

# A novel approach reveals genomic landscapes of single-strand DNA breaks with nucleotide resolution in human cells

Fan Gao  
Ye Cai  
Lu Tang  
Chun-Lin Wu  
Huifen Cao  
Dongyang Xu  
Lorena Salazar-García  
Xueer Han  
Fang Wang  
Natalie Ricciardi  
Thor Wahlestedt  
DingDing Cai  
Huifang Wang  
Mario P. S. Chin  
James A. Timmons  
Claes Wahlestedt  
Philipp Kapranov (✉ [philippk08@hotmail.com](mailto:philippk08@hotmail.com))

---

## Method Article

**Keywords:** single-strand DNA break, single-strand DNA break mapping, next generation sequencing, DNA repair, molecular biology, genome stability

**Posted Date:** August 27th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.pex-920/v2>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

## Abstract

Single strand breaks (SSBs) represent the major form of DNA damage, yet no technique exists to map these lesions genome-wide with nucleotide-level precision. Herein, we present a method, termed SSiNGLe, and demonstrate its utility to explore the distribution and dynamic changes of genome-wide SSBs in response to different biological and environmental stimuli. We validate SSiNGLe using two very distinct sequencing techniques and apply it to derive global profiles of SSBs in different biological states. Strikingly, we show that patterns of SSBs in the genome are non-random, specific to different biological states, enriched in regulatory elements, exons, introns, specific types of repeats and exhibit differential preference for the template strand between exons and introns. Furthermore, we show that breaks likely contribute to naturally occurring sequence variants. Finally, we demonstrate strong links between SSB patterns and age. Overall, SSiNGLe provides access to unexplored realm of cellular biology, not obtainable with current approaches.

## Introduction

### Introduction

DNA damage is now widely recognized as a major reason behind cancer and many other aging-associated diseases and as such represents a very important issue for human health<sup>1,2</sup>. While multiple types of DNA lesions exist, SSBs are considered the most common type of DNA damage<sup>3</sup>. These lesions can represent sites of oxygen radical DNA damage, intermediates in excision DNA repair pathway and products of unresolved intermediates of enzymes such as topoisomerases<sup>3</sup>. SSBs can further deteriorate into highly toxic double-strand breaks (DSBs) by stalling or collapsing replication fork<sup>4</sup>. However, by themselves SSBs can also represent a major issue for cell as they can inhibit progression of RNA polymerase<sup>5</sup> and in some cases cause apoptosis<sup>6,7</sup>. The importance of this type of lesion is underscored by existence of dedicated cellular pathways that deal with every step of fixing SSBs from detection to processing to repair<sup>3</sup>. Defects in these pathways can lead to cellular sensitivity to genotoxic stress, embryonic lethality and a number of neurodegenerative diseases<sup>8</sup>.

The remarkable progress in appreciation of the fine details of the SSB repair machinery however stands in stark contrast with total absence of methods to map endogenous SSBs in a global, unbiased and genome-wide fashion with nucleotide precision. This gap in available methodologies also contrasts with a suite of comprehensive approaches developed for mapping DSBs with nucleotide-level resolution, such as BLESS<sup>9</sup>, BLISS<sup>10</sup> and others (reviewed in Ref<sup>11</sup>). To our knowledge, only one SSB genome-wide mapping method can provide comprehensive and unbiased data<sup>12</sup>. The procedure relies on the 3'OH group of an SSB to prime a DNA polymerase I nick-translation reaction that labels downstream DNA with

a biotinylated nucleotide<sup>12</sup>. The labeled DNA is then purified and subjected to next-generation sequencing<sup>12</sup>. However, this approach maps a region of DNA, quite possibly extending thousands of bases from the original SSB, thus precluding identification of the lesion with nucleotide precision. On the other hand, a nucleotide-level method to map sites of excision repair has been developed<sup>13</sup>, however, it cannot provide information on breaks generated by other mechanisms.

Thus, all this leads to a total dearth of knowledge of nucleotide-level genome-wide patterns for this critical type of DNA lesion. Here, we develop and validate an approach, SSiNGLe (single-strand break mapping at nucleotide genome level, **Fig.1**) that can provide nucleotide level maps of native SSBs genome-wide. We implement this approach to work with two next-generation sequencing (NGS) platforms: (1) a 3<sup>rd</sup> generation single molecule sequencing (SMS) platform (Helicos/SeqLL) whose unique capabilities obviate the need for lengthy sample preparation and PCR amplification, and (2) the more commonly used Illumina platform. We show highly consistent and striking patterns of SSBs using both sequencing platforms. Furthermore, the results show that the genomic pattern of breaks – the SSB “breakome” - has strong potential to represent a novel dimension describing state of a biological system and a novel source of blood-based biomarkers, with potentially yet undiscovered connections to aging.

## Reagents

### Reagents

- 1× PBS, pH7.2-7.4 (Solarbio, cat. no. P1020)
- Trypsin-EDTA (0.05%), phenol red (Gibco, cat. no. 25300054)
- Solarbio Lymphocyte Separation Medium (Solarbio, cat. no. P8610)
- 37% formaldehyde (Amresco, cat. no. C000T93)
- UltraPure™ Glycine (Invitrogen, cat. no. 15527013)
- Cell lysis buffer: 5 mM PIPES (pH8), 85 mM KCl, 0.5% NP-40
- UltraPure™ SDS Solution, 10% (SDS) (Invitrogen, cat. no. 15553027)
- Micrococcal Nuclease (MNase), 2× 10<sup>6</sup> gel units/ml (NEB, cat. no. M0247S)
- 10 mg/ml bovine serum albumin (BSA), part of the MNase package (NEB, cat. no. M0247S)
- 1× MNase buffer (50 mM Tris-HCl pH8, 5 mM CaCl<sub>2</sub>)



-P5\_N10G10: CTACACGACGCTCTTCCGATCTAGTTGCGGATGGGGGGGGGGHN

-P7\_T12: GTTCAGACGTGTGCTCTTCCGATCTTTTTTTTTTTTTTTVN

-Illumina-P5:

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT

-Illumina-P7:

CAAGCAGAAGACGGCATACGAGAT(CGTGAT)GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

**Note:** the sequence in brackets corresponds to the library index sequence ATCACG. Different Illumina-P7 primer oligonucleotides could be designed by varying the index sequences depending on the experimental design. Samples with different indexes can be sequenced on the same flow cell lane. However, we always use the same index for control and treatment samples that are to be directly compared.

## Equipment

### Equipment

-Manual Pipettes (Eppendorf)

-Coverslip (Citotest Scientific Co., Ltd, cat. no.10212432C)

-Blood counting chamber (Shanghai precision biochemical reagent instrument Co., Ltd)

-Inverted microscope (Eclipse TS100, Nikon)

-Tabletop centrifuge to precipitate cells (5702R centrifuge, Eppendorf)

-1.5/2 ml Eppendorf tubes (DNase/RNase-free, ultra-low attachment) (Axygen, cat. no. MCT-150-L-C/MCT-200-C).

