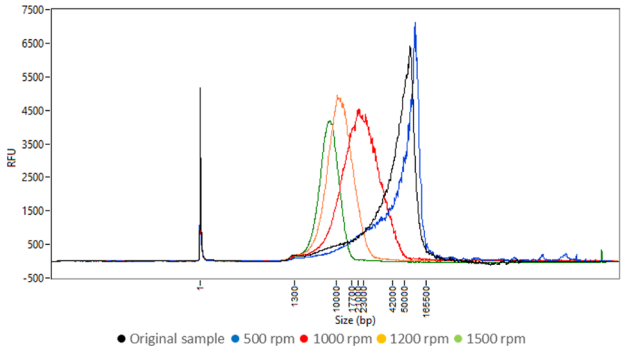

Figure 3. DNA size distributions of plate-sheared *E. coli* gDNA at different shear speeds. The modal length of the samples (as seen on the graph) with increasing speed starting from the original sample, is 80.6 kb, 105 kb, 21.0 kb, 10.5 kb, and 8.7 kb, respectively. This experiment was done in triplicate, but representative traces are used for this figure. DNA sizing was performed on the *Femto Pulse* system.

Volume

Increasing the volume of the sample leads to a decrease in the sheared size of the DNA. Increasing the total sample volume from 50–300 µL resulted in a decrease in the sheared DNA size from ~20 kb to ~8 kb for *E. coli* and *S. enterica*. For microbial applications, we recommend a volume of 300 µL to ensure a sheared size that is 7–10 kb, the recommend range for microbial whole genome sequencing.

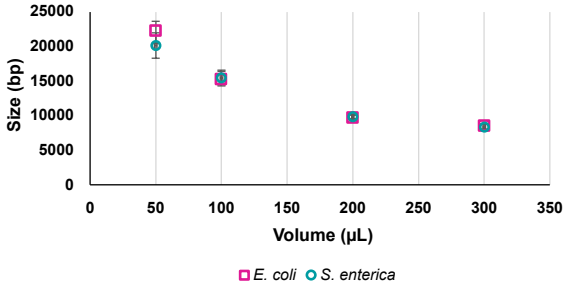

Figure 4. Size of sheared DNA samples of *E. coli* and *S. enterica* with increasing volume. Three micrograms of gDNA were diluted to the indicated volumes using low TE buffer and sheared at 1,500 rpm for 3 minutes. Each sample was done in triplicate. Errors bars show the standard deviation. Sizes are modal sizes calculated from DNA size distributions on the *Femto Pulse* system.

Spatial dependence

With plate-based shearing, the sheared DNA size is independent of the spatial position of the sample. Fifteen samples were randomly distributed across a 96-well plate and sheared. The samples sheared to an average modal size of 7.1 ± 0.3 kb with no discernable correlation between sheared size and plate position.

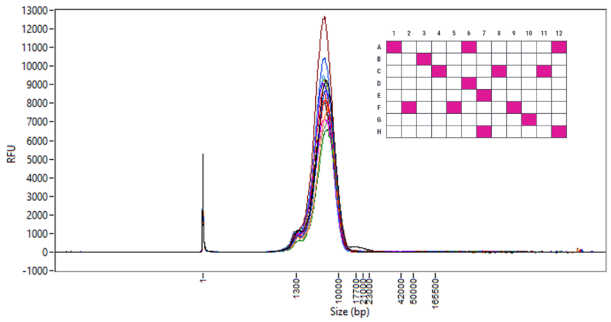

Figure 5. DNA size distributions of plate-sheared *E. coli* gDNA from 15 locations on a 96-well plate. Three micrograms of gDNA for each of the 15 samples were diluted with low TE buffer to a total volume of 300 µL and distributed into 15 randomly selected wells across the 96-well plate. The samples were sheared at 1,500 rpm for 3 minutes. DNA sizing was performed on the *Femto Pulse* system. The average modal length of all samples was 7.1 kb. Inset illustrates the positions of 15 randomly selected samples.

Comparison of plate-based shearing to shearing with the Diagenode *Megaruptor 3* system

E. coli gDNA was sheared with either the *Megaruptor 3* system or the plate-based shearing method to a target size of 7–10 kb. SMRTbell templates were made using the SMRTbell prep kit 3.0 (PacBio 102-182-700) and sized on a *Femto Pulse* system. The two shearing methods performed equivalently and yielded SMRTbell libraries with similar size distributions.

