

Journal of Visualized Experiments

Ultra-long Read Sequencing for Whole Genomic DNA Analysis

--Manuscript Draft--

Article Type:	Invited Methods Article - JoVE Produced Video
Manuscript Number:	JoVE58954R2
Full Title:	Ultra-long Read Sequencing for Whole Genomic DNA Analysis
Keywords:	Nanopore DNA sequencing; Ultra-long reads; Third-generation DNA sequencing; Whole genome sequence analysis; High Molecular Weight DNA extraction; Tagmentation based DNA library construction
Corresponding Author:	Chia-Lin Wei Jackson Laboratory Farmington, CT UNITED STATES
Corresponding Author's Institution:	Jackson Laboratory
Corresponding Author E-Mail:	Chia-lin.Wei@jax.org
Order of Authors:	Chia-Lin Wei Liang Gong Chee-Hong Wong Jennifer Idol Chew Yee Ngan
Additional Information:	
Question	Response
Please indicate whether this article will be Standard Access or Open Access.	Open Access (US\$4,200)
Please indicate the city, state/province, and country where this article will be filmed . Please do not use abbreviations.	Farmington, Connecticut, the United States

TITLE:

Ultra-Long Read Sequencing for Whole Genomic DNA Analysis

AUTHORS AND AFFILIATIONS:

Liang Gong, Chee-Hong Wong, Jennifer Idol, Chew Yee Ngan, Chia-Lin Wei

Genome Technologies, The Jackson Laboratory for Genomic Medicine, Farmington, CT, USA

Corresponding Author:

Chia-Lin Wei

chia-lin.wei@jax.org

Email Addresses of Co-authors:

liang.gong@jax.org

cheehong.wong@jax.org

jennifer.idol@jax.org

chewyee.ngan@jax.org

KEYWORDS:

Nanopore DNA sequencing, ultra-long reads, third-generation DNA sequencing, whole genome sequence analysis, high molecular weight DNA extraction, tagmentation based DNA library construction

SUMMARY:

Long-read sequences greatly facilitate the assembly of complex genomes and characterization of structural variation. We describe a method to generate ultra-long sequences by nanopore-based sequencing platforms. The approach adopts an optimized DNA extraction followed by modified library preparations to generate hundreds of kilobase reads with moderate coverage from human cells.

ABSTRACT:

Third generation single-molecule DNA sequencing technologies offer significantly longer read length that can facilitate the assembly of complex genomes and analysis of complex structural variants. Nanopore platforms perform single-molecule sequencing by directly measuring the current changes mediated by DNA passage through the pores and can generate hundreds of kilobase (kb) reads with minimal capital cost. This platform has been adopted by many researchers for a variety of applications. Achieving longer sequencing read lengths is the most critical factor to leverage the value of nanopore sequencing platforms. To generate ultra-long reads, special consideration is required to avoid DNA breakages and gain efficiency to generate productive sequencing templates. Here, we provide the detailed protocol of ultra-long DNA sequencing including high molecular weight (HMW) DNA extraction from fresh or frozen cells, library construction by mechanical shearing or transposase fragmentation, and sequencing on a nanopore device. From 20-25 µg of HMW DNA, the method can achieve N50 read length of 50-70 kb with mechanical shearing and N50 of 90-100 kb read length with transposase mediated

fragmentation. The protocol can be applied to DNA extracted from mammalian cells to perform whole genome sequencing for the detection of structural variants and genome assembly. Additional improvements on the DNA extraction and enzymatic reactions will further increase the read length and expand its utility.

INTRODUCTION:

Over the past decade, massively parallel and highly accurate second-generation high-throughput sequencing technologies have driven an explosion of biomedical discovery and technological innovation¹⁻³. Despite the technical advances, the short-read data generated by the second-generation platforms are ineffective in resolving complex genomic regions and are limited in the detection of genomic structural variants (SVs), which play important roles in human evolution and diseases^{4,5}. Furthermore, short-read data are unable to resolve repeat variation and are unsuitable for discerning haplotype phasing of genetic variants⁶.

Recent progress in single-molecule sequencing offers significantly longer read length, which can facilitate the detection of the full spectrum of SVs⁷⁻⁹, and offers accurate and complete assembly of complex microbial and mammalian genomes^{6,10}. The nanopore platform performs single-molecule sequencing by directly measuring the current changes mediated by DNA passage through the pores¹¹⁻¹³. Unlike any existing DNA sequencing chemistry, nanopore sequencing can generate long (tens to thousands of kilobases) reads in real-time without relying on polymerase kinetics or artificial amplification of the DNA sample. Therefore, nanopore long-read sequencing (NLR-seq) holds great promise for generating ultra-long read lengths well beyond 100 kb, which would greatly advance genomic and biomedical analyses¹⁴, particularly in the low-complexity or repeat-rich regions of the genomes¹⁵.

The unique feature of nanopore sequencing is its potential to generate long reads without a theoretical length limitation. Therefore, the read length is dependent on the physical length of the DNA which is directly affected by the DNA integrity and sequencing template quality. Moreover, depending on the extent of manipulation and the number of steps involved, such as pipetting forces and extraction conditions, the quality of the DNA is highly variable. Therefore, it is challenging for one to yield long reads by just applying the standard DNA extraction protocols and manufacturer's supplied library construction methods. Toward this end, we have developed a robust method to generate ultra-long read (hundreds of kilobases) sequencing data starting from harvested cell pellets. We adopted multiple improvements in the DNA extraction and library preparation procedures. We streamlined the protocol to exclude unnecessary procedures that cause DNA degradation and damages. This protocol is composed of high molecular weight (HMW) DNA extraction, ultra-long DNA library construction, and sequencing on a nanopore platform. For a well-trained molecular biologist, it typically takes 6 h from cell harvesting to the completion of HMW DNA extraction, 90 min or 8 h for library construction depending on the shearing method, and up to a further 48 h for DNA sequencing. The use of the protocol will empower the genomics community to improve our understanding of genome complexity and gain new insight into genome variation in human diseases.

PROTOCOL:

NOTE: The NLR-seq protocol consists of three consecutive steps: 1) extraction of high-molecular weight (HMW) genomic DNA; 2) ultra-long DNA library construction, which includes fragmentation of the HMW DNA into the desired sizes and ligation of sequencing adapters to the DNA ends; and 3) loading of the adapter-ligated DNA onto the arrays of nanopores (**Figure 1**).

1. HMW DNA extraction

1.1 Reagent setup. Make 1x phosphate buffered saline (PBS) buffer (1,000 mL) by adding 100 mL of PBS (10x) to 900 mL of water and mix well. Make lysis buffer (500 mL) by adding 43.5 mL of water to a 50 mL tube. Add 500 μ L of Tris (1 M, pH 8.0), 1 mL of sodium chloride (NaCl) (5 M), 2.5 mL of ethylenediaminetetraacetic acid (EDTA) (0.5 M, pH 8.0) and 2.5 mL of sodium dodecyl sulfate (SDS) (10%, wt/vol) to the tube and mix well.

NOTE: This PBS buffer can be stored at 4 °C for up to 6 months. The premade lysis buffer can be stored at RT for up to 2 months.

1.2 Check the cell mortality and count the cells. Ensure that the live ratio is > 85% and the total cell number is 30×10^6 .

NOTE: The cells used in this protocol are from the HG00733 cell line, a human lymphoblastoid cell line of Puerto Rican origin widely used in the 1000 Genome consortium for structural variation analysis (see table of materials for ordering information), which belongs to International Genome Sample Resource.

