bio-protocol

Published: Mar 20, 2022

Single-cell Damagenome Profiling by Linear Copying and Splitting based Whole Genome Amplification (LCS-WGA)

Yichi Niu[1, 2, #], Qiangyuan Zhu[1, #] and Chenghang Zong[1, 2, 3, 4, 5, *]

[1]Department of Molecular and Human Genetics, Baylor College of Medicine, TX, USA
[2]Genetics & Genomics Program, Baylor College of Medicine, TX, USA
[3]Cancer and Cell Biology Program, Baylor College of Medicine, TX, USA
[4]Dan L Duncan Comprehensive Cancer Center, Baylor College of Medicine, TX, USA
[5]McNair Medical Institute, Baylor College of Medicine, TX, USA
[#]Contributed equally to this work
*For correspondence: czong@bcm.edu

Abstract

Spontaneous DNA damage frequently occurs on the human genome, and it could alter gene expression by inducing mutagenesis or epigenetic changes. Therefore, it is highly desired to profile DNA damage distribution on the human genome and identify the genes that are prone to DNA damage. Here, we present a novel single-cell whole-genome amplification method which employs linear-copying followed by a split-amplification scheme, to efficiently remove amplification errors and achieve accurate detection of DNA damage in individual cells. In comparison to previous methods that measure DNA damage, our method uses a next-generation sequencing platform to detect misincorporated bases derived from spontaneous DNA damage with single-cell resolution.

Keywords: DNA damage, Whole genome amplification (WGA), Single cell whole genome amplification (scWGA), Linear amplification, Damagenome, Genome vulnerability, Vulnerable genes, High-damage genes

This protocol was validated in: Sci Adv (2021), DOI: 10.1126/sciadv.abf3329

Background

Spontaneous DNA damage, which is usually caused by exposure to environmental chemicals or the free radicals generated from metabolic processes, frequently occurs in our genome. If the damage is not repaired correctly, it could lead to permanent change in the genomic sequence and sabotage gene functions (Ames *et al.*, 1993; De Bont and van Larebeke, 2004; Tudek *et al.*, 2010; Dizdaroglu *et al.*, 2015). Besides mutagenesis, DNA damage could also perturb the local epigenetic stability and, as a result, alter gene expression patterns (Dabin *et al.*, 2016). The non-uniform distribution of DNA damage on our genome could result from various levels of vulnerability to DNA damage in different genes. With high-damage genes being the potential Achilles' heels of our genome, the alteration in expression of these genes could directly contribute to the development of human multifactorial diseases. Therefore, the accurate examination of the distribution of DNA damage and identification of high-damage genes are greatly desirable.

So far, high-performance liquid chromatography (HPLC)-based methods, gas chromatography/mass spectrometry (GC/MS)-based methods, and chromatin immunoprecipitation (ChIP)-based methods have been the primary approaches in measuring DNA damage. HPLC and GC/MS based methods are mainly used to measure the global levels of various DNA modifications (Wagner *et al.*, 1992; De Bont and van Larebeke, 2004; Galashevskaya *et al.*, 2013). However, since HPLC and GC/MS-based approaches require the DNA samples to be hydrolyzed into mononucleotides, the sequence context information cannot be recovered. Furthermore, the extensive treatment at the DNA hydrolysis step could introduce a significant amount of artificial DNA damage, which leads to large technical variations. On the other hand, ChIP-based methods could provide a sequence-based measurement of DNA damage by using antibodies to pull down the DNA fragments with damaged DNA (Yoshihara *et al.*, 2014; Garcia-Nieto *et al.*, 2017; Ding *et al.*, 2017; Poetsch *et al.*, 2018; Wu *et al.*, 2018; Amente *et al.*, 2019). However, such methods largely depend on the availability of antibodies that could recognize specific types of DNA damage, and do not have single-base resolution. The artificial damage induced during the DNA shearing step could lead to the overestimation of DNA damage levels.

Recently, we have reported a new single-cell whole-genome amplification (WGA) method, called linear copying and splitting based whole genome amplification (LCS-WGA), which can allow us to achieve genome-wide profiling of DNA damage (Zhu *et al.*, 2021), since DNA damage can induce base misincorporation during amplification (Purmal *et al.*, 1994, 1998; Gehrke *et al.*, 2013). These misincorporated bases can then be directly identified in the sequencing data. We refer to damage-associated single-nucleotide variants as damSNVs. The main advantage of the single-cell WGA-based approach is that, without the requirement of either DNA hydrolysis or shearing, the artificial DNA damage is significantly reduced. Meanwhile, the genomic context of DNA damage can be effectively detected.

The only major technical challenge to single-cell WGA is the large number of amplification errors, which could limit the accuracy in single-cell WGA based approaches (Zong *et al.*, 2012; Chen *et al.*, 2017; Luquette *et al.*, 2019; Bohrson *et al.*, 2019; Xing *et al.*, 2021). In LCS-WGA, we have overcome this technical hurdle by employing linear copying and then a split-amplification scheme. Briefly, we first perform three cycles of low-temperature annealing and extension with temperature ramping to generate multiple copies of the genomic DNA of the single cell as the preamplification step (Figure 1A). In this process, the DNA copies directly generated from the original genomic DNA are termed as semiamplicons, and the DNA products that are copied from the semiamplicons are termed full amplicons. It is worth pointing out that the semiamplicons are linearly produced original genomic DNA of the single cell. The full amplicons are nonlinearly copied from semiamplicons. To quench the nonlinearly produced full amplicons from amplification in the downstream multiple displacement amplification (MDA) reaction, we perform a double-stranded conversion, to convert all the full amplicons into double-stranded DNA, while the semiamplicons are kept as single-stranded DNA. Next, we split the preamplification products into three tubes, to perform independent MDA reactions. As MDA only amplifies single-stranded DNA, the double-stranded full amplicons cannot be amplified and, as a result, only linearly copied DNA products and original genomic DNA are amplified in these MDA reactions. Importantly, since the amplification errors in semiamplicons are independently produced during the preamplification step, we can cross-compare the sequencing result of different splits, to effectively filter out amplification errors. Here, we employ a two-split detection criterion, by which we only keep the variants detected in at least two split samples in the variant calling step (Figure 1B). Based on amplification errors in single

split reaction, we estimate that the false positive rate due to amplification errors is 8.3×10^{-10} per base when we apply two-split detection criterion (Zhu *et al.*, 2021).

With this new method, we have successfully examined the DNA damage in human neurons and identified the association between the non-uniform distribution of DNA damage and 3D chromatin structure, as well as the connection between high-damage genes and the altered gene expressions in Alzheimer's disease and autism. This method can be readily applied to different biological contexts, to quantify damage levels in cells and, more importantly, identify high-damage genes. We believe that the characterization of the damagenome in different cell types from different tissues could shed new light on complex human diseases.