-15/50 ml conical sterile tubes (Thermo Nunc, cat. no. 339651/339653)

-0.2 ml PCR tube (DNase/RNase-free, ultra-low attachment) (Axygen, cat. no. PCR-02-C)

-10/20/200/1000 µl pipette tips (DNase/RNase-free, Ultra-low attachment) (Axygen, cat. no. TF-300/20/200/1000)

-Agarose gel electrophoresis set up (DLAB Scientific Co., Ltd)

-Magnetic rack for 1.5 ml tubes (Promega, cat. no. Z5342)

-Vortex (DLAB Scientific Co., Ltd cat. no. D1008)

- Eppendorf Concentrator plus (Eppendorf)
- Tabletop centrifuge (5804R centrifuge, Eppendorf)
- High-speed refrigerated centrifuge (5424R centrifuge, Eppendorf)
- PCR thermocycle (C1000 PCR amplifier, Bio-Rad)
- Thermostatic metal bath (Thermomixer C, Eppendorf)
- Qubit® 3.0 Fluorometer (Thermo, cat. no. Q33216)
- Agilent 2100 bioanalyzer (Agilent, cat. no. G2939A)
- Dell R540 Server, 2\*CPU Intel Xenon 4116 2.1GHz, 2\*20 threads, 10\*32GB RDIMM, 10\*10TB 7.2K RPM SATA 6Gbps 512e hard drives for Illumina reads processing and extraction of coordinates of breaks

### **Illumina reads processing software**

- Ubuntu (v14.04)
- BWA (v0.7.12)
- BEDtools suite (v2)
- FASTX-Toolkit (v0.013)
- Custom scripts used in the pipeline are available from GitHub at [https://github.com/huifencao/DB-Nature\\_Communications-2019](https://github.com/huifencao/DB-Nature_Communications-2019).

## **Procedure**

### **Procedure**

#### **Day 1:**

**Note:** this procedure has been used for cultured adherent and suspension cells as well as for Peripheral Blood Mononuclear Cells (PBMCs). The cultured cells can be grown using standard conditions appropriate for particular cells lines being assayed and processed as described below starting with the “Preparation of crosslinked nuclei” section. The procedure for preparation of PBMCs is given below.

### **Preparation of PBMCs (optional), 1 hour:**

**Note:** here and below the times in hours are approximate and given for 1-10 samples processed in parallel.

1. Collect 2-4 ml of peripheral blood in EDTA anti-coagulation tubes and mix well by gently inverting the tube several times. Keep on ice and process as soon as possible. Do not freeze.
2. Dilute the samples with equal volume of 1× PBS at room temperature.
3. Take several of the Solarbio Lymphocyte Separation Medium into a 15 ml centrifuge tube (add 3 ml when the volume of the diluted blood is less than 3 ml, add equal volume of Solarbio Lymphocyte Separation Medium when the volume of the diluted blood is more than 3 ml).
4. Gently layer the diluted blood on the top of Solarbio Lymphocyte Separation Medium. The layering should be done very slowly that diluted blood and Solarbio Lymphocyte Separation Medium should stay as two different layers.

**Note:** the total volume of diluted blood and Solarbio Lymphocyte Separation Medium shouldn't more than 10 ml, otherwise the separation will be affected. Make sure that the samples are at room temperature before proceeding to the next step.

5. Centrifuge the samples for 30 min at 1000g at 25°C in the 5804R centrifuge with acceleration and brake settings of 6 and 0 respectively.
6. Carefully aspirate the PBMCs formed in the interphase into a new 15 ml tube. The cells in the interphase need to be aspirated without delay.
7. Add 10 ml of sterile 1× PBS into the tube with the aspirate, gently resuspend the aspirate and then centrifuge at 250g at 25°C for 10 min.
8. Remove the supernatant and repeat the last step.
9. After the centrifugation, resuspend the aspirate in 1-2 ml of sterile 1× PBS, count the PBMCs while keeping cells on ice and proceed to the crosslinking immediately as described in the section below at step 8.

### **Preparation of crosslinked nuclei, 2.5 hours:**

Keep samples on ice unless specified otherwise. Samples processed with high-speed refrigerated centrifuge below.

1. For non-adherent cells, proceed directly to step 5 of this procedure.

2. For adherent cells, remove the growth medium carefully and wash the cells twice with 1× PBS at room temperature.
3. Clearly remove the 1× PBS and digest with trypsin until the cells detach.
4. Inactivate trypsin with the growth medium with serum.
5. Resuspend and count the cells.
6. Collect the cells, centrifuge to remove the growth medium with serum at 100 g for 5 min at 4°C.
7. Remove the supernatant, completely resuspend the cells in the fresh growth medium with serum.
8. Start with 1-3 million non-adherent or adherent cells or 4-6 million PBMCs resuspended in 1-2 ml of either the growth medium with serum or 1× PBS (PBMCs) in a new 2 ml Eppendorf tube.
9. Add 27 µl of 37% formaldehyde per 1 ml of cell culture medium or PBS while slowly shaking the cells for 10 min at room temperature. The final concentration of formaldehyde is 1%.
10. Add 100 µl of 1.375 M glycine (final concentration 0.125 M) per milliliter to quench formaldehyde for 5 min at room temperature. Gently pipette with tip and invert several times with the tube to mix them well.
11. Centrifuge the cells at 220 g for 10 min at 4°C.
12. Remove the supernatant and wash the cells twice with 2 ml of cold 1× PBS at 220 g for 10 min at 4°C.
13. Carefully remove the supernatant and resuspend the cell pellet in 2 ml of cold cell lysis buffer.  
**Note:** the pelleted cells would stick to each other after adding the cell lysis buffer, therefore pipette with up and down to separate them completely. Setting the volume to one third of the pipette range can effectively avoid the pellet loss.
14. Incubate for 10 min on ice.
15. After the lysis, centrifuge the nuclei at 300 g in a benchtop centrifuge for 10 min at 4°C.
16. Aspirate the supernatant carefully.
17. Wash the pellet with 500 µl cold 1× MNase buffer, centrifuge for 10 min at 4°C at 300 g.
18. Repeat steps 16-17 and remove the supernatant.
19. At this step, the nuclei can be stored at -80°C for up to one month. If this is desirable, resuspend the pellet by adding 100 µl 1× MNase buffer and move the samples directly into -80°C. When ready to proceed, thaw them on the ice around 5 min until completely thawed, centrifuge for 10 min at 4°C at 300

g, remove the supernatant and proceed to the sections below. If storage is not necessary, proceed with the nuclear pellet directly to the sections below.

### **SDS permeabilization (optional), 2 hours:**

**Note:** this step is only necessary if nuclei are to be treated with enzymes that cannot easily diffuse into the nuclei (having molecular weight >30 kDa) as shown in the example below for the treatment with shrimp alkaline phosphatase (SAP) that can make breaks with 3'-phosphate termini detectable in the SSiNGLe procedure. If no such treatments are necessary then proceed directly to the section "MNase fragmentation and DNA purification".

