nature protocols

PROTOCOL

https://doi.org/10.1038/s41596-022-00783-7

Check for updates

Purification, full-length sequencing and genomic origin mapping of eccDNA

Yuangao Wang[1,2], Meng Wang[1,2] and Yi Zhang[1,2,3,4]✉

Extrachromosomal circular DNA (eccDNA) was discovered more than half a century ago. However, its biogenesis and function have just begun to be elucidated. One hurdle that has prevented our understanding of eccDNA is the difficulty in obtaining pure eccDNA from cells. The current eccDNA purification methods mainly rely on depleting linear DNAs by exonuclease digestion after obtaining crude circles by alkaline lysis. Owing to eccDNA's low abundance and heterogeneous size, the current purification methods are not efficient in obtaining pure eccDNA. Here we describe a new three-step eccDNA purification (3SEP) procedure that adds a step to recover circular DNA, but not linear DNA that escape from the exonuclease digestion, whereby 3SEP results in eccDNA preparations with high purity and reproducibility. Additionally, we developed a full-length eccDNA sequencing technique by combining rolling-circle amplification with Nanopore sequencing. Accordingly, we developed a full-length eccDNA caller (Flec) to call the consensus sequence of multiple tandem copies of eccDNA contained within the debranched rolling-circle amplification product and map the consensus to its genomic origin. Collectively, our protocol will facilitate eccDNA identification and characterization, and has the potential for diagnostic and clinical applications. For a well-trained molecular biologist, it takes ~1–2 d to purify eccDNAs, another 5–6 d to carry out Nanopore library preparation and sequencing, and 1–5 d for an experienced bioinformatic scientist to analyze the data.

Introduction

Extrachromosomal circular DNAs (eccDNAs) have been discovered in almost all cell lines, tissues and organs across species since their first description in 1964 (ref. [1]). Robust eccDNA purification methods are crucial for identifying eccDNAs' genomic origin and understanding their biogenesis and biological function[2–4], including their potent immunostimulatory activity[2]. However, the size heterogeneity (ranging from hundreds of base pairs to hundreds of thousands of base pairs) and low abundance of eccDNAs relative to their linear chromosomal counterparts[3,4] make obtaining pure eccDNAs extremely difficult. Thus, a successful strategy for obtaining pure eccDNAs requires combining different principles to separate circular DNA from their linear counterparts[4].

Thus far, Hirt extraction[5,6], alkaline lysis[7], exonuclease digestion[8–10] and buoyant density[5,6,8,11] have been used to purify eccDNAs. In general, two rounds of enrichment are used to obtain eccDNAs. Crude DNA circles are first extracted from biological samples after large chromosomal DNAs, and their associated proteins are precipitated and removed by either Hirt[5,6] or alkaline lysis[7]. Then, linear contaminant DNAs within the crude circular DNA are separated either by buoyant density[5,6,8,11] or digested with an exonuclease such as ATP-dependent Plasmid-Safe DNase (PS DNase)[9], exonuclease V (RecBCD)[10] or exonuclease III (ref. [12]). Among these methods, Circle-seq, initially developed by Dutta's lab[13], then named by Regenberg' lab[14], became popular, and detailed protocols with slight modifications are available[15–17]. Briefly, in Circle-seq, crude circular DNAs are isolated by alkaline lysis, then the exonuclease PS DNase is used to digest linear contaminating DNA. However, two precautions have to be taken when using exonuclease treatment. First, the specificity and efficacy of the exonuclease used have to be examined because trace amount of endonuclease activity in the exonuclease can cause eccDNA loss, particularly considering that extended multiple rounds of exonuclease digestion are used in Circle-seq[13–17]; second, the exonuclease incision can be abolished by different blockades at the ends of endogenous linear DNA such as protruding termini[18], certain abnormal and damaged nucleotides[19], cross-linkages[20], special structures[21], etc., or even inhibited by the nuclease digestion products[9]. Thus, for the crude circular DNA from alkaline lysis, digesting solely by

[1]Howard Hughes Medical Institute, Boston, MA, USA. [2]Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA, USA. [3]Department of Genetics, Boston, MA, USA. [4]Harvard Stem Cell Institute WAB-149G, Boston, MA, USA. ✉e-mail: yzhang@genetics.med.harvard.edu

exonuclease may not be sufficient to obtain pure eccDNA, as demonstrated by direct microscopy imaging[22]. Although the crude circular DNA from Hirt, alkaline lysis or even exonuclease treatment could be further purified by cesium chloride–ethidium bromide gradient ultracentrifuge[6], given its laborious nature[23], this method does not satisfy the needs of most studies.