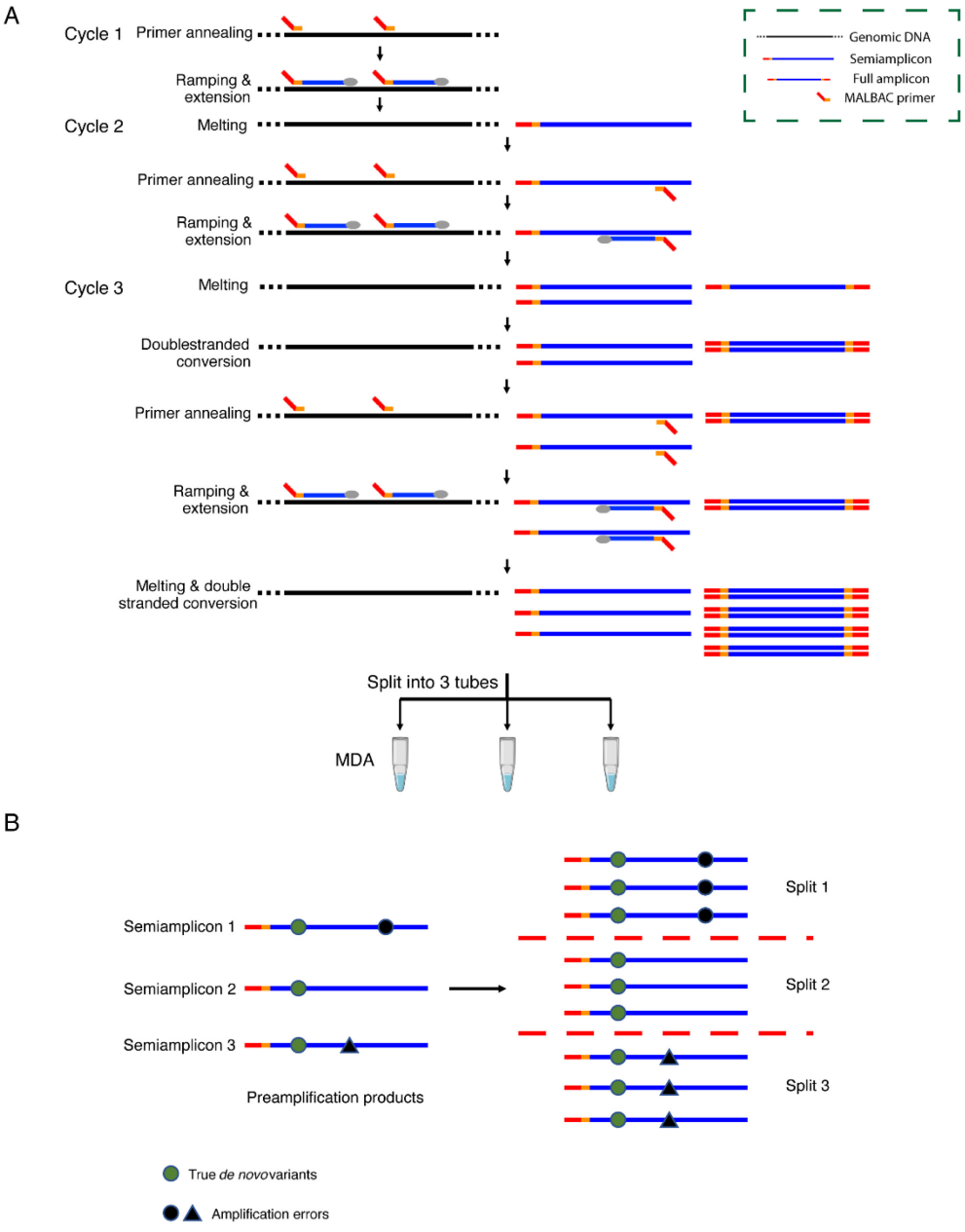


Figure 1. Scheme of LCS-WGA.
(A) The scheme of pre-amplification and split-MDA. (B) The scheme to distinguish true *de novo* variants from amplification errors.

Materials and Reagents

1. VWR Signature™ micro aerosol filter pipette tips, Volume: 0.1–10 µL (VWR, catalog number: 53509-138)
2. 0.2 mL PCR tubes with Flat Cap (Axygen, catalog number: PCR-02-L-C)
3. 1.7 mL microtube (Olympus plastics, catalog number: 24-282LR)

Notes:
a. Pipette tips used in this protocol should be low-retention, RNase-free, DNase-free, with aerosol filter.
b. PCR tubes and microtubes used in this protocol should be low-retention, RNase-free, and DNase-free.

4. Potassium hydroxide solution, 8 N (Sigma-Aldrich, catalog number: P4494)
5. DTT, 5 g (Sigma-Aldrich, catalog number: GE17-1318-02)
6. Ethylenediaminetetraacetic acid (EDTA) solution, 0.5 M in H_2O (Sigma-Aldrich, catalog number: E7889)
7. Tris-HCl, pH = 7.5, 1 M (Invitrogen, catalog number: 15567-027)
8. Hydrochloric Acid Solution, 2 N (Fisher Scientific, catalog number: SA431)
9. Invitrogen™ UltraPure™ DEPC-Treated Water (Invitrogen, catalog number: 750024)
10. ThermoPol Reaction Buffer, 10× (New England BioLabs, catalog number: B9004S)
11. Uracil-DNA Glycosylase, 5,000 units/mL (New England BioLabs, catalog number: M0280S)
12. Deoxynucleotide (dNTP) Solution Mix (10 mM) (New England BioLabs, catalog number: N0447S)
13. *Bst* DNA Polymerase, Large Fragment (New England BioLabs, catalog number: M0275S)
14. iTaq Universal SYBR Green Supermix (Bio-Rad, catalog number: 1725124)
15. phi29 DNA Polymerase, supplemented with phi29 DNA polymerase reaction buffer (10×) and BSA, Molecular Biology Grade (New England BioLabs, catalog number: M0269S)
16. TWEEN® 20, for molecular biology, viscous liquid (Sigma-Aldrich, catalog number: P9416)
17. Agencourt AMPure XP (Beckman Coulter, catalog number: A63880)
18. Illumina Tagment DNA TDE1 Enzyme and Buffer Kit (Illumina, catalog number: 20034197)
19. Magnesium acetate ($Mg(Ac)_2$) solution, BioUltra, for molecular biology (Sigma-Aldrich, catalog number: 63052)
20. NEBNext® Ultra™ II Q5® Master Mix (New England BioLabs, catalog number: M0544S)
21. Agilent D1000 ScreenTape (Agilent, catalog number: 5067-5582)
22. Agilent D1000 Reagents (Agilent, catalog number: 5067-5583)
23. Qubit™ dsDNA HS Assay Kit (Invitrogen, catalog number: Q32851)
24. Phosphate-Buffered Saline (PBS), 1× without calcium and magnesium (Corning, catalog number: 21-040-CV)
25. DNA oligos (Ordered from Integrated DNA Technologies)
26. MCF 10A cell line (ATCC CRL-10317)
27. Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F-12), HEPES (Gibco, catalog number: 11330032)
28. Horse serum (Gibco, catalog number: 16050122)
29. Hydrocortisone (Sigma-Aldrich, catalog number: H0888)
30. Cholera toxin (Sigma-Aldrich, catalog number: C8052)
31. Insulin (Sigma-Aldrich, catalog number: I1882)
32. Recombinant human epidermal growth factor (ThermoFisher Scientific, catalog number: PHG0311)
33. Frozen human brain tissue (Obtained from NIH NeuroBioBank: https://neurobiobank.nih.gov/)
34. Lysis buffer (400 mM KOH, 100 mM DTT, 2 mM EDTA) (see Recipes)
35. Stop solution (400 mM HCl, 600 mM Tris-HCl) (see Recipes)
36. UDG reaction mix (see Recipes)
37. Pre-amplification primer mix (see Recipes)
38. MDA reaction mix (see Recipes)
39. Library amplification mix (see Recipes)
40. UV-treated UltraPure H_2O (see Recipes)