1. Resuspend the pellet completely in 100  $\mu$ l of 1 $\times$  MNase buffer and transfer the sample to a 1.5 ml Eppendorf tube.

2. Add 3  $\mu$ l of 10% SDS into the tube, gently pipette up and down while moving around the tip to mix the SDS well, the final concentration of SDS is 0.3%. Slowly shake (300 rpm) at 37°C for 1h in the thermostatic metal bath.

**Note:** the nuclei would stick to the pipette tip at the following step easily, therefore use ultra-low attachment tip and pipette gently to avoid loss.

3. Slowly add 1.8  $\mu$ l of Triton X-100 into the solution, gently pipette up and down while moving around the tip to mix Triton X-100 well, the final concentration of Triton X-100 is 1.8%. Incubate for 5 min at room temperature to sequester SDS.

4. Centrifuge 15 min at 1500 g with the temperature setting of 27°C in a high-speed refrigerated centrifuge.

5. Aspirate the supernatant carefully and completely. Wash the pellet with 500  $\mu$ l of 1 $\times$  MNase buffer, centrifuge 15 min at 1500 g with the temperature setting of 27°C in a high-speed refrigerated centrifuge.

**Note:** prolong the centrifugation time if complete precipitation is not achieved initially.

6. Repeat step 5.

7. Aspirate the supernatant carefully and use the nuclei for the following step.

### **Treatment with SAP (optional), 3.5 hours:**

1. Wash the permeabilized nuclei with 500  $\mu$ l of 1 $\times$  NEB buffer 2.1, gently resuspend the permeabilized nuclei and centrifuge at 1500 g with the temperature setting of 27°C in a high-speed refrigerated

centrifuge for 10 min.

2. Carefully remove the supernatant and repeat step 1.
3. Remove the supernatant, resuspend the nuclei in 50  $\mu$ l of 1 $\times$  NEB buffer 2.1.
4. Add 1  $\mu$ l SAP into the solution, treat the nuclei in 50  $\mu$ l of 1 $\times$  NEB buffer 2.1 with 1U of SAP for 2h at 37°C in thermostatic metal bath.
5. Inactivate enzyme by incubation at 70°C for 20 min in thermostatic metal bath.
6. Collect the nuclei by centrifugation at 1500 g with the temperature setting of 27°C in a high-speed refrigerated centrifuge for 10 min.
7. Aspirate the supernatant carefully and completely.
8. Wash the pellet twice with 500  $\mu$ l of 1 $\times$  MNase buffer at 1500 g with the temperature setting of 27°C in a high-speed refrigerated centrifuge for 10 min.
9. Remove the supernatant, use the nuclei pellet for the following step.

#### **MNase fragmentation and DNA purification, 4 hours:**

1. Completely resuspend the nuclei in 50  $\mu$ l of cold 1 $\times$  MNase buffer supplemented with 200  $\mu$ g/ml BSA (1  $\mu$ l of BSA) per 1 million nuclei (based on the original number of cells) in a 1.5 ml Eppendorf tube.

**Note:** the nuclei would stick to the pipette tip from this step easily, therefore use Ultraultra-low attachment tip and pipette gently to avoid loss.

2. Split the sample into aliquots of 1 million nuclei per 0.2 ml PCR tube.
3. Add 1200-3000 units of MNase and 25 units of RNase If per aliquot.
4. Mix them well by pipetting up and down. Incubate for 30 minutes on ice.
5. Add 5.6  $\mu$ l of 0.5 M EDTA into the solution and slowly mix by pipetting up and down to stop the digestion.
6. Add 150  $\mu$ l of nuclei lysis buffer to the tube and incubate for 5 min at room temperature.
7. Add 1  $\mu$ l of 20 mg/ml proteinase K into the mixture and incubate for 45 min at 55°C and 45 min at 65°C in the PCR thermocycler.
8. Transfer the samples to 1.5 ml Eppendorf tubes.

9. Add 20  $\mu$ l of 3 M KCl solution and 176  $\mu$ l of cold isopropanol.
  10. Vortex 5s, then incubate for 30 min at -80°C.
- Note:** the samples can be stored at -80°C for up to several months.
11. Thaw the samples at room temperature and centrifuge for 15 min at 4°C at 12,000g (10 min at max speed at 4 °C).
  12. Remove the supernatant, then add 100  $\mu$ l of ice-cold 70% ethanol into the tube.
  13. Vortex 5s, then centrifuge the tube for 5 min at max speed at room temperature.
  14. Repeat steps 12-13.
  15. Aspirate the supernatant completely and carefully, vacuum-dry the DNA pellet for 2 min in the Eppendorf Concentrator plus or at room temperature for 5-10 min.
  16. Dissolve the DNA pellet in 20  $\mu$ l UltraPure™ DNase/RNase-Free Distilled Water and transfer to a new 1.5 ml Eppendorf tube.
  17. Pre-warm VAHTS DNA Clean Beads for 30 min at room temperature.
  18. Add double volume (40  $\mu$ l) ratio of VAHTS DNA Clean Beads suspension to the DNA.
  19. Mix thoroughly by pipetting up and down ~20 times.
  20. Incubate for 5 min at room temperature.
  21. Place the tube on the 1.5 ml magnetic rack to separate beads from the supernatant.
  22. When the solution is clear (~ 5 min), aspirate and discard the supernatant.
  23. Add 200  $\mu$ l of freshly prepared 80% ethanol to the tube with the sample while on the magnetic stand (keep the beads together).
  24. Incubate at room temperature for 30 seconds, and then carefully remove and discard the supernatant.
  25. Repeat steps 23-24 once for a total of twice washing.
  26. Briefly spin the tube, and put the tube back in the magnetic rack.
  27. Remove the residual ethanol completely with a 10  $\mu$ l tip.
  28. Air dry the beads (~5 min).
  29. Remove the tube from the magnetic rack and add 17  $\mu$ l of nuclease-free water.

30. Mix well by pipetting up and down 30-40 times, incubate at room temperature for 5 min.
31. Put the tube in the magnetic rack until the solution is clear (~5 min).
32. Carefully remove 16  $\mu$ l of the supernatant and transfer to a clean 1.5 ml Eppendorf tube.
33. Measure the DNA concentration using Qubit 3.0 fluorometer and "dsDNA HS Assay" kit.
34. Check the DNA products on 1% agarose gel or BioAnalyzer to ensure that the majority of DNA is in the range of 150-500 bp.