1.3 Collect the cells by centrifuging at $200 \times g$ for 5 min at RT. Discard the medium and resuspend the cell pellet (30×10^6 cells) with 5 mL of 1x PBS buffer. Centrifuge again at $200 \times g$ for 5 min at RT and discard the supernatant.

NOTE: $25\text{--}35 \times 10^6$ of cells are acceptable for this approach. Further variation in the amount of cells used will need further optimization. The cell pellet can be stored at -80°C for up to 6 months.

1.4 Resuspend the cell pellet in 200 μ L of 1x PBS buffer. If using a frozen cell pellet, wash with 5 mL of 1x PBS buffer. Centrifuge the solution at $200 \times g$ for 5 min at RT, discard the supernatant and resuspend the cells in 200 μ L of 1x PBS buffer.

1.5 Prepare 10 mL of lysis buffer in a 50 mL tube. Add the 200 μ L cell suspension to the lysis buffer and vortex at the highest speed for 3 s. Incubate the solution at 37°C for 1 h.

1.6 Add 2 μ L of RNase A (100 mg/mL) to the lysate. Gently rotate the 50 mL tube to mix the sample. Incubate the solution at 37°C for 1 h.

1.7 Add 50 μ L proteinase K (20 mg/mL) to the lysate. Gently rotate the 50 mL tube to mix the

sample. Incubate the solution at 50 °C for 2 h. During incubation, gently mix the sample every 30 min.

1.8 Remove the 50 mL tube from 50 °C and let stand at RT for 5 min.

1.9 Add 10 mL of the phenol layer of phenol:chloroform:isoamyl alcohol (25:24:1, vol/vol/vol) to the lysate and rotate the tube on a rotator mixer (see **Table of Materials**) at RT in a fume hood at 20 rpm for 10 min. Wrap the tube cap with parafilm to prevent leakage during rotation.

1.10 Prepare two 50 mL gel tubes (see **Table of Materials**) by centrifuging at 1,500 x *g* for 2 min at RT.

NOTE: The gel forms a stable barrier between the nucleic acid-containing aqueous phase and the organic solvent.

1.11 Pour the sample/phenol solution into one of the prepared 50 mL gel tubes from Step 1.10. Centrifuge the solution at 3,000 x *g* for 10 min at RT.

1.12 Pour the supernatant into a new 50 mL tube. Add 10 mL of the phenol layer of phenol:chloroform:isoamyl alcohol (25:24:1, vol/vol/vol) and rotate the tube on a rotator mixer at RT in a fume hood at 20 rpm for 10 min.

1.13 Repeat step 1.11 once with the second prepared gel tube.

1.14 Pour the supernatant into a new 50 mL tube. Add 25 mL of ice-cold 100% ethanol and gently rotate the tube by hand until the DNA precipitates (**Figure 2**).

NOTE: The precipitation approach helps to stabilize the HMW DNA.

1.15 Bend a 20 µL tip to make a hook. Carefully take out the HMW DNA with the hook and let the liquid drop off.

1.16 Place the HMW DNA into a 50 mL tube containing 40 mL of 70% ethanol. Wash the DNA by gently inverting the tube 3 times.

1.17 Repeat step 1.15 once to collect DNA from the 70% ethanol tube.

1.18 Place the HMW DNA into a 2 mL tube containing 1.8 mL of 70% ethanol.

1.19 Centrifuge the washed HMW DNA at 10,000 x *g* for 3 s at RT. Remove as much of the residual ethanol as possible by pipetting.

NOTE: Do not disturb the DNA pellet when pipetting the residual ethanol.

1.20 Incubate the 2 mL tube at 37 °C for 10 min with the lid open to dry the sample.

1.21.1 If continuing with step 2.1 (with mechanical shearing and 1D Ligation Sequencing Kit), add 1 mL of TE (10 mM Tris and 1 mM EDTA, pH 8.0) to the 2 mL tube.

1.21.2 If continuing with Step 2.2 (with transposase-based fragmentation and Rapid Sequencing Kit), add 200 µL of 10 mM Tris (pH 8.0) with 0.02% Triton X-100.

NOTE: Do not disturb the DNA pellet. Letting the tube stand at 4 °C in dark for 48 h will help the sample fully resuspend. The HMW DNA can be stored at 4 °C for up to 2 weeks. Longer storage time or other storage conditions may introduce more short fragments.

2. Ultra-long DNA library construction

NOTE: There are two ways to construct the ultra-long DNA libraries based on two different shearing methods coupled with nanopore sequencing kits. A mechanical shearing-based library produces data with an N50 of 50-70 kb, taking about 8 h for the library construction. A transposase fragmentation-based library produces an N50 of 90-100 kb data, taking only 90 min for the library construction. The mechanical shearing protocol gives higher yield from the same DNA input using identical versions of sequencing adapter and quality of nanopore flow cells.

2.1 Mechanical shearing-based library construction

2.1.1 Thaw and mix the reagents from the ligation kit (see **Table of Materials**). Thaw FFPE DNA repair buffer and end repair/dA-tailing buffer on ice, then vortex and spin down to mix. Thaw adapter mix (AMX) and adapter bead binding buffer (ABB) on ice, then pipette and spin down to mix. Thaw running buffer with fuel mix (RBF) and elution buffer (ELB) at RT, then vortex and spin down to mix. Thaw library loading beads (LLB) at RT and pipette to mix before use.

2.1.1.1. Once thawed, keep all kit components on ice. Take out the enzymes only when needed. Bring the magnetic beads to RT for use.

NOTE: For recommendations on the magnetic beads to use see the table of materials.

2.1.2 Check the quality and quantity of the HMW DNA from step 1.21.1. Pipette out 20 µL of DNA into new 1.5 mL tubes from three different locations in the HMW DNA tube using P200 wide bore tips. Take 1 µL from the three aliquots to detect the concentration using a fluorometer and the quality using a UV reading. Check multiple times to confirm the results.

NOTE: The expected results are shown in **Figure 3A**. The OD_{260/280} value is approximately 1.9 and the OD_{260/230} value is approximately 2.3.

2.1.3 Transfer the remaining 940 µL of HMW DNA into a 50 mL tube cap with a P1000 wide bore tip.

221
222 2.1.4 Aspirate all DNA into a 1 mL syringe without the needle.

223
224 2.1.5 Put the 27 G needle onto the syringe and eject all DNA into the cap gently and slowly (~10
225 s). Take off the 27 G needle from the syringe.

226
227 2.1.6 Repeat steps 2.1.4 and 2.1.5 for 29 times for a total of 30 passes through the needle.

228
229 NOTE: The sheared HMW DNA can be stored at 4 °C in the dark for up to 24 h. Quality control
230 (QC) is highly recommended by pulsed-field gel electrophoresis, but it is costly and time
231 consuming. If performing QC on an automated pulse field gel electrophoresis machine use a 5-
232 150 kb protocol for a 20 h run. The expected results are shown in **Figure 4**.

233
234 2.1.7 Prepare the DNA repair reaction in a 0.2 mL tube by adding 100 µL of sheared HMW DNA
235 (20 µg), 15 µL of FFPE DNA repair buffer, 12 µL of FFPE DNA repair mix, and 16 µL of nuclease-
236 free water. Mix the reaction by flicking gently 6 times and spin down to remove bubbles.

237
238 2.1.8 Incubate the reaction at 20 °C for 60 min. Transfer the sample into a new 1.5 mL tube with
239 a P200 wide bore tip.

240
241 2.1.9 Resuspend the magnetic beads by pipetting or vortexing. Add 143 µL beads (1x) to the DNA
242 repair reaction and mix gently by flicking the tube 6 times. Rotate the tube on a rotator mixer at
243 RT at 20 rpm for 30 min.