41. Growth medium for the MCF 10A cell line (see Recipes)

Equipment

1. S1000™ Thermal Cycler (Bio-Rad)
2. CFX96 Touch Real-Time PCR Detection System (Bio-Rad)
3. Qubit® 2.0 fluorometer (ThermoFisher)
4. 4200 TapeStation System (Agilent)
5. CL-1000 Ultraviolet Crosslinker
6. Microscope Slides (VWR VistaVision™ HistoBond®, catalog number: 16004-406)
7. Motorized inverted research microscope (Olympus IX81)
8. VWR® Mini Centrifuge
9. GILSON® Pipetman L P2L, 0.2–2 µL, Metal Ejector (for small volume pipetting)
10. Bunsen burner (natural gas)
11. Borosilicate glass with filament (Sutter Instrument, catalog number: BF100-58-15)
12. Aspirator tube assemblies for calibrated microcapillary pipettes (Sigma-Aldrich, catalog number: A5177-5EA)
13. High-performance computing cluster (Memory: 256 GB, Nodes: 32, Iris 2281T nodes from or similar nodes from Penguin Computing)

Software

1. Burrows-Wheeler Aligner (BWA, version: 0.7.12-r1039, http://bio-bwa.sourceforge.net/, free)
2. Genome Analysis Toolkit (GATK, version: 3.8, https://gatk.broadinstitute.org/hc/en-us, free)
3. Samtools (v1.7, http://www.htslib.org/, free)
4. Bedtools (v2.17.0, https://bedtools.readthedocs.io/en/latest/, free)

Procedure

A. Preparation of a single cell suspension from cell culture

1. Trypsinize MCF10A cells by adding 1 mL of 0.05% trypsin to a T25 flask. Incubate in a 5% CO_2 humidified incubator at 37°C for 10 min.
2. Add 1 mL of growth medium (see Recipes) to neutralize trypsin and transfer the cells to a 15 mL centrifuge tube.
3. Centrifuge the cells at 300 × *g* and room temperature (RT) for 3 min.
4. Carefully aspirate the supernatant and resuspend the cells in 1 mL of 1× PBS.

Note: Here, we take MCF10A cell as an example. These steps can be adjusted based on different cell lines.

B. Hydrogen peroxide (H_2O_2) treatment of MCF10A cells

Note: H_2O_2-treated MCF 10A cells are used as a positive control, as described in Zhu et al. (2021). Compared to the non-treated cells, we detected a significantly higher number of damage-associated single-nucleotide variants in H_2O_2-treated MCF 10A cells, which demonstrates the ability of our method to detect DNA damage.

1. Trypsinize MCF10A cells and resuspend them in 1× PBS, as described in procedure A.
2. Centrifuge the cells at 300 × *g* for 3 min. Discard the supernatant.
3. Resuspend the cells with 2 mL of freshly prepared H_2O_2 (300 μM in PBS).
4. Incubate the cells on ice for 5 min.
5. Dilute the cells 1:100 with cold 1× PBS. The cells then directly proceed to picking and lysis.

C. Preparation of neuronal nuclei from frozen human brain tissue

The preparation of single nuclei suspension and staining with anti-NeuN antibody are performed according to the protocol in Krishnaswami *et al.* (2016). After staining, single nuclei are resuspended in 500 μL of cold 1× PBS.

Note: Here, we take the processing of frozen human brain tissue as an example. This step can be adjusted based on different sources of tissue samples.

D. Single cell isolation by mouth pipetting

1. Preparation of glass capillaries: Flame the center until the glass pipette softens. Then, remove the glass pipette from the flame and quickly pull to generate the capillaries, usually ~3-fold longer than the original glass pipette length. After pulling, bend the glass pipette until the thin middle part breaks.
2. Prepare the aspirator tube assembly using the glass capillary (Figure 2).

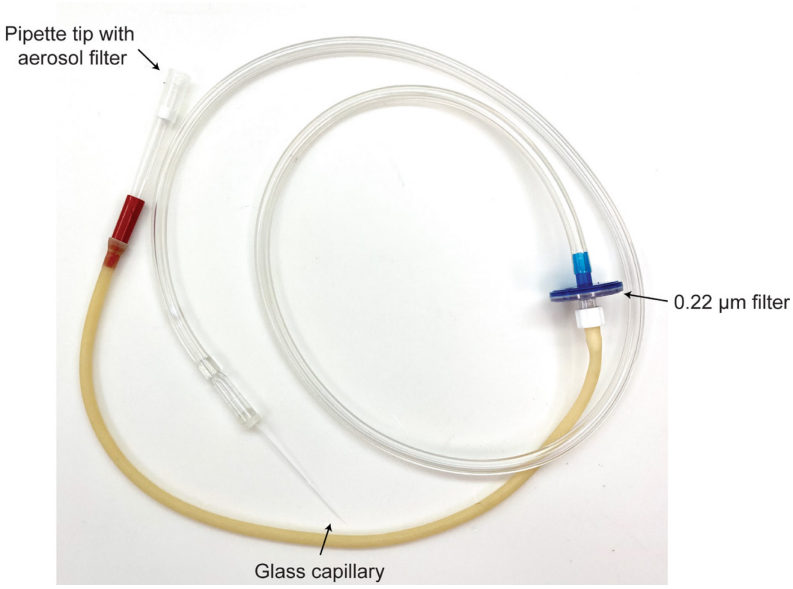
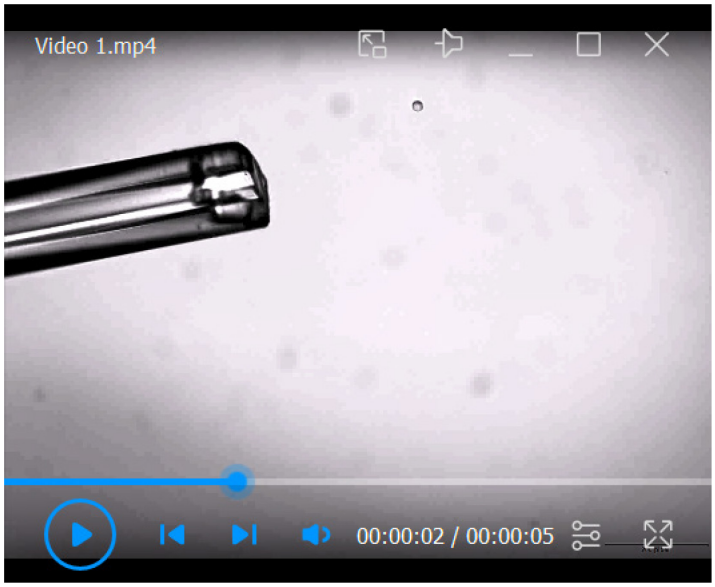


Figure 2. The aspirator tube assembly with the glass capillary for mouth pipetting.