## **Day 2:**

### **PolyA-tailing, 2 hours:**

1. Transfer 100 ng of the purified fragmented DNA into a 0.2 ml Axygen RNase/DNase free PCR tube.

2. Add the following:

-10 $\times$  TdT buffer: 2  $\mu$ l

-2.5 mM  $\text{CoCl}_2$ : 2  $\mu$ l

-Nuclease free water to make the total volume of 19  $\mu$ l

3. Denature the mixture at 95 $^\circ\text{C}$  for 5 min followed by rapid snap-cooling on ice.

4. Dilute TdT (20 U/ $\mu$ l) to 4 U/ $\mu$ l with 1 $\times$  TdT buffer.

5. Add the following to the tube:

-4 U/ $\mu$ l TdT: 1  $\mu$ l

-10 mM dATP: 2  $\mu$ l

6. Mix gently (do not vortex) to ensure that the enzyme is mixed completely and incubate the mixture at 37 $^\circ\text{C}$  for 30 min and keep at 4 $^\circ\text{C}$  before going to the next step in a PCR thermocycler.

7. Take the sample out, add 2  $\mu$ l of 10 mM of either ddCTP, ddGTP, ddATP or ddTTP to the mixture to block the tailing.

**Note:** we used any one of these dideoxy nucleotides at the blocking step.

8. Mix gently but thoroughly (do not vortex) and incubate at 37  $^\circ\text{C}$  30 min and at 70 $^\circ\text{C}$  10 min and keep at 4 $^\circ\text{C}$  before going to the next step in a PCR thermocycler.

9. At this stage, the tailed material can be used directly for single-molecule sequencing using the Helicos/SeqLL platform without any further purification (the SSiNGLe-SMS approach). Alternatively, it can be used to prepare libraries for Illumina-based sequencing using the steps below.
10. Pre-warm VAHTS DNA Clean Beads for 30 min at room temperature.
11. Add double volume (48  $\mu$ l) ratio of VAHTS DNA Clean Beads suspension to the polyA-tailing mixture.
12. Mix thoroughly by pipetting up and down ~20 times.
13. Incubate for 5 min at room temperature.
14. Place the tube on the 1.5 ml magnetic rack to separate beads from the supernatant.
15. When the solution is clear (~ 5 min), aspirate and discard the supernatant.
16. Add 200  $\mu$ l of freshly prepared 80% ethanol to the tube with the sample while on the magnetic stand (keep the beads together).
17. Incubate at room temperature for 30 seconds, and then carefully remove and discard the supernatant.
18. Repeat steps 16-17 once for a total of twice washing.
19. Briefly spin the tube, and put the tube back in the magnetic rack.
20. Remove the residual ethanol completely with a 10  $\mu$ l tip.
21. Air dry the beads (~5 min).
22. Remove the tube from the magnetic rack and add 18  $\mu$ l of nuclease-free water.
23. Mix well by pipetting up and down 30-40 times, incubate at room temperature for 5 min.
24. Put the tube in the magnetic rack until the solution is clear (~5 min).
25. Transfer 16.2  $\mu$ l of the supernatant to a new 0.2 ml PCR tube carefully.

### **SSiNGLe-ILM sequencing library construction:**

#### **Linear amplification, 1.5 hours:**

1. Add the following to the 16.2  $\mu$ l of the eluted DNA from the step above:

-10 $\times$  PCR buffer: 2  $\mu$ l

-10 mM dNTP mix: 0.4  $\mu$ l

-10  $\mu$ M of a chimeric DNA-RNA oligo-d(T)50-r(T)3 oligonucleotide: 1  $\mu$ l

-Taq DNA polymerase: 0.4  $\mu$ l

2. Cycling parameters in a thermocycler with the lid set at 105 °C:

1). 94°C for 30 sec

2). 94°C for 1 min

3). 55°C for 30 sec

4). 72°C for 30 sec

5). Loop back to step 2) 9 times, for a total of 10 cycles.

6). Rapid snap-cooling on ice to prevent non-specific extension.

3. Pre-warm VAHTS DNA Clean Beads for 30 min at room temperature.

4. Add double volume (40  $\mu$ l) ratio of VAHTS DNA Clean Beads suspension to the polyA-tailing mixture.

5. Mix thoroughly by pipetting up and down ~20 times.

6. Incubate for 5 min at room temperature.

7. Place the tube on the 1.5 ml magnetic rack to separate beads from the supernatant.

8. When the solution is clear (~ 5 min), aspirate and discard the supernatant.

9. Add 200  $\mu$ l of freshly prepared 80% ethanol to the tube with the sample while on the magnetic stand (keep the beads together).

10. Incubate at room temperature for 30 seconds, and then carefully remove and discard the supernatant.

11. Repeat steps 9-10 once for a total of twice washing.

12. Briefly spin the tube, and put the tube back in the magnetic rack.

13. Remove the residual ethanol completely with a 10  $\mu$ l tip.

14. Air dry the beads (~5 min).

15. Remove the tube from the magnetic rack and add 16  $\mu$ l of nuclease-free water.

16. Mix well by pipetting up and down 30-40 times, incubate at room temperature for 5 min.

17. Put the tube in the magnetic rack until the solution is clear (~5 min).
18. Transfer 15  $\mu$ l of the supernatant to a new 0.2 ml PCR tube carefully.

### **PolyC-tailing, 2 hours:**

1. Add the following to the 15  $\mu$ l of the polyA-tailed DNA from above:

-10 $\times$  TdT buffer: 2  $\mu$ l

-2.5 mM  $\text{CoCl}_2$ : 2  $\mu$ l

2. Denature the mixture at 95°C for 5 min, and crash on ice for 2 min.

3. Dilute TdT (20 U/ $\mu$ l) to 4 U/ $\mu$ l with 1 $\times$  buffer.

4. Add the following:

-4 U/ $\mu$ l TdT: 1  $\mu$ l

-10 mM dCTP: 2  $\mu$ l

5. Mix gently (do not vortex) to ensure that the enzyme is mixed completely and incubated at 37°C for 30 min, 4°C forever in a PCR thermocycler.

6. Add 2  $\mu$ l of 10 mM of either ddCTP, ddGTP, ddATP or ddTTP to the mixture to block the tails.

**Note:** we used any one of these dideoxy nucleotides at the blocking step

7. Mix gently but thoroughly (do not vortex) and incubate at 37 °C 30 min and at 70°C 10 min and keep at 4°C before going to the next step.

8. Pre-warm VAHTS DNA Clean Beads for 30 min at room temperature.

9. Add double volume (48  $\mu$ l) ratio of VAHTS DNA Clean Beads suspension to the polyA-tailing mixture.

10. Mix thoroughly by pipetting up and down ~20 times.

11. Incubate for 5 min at room temperature.

12. Place the tube on the 1.5 ml magnetic rack to separate beads from the supernatant.

13. When the solution is clear (~ 5 min), aspirate and discard the supernatant.

14. Add 200  $\mu$ l of freshly prepared 80% ethanol to the tube with the sample while on the magnetic stand (keep the beads together).
15. Incubate at room temperature for 30 seconds, and then carefully remove and discard the supernatant.
16. Repeat steps 14-15 once for a total of twice washing.
17. Briefly spin the tube, and put the tube back in the magnetic rack.
18. Remove the residual ethanol completely with a 10  $\mu$ l tip.
19. Air dry the beads (~5 min).
20. Remove the tube from the magnetic rack and add 16.5  $\mu$ l of nuclease-free water.
21. Mix well by pipetting up and down 30-40 times, incubate at room temperature for 5 min.
22. Put the tube in the magnetic rack until the solution is clear (~5 min).
23. Transfer 15.2  $\mu$ l of the supernatant to a new 0.2 ml PCR tube carefully.