244
245 2.1.10 Spin down the sample at 1,000 x *g* for 2 s at RT. Place the tube on a magnetic rack for 10
246 min. Keep the tube on the magnetic rack and discard the supernatant.

247
248 2.1.11 Keeping the tube on the magnetic rack, add 400 µL of freshly prepared 70% ethanol
249 without disturbing the pellet. Remove the 70% ethanol after 30 s.

250
251 2.1.12 Repeat step 2.1.11 once.

252
253 2.1.13 Spin down the sample at 1,000 x *g* for 2 s at RT. Place the tube back on the magnetic rack.
254 Remove any residual ethanol and air dry for 30 s. Do not over dry the pellet.

255
256 2.1.14 Remove the tube from the magnetic rack and add 103 µL of TE (10 mM Tris and 1 mM
257 EDTA, pH 8.0). Gently flick the tube to ensure that beads are covered in the buffer, and incubate
258 on a rotator mixer at RT for 30 min. Gently flick the tube every 5 min to aid resuspension of the
259 pellet.

260
261 2.1.15 Pellet the beads on the magnetic rack for at least 10 min. Transfer 100 µL of eluate with a
262 P200 wide bore tip into a 0.2 mL tube.

263
264 2.1.16 Prepare the end repair and dA-tailing reaction in a 0.2 mL tube by adding 100 µL of

repaired HMW DNA, 14 μ L of end repair/dA-tailing buffer and 7 μ L of end repair/dA-tailing mix. Mix the reaction by flicking gently 6 times, and spin down to remove bubbles.

2.1.17 Incubate the reaction at 20 °C for 60 min followed by 65 °C for 20 min, and then hold at 22°C. Transfer the sample into a new 1.5 mL tube using a P200 wide bore tip.

2.1.18 Resuspend the magnetic beads by pipetting or vortexing. Add 48 μ L of beads (0.4x) to the end repair/dA-tailing reaction and mix gently by flicking the tube 6 times. Rotate the tube on a rotator mixer at RT at 20 rpm for 30 min.

2.1.19 Repeat steps 2.1.10-2.1.13 once.

2.1.20 Remove the tube from the magnetic rack and add 33 μ L of TE (10 mM Tris and 1 mM EDTA, pH 8.0). Gently flick the tube to ensure that beads are covered in the buffer, and incubate on a rotator mixer at RT for 30 min. Gently flick the tube every 5 min to aid resuspension of the pellet.

2.1.21 Pellet the beads on the magnetic rack for at least 10 min. Transfer 30 μ L of eluate with a P200 wide bore tip into a new 1.5 mL tube. Take the extra 1-2 μ L to detect the concentration using a fluorometer.

NOTE: Recovery of 5-6 μ g at this step is expected.

2.1.22 Prepare the ligation reaction in the 1.5 mL sample tube by adding 30 μ L of end-repaired HMW DNA, 20 μ L of adapter mix (AMX 1D), and 50 μ L of blunt/TA ligation master mix. Mix the reaction by flicking gently 6 times between each sequential addition and spin down to remove bubbles.

2.1.23 Incubate the reaction at RT for 60 min.

2.1.24 Resuspend the magnetic beads by pipetting or vortexing. Add 40 μ L beads (0.4x) to the ligation reaction and mix gently by flicking the tube 6 times. Rotate the tube on a rotator mixer at RT at 20 rpm for 30 min.

2.1.25 Repeat step 2.1.10 once.

2.1.26 Add 400 μ L of adapter bead binding (ABB) buffer into the tube. Flick the tube gently 6 times to resuspend the beads. Place the tube back on the magnetic rack to separate the beads from the buffer and discard supernatant.

2.1.27 Repeat step 2.1.26 once.

2.1.28 Spin down the sample at 1,000 x g for 2 s at RT. Place the tube back on the magnetic rack. Remove any residual buffer and air dry for 30 s. Do not over dry the pellet.

2.1.29 Remove the tube from the magnetic rack and resuspend the pellet in 43 μ L of elution buffer. Gently flick the tube to ensure that beads are covered in the buffer and incubate on a rotator mixer at RT for 30 min. Gently flick the tube every 5 min to aid resuspension of the pellet.

2.1.30 Pellet the beads on the magnetic rack for at least 10 min. Transfer 40 μ L of eluate with a P200 wide bore tip into a new 1.5 mL tube. Take the extra 1-2 μ L to detect the concentration using a fluorometer.

NOTE: Recovery of 1-2 μ g at this step is expected. The mechanical shearing-based library is ready for loading. The library can be stored on ice for up to 2 h until loading for sequencing if needed.

2.2 Transposase fragmentation-based library construction

2.2.1 Thaw the reagents from the transposase kit (see **Table of Materials**). Thaw fragmentation mix (FRA) and rapid adapter (RAP) on ice and pipette to mix. Thaw sequencing buffer (SQB), loading beads (LB), flush buffer (FLB) and flush tether (FLT) at RT and pipette to mix. Thaw loading beads (LB) at RT and pipette to mix before use. Once thawed, keep all kit components on ice. Take out the enzymes only when needed.

2.2.2 Check the quality and quantity of the HMW DNA from step 1.21.2. Pipette out 20 μ L of DNA into new 1.5 mL tubes from three different locations in the HMW DNA tube using P200 wide bore tips. Take 1 μ L from the three aliquots to detect the concentration using a fluorometer and the quality using a UV reading. Check multiple times to confirm the results.

NOTE: The expected results are shown in **Figure 3B**. The OD_{260/280} value is approximately 1.9 and the OD_{260/230} value is approximately 2.3.

2.2.3 Prepare the DNA fragmentation reaction in a 0.2 mL tube by adding 22 μ L of HMW DNA, 1 μ L of 10 mM Tris (pH 8.0) with 0.02% Triton X-100 and 1 μ L of fragmentation mix (FRA). Mix by pipetting with a P200 wide bore tip as slowly as possible 6 times, taking care not to introduce bubbles.

2.2.4 Incubate the reaction at 30 °C for 1 min followed by 80 °C for 1 min, and then hold at 4 °C. Transfer the mix into a new 1.5 mL tube with a P200 wide bore tip and go to next step immediately.

2.2.5 Add 1 μ L of rapid adapter (RAP) to the 1.5 mL sample tube. Mix by pipetting with a P200 wide bore tip as slowly as possible 6 times, taking care not to introduce bubbles.

2.2.6 Incubate the reaction at RT for 60 min.

NOTE: The transposase fragmentation-based library is ready for loading. The library can be stored on ice for up to 2 h until loading for sequencing if needed.

3. Sequencing on the nanopore device

3.1 Check the nanopore sequencing device (see **Table of Materials**). Make sure both the software and hardware are working and there is enough storage space.

3.2 Check the flow cell. Open a new flow cell and insert the flow cell into the nanopore device. Check the box of the location the flow cell was inserted into (X1-X5). Select the correct flow cell type. Click on the **Check Flow Cells** workflow. Click on the **Start Test** button to start the flow cell QC analysis.

NOTE: If the reported total active pore number is less than 800, use a different new flow cell for sequencing.

3.3 Prepare the priming buffer. For a mechanical shearing-based library, add 576 μL of running buffer with fuel mix (RBF) and 624 μL of nuclease-free water into a 1.5 mL tube. Vortex and spin down to mix the priming buffer. For a transposase fragmentation-based library, add 30 μL of flush tether (FLT) to the tube of flush buffer (FLB). Vortex and spin down to mix the priming buffer.