3. Prepare the single cell suspension with 1× PBS, as described above.
4. Drip a few PBS drops (3 μL per drop) on the microscope slide (4 drops per slide) and place the slide under the microscope (10× field).
5. Add the cells to one PBS drop and perform a serial dilution, by transferring the cells from one drop to another via mouth pipetting.
6. When only one cell is in the field, pick the cell and deposit it into the PCR tube with 2 μL of lysis buffer (see Recipe 1) via mouth pipetting (Video 1).



Video 1. Single cell picking under microscopy.

E. (Alternative) Single cell isolation by fluorescence-activated cell sorting (FACS)

1. Prepare the single cell suspension with 1× PBS.
2. (Optional) Cells can be stained with Hoechst 33342, to facilitate gating.
3. Set up the FACS machine in plate mode and single-cell purity.
4. Sort the single cells into PCR tubes with 2 µL of lysis buffer (the PCR tubes can be placed on a full skirted 96-well PCR plate, or a regular rack for PCR tubes).

Note: As downstream reactions require multiple steps of operation on each single cell, we recommend sorting single cells into individual PCR tubes.

F. Cell lysis (estimated time: 90 min)

1. Lyse cells in 2 µL of lysis buffer at 30°C for 90 min.
2. Add 2 µL of stop buffer (see Recipe 2) to each cell, to neutralize the lysis buffer.
3. After briefly spinning down, the single-cell lysate can be stored at -80°C for future use.
   Safe stopping point: The single-cell lysate can be stored at -80°C for up to 3 months.

G. UDG treatment (estimated time: 30 min)

Note: This step is designed to remove artificial DNA damage (cytosine deamination) introduced during the cell lysis step.

1. Add 5.5 µL of UDG reaction mix (see Recipe 3) to the single-cell lysate.
2. Incubate at 37°C for 30 min in a thermal cycler.

H. Pre-amplification (estimated time: ~2 h)

1. Add 0.3 µL of deoxynucleotide (dNTP) solution mix and 0.38 µL of pre-amplification primer mix (see

Recipe 4) to each sample. Gently flick to mix and briefly spin down.

2. Perform the following reaction on a PCR machine:

   First cycle:
   a. Heat the samples at 94°C for 50 s.
   b. Pause the PCR machine at 65°C and add 2.8 U Bst DNA polymerase (large fragment). Adding the enzyme will take 10–20 s.
   c. Quench the reaction by transferring the samples to ice and then keep them on ice for at least 30 s. After the PCR machine cools down to 10°C, transfer the samples back to the PCR block.
   d. Resume the PCR program to finish the first round of annealing and extension: 10°C for 40 s, 20°C for 40 s, 25°C for 40 s, 30°C for 40 s, 40°C for 1 min, 45°C for 1 min, 55°C for 40 s, 65°C for 4 min.

   Second cycle:
   a. Heat the sample at 94°C for 20 s.
   e. Pause the PCR machine at 65°C and add 2.8 U Bst DNA polymerase (large fragment) and 0.25 µL of 10 µM GAT27 primer to the reaction. Adding the two reagents will take 20–30 s.
   b. Perform double-strand conversion with 8 cycles of 63°C for 15 s and 65°C for 20 s, followed by 65°C for 1 min.
   c. Quench the reaction by transferring the samples to ice and keep them on ice for at least 30 s. After the PCR machine cools down to 10°C, transfer the samples back to the PCR block.

   Note: The efficient primer hybridization requires fast cooling of the reaction to a lower temperature. The ramping process on a regular thermal cycler is slower than directly placing the sample tubes on ice, which leads to less efficient primer hybridization.

   d. Resume the program to finish the second round of annealing and extension: 10°C for 40 s, 20°C for 40 s, 25°C for 40 s, 30°C for 40 s, 40°C for 1 min, 45 °C for 1 min, 55°C for 40 s and 65°C for 4 min 30 s.

   Third cycle:
   a. Heat the samples at 94°C for 20 s.
   b. Pause the PCR machine at 65°C and add 2.8 U Bst DNA polymerase (large fragment) to the reaction. Adding the enzyme will take 10–20 s.
   c. Perform double-strand conversion with 8 cycles of 63°C for 15 s and 65°C for 20 s, followed by 65°C for 1 min.
   d. Quench the reaction by transferring the samples to ice and keep them on ice for at least 30 s. After the PCR machine cools down to 10°C, transfer the samples back to the PCR block.
   e. Resume the program to finish the last round of annealing and extension: 10°C for 40 s, 20°C for 40 s, 25°C for 40 s, 30°C for 40 s, 40°C for 1 min, 45°C for 1 min, 55°C for 40 s and 65°C for 5 min.

   Extensive double-strand conversion:
   a. Pause the PCR machine at 78°C and add 0.2 µL of 10 µM GAT27 primer and 3.8 µL of Ultra-pure H_2O.
   b. Heat the sample at 94°C for 20 s.
   c. Pause the PCR machine at 65°C and add 3.6 U Bst DNA polymerase (large fragment).
   d. Resume the program and perform double-strand conversion with 30 cycles of 63°C for 15 s and 65°C for 20 s, followed by 65°C for 2 min. Next, inactivate the Bst DNA polymerase (large fragment) by incubating the sample at 72°C for 25 min.
      The preamplification products can be temporarily stored at -20°C for a few days.

3. Yield test (estimated time: ~2 h)

For each sample, prepare the qPCR reaction mix (5 µL of iTaq Universal SYBR Green Supermix, 0.5 µL of 10 µM GAT27 primer, and 4 µL of H_2O), and add 0.5 µL of preamplification products for yield test. The qPCR program is as follows: initial activation at 94°C for 2 min, followed by 28 cycles of denaturation at 94°C for 20 s, primer annealing at 60°C for 25 s, and primer extension at 72°C for 2 min 20 s. Discard the tubes without single cells (Figure 3).

Safe stopping point: The pre-amplification products can be stored at -20°C for up to 3 days.