#### **Illumina library construction, 4 hours:**

**Note:** this procedure differs from the one published in the original paper. Specifically, the version presented here employs 2 rounds of PCR to construct the library which significantly increase the yield of the amplified products.

**Note:** the original protocol was tested on the HiSeq X Ten Illumina platform. However, the NovaSeq Illumina platform can not sequence DNA fragments starting with 10 G residues. Therefore, in order to build libraries for the NovaSeq platform, the P5\_G10 oligonucleotide has to be substituted for P5\_N10G10 in the PCR reaction below. The rest of the procedure is the same.

1. Add the following to the 15.2  $\mu$ l of the polyC-tailed DNA from above:

-10 $\times$  PCR buffer: 2  $\mu$ l

-10 mM dNTP mix: 0.4  $\mu$ l

-10  $\mu$ M P5\_G10 or P5\_N10G10 oligonucleotide: 1  $\mu$ l

-10  $\mu$ M P7\_T12 oligonucleotide: 1  $\mu$ l

-Taq DNA polymerase: 0.4  $\mu$ l

2. Cycling parameters with the lid set at 105  $^{\circ}$ C:

- 1). 94°C, 3 min
  - 2). 94°C, 30 sec
  - 3). 55°C, 1 min
  - 4). 72°C, 1min
  - 5). 94°C, 30 sec
  - 6). 37°C, 1 min
  - 7). Slow ramp at 2°C per minute to 72°C followed by 2 min incubation
  - 8). 94°C, 30 sec
  - 9). 60°C, 30 sec
  - 10). 72°C, 30 sec
  - 11). Loop back to step 8) 29 times for a total of 30 cycles
  - 12). 72°C, 7min
  - 13). 4 °C, hold
3. Pre-warm VAHTS DNA Clean Beads for 30 min at room temperature.
  4. Add double volume (40 µl) ratio of VAHTS DNA Clean Beads suspension to the polyA-tailing mixture.
  5. Mix thoroughly by pipetting up and down ~20 times.
  6. Incubate for 5 min at room temperature.
  7. Place the tube on the 1.5 ml magnetic rack to separate beads from the supernatant.
  8. When the solution is clear (~ 5 min), aspirate and discard the supernatant.
  9. Add 200 µl of freshly prepared 80% ethanol to the tube with the sample while on the magnetic stand (keep the beads together).
  10. Incubate at room temperature for 30 seconds, and then carefully remove and discard the supernatant.
  11. Repeat steps 9-10 once for a total of twice washing.
  12. Briefly spin the tube, and put the tube back in the magnetic rack.
  13. Remove the residual ethanol completely with a 10 µl tip.

14. Air dry the beads (~5 min).
15. Remove the tube from the magnetic rack and add 18  $\mu$ l of nuclease-free water.
16. Mix well by pipetting up and down 30-40 times, incubate at room temperature for 5 min.
17. Put the tube in the magnetic rack until the solution is clear (~5 min).
18. Transfer 17  $\mu$ l of the supernatant to a 1.5 ml Eppendorf tube carefully.
19. Prepare the following PCR reaction:
  - The eluted DNA from the step above: 4  $\mu$ l
  - 10 $\times$  PCR buffer: 2  $\mu$ l
  - 2.5 mM dNTP mix: 1.6  $\mu$ l
  - 10  $\mu$ M Illumina-P5 primer oligonucleotide: 1  $\mu$ l
  - 10  $\mu$ M Illumina-P7 primer oligonucleotide: 1  $\mu$ l
  - Taq DNA polymerase: 0.4  $\mu$ l
  - Nuclease free water: 10  $\mu$ l
20. Cycling parameters with the lid set at 105  $^{\circ}$ C:
  - 1). 94 $^{\circ}$ C, 3 min
  - 2). 94 $^{\circ}$ C, 30 sec
  - 3). 55 $^{\circ}$ C, 30 sec
  - 4). 72 $^{\circ}$ C, 30 sec
  - 5). Loop back to step 2) 5 times for a total of 6 cycles
  - 6). 72 $^{\circ}$ C, 7min
  - 7). 4  $^{\circ}$ C, hold
21. Pre-warm VAHTS DNA Clean Beads for 30 min at room temperature.
22. Add double volume (40  $\mu$ l) ratio of VAHTS DNA Clean Beads suspension to the polyA-tailing mixture.
23. Mix thoroughly by pipetting up and down ~20 times.

24. Incubate for 5 min at room temperature.
25. Place the tube on the 1.5 ml magnetic rack to separate beads from the supernatant.
26. When the solution is clear (~ 5 min), aspirate and discard the supernatant.
27. Add 200  $\mu$ l of freshly prepared 80% ethanol to the tube with the sample while on the magnetic stand (keep the beads together).
28. Incubate at room temperature for 30 seconds, and then carefully remove and discard the supernatant.
29. Repeat steps 27-28 once for a total of twice washing.
30. Briefly spin the tube, and put the tube back in the magnetic rack.
31. Remove the residual ethanol completely with a 10  $\mu$ l tip.
32. Air dry the beads (~5 min).
33. Remove the tube from the magnetic rack and add 22  $\mu$ l of nuclease-free water.
34. Mix well by pipetting up and down 30-40 times, incubate at room temperature for 5 min.
35. Put the tube in the magnetic rack until the solution is clear (~5 min).
36. Transfer 21  $\mu$ l of the supernatant to a 1.5 ml Eppendorf tube carefully.
37. Measure the DNA concentration using Qubit 3.0 fluorometer and "dsDNA HS Assay" kit.
38. At this stage, the samples can be stored for at least 1 year in -80°C, however, we usually send them for sequencing as soon as possible.

### **Day 3:**

Shipping the samples for NGS. We outsource NGS to Novogene Corporation (Beijing). The samples from above are typically shipped on dry ice in the concentration and volume required by the vendor. We have so far successfully sequenced these libraries on the HiSeq X Ten or NovaSeq platforms using the paired-end 150 bp strategy.

### **Processing of the Illumina reads and extracting the coordinates of breaks:**

1. Filter the reads such that each read has Phred quality score >20 and length >30:

```
fastq_quality_trimmer -v -t 20 -l 30 -i read_1.clean.fq -o read_1.clean.fq.1 -Q33;
```

```
fastq_quality_trimmer -v -t 20 -l 30 -i read_2.clean.fq -o read_2.clean.fq.1 -Q33;
```

2. Select read pairs where reads 1 start with 10 G's or AGTTGCGGATGGGGGGGGGG in the NovaSeq protocol and reads 2 start with 12T's.