3.4 On the flow cell, move the priming port cover clockwise to expose the priming port.

3.5 Set a P1000 pipette to 100 μL and insert the tip into the priming port. Draw back a small volume of buffer (less than 30 μL) to remove any bubbles from the flow cell. Stop pipetting once a small amount of yellow fluid enters the tip.

3.6 Use a P1000 pipette to load 800 μL of the priming mix into the flow cell via the priming port. To avoid introducing bubbles, add 30 μL of the priming mix to cover the top of the priming port first, then insert the tip into the priming port and slowly add the rest of the priming mix. Take out the tip when there is about 50 μL left. Add the rest of priming mix on the top of the priming port. The fluid will go inside by itself.

3.7 Leave the setup to incubate for 5 min. In the meantime, prepare the library mix in the 1.5 mL tube containing the library.

NOTE: For a mechanical shearing-based library add 35 μL of running buffer with fuel mix (RBF) to 40 μL of the DNA library. For a transposase fragmentation-based library add 34 μL of sequencing buffer (SQB) and 16 μL of nuclease-free water to 25 μL of the DNA library.

3.8 Open the flow cell sample port cover gently to expose the sample port. Use a P1000 pipette to add 200 μL of the priming mix through the priming port into the flow cell as described in step 3.5. Make sure that the priming mix is not loaded into the flow cell through the sample port.

3.9 Set a P200 pipette to 80 μL . Mix the library gently with a wide bore tip by pipetting up and down 6 times just prior to loading.

3.10 Load the library mix dropwise through the sample port into the flow cell. Add each drop only after the previous drop is completely loaded into the port.

3.11 Put back the sample port cover gently and make sure the sample port is fully covered. Move the priming port cover anticlockwise to cover the priming port. Close the device lid.

3.12 Click on the **New Experiment** workflow. Type the library name, select the correct kit according to procedures used, and check that the settings are correct (48 h run, real-time base-calling ON).

3.13 Click **Start Run**. After 10 min, record the flow cell ID and the active nanopore numbers (total number and each four groups' numbers) from the run information.

3.14 Data analysis. Copy the data to a local computer or a cluster at any time of the sequencing or when the run is complete. Use Minimap2¹⁶ (<https://github.com/lh3/minimap2>) to align the sequence data to the reference genome. Summarize the sequencing performance from the raw sequence data and the alignments by NanoPlot¹⁷ (<https://github.com/wdecoster/NanoPlot>).

REPRESENTATIVE RESULTS:

The ultra-long DNA sequencing protocol applies HMW DNA for library construction. Therefore, it is critical to choose well-cultured cells with the live ratio >85% at the cell harvesting step. The amount of cells used for DNA extraction will affect the quality and the quantity of the HMW DNA. The cell lysis does not work well if starting with too many cells. Using too few cells does not generate enough DNA for library construction because the HMW DNA precipitation is performed using gentle rotation by hand instead of high-speed centrifugation. An example of the HMW DNA after adding ice-cold 100% ethanol and rotating is shown as the white cotton-like precipitate in **Figure 2**.

It is important to check the quality of the input DNA before beginning the library construction. Degradation, incorrect quantification, contamination (e.g., proteins, RNAs, detergents, surfactant, and residual phenol or ethanol) and low molecular weight DNA can have a significant effect on the subsequent procedures and on the final read length. We recommend performing the QC analysis using the DNA from three different locations in the tube containing HMW DNA. From UV reading results for the HMW DNA, the OD₂₆₀/OD₂₈₀ value is approximately 1.9 and the OD₂₆₀/OD₂₃₀ value is approximately 2.3 (**Figure 3A,B**). These ratio values are consistent among the three tests for a good HMW DNA sample. Different shearing methods requires different volumes of input DNA. The concentration of HMW DNA needs to be >200 ng/μL for mechanical shearing while it needs to be >1 μg/μL for transposase fragmentation. The concentration detected by a fluorometer is a little lower than UV reading. However, the coefficient of variation of the concentration of the same HMW DNA sample is required to be less than 15% with both the fluorometer and the UV reading assays. Mechanical shearing applies a syringe with a needle to break the HMW DNA so that the number of passes through the needle will impact the size of the sheared DNA and the final read length. It is recommended to perform size QC after needle shearing to ensure the majority of the HMW DNA is larger than 50 kb as illustrated in **Figure 4**. In

the mechanical shearing method, 30 passes generated the best sequencing results considering both the length and output.

The N50 of a mechanical shearing-based library is 50-70 kb while a transposase fragmentation-based library is 90-100 kb. The results of four runs using the HG00733 cell line are shown in **Table 1**. All four runs have over 2,300 reads with length longer than 100 kb. The maximum length is longer in the transposase fragmentation-based libraries (455 kb and 489 kb) compared with the mechanical shearing-based libraries (348 kb and 387 kb) while the latter produced more total reads, indicating a higher yield. The transposase fragmentation-based library construction has fewer steps and shorter preparation time so that it will introduce fewer short fragments. The two runs using transposase have a longer mean length (>30 kb) and median length (>10 kb). In addition, the data shows consistent high quality in all runs (mean quality score is approximately 10.0, ~90% base accuracy). More than 97% of the total bases were aligned to the human reference genome (hg19) using Minimap2¹⁶ with the default settings. The expected size distributions of the raw reads are shown in **Figure 5**. All runs have a large proportion of data above 50 kb while transposase fragmentation-based libraries have a higher ratio of ultra-long reads (e.g. > 100 kb). This protocol has been successfully applied in multiple human cell lines (Supplementary Table 1).

FIGURE AND TABLE LEGENDS:

Figure 1: Schematic overview of the nanopore long-read sequencing (NLR-seq) workflow. Orange, the transposase complex. Yellow-green, the nanopore adapter.

Figure 2: Representative DNA precipitation from phenol-chloroform extraction method. The white arrow indicates the HMW DNA.

Figure 3: Example QC results of the HMW DNA from UV reading. (A) HMW DNA from step 1.21.1 ready for mechanical shearing-based library construction. (B) HMW DNA from step 1.21.2 for transposase fragmentation-based library construction.

Figure 4: QC results of the needle sheared HMW DNA by pulsed-field gel electrophoresis. L1: Quick-load 1 kb DNA ladder; L2: Quick-load 1 kb extend DNA ladder. 1-8: DNA with different passing times through the needle shearing. 1-3, no shearing; 4, 10 times; 5, 20 times; 6, 30 times; 7, 40 times; 8, 50 times. This QC step is optional.

Figure 5: Expected size distributions of the nanopore ultra-long DNA libraries. MS, mechanical shearing-based libraries. TF, transposase fragmentation-based libraries.

Table 1: Performance metrics summary from runs with different shearing protocols.

Supplementary Table 1: Summary of two NLR-seq runs using other cell lines with the mechanical shearing protocol.

DISCUSSION:

In principle, nanopore sequencing is able to generate 100 kb to megabase reads in length¹¹⁻¹³. Four major factors will affect the performance of the sequencing run and data quality: 1) active pore numbers and the activity of the pores; 2) motor protein, which controls the speed of DNA passing through the nanopore; 3) DNA template (length, purity, quality, mass); 4) sequencing adapter ligation efficiency, which determines the useable DNA from the input sample. The first two factors depend on the version of the flow cell and the sequencing kit provided by the manufacturer. The second two factors are critical steps in this protocol (HMW DNA extraction, shearing and ligation).