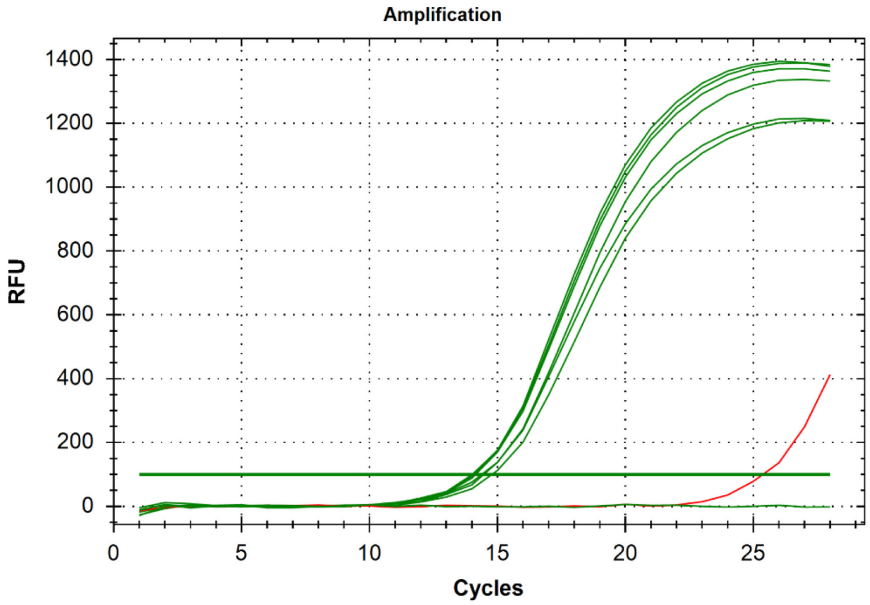


Figure 3. The example result of the yield test of single human neurons.
The negative control (tube without single cell) is colored in red.

I. Split amplification by multiple displacement amplification (MDA) with frequent pipetting mixing (estimated time: ~2 h)

Note: The limited reaction time and the inclusion of pipetting mixing steps help to reduce the amplification bias and increase the robustness of MDA reactions.

1. Briefly vortex the preamplification product and split it into three tubes (~4.7 µL per tube) to perform three independent MDA reactions.
2. Add 24.3 µL of MDA reaction mix (see Recipe 5) to each reaction.
3. Add 10 U phi29 DNA polymerase to each reaction on ice, followed by incubation on ice for 3 min.
4. After the incubation at 4°C for 1 min on a PCR block, start the MDA reaction by incubating the reactions at 30°C for 10 min (ramping at 0.2°C/s).
5. Pause at 30°C and mix each reaction by quickly pipetting 30–40 times with a 10 µL pipettor.
6. Incubate the reactions on ice for 1 min, and cool down the PCR machine to 4°C.
7. After transferring the reactions back to the PCR machine, incubate them at 30°C for 10 min (ramping at 0.2°C/s).
8. Pause at 30°C and mix each reaction by quickly pipetting 30–40 times with a 10 µL pipettor.
9. Incubate the reactions on ice for 1 min and cool down the PCR machine to 4°C.
10. After transferring the reactions back to the PCR machine, incubate them at 30°C for 5 min (ramping at 0.2°C/s).

11. Inactivate phi29 DNA polymerase at 65°C for 10 min.
12. Purify the MDA products:
    a. Add 30 µL (1× volume) of AMPure XP beads and mix well by pipetting up and down 15 times.
    b. Incubate at room temperature for 7 min and then place the tube on a magnetic stand for 0.2 mL tubes.
    c. After the solution becomes clear, discard the supernatant without disturbing the beads.
    d. Add 100 µL of 80% ethanol to the tube, while keeping it on the magnetic stand. After incubating at RT for 30 s, discard the supernatant without disturbing the beads.
    e. Repeat the last step.
    f. Remove the residual ethanol.
    g. Keep the tube on the magnetic stand with its lid open and air dry the beads until there is no ethanol left (which usually takes 5–10 min).
    h. Remove the tube from the magnetic stand.
    i. Add 6.5 µL of UV-treated ultrapure H_2O for elution. As BSA makes beads very sticky, avoid any pipetting after adding the water. Instead, mix the beads with water by frequently flicking the tube.
    j. After mixing well, incubate at room temperature for 5 minutes.
    k. Place the tube back on the magnetic stand.
    l. After the solution becomes clear, transfer the supernatant to a new PCR tube.

    Safe stopping point: The purified MDA products can be stored at -20°C for up to 7 days.

J. Library construction (estimated time: ~2 h)

1. Add 6.3 µL of 2× Illumina Tagment DNA buffer and 0.2 µL of 10-fold diluted Illumina TDE1 to each reaction. Incubate at 55°C for 2 min to tagment the DNA.
2. To stop the tagmentation, pause at 4°C and add 1.5 µL of 0.2 M EDTA to each reaction. Then, incubate at 50°C for 30 min.
3. Add 23.3 µL of library amplification mix (see Recipes) to each reaction, and amplify the library with the following program: incubation at 72°C for 5 min, initial activation at 98°C for 30 s, 12–14 cycles of denaturation at 98°C for 10 s, primer annealing at 58°C for 30 s, and primer extension at 72°C for 1 min, followed by a final extension at 72°C for 5 min.
4. Purify the amplified library with AMPure XP beads:
    a. Add 30.2 µL (0.8× volume) of AMPure XP beads, and mix well by pipetting up and down 15 times.
    b. Incubate at room temperature for 5 min, and then place the tube on a magnetic stand for 0.2 mL tubes.
    c. After the solution becomes clear, discard the supernatant.
    d. Add 100 µL of 80% ethanol, while keeping the tube on the magnetic stand. Incubate at RT for 30 s, and then discard the supernatant.
    e. Repeat the last step.
    f. Remove the residual ethanol.
    g. Keep the tube on the magnetic stand with its lid open. Air dry the beads until there is no ethanol left (which usually takes 5–10 min).
    h. Remove the tube from the magnetic stand.
    i. Add 20 µL of UV-treated ultrapure H_2O. Mix well by pipetting up and down 15 times.
    j. Incubate at RT for 5 min.
    k. Place the tube back on the magnetic stand.
    l. After the solution becomes clear, transfer the supernatant to a new PCR tube.
5. Measure the library concentration using the Qubit™ dsDNA HS Assay Kit. Briefly, mix 1 µL of library DNA, 1 µL of Qubit dsDNA HS Reagent, and 198 µL of Qubit dsDNA HS Buffer in a clean Qubit assay tube. Incubate at room temperature for 2 min, and then measure the concentration in a Qubit 2.0 fluorometer. Usually, the concentration ranges from 4.5 ng/µL to 6 ng/µL, which corresponds to 90–120 ng in total.
6. Measure the library size distribution by using TapeStation. Briefly, mix 1 µL of library DNA with 3 µL of Agilent D1000 Sample Buffer in a 0.2 mL tube of an 8-Tube PCR strip without caps. Run the sample on a Agilent D1000 ScreenTape on the 4200 TapeStation System (Figure 4A).

Note: Library size after size selection is expected to be between 300–700 bp.

K. Loci test (estimated time: ~1 h)

We randomly chose 6 loci (see Table 1) to verify an even amplification using qPCR (Figure 4B–C). For each locus, 5 ng of library DNA is used as the template. The qPCR program is as follows: initial activation with 94°C for 2 min, 35 cycles of denaturation at 94°C for 12 s, primer annealing and extension at 58°C for 25 s. The cells with no more than two loci drop out in each split reaction proceed for paired-end (150 bp × 2) sequencing.