**Note:** the selection criteria could be relaxed to include reads starting with 9-11G's and reads 2 starting with 11-13 T's by selecting reads starting with the desired sequences and then merging the filtering results. For example, the data below were filtered for 10 G's and 12 T's (2.1) or AGTTGCGGATGGGGGGGGGG and 12 T's (2.2). The data can also be for example selected for reads starting with 9 G's and 12 T's, 10 G's and 12 T's, 11 G's and 12 T's and so on and merged together.

```
(2.1) perl 1_filter_paired.flags.pl read_1.clean.fq.1 read_2.clean.fq.1 GGGGGGGGGG TTTTTTTTTTTT
```

```
(2.2) perl 1_filter_paired.flags.pl read_1.clean.fq.1 read_2.clean.fq.1 AGTTGCGGATGGGGGGGGGG  
TTTTTTTTTTTTT
```

3. Map the reads to the desired reference genome, which should be downloaded in advance, for example the human GRCh37/hg19 assembly (Human19.fasta) downloaded from the UCSC genome browser (<http://hgdownload.soe.ucsc.edu/goldenPath/hg19/database/>).

**Note:** to determine a profile of mitochondrial breaks, we align reads to the chrM genome separately and use the same downstream analysis steps below.

Build the genome index:

```
bwa index Human19.fasta -p HG19
```

Map the reads:

```
bwa mem -t 8 -M -R '@RG\tID:foo\tSM:bar\tLB:library1' HG19 read_1.clean.fq.1.trim
read_2.clean.fq.1.trim > read.sam
```

4. Filter the alignments by extracting only unique and proper mapped paired reads, and then extract the alignment records of reads 2 containing the coordinates of breaks:

```
perl 3_filter.sam.pl read.sam >read.proper.sam 2>log.3
```

5. Convert the SAM file to a BED file and then extract the very first position of each aligned record of reads 2:

```
sh 4_sam2bed.right.proper.sh read.proper.sam read.candidates.bed 2>log.4
```

6. Filter the alignments based on the fraction of T's in the 20-base genomic sequence 5' upstream of the reads 2 alignments:

```
perl 5_filtering.internal.polyA.pl read.candidates.bed Human19.fasta 20 0.4 >read.bed 3>log.5
```

7. Extract the coordinates of breaks:

```
cat read.bed | awk '$6 == "+" | awk '{print $1"\t"$2"\t"$2+1"\t"$4"\t"$5"\t-"}' >read.break.bed.minus
```

```
cat read.bed | awk '$6 == "-" | awk '{print $1"\t"$2"\t"$2+1"\t"$4"\t"$5"\t+'}' >read.break.bed.plus
```

```
cat read.break.bed.plus read.break.bed.minus | sort -k1,1 -k2n,2 >read.break.bed
```

```
rm -f read.break.bed.minus read.break.bed.plus
```

# Troubleshooting

## Troubleshooting

We noticed variation in MNase activity from lot to lot and also occurring with time of storage. Therefore, pre-testing to ensure fragmentation in the desired range can be helpful when using new lot or if they enzyme has not been used for some time.

## Time Taken

### Timing

From the time that samples are ready to be processed, the whole wet lab procedure takes around 2 days for 1-10 samples and 3 days for 11-24 samples. Analysis from raw reads to the coordinates of breaks takes around 1 and 2 days respectively for such numbers of samples with 1 Gb sequencing scale per sample and using the server provided in this protocol. However, the processing times would vary depending on the amount of sequencing being processed and the computer hardware used.

## Anticipated Results

### Anticipated Results

QC metrics of a library ready for sequencing are shown in the Fig. 2.

## References

### References

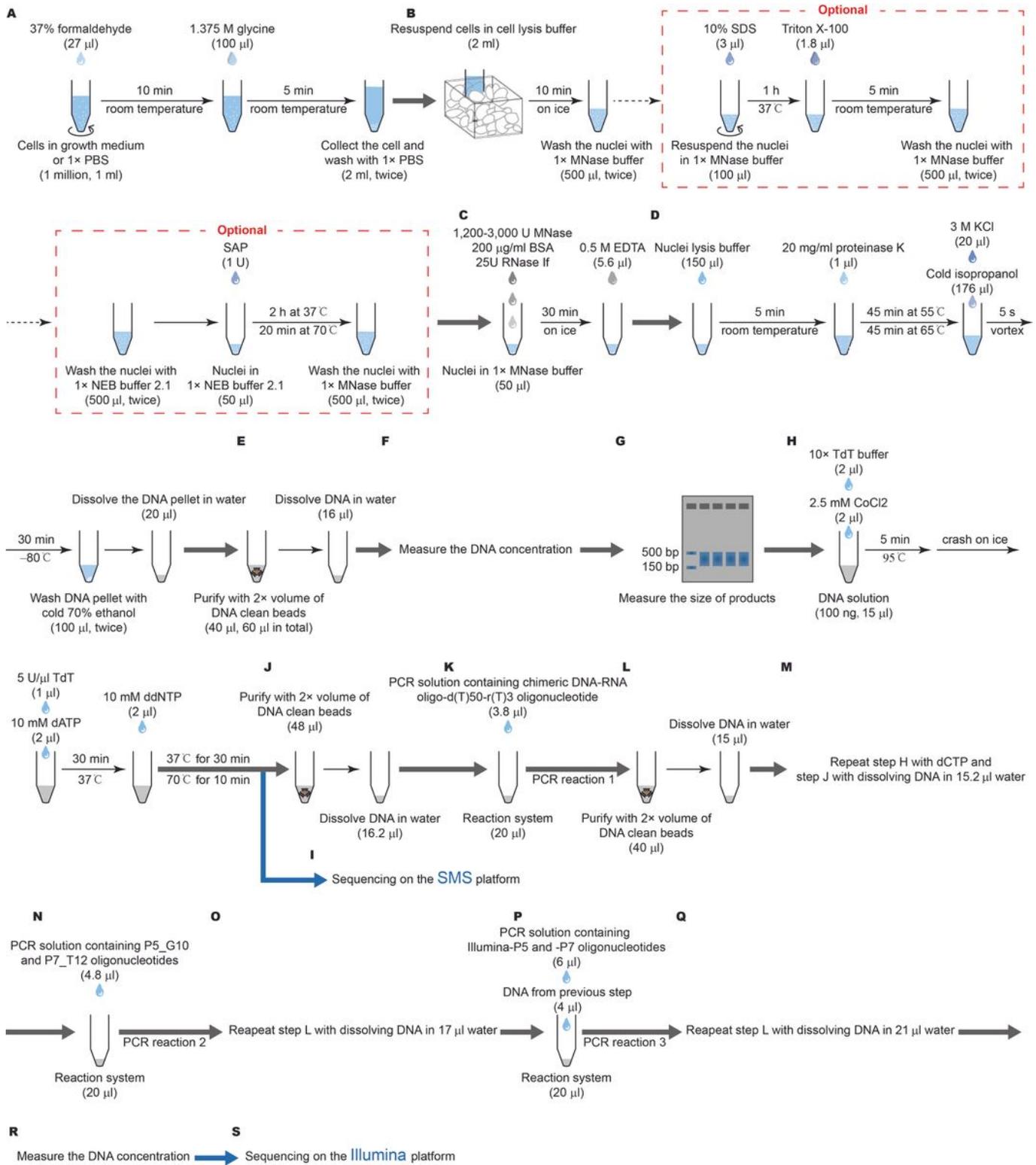
1. Hoeijmakers, J.H. DNA damage, aging, and cancer. *N Engl J Med* **361**, 1475-1485 (2009).
2. Jackson, S.P. & Bartek, J. The DNA-damage response in human biology and disease. *Nature* **461**, 1071-1078 (2009).
3. Caldecott, K.W. Single-strand break repair and genetic disease. *Nat Rev Genet* **9**, 619-631 (2008).
4. Kuzminov, A. Single-strand interruptions in replicating chromosomes cause double-strand breaks. *Proc Natl Acad Sci U S A* **98**, 8241-8246 (2001).
5. Kathe, S.D., Shen, G.P. & Wallace, S.S. Single-stranded breaks in DNA but not oxidative DNA base damages block transcriptional elongation by RNA polymerase II in HeLa cell nuclear extracts. *The Journal*