This protocol requires patience and practice. The quality of HMW DNA is important for ultra-long DNA libraries⁶. The protocol starts with cells collected with high viability (>85% viable cell preferred), limiting the degraded DNA from dead cells. Any harsh process which may introduce damages to the DNA (e.g., strong disturbing, shaking, vortex, multiple pipetting, repeated freezing and thawing) should be avoided. In the design of the protocol, we omit pipetting in the entire process of DNA extraction. Wide bore tips need to be used when pipetting is necessary after the mechanical shearing during library construction and sequencing. As the nanopores are sensitive to the chemistries in the chamber buffer¹², there should be as few residual contaminants (e.g., the detergents, surfactants, phenol, ethanol, proteins RNAs, etc.) as possible in the DNA. Considering the length and yield, the phenol extraction method shows the best and most reproducible results compared with multiple different extraction methods tested so far.

Despite the ability of this protocol to produce long-read sequences, several limitations still remain. First, this protocol was optimized based on the nanopore sequencing device available at time of publication; thus, it is limited to the selective nanopore-based sequencing chemistry and could be suboptimal when performed in other types of long-read sequencing devices. Second, the outcome is highly dependent on the quality of the DNA extracted from the starting material (tissues or cells). Read length will be compromised if the starting DNA is already degraded or damaged. Third, although multiple QC steps are incorporated in the protocol to check the DNA quality, the final yield and length of the reads can be affected by the flow cell and pore activity, which could be variable at this early stage of nanopore sequencing platform development.

The protocol described here uses human suspension cell line samples for DNA extraction. We have optimized the passing times in needle shearing, the ratio of HMW DNA to transposase and the ligation time to produce the described results. The protocol can be expanded in four ways. First, users can start with other cultured mammalian cells and with different amount of cells, tissues, clinical samples, or other organisms. Further optimization on lysis incubation time, reaction volume and centrifugation will be needed. Second, it is hard to predict the target size for ultra-long read sequencing. If the read lengths are shorter than expected, the users can adjust the passing times in the mechanical shearing-based method or change the ratio of the HMW DNA to transposase in the transposase fragmentation-based method. Longer binding and elution time during cleanup steps are helpful because the HMW DNA is highly viscous. Third, with different nanopore sequencing devices, one can adjust the amount and volume of the DNA to meet the criteria of the sequencer. Fourth, only those DNA ligated to sequencing adapters will be sequenced. To further improve ligation efficiency, one can attempt to titrate the adapter and

ligase concentrations. Modified ligation time and molecular crowding agents such as PEG¹⁸ can be applied in future. The ultra-long DNA sequencing protocol combined with CRISPR^{19,20} may offer an effective tool for target enrichment sequencing.

ACKNOWLEDGMENTS:

The authors thank Y. Zhu for her comments on the manuscript. Research reported in this publication was partially supported by the National Cancer Institute of the National Institutes of Health under Award Number P30CA034196. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

DISCLOSURES:

The authors declare that they have no competing financial interests.

REFERENCES:

- 1 Mardis, E. R. Next-generation sequencing platforms. *Annual Review of Analytical Chemistry*. **6** 287-303 (2013).
- 2 Goodwin, S., McPherson, J. D., McCombie, W. R. Coming of age: ten years of next-generation sequencing technologies. *Nature Reviews Genetics*. **17** (6), 333-351 (2016).
- 3 Shendure, J. et al. DNA sequencing at 40: past, present and future. *Nature*. **550** (7676), 345-353 (2017).
- 4 Alkan, C., Coe, B. P., Eichler, E. E. Genome structural variation discovery and genotyping. *Nature Reviews Genetics*. **12** (5), 363-376 (2011).
- 5 Weischenfeldt, J., Symmons, O., Spitz, F., Korbel, J. O. Phenotypic impact of genomic structural variation: insights from and for human disease. *Nature Reviews Genetics*. **14** (2), 125-138 (2013).
- 6 Pollard, M. O., Gurdasani, D., Mentzer, A. J., Porter, T., Sandhu, M. S. Long reads: their purpose and place. *Human Molecular Genetics*. **27** (R2), R234-r241 (2018).
- 7 Cretu Stancu, M. et al. Mapping and phasing of structural variation in patient genomes using nanopore sequencing. *Nature Communications*. **8** (1), 1326 (2017).
- 8 Gong, L. et al. Picky comprehensively detects high-resolution structural variants in nanopore long reads. *Nature Methods*. **15** (6), 455-460 (2018).
- 9 Sedlazeck, F. J. et al. Accurate detection of complex structural variations using single-molecule sequencing. *Nature Methods*. **15** (6), 461-468 (2018).
- 10 Jain, M. et al. Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nature Biotechnology*. **36** (4), 338-345 (2018).
- 11 Jain, M. et al. Improved data analysis for the MinION nanopore sequencer. *Nature Methods*. **12** (4), 351-356 (2015).
- 12 Deamer, D., Akeson, M., Branton, D. Three decades of nanopore sequencing. *Nature Biotechnology*. **34** (5), 518-524 (2016).
- 13 Jain, M., Olsen, H. E., Paten, B., Akeson, M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biology*. **17** (1), 239 (2016).
- 14 Editorial. The long view on sequencing. *Nature Biotechnology*. **36** (4), 287 (2018).
- 15 Jain, M. et al. Linear assembly of a human centromere on the Y chromosome. *Nature Biotechnology*. **36** (4), 321-323 (2018).

573 16 Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. **34**, 3094-
574 3100 (2018).

575 17 De Coster, W., D'Hert, S., Schultz, D.T., Cruts, M., Van Broeckhoven, C. NanoPack:
576 visualizing and processing long-read sequencing data. *Bioinformatics* . **34**, 2666-2669
577 (2018).

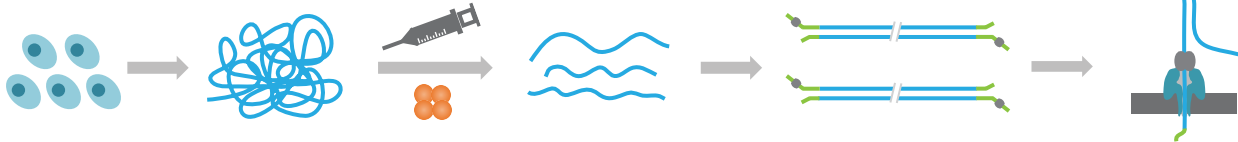
578 18 Akabayov, B., Akabayov, S. R., Lee, S. J., Wagner, G., Richardson, C. C. Impact of
579 macromolecular crowding on DNA replication. *Nature Communications*. **4**, 1615 (2013).

580 19 Gabrieli, T., Sharim, H., Michaeli, Y., Ebenstein, Y. Cas9-Assisted Targeting of
581 CHromosome segments (CATCH) for targeted nanopore sequencing and optical genome
582 mapping. *bioRxiv*. [10.1101/110163](https://doi.org/10.1101/110163) (2017).