Here we describe a robust three-step eccDNA purification (3SEP) technique that includes another positive selection step to further enrich eccDNA after modified alkaline lysis and PS DNase digestion[2] (Fig. 1a). In the first step, the crude DNA circles are extracted with buffered (pH 11.8) alkaline lysis, a condition that causes less circular DNA loss than the conventional 0.2 M sodium hydroxide[24]. In the second step, linear DNA in the crude circular DNA product is reduced by the treatment of ATP-dependent PS DNase; meanwhile the circular mitochondrial DNA can be linearized by restriction digestion of an 8 bp cutter PacI to make our method compatible for eccDNA extraction from whole cells. Lastly, and most importantly, we make further positively selection of eccDNA with solution A, a proprietary reagent that allows selective binding to and recovery of circular DNA by silica beads while leaving linear DNAs that have escaped PS DNase digestion in solution.

The 3SEP protocol is designed for eccDNA purification from cultured mammalian cells. To adapt the protocol for eccDNA purification from tissue samples or other types of sample, sample processing, such as cell dissociation or tissue grinding, may be needed before starting 3SEP.

To obtain the full-length sequence and genomic origin of eccDNA, we use Oxford Nanopore to sequence the long DNA molecules that are generated by T7 endonuclease I (T7I) debranching the rolling-circle amplification (RCA) products of the purified eccDNAs (Fig. 1b), then we developed a novel computational method named Flec (full-length eccDNA caller) that can reconstruct eccDNA full-length sequences directly from Nanopore reads alone, without the need for short split reads obtained from second-generation sequencing techniques such as Illumina that are used in Circle-seq. The Nanopore reads of long DNA molecules consist of tandem multiple copies of the template eccDNA (Fig. 1b). To obtain the full-length consensus sequence and determine the genomic origin(s) of each eccDNA, each Nanopore read is mapped to the reference genome and divided into subreads according to the location of each mapped portion (Fig. 1b). If all the subreads are mapped to a single