*of biological chemistry* **279**, 18511-18520 (2004).

6. Ljungman, M. & Zhang, F. Blockage of RNA polymerase as a possible trigger for u.v. light-induced apoptosis. *Oncogene* **13**, 823-831 (1996).
7. Ljungman, M., Zhang, F., Chen, F., Rainbow, A.J. & McKay, B.C. Inhibition of RNA polymerase II as a trigger for the p53 response. *Oncogene* **18**, 583-592 (1999).
8. Rulten, S.L. & Caldecott, K.W. DNA strand break repair and neurodegeneration. *DNA repair* **12**, 558-567 (2013).
9. Crosetto, N. et al. Nucleotide-resolution DNA double-strand break mapping by next-generation sequencing. *Nat Methods* **10**, 361-365 (2013).
10. Yan, W.X. et al. BLISS is a versatile and quantitative method for genome-wide profiling of DNA double-strand breaks. *Nature communications* **8**, 15058 (2017).
11. Vitelli, V. et al. Recent Advancements in DNA Damage-Transcription Crosstalk and High-Resolution Mapping of DNA Breaks. *Annual review of genomics and human genetics* **18**, 87-113 (2017).
12. Baranello, L. et al. DNA break mapping reveals topoisomerase II activity genome-wide. *International journal of molecular sciences* **15**, 13111-13122 (2014).
13. Hu, J., Adar, S., Selby, C.P., Lieb, J.D. & Sancar, A. Genome-wide analysis of human global and transcription-coupled excision repair of UV damage at single-nucleotide resolution. *Genes Dev* **29**, 948-960 (2015).

## Acknowledgements

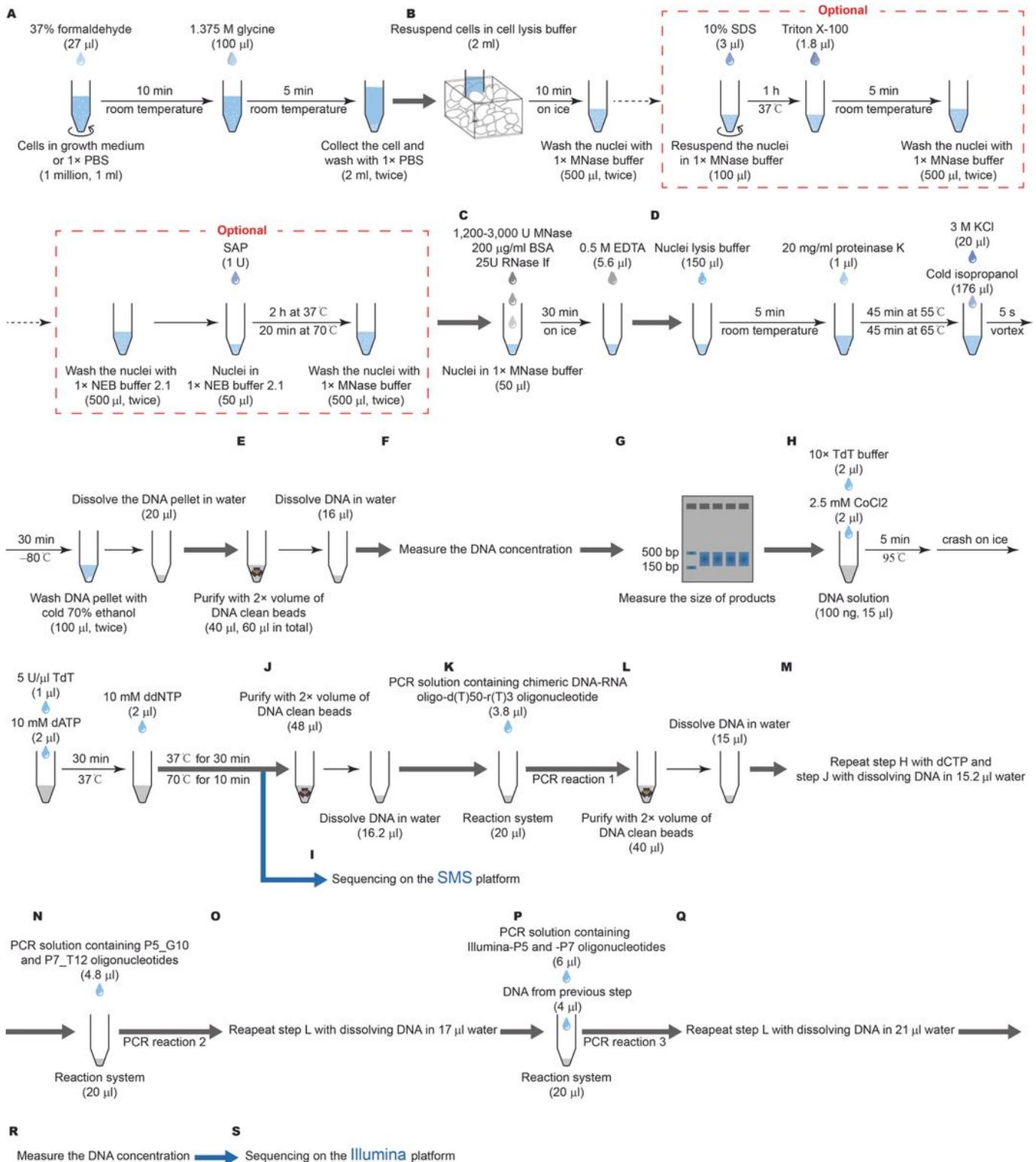
## Figures



**Figure 1**

Figure1. Flowchart of the SSiNGLe procedure. A. crosslinking; B. cell lysis; C. DNA fragmentation; D. DNA extraction; E. J. L. O. Q. DNA purification; F. R. concentration measuring; G. DNA fragment size measurement; H. polyA tailing; I. sequencing on the SMS platform; K. linear amplification; M. polyC tailing and DNA purification; N. first step PCR of the Illumina library construction with P5\_G10 and P7\_T12 oligonucleotides; P. second step PCR of the Illumina library construction with Illumina-P5 and -P7