583 20 Gabrieli, T. et al. Selective nanopore sequencing of human BRCA1 by Cas9-assisted
584 targeting of chromosome segments (CATCH). *Nucleic Acids Research*.
585 10.1093/nar/gky411 (2018).
586

Figure 1

[Click here to access/download;Figure;Figure 1.ai](#) 



HMW DNA Extraction

DNA Fragmentation

Adapter Ligation

Sequencing

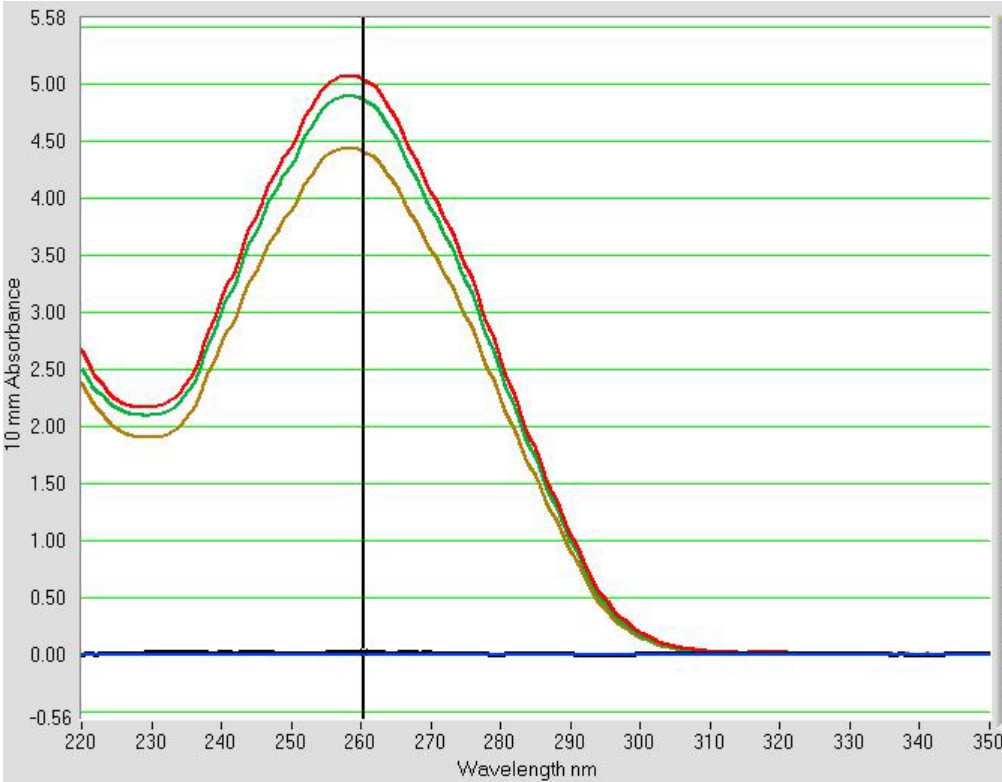
Figure 2

Click here to
access/download:Figur



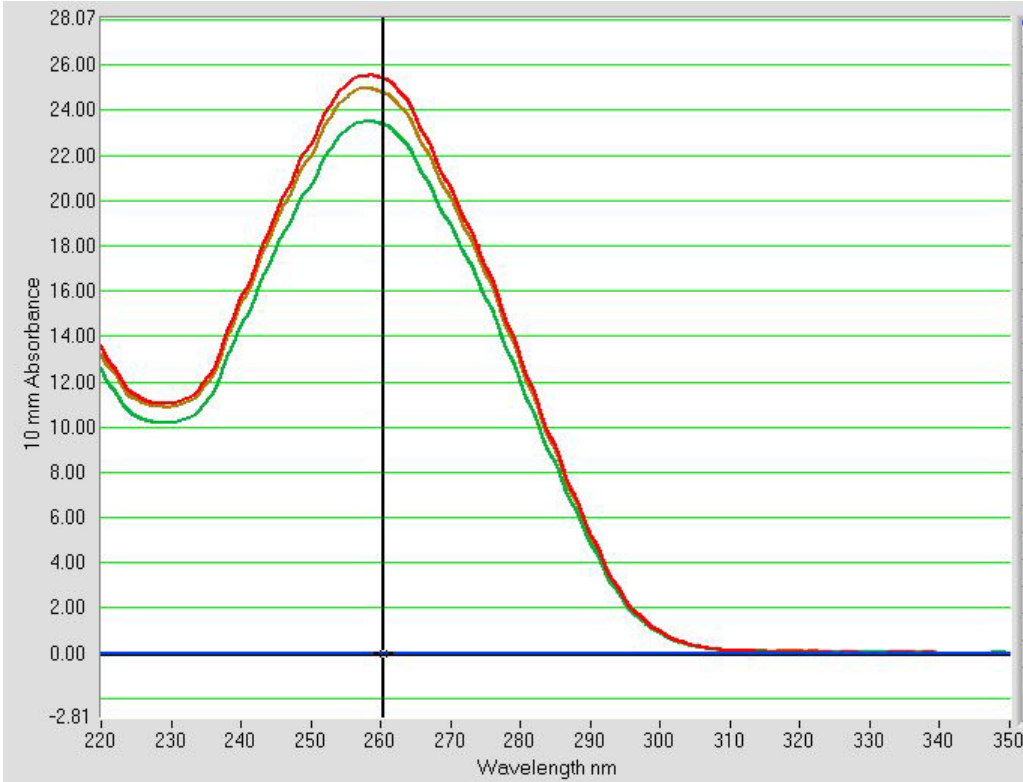
Figure 3

A



Sample	Con. (ng/ μ L)	260/280	260/230
DNA_1D_1	252.48	1.94	2.33
DNA_1D_2	243.69	1.96	2.32
DNA_1D_3	220.86	1.96	2.32

B



Sample	Con. (ng/ μ L)	260/280	260/230
DNA_RAD_1	1271.62	1.94	2.31
DNA_RAD_2	1170.02	1.94	2.30
DNA_RAD_3	1239.83	1.92	2.28

Figure 4

[Click here to access/download;Figure;Figure 4.ai](#)