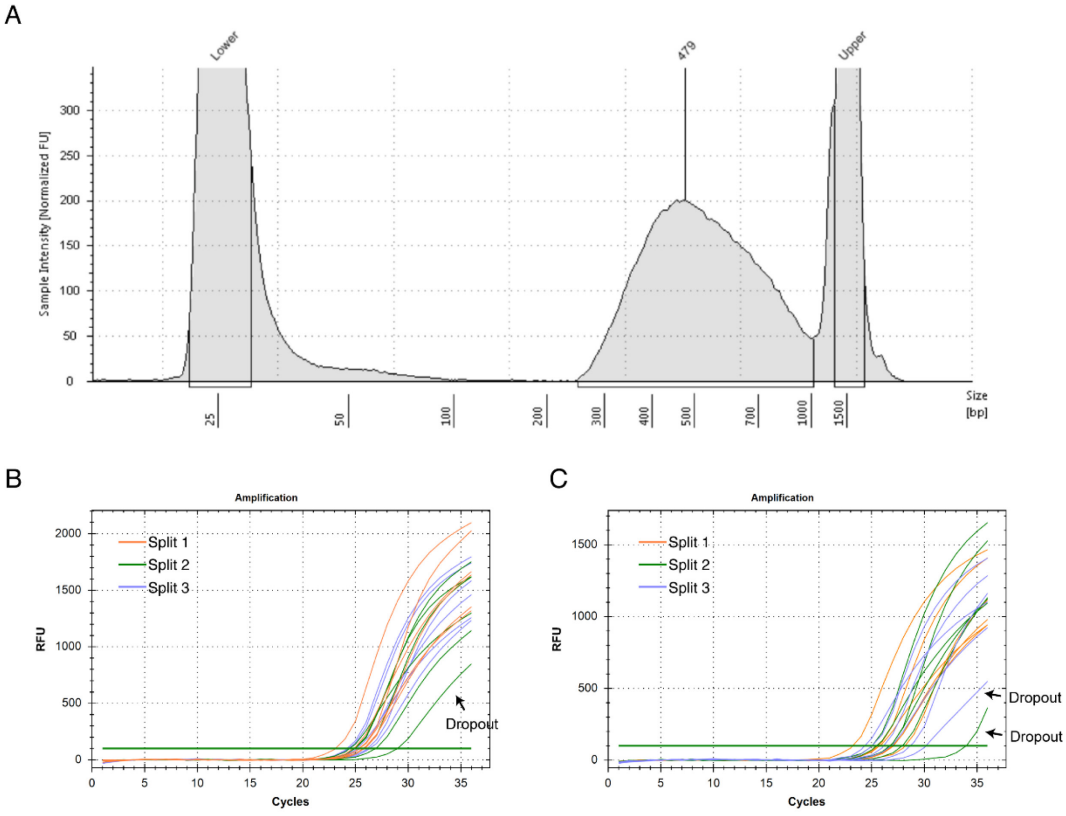


Figure 4. Quality control of the sequencing libraries.
(A) The size distribution of the sequencing library generated from a single human neuron by LCS-WGA. (B–C) The loci tests on the sequencing libraries generated from single human neurons. The example result in panel B represents the average performance of LCS-WGA on single human cells. The result in panel C shows the cell with the most biased amplification in our reported dataset.

L. Sequencing

We recommend sequencing on a HiSeq X Ten or NovaSeq platforms with 400 million paired-end (150 bp × 2) reads per single cell (three libraries).

Table 1. List of DNA oligos.

Note: All DNA oligos are dissolved in UV-treated UltraPure DEPC-treated H_2O.

Oligo names	Sequences
GAT27-5N-3T	GTG AGT GAT GGT TGA GGA TGT GTG GAG NNNNN TTT
GAT27-5N-3G	GTG AGT GAT GGT TGA GGA TGT GTG GAG NNNNN GGG
GAT27	GTG AGT GAT GGT TGA GGA TGT GTG GAG
Random hexamer	NNNN*N*N (*:phosphorothioate bond)
Loci_human_chr1_F	AGGAAAGGCATACTGGAGGGACAT
Loci_human_chr1_R	TTAGGGATGGCACCACACTCTTGA
Loci_human_chr2_F	TCCCAGAGAAGCATCCTCCATGTT
Loci_human_chr2_R	CACCACACTGCCTCAAATGTTGCT
Loci_human_chr4_F	ATGGGCAAATCCAGAAGAGTCCAG
Loci_human_chr4_R	CCATTCACTTCCTTGGAAAGGTAGCC
Loci_human_chr5_F	AATAGCGTGCAGTTCTGGGTAGCA
Loci_human_chr5_R	TTCACATCCTGGGAGGAACAGCAT
Loci_human_chr6_F	TGAATGCCAGGGTGAGACCTTTGA
Loci_human_chr6_R	TGTTCATTATCCCACGCCAGGACT
Loci_human_chr7_F	ACCAAAGGAAAGCCAGCCAGTCTA
Loci_human_chr7_R	ACTCCACAGCTCCCAAGCATACAA

Note: 5N represents 5 random bases. The 3T or 3G design allows the efficient priming at the 3' end, which is critical for initiating extension.

Data analysis

1. Map the sequencing reads to GRCh37 (hg19) by BWA-MEM with default settings. This step is performed on each of three split libraries independently.
2. Call variants by using GATK 3.8 (HaplotypeCaller) with the BAM files from all three split libraries as the input. Duplicate reads are kept by using the *-drf* DuplicateRead option. The reads with no stored bases are filtered out with the *-filterNoBases* option. Discard the reads with mapping quality score lower than 50 with the *-mmq* option, as well as the reads with base quality lower than 20, with the *-mbq* option. The ploidy number (*-ploidy*) is set as 10, for detecting variants at low allele frequency. The output variants include both germline variants and *de novo* variants.
3. Apply a python script to filter out the variants that are only detected in one split library (Figure 5A), the majority of which are the false positives caused by amplification errors.
4. Identify *de novo* variants with the following criteria:
   a. The variant is not detected (zero reads) in the bulk sequencing data.
   b. The corresponding locus is covered by at least ten reads in the bulk sequencing data.
5. To avoid potential false positives caused by misalignment, discard the *de novo* variants 1) within the tandem repeats, 2) within the regions close to centromeres/telomeres, 3) three-base next to the indels, 4) within the homopolymers or the regions with low complexity (Figure 5B). The related filter files: cento.bed and hg19_tandem.bed file are provided in the Github website. After the effective filtering of amplification errors, the *de novo* variants are mainly composed of damage-associated single-nucleotide variants (damSNVs).
6. The related customized scripts are available at https://github.com/zonglab/LCS-WGA-code.