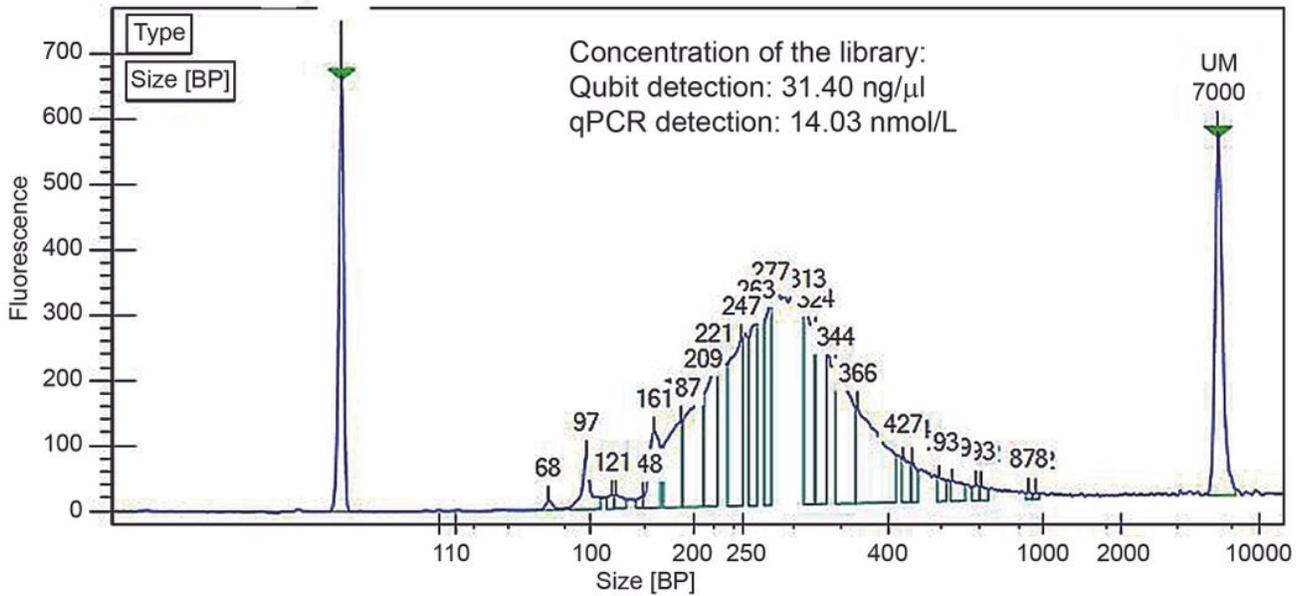
oligonucleotides; S. sequencing on the Illumina platform; Optional. SDS permeabilization; Optional. SAP digestion.



**Figure 1**

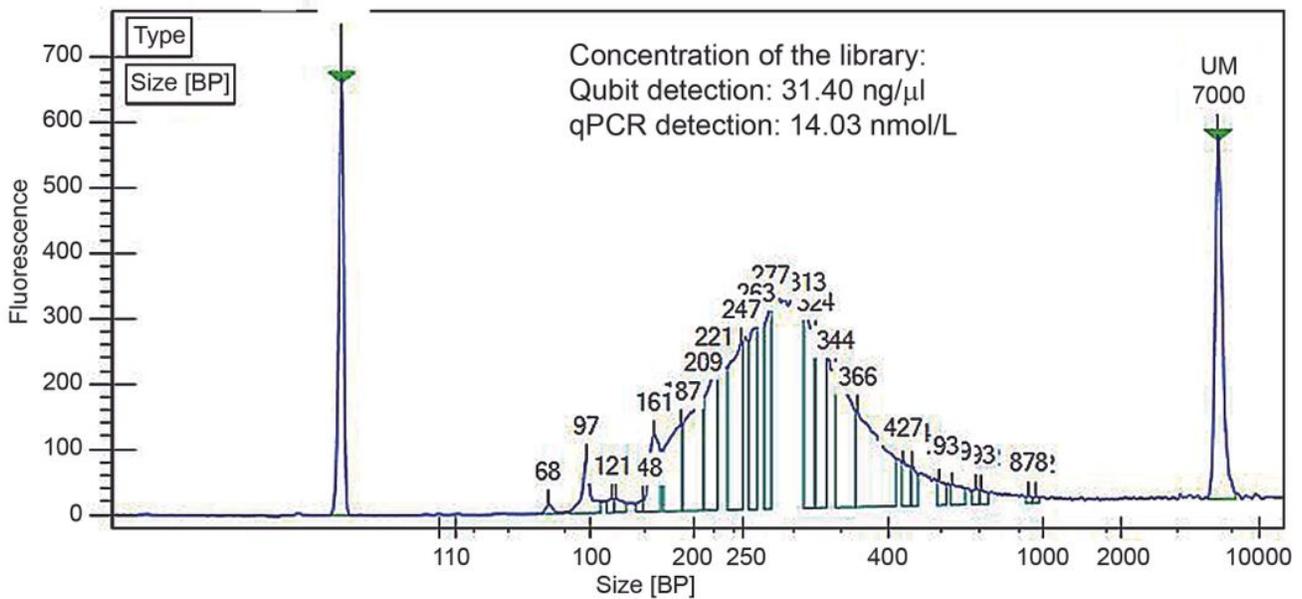
Figure1. Flowchart of the SSiNGLe procedure. A. crosslinking; B. cell lysis; C. DNA fragmentation; D. DNA extraction; E. J. L. O. Q. DNA purification; F. R. concentration measuring; G. DNA fragment size measurement; H. polyA tailing; I. sequencing on the SMS platform; K. linear amplification; M. polyC tailing

and DNA purification; N. first step PCR of the Illumina library construction with P5\_G10 and P7\_T12 oligonucleotides; P. second step PCR of the Illumina library construction with Illumina-P5 and -P7 oligonucleotides; S. sequencing on the Illumina platform; Optional. SDS permeabilization; Optional. SAP digestion.



**Figure 2**

Figure 2. QC metrics of a SSiNGLe-ILM library. The fragment size (base pairs, bp, X-axis) distribution was analyzed by LabChip GX Touch. LM and UM – lower and upper size standards. The QC was performed by the Novogene Corporation.



**Figure 2**

Figure 2. QC metrics of a SSiNGLe-ILM library. The fragment size (base pairs, bp, X-axis) distribution was analyzed by LabChip GX Touch. LM and UM – lower and upper size standards. The QC was performed by the Novogene Corporation.

the Novogene Corporation.