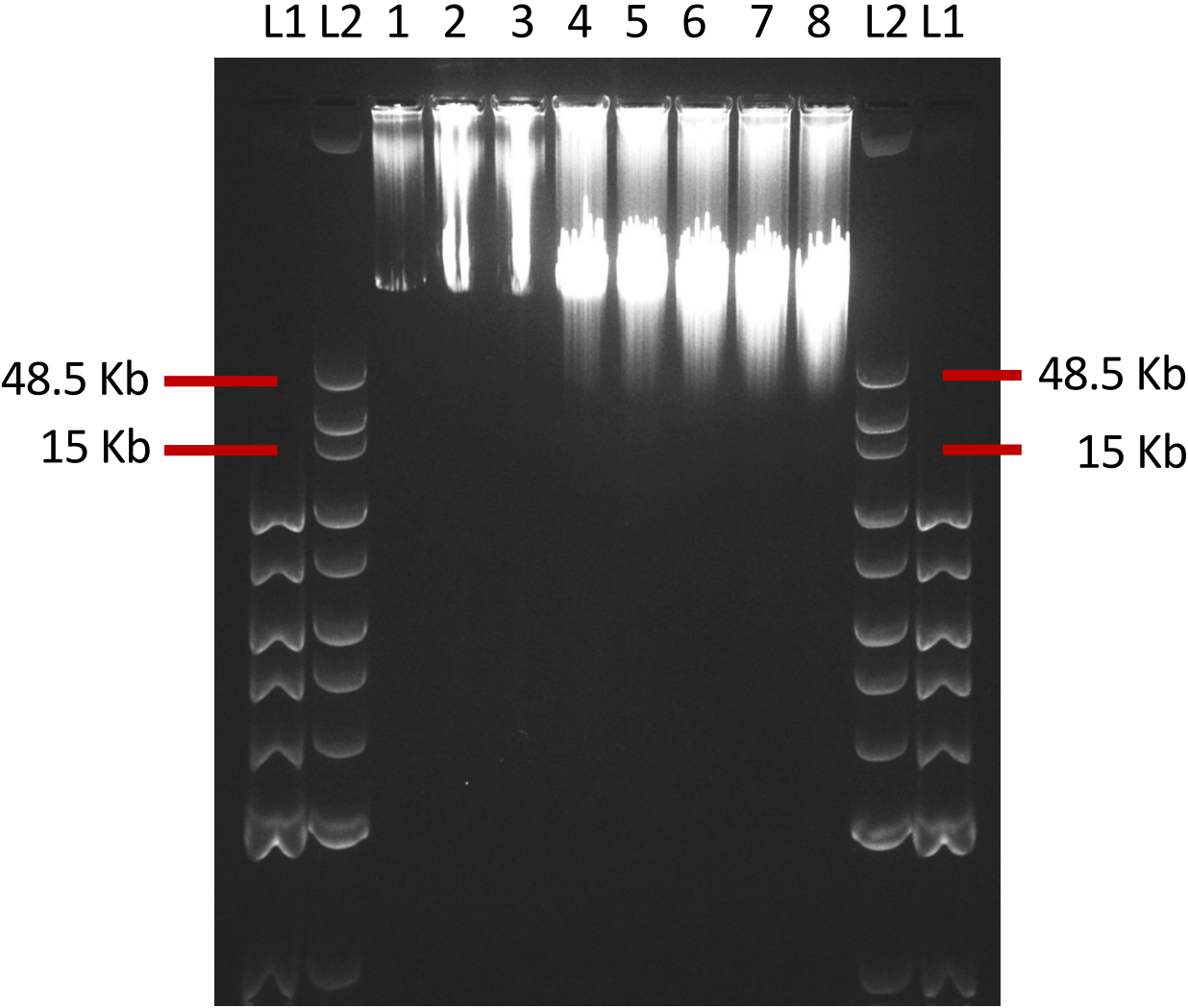
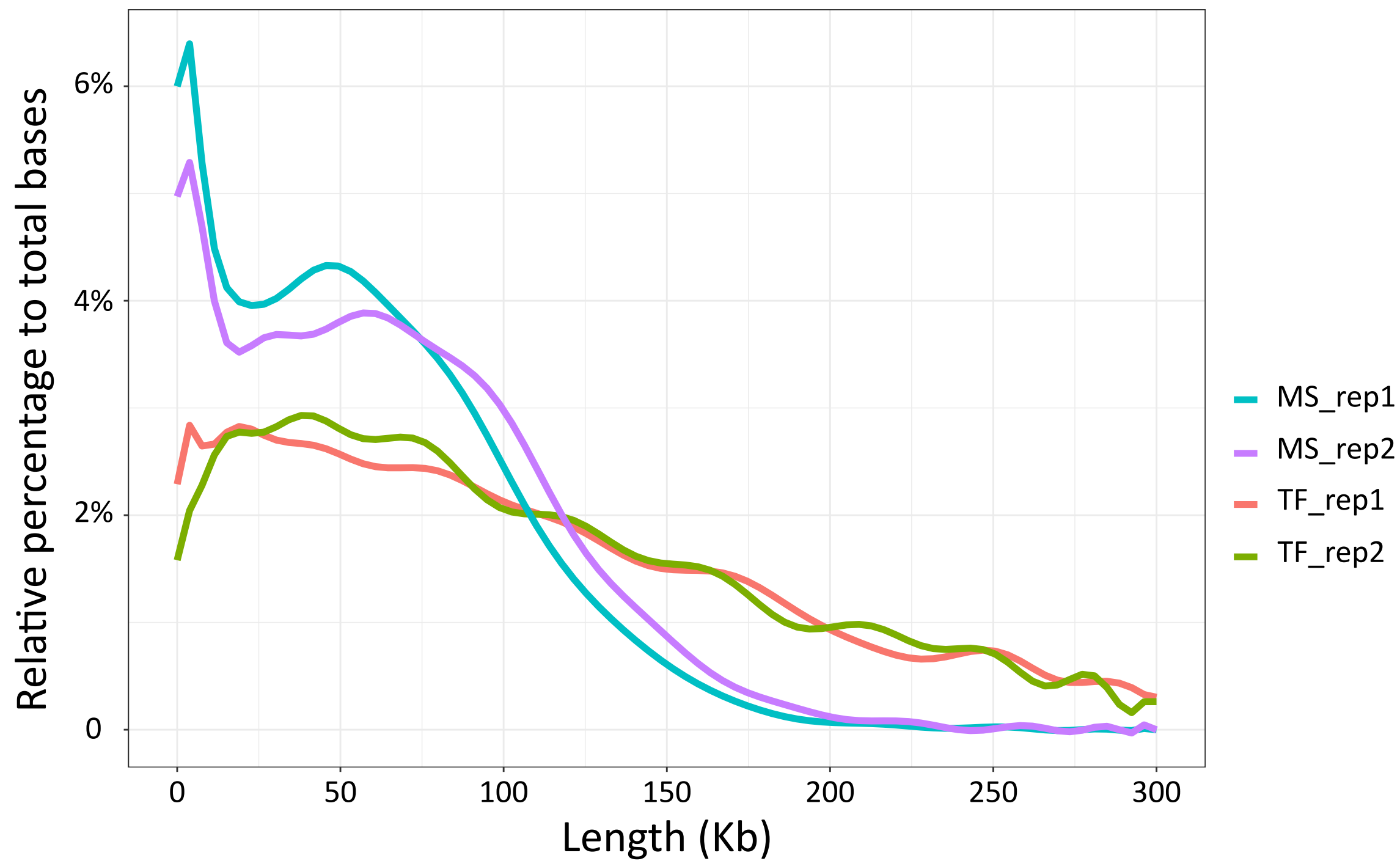


Figure 5



Mechanical shearing_rep1	
Cell line	HG00733
N50 of the reads	55,180
Number of reads longer than 100 Kb	2,500
Number of total reads	97,859
Maximum length (bp)	348,482
Mean length (bp)	17,861
Median length (bp)	5,335
Mean quality of the reads	10.0
Total bases of raw reads	1,747,849,822
Total bases of aligned reads	1,693,300,832
Mapped ratio of total bases (hg19, Minimap2)	96.9%
Number of active pores	1225: 480, 402, 254, 89

Mechanical shearing_rep2	Transposase fragmentation_rep1	Transposase fragmentation_rep2
HG00733	HG00733	HG00733
63,007	98,237	95,629
3,082	2,386	2,355
80,465	24,166	21,032
387,113	454,660	489,426
20,395	33,528	38,175
5,894	10,249	15,656
10.1	9.9	10.0
1,641,058,932	810,229,733	802,886,304
1,607,975,925	791,422,077	778,417,627
98.0%	97.7%	97.0%
1058: 480, 356, 176, 46	958: 452, 328, 148, 30	1092: 487, 367, 195, 43

Name of Reagent/ Equipment	Company	Catalog Number	Comments/Description
Reagents			
Absolute ethanol	Sigma-Aldrich	E7023	
Agencourt AMPure XPbeads	Beckman	A63881	magnetic beads for cleanup
BD conventional needles	Becton Dickinson	305136	27G, for mechanical shearing
BD Luer-Lok syringe	Becton Dickinson	309628	for mechanical shearing
Blunt/TA Ligase Master Mix	NEB	M0367S	
Countess Cell Counting Chamber Slides	Invitrogen	C10228	for cell counting
EDTA	Invitrogen	AM9261	pH 8.0, 0.5 M, 500 mL
Flow Cell	Oxford Nanopore Technologies	FLO-MIN106	R9.4.1
HG00773 cells	Coriell Institute	HG00733	cells used in this protocol
	Oxford Nanopore Technologies		
Ligation Sequencing Kit 1D	Technologies	SQK-LSK108	nanopore ligation kit
MaXtract High Density tubes	Qiagen	129073	gel tubes
NEBNext FFPE DNA Repair Mix	NEB	M6630S	
NEBNext Ultra II End Repair/dA-Tailing Module	NEB	M7546S	
Nuclease-free water	Invitrogen	AM9937	
Phosphate-Buffered Saline, PBS	Gibco	70011044	10X, pH 7.4
Phenol:chloroform:IAA	Invitrogen	AM9730	
Proteinase K	Qiagen	19131	20 mg/mL
Qubit dsDNA BR Assay Kit	Invitrogen	Q32850	fluorometer assays for DNA quantification
	Oxford Nanopore Technologies		
Rapid Sequencing Kit	Technologies	SQK-RAD004	nanopore transposase kit
RNase A	Qiagen	19101	100 mg/mL
SDS	Invitrogen	AM9822	10% (wt/vol)
Sodium chloride solution	Invitrogen	AM9759	5.0 M
TE buffer	Invitrogen	AM9849	pH 8.0
Tris	Invitrogen	AM9856	pH 8.0, 1 M
Triton X-100 solution	Sigma-Aldrich	93443	~10%
Equipment			