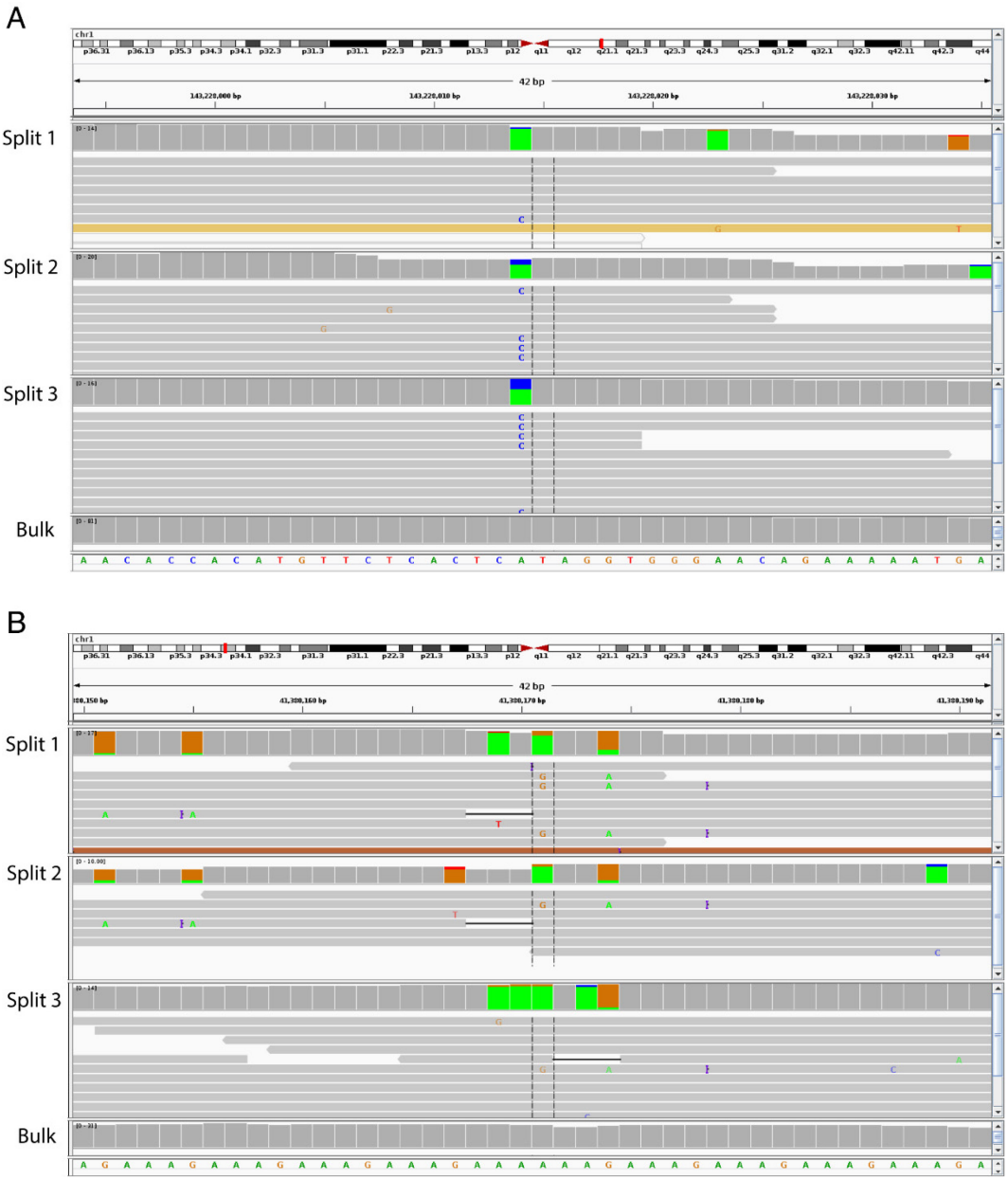


Figure 5. The IGV screenshots show the criteria of variant calling.
(A) The variants detected in at least two splits and not detected in the bulk data are called as true *de novo* variants. (B) The examples of potential false positives caused by misalignment or sequencing errors in the genomic regions with repetitive sequences. Such variants are filtered out in the analysis.

Recipes

UV-treated UltraPure H_2O
All the UltraPure DEPC-treated H_2O used in this protocol will be subjected to UV treatment before usage to prevent any potential DNA contamination. The UV treatment is performed in a CL-1000 Ultraviolet Crosslinker with 2000 × 100 μJ/CM^2 for 5 min.

1. **Lysis buffer (400 mM KOH, 100 mM DTT, 2mM EDTA)**
   200 µL of 1 M KOH
   50 µL of 1 M DTT
   2 µL of 0.5 M EDTA
   248 µL of UV-treated H_2O

2. **Stop solution (400 mM HCl, 600 mM Tris-HCl)**
   100 µL of 2 M HCl
   300 µL of 1 M Tris-HCl (pH=7.5)
   100 µL of UV-treated H_2O

3. **UDG reaction mix**
   1.5 µL of ThermoPol Reaction Buffer (10×)
   0.2 µL of Uracil-DNA Glycosylase
   3.8 µL of UV-treated H_2O

4. **Pre-amplification primer mix**
   5 µM GAT27-5N-3G
   5 µM GAT27-5N-3T

5. **MDA reaction mix**
   3 µL of phi29 DNA polymerase reaction buffer (10×)
   3 µL of dNTP solution mix
   3 µL of 1 mM random hexamer
   3 µL of TWEEN 20 (5%)
   0.3 µL of BSA solution (20 mg/mL) that comes with Phi29 polymerase
   12 µL of UV-treated H_2O

6. **Library amplification mix**
   1.5 µL of 0.2 M $Mg(AC)_2$ Solution
   1.5 µL of 10 µM Nextera i5 index primer (Table 2)
   1.5 µL of 10 µM Nextera i7 index primer (Table 2)
   18.8 µL of NEBNext® Ultra™ II Q5® Master Mix

Table 2. Examples of index primers (Illumina Nextera kit)

Nextera index	Sequences (The index sequences are underlined.)
N701	CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGGGCTCGG
N702	CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTCGTGGGCTCGG
N703	CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTCTCGTGGGCTCGG
N501	AATGATACGGCGACCACCGAGATCTACACTAGATCGCTCGTCGGCAGCGTC
N502	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTCGTCGGCAGCGTC
N503	AATGATACGGCGACCACCGAGATCTACACTATCCTCTTCGTCGGCAGCGTC

*Note: All index primers are ordered from Integrated DNA Technologies (IDT). The primers are dissolved in UV-treated UltraPure DEPC-treated H_2O. The index primer sequences are obtained from Illumina Nextera kit. The i7 index primer follows the pattern of CAAGCAGAAGACGGCATACGAGAT[i7]GTCTCGTGGGCTCGG. The i5 index primer follows the pattern of AATGATACGGCGACCACCGAGATCTACAC[i5]TCGTCGGCAGCGTC. Here, we only list a few index primers as examples. To obtain the full information, please refer to '**Illumina Adapter Sequences**' (https://support.illumina.com/downloads/illumina-adapter-sequences-document-1000000002694.html).*

7. Growth medium for MCF 10A cells

100 mL of DMEM/F12
5 mL of Horse serum
20 µL of Recombinant human EGF (100 µg/mL in sterile H_2O)
50 µL of Hydrocortisone (1 mg/mL in 200 proof ethanol)
10 µL of Cholera toxin (1 mg/mL in sterile H_2O)
100 µL of Insulin (10 mg/mL in sterile H_2O with 1% glacial acetic acid)

Acknowledgments

We are grateful to the McNair family, Dr. Arthur Beaudet and Dr. Brendan Lee, and Dr. C. Neblett, for their support. We would like to acknowledge Dr. Susan Rosenberg, Dr. Christophe Herman, Dr. Herman Dierick, Dr. Greg Ira, Dr. Hui Zheng, Dr. Philip Hastings and Catherine Bradley for their helpful discussions. We would also thank other Zong lab members for their kind support. We also acknowledge NIH NeuroBioBank for providing brain tissue samples. This study was supported by NIH Director's New Innovator Award (1DP2EB020399) and McNair Scholarship. The original paper has been published in Zhu *et al*. (2021; DOI: 10.1126/sciadv.abf3329).