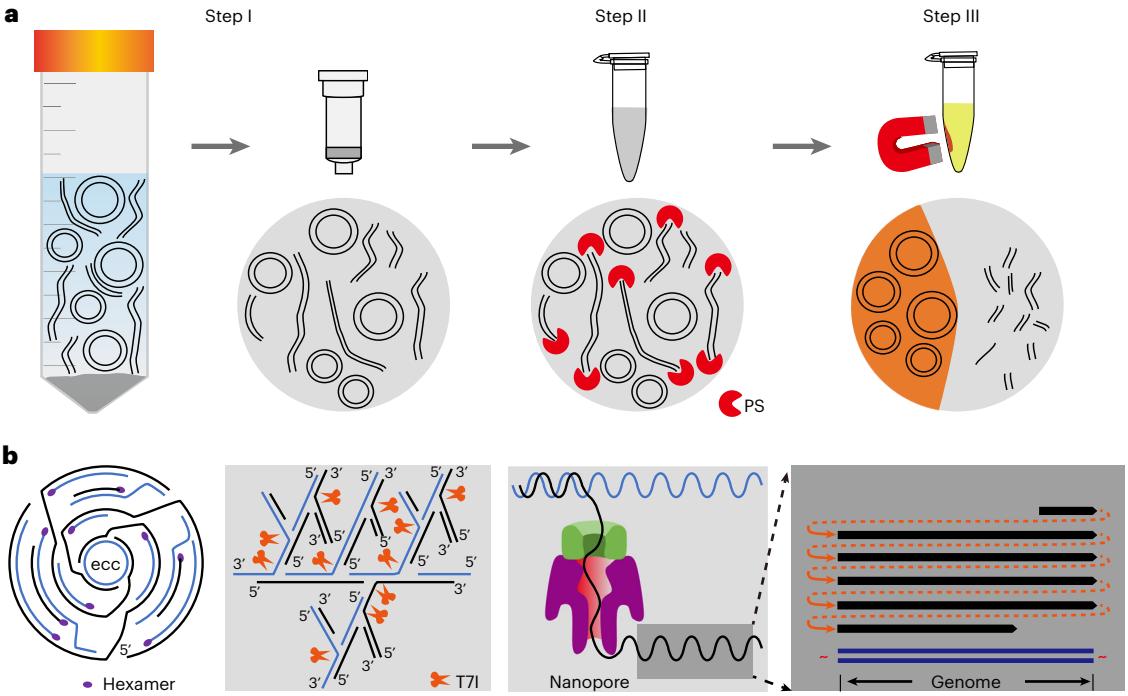

Fig. 1 | Scheme of the eccDNA purification and sequencing procedure. a, The 3SEP scheme. Step I, crude DNA circles are extracted from whole cells with buffered alkaline lysis at pH 11.8, and are then bound to and eluted from a silica column; Step II, linear DNA and PacI linearized mitochondrial DNA are digested with PS DNase; Step III, eccDNAs are selectively recovered with magnetic beads and solution A. **b**, EccDNA sequencing procedure. EccDNAs are first amplified by random hexamers priming rolling-circle amplification and debranched with T7I, then read through by Oxford Nanopore sequencing. The full-length sequence of the eccDNA is called by Flec as a consensus sequence (represented by blue bar) of multiple tandem copies (represented by black bars) within a long read. Reprinted from ref. [2], Springer Nature Limited.

Box 1 | Example of detailed eccDNA reconstruction process from Nanopore reads of rolling-circle amplified eccDNA

Self-ligation of single-fragment eccDNA:

de6f925d-010a-4ab5-a2ef-699d815d9904
#Fragment: 1 Full Pass: 5 Read Length: 3790
1 - chr1:45731744-45732099 (+) - 365
366 - chr1:45731400-45732099 (+) - 1051
1052 - chr1:45731400-45732097 (+) - 1743
1744 - chr1:45731400-45732099 (+) - 2436
2437 - chr1:45731400-45732099 (+) - 3131
3132 - chr1:45731400-45732055 (+) - 3782
Location:
chr1:45731400-45732099 (+)

Multiple ligation of fragments from different genomic regions:

7573a6ef-1930-4d65-ad73-a03edf6c862c
#Fragment: 2 Full Pass: 4 Read Length: 4108

	7 - chr19:16574531-16574838 (+) - 316
317 - chr5:99484102-99484687 (-) - 886	887 - chr19:16574470-16574838 (+) - 1248
1249 - chr5:99484102-99484687 (-) - 1828	1829 - chr19:16574470-16574837 (+) - 2181
2182 - chr5:99484102-99484687 (-) - 2757	2758 - chr19:16574471-16574838 (+) - 3111
3112 - chr5:99484113-99484687 (-) - 3676	3677 - chr19:16574463-16574839 (+) - 4053

Location:
chr5:99484102-99484687 (-) chr19:16574470-16574838 (+)

locus, this suggests that the eccDNA is from self-ligation of one genomic fragment[2]. If the subreads are mapped to two or more loci of the same or different chromosomes, the eccDNA is formed by ligation and circularization of multiple genomic fragments[2]. To determine the boundary (joining ends) and consensus sequence of eccDNAs composed of either single or multiple fragments, we developed a unified threading method to reconstruct the structure of eccDNA based on the order and location of mapped subreads in each Nanopore read (Box 1). We implemented the eccDNA calling method in the Flec tool and provided an integrated computational pipeline from raw sequencing signal processing to generating the final eccDNA consensus.

Compared with the previous eccDNA purification methods[22], our method not only results in eccDNA of high purity, as revealed by microscopy imaging (Fig. 2b,d right), but also saves time, as the purification procedure can be completed within 1 d instead of 1 week (refs. [14,25]). Given its high reproducibility in yield among parallel preparations, as revealed by agarose gel[2], our method can be used for evaluating the eccDNA yield by agarose gel quantification (Fig. 2a,c), as shown in our previous study[2]. Additionally, Solution A and the magnetic silica beads used to enrich circular DNAs in 3SEP could be applied to recover double-stranded circular