Bio-Rad C1000 Thermal Cyclor	Bio-Rad	1851196EDU	
Centrifuge 5810R	Eppendorf	22628180	
Countess II FL Automated Cell Counter	Life Technologies	AMQAF1000	for cell counting
DynaMag-2 Magnet	Life Technologies	12321D	magnetic rack
Eppendorf ThermoMixer	Eppendorf	5382000023	for incubation
Freezer	LabRepCo	LHP-5-UFMB	
	Oxford Nanopore		
GridION	Technologies	GridION X5	nanopore device used in this protocol
HulaMixer Sample Mixer	Thermo Fisher Scientific	15920D	rotator mixer
MicroCentrifuge	Benchmark Scientific	C1012	
NanoDrop ND-1000 Spectrophotometer	Thermo Fisher Scientific	ND-1000	for UV reading
Pippin Pulse	Sage Science	PPI0200	pulsed-field gel electrophoresis instrument
Qubit 3.0 Fluorometer	Invitrogen	Q33216	fluorometer
Refrigerator	LabRepCo	LABHP-5-URBSS	
Vortex-Genie 2	Scientific Industries	SI-A236	
Water bath	VWR	89501-464	



1 Alewife Center #209
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT

Title of Article:

Ultra-long Read Sequencing Analysis by Oxford Nanopore Devices

Author(s):

Liang Gong, Chee-Hong Wong, Jennifer Idol, Chew Yee Ngan, Chia-Lin Wei

Item 1: The Author elects to have the Materials be made available (as described at <http://www.jove.com/publish>) via:

☐

Standard Access

☒

Open Access

Item 2: Please select one of the following items:

☒

The Author is **NOT** a United States government employee.

☐

The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.

☐

The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

ARTICLE AND VIDEO LICENSE AGREEMENT

1. **Defined Terms.** As used in this Article and Video License Agreement, the following terms shall have the following meanings: **"Agreement"** means this Article and Video License Agreement; **"Article"** means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; **"Author"** means the author who is a signatory to this Agreement; **"Collective Work"** means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; **"CRC License"** means the Creative Commons Attribution-Non Commercial-No Derivs 3.0 Unported Agreement, the terms and conditions of which can be found at: <http://creativecommons.org/licenses/by-nc-nd/3.0/legalcode>; **"Derivative Work"** means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; **"Institution"** means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; **"JoVE"** means MyJoVE Corporation, a Massachusetts corporation and the publisher of The Journal of Visualized Experiments; **"Materials"** means the Article and / or the Video; **"Parties"** means the Author and JoVE; **"Video"** means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion

of the Article, and in which the Author may or may not appear.

2. **Background.** The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3. **Grant of Rights in Article.** In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to **Sections 4 and 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License.

4. **Retention of Rights in Article.** Notwithstanding the exclusive license granted to JoVE in **Section 3** above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5. **Grant of Rights in Video – Standard Access.** This **Section 5** applies if the "Standard Access" box has been checked in **Item 1** above or if no box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, Subject to **Section 7** below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

6. **Grant of Rights in Video – Open Access.** This **Section 6** applies only if the "Open Access" box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to **Section 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. For any Video to which this **Section 6** is applicable, JoVE and the Author hereby grant to the public all such rights in the Video as provided in, but subject to all limitations and requirements set forth in, the CRC License.

7. **Government Employees.** If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in **Item 2** above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum

rights permitted under such statute. In such case, all provisions contained herein that are not in conflict with such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

8. **Protection of the Work.** The Author(s) authorize JoVE to take steps in the Author(s) name and on their behalf if JoVE believes some third party could be infringing or might infringe the copyright of either the Author's Article and/or Video.

9. **Likeness, Privacy, Personality.** The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

10. **Author Warranties.** The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

11. **JoVE Discretion.** If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole

ARTICLE AND VIDEO LICENSE AGREEMENT

discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole discretion and without giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

12. **Indemnification.** The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in JoVE or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contamination due to


the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

13. **Fees.** To cover the cost incurred for publication, JoVE must receive payment before production and publication the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US\$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

14. **Transfer, Governing Law.** This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to be one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement is required per submission.

CORRESPONDING AUTHOR

Name:	Chia-Lin Wei	
Department:	Genome Technologies	
Institution:	The Jackson Laboratory for Genomic Medicine	
Title:	Director	
Signature:		Date: 8/15/2018

Please submit a **signed** and **dated** copy of this license by one of the following three methods:

1. Upload an electronic version on the JoVE submission site
2. Fax the document to +1.866.381.2236
3. Mail the document to JoVE / Attn: JoVE Editorial / 1 Alewife Center #200 / Cambridge, MA 02140

Editorial comments:

The manuscript has been modified and the updated manuscript, 58954_R1.docx, is attached and located in your Editorial Manager account. Please use the updated version to make your revisions.

1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues.

[As suggested by the editors, we have proofread the entire manuscript and corrected spelling and grammatic errors.](#)

2. Please do not highlight notes for filming.

[We have removed the highlight for the notes as suggested by the editor to adhere JoVE journal protocol style \(Line 156, Page 4\).](#)

3. JoVE cannot publish manuscripts containing commercial language. This includes company names or models of an instrument or reagent. Please remove all commercial language from your manuscript and use generic terms instead. All commercial products should be sufficiently referenced in the Table of Materials and Reagents. Examples of commercial language in your manuscript include GridION, MinION, SQK-LSK108, SQK-RAD004, etc.

[We have removed all instruments and reagents labeled with commercial language in the revised manuscript \(Lines 196-202, Page 5; Lines 318, 333 and 351, Page 8\).](#)

4. Please use h, min, s for time units.

[We have revised the time units in the revised manuscript \(Line 123, Page 3; Line 403, Page 10\).](#)

Supplementary Table 1. Summary of two NLR-seq runs using other cell lines with the

	Library 1
Cell line	K562
Cell Ordering Information	ATCC, cat. No. CCL-243
Protocol	mechanical shearing
N50 of the reads	60,063
Number of total reads	193,783
Median length (bp)	1,843
Mean length (bp)	9,825
Maximum length (bp)	548,780
Total bases of raw reads	1,903,989,686
Total bases of aligned reads	1,837,350,047
Mapped ratio of total bases (hg19, Minimap2)	96.6%
Number of active pores	1111: 482, 371, 203, 55

mechanical shearing protocol.

Library 2
GM19240
Coriell Institute, cat. No. GM19240
mechanical shearing
55,295
120,807
4,688
17,408
212,338
2,103,015,331
1,997,419,761
95.0%
1032: 447, 333, 196, 56