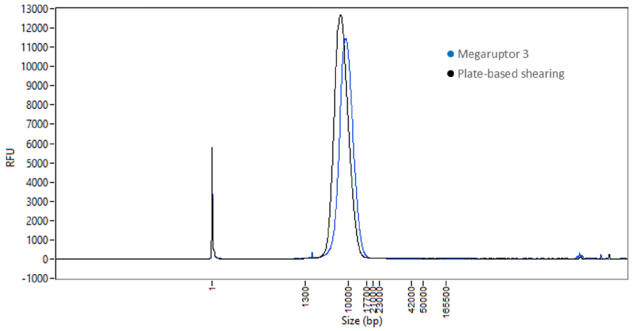

Figure 6. SMRTbell library size distributions for *E. coli* gDNA sheared on the *Megaruptor 3* system and with plate-based shearing. The *Megaruptor 3* sheared DNA (blue) and the plate-sheared DNA (black) yielded similar size distributions with modal sizes of 8.4 kb and 9.3 kb, respectively. DNA sizing was performed on the *Femto Pulse* system.

96-plex sequencing

DNA from 8 microbial species with 12 replicates each were extracted using the PacBio Nanobind® CBB kit (PacBio 102-301-900). All were sheared using the recommended plate shearing method on the SPEX *MiniG* homogenizer, barcoded with the SMRTbell barcoded adapter plate 3.0 (PacBio 102-009-200), and prepared for sequencing with the SMRTbell prep kit 3.0 (PacBio 102-182-700) and Sequel II binding kit 3.2 (PacBio 102-194-100). The pooled sample was sequenced on the Sequel IIe system in a 30-hour movie, demultiplexed on the instrument, and assembled using the *Microbial Genome Assembly* application in SMRT® Link using the default parameters with *Run Base Modification Analysis* and *Find Modified Base Motifs* turned off.

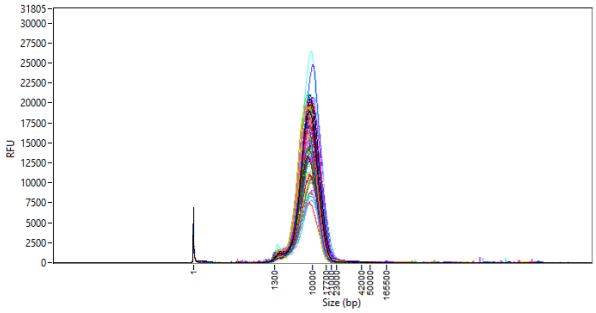

Figure 7. DNA size distributions of the individual SMRTbell libraries for the 96-plex sequencing run. The panel contains 8 bacterial species with 12 replicates. The average modal length of all samples was 9.4 kb. Irrespective of bacteria species, all samples sheared to roughly the same size. DNA sizing was performed on the *Femto Pulse* system.

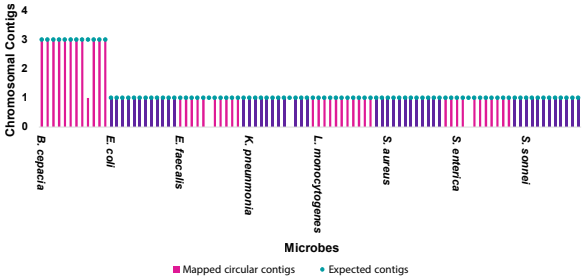

Figure 8. Mapped chromosomal contigs. Assembled circular chromosomal contigs produced by the *Microbial Genome Assembly* application using default parameters in SMRT Link without any manual curation. *B. cepacia* has 3 chromosomes while the other species in the panel all have 1. All samples aside from one replicate of *B. cepacia*, *E. faecalis*, *K. pneumonia*, and *S. enteria* assembled to the expected number of chromosomes.

Discussion

Plate-based shearing using the SPEX SamplePrep *1600 MiniG* homogenizer is a fast, inexpensive method to shear gDNA for highly accurate, complete microbial assemblies using PacBio HiFi reads. This technical note demonstrates (I) the speed, simplicity, and robustness of this method and (II) its flexibility of operating conditions in order to refine the mean size of sheared DNA depending on individual application and use.

KEY REFERENCES

1. Procedure & checklist — Preparing whole genome and metagenome libraries using SMRTbell prep kit 3.0 (102-166-600)

Questions? Visit **pacb.com/contact** or contact your local customer service representative

102-326-575 REV01 19MAY2023