Competing interests

C.Z. is the inventor on two patents related to this work filed by the Baylor College of Medicine (no. 16/407,032, filed 8 May 2019, published 29 August 2019; no. 15/308,592, filed 5 May 2015, published 28 May 2019). The authors declare that they have no other competing interests.

References

Amente, S., Di Palo, G., Scala, G., Castrignano, T., Gorini, F., Cocozza, S., Moresano, A., Pucci, P., Ma, B., Stepanov, I., *et al*. (2019). Genome-wide mapping of 8-oxo-7,8-dihydro-2'-deoxyguanosine reveals accumulation of oxidatively-generated damage at DNA replication origins within transcribed long genes of mammalian cells. *Nucleic Acids Res* 47(1): 221-236.

Ames, B. N., Shigenaga, M. K. and Hagen, T. M. (1993). Oxidants, antioxidants, and the degenerative diseases of aging. *Proc Natl Acad Sci U S A* 90(17): 7915-7922.

Bohrson, C. L., Barton, A. R., Lodato, M. A., Rodin, R. E., Luquette, L. J., Viswanadham, V. V., Gulhan, D. C., Cortes-Ciriano, I., Sherman, M. A., Kwon, M., *et al*. (2019). Linked-read analysis identifies mutations in single-cell DNA-sequencing data. *Nat Genet* 51(4): 749-754.

Chen, C., Xing, D., Tan, L., Li, H., Zhou, G., Huang, L. and Xie, X. S. (2017). Single-cell whole-genome analyses by Linear Amplification via Transposon Insertion (LIANTI). *Science* 356(6334): 189-194.

Dabin, J., Fortuny, A. and Polo, S. E. (2016). Epigenome Maintenance in Response to DNA Damage. *Mol Cell* 62(5): 712-727.

De Bont, R. and van Larebeke, N. (2004). Endogenous DNA damage in humans: a review of quantitative data. *Mutagenesis* 19(3): 169-185.

Ding, Y., Fleming, A. M. and Burrows, C. J. (2017). Sequencing the Mouse Genome for the Oxidatively Modified Base 8-Oxo-7,8-dihydroguanine by OG-Seq. *J Am Chem Soc* 139(7): 2569-2572.

Dizdaroglu, M., Coskun, E. and Jaruga, P. (2015). Measurement of oxidatively induced DNA damage and its repair, by mass spectrometric techniques. *Free Radic Res* 49(5): 525-548.

Galashevskaya, A., Sarno, A., Vagbo, C. B., Aas, P. A., Hagen, L., Slupphaug, G. and Krokan, H. E. (2013). A robust, sensitive assay for genomic uracil determination by LC/MS/MS reveals lower levels than previously reported. *DNA Repair (Amst)* 12(9): 699-706.

Garcia-Nieto, P. E., Schwartz, E. K., King, D. A., Paulsen, J., Collas, P., Herrera, R. E. and Morrison, A. J. (2017). Carcinogen susceptibility is regulated by genome architecture and predicts cancer mutagenesis. *EMBO J* 36(19): 2829-2843.

Gehrke, T. H., Lischke, U., Gasteiger, K. L., Schneider, S., Arnold, S., Muller, H. C., Stephenson, D. S., Zipse, H. and Carell, T. (2013). Unexpected non-Hoogsteen-based mutagenicity mechanism of FaPy-DNA lesions. *Nat Chem Biol* 9(7): 455-461.

Krishnaswami, S. R., Grindberg, R. V., Novotny, M., Venepally, P., Lacar, B., Bhutani, K., Linker, S. B., Pham, S., Erwin, J. A., Miller, J. A., *et al*. (2016). Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat Protoc* 11(3): 499-524.

Luquette, L. J., Bohrson, C. L., Sherman, M. A. and Park, P. J. (2019). Identification of somatic mutations in single cell DNA-seq using a spatial model of allelic imbalance. *Nat Commun* 10(1): 3908.

Poetsch, A. R., Boulton, S. J. and Luscombe, N. M. (2018). Genomic landscape of oxidative DNA damage and repair reveals regioselective protection from mutagenesis. *Genome Biol* 19(1): 215.

Purmal, A. A., Kow, Y. W. and Wallace, S. S. (1994). Major oxidative products of cytosine, 5-hydroxycytosine and 5-hydroxyuracil, exhibit sequence context-dependent mispairing *in vitro*. *Nucleic Acids Res* 22(1): 72-78.

Purmal, A. A., Lampman, G. W., Bond, J. P., Hatahet, Z. and Wallace, S. S. (1998). Enzymatic processing of uracil glycol, a major oxidative product of DNA cytosine. *J Biol Chem* 273(16): 10026-10035.

Tudek, B., Winczura, A., Janik, J., Siomek, A., Foksinski, M. and Olinski, R. (2010). Involvement of oxidatively damaged DNA and repair in cancer development and aging. *Am J Transl Res* 2(3): 254-284.

Wagner, J. R., Hu, C. C. and Ames, B. N. (1992). Endogenous oxidative damage of deoxycytidine in DNA. *Proc Natl Acad Sci U S A* 89(8): 3380-3384.

Wu, J., McKeague, M. and Sturla, S. J. (2018). Nucleotide-Resolution Genome-Wide Mapping of Oxidative DNA Damage by Click-Code-Seq. *J Am Chem Soc* 140(31): 9783-9787.

Xing, D., Tan, L., Chang, C. H., Li, H. and Xie, X. S. (2021). Accurate SNV detection in single cells by transposon-based whole-genome amplification of complementary strands. *Proc Natl Acad Sci U S A* 118(8).

Yoshihara, M., Jiang, L., Akatsuka, S., Suyama, M. and Toyokuni, S. (2014). Genome-wide profiling of 8-oxoguanine reveals its association with spatial positioning in nucleus. *DNA Res* 21(6): 603-612.

Zhu, Q., Niu, Y., Gundry, M. and Zong, C. (2021). Single-cell damagenome profiling unveils vulnerable genes and functional pathways in human genome toward DNA damage. *Sci Adv* 7(27). DOI: 10.1126/sciadv.abf3329.

Zong, C., Lu, S., Chapman, A. R. and Xie, X. S. (2012). Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science* 338(6114): 1622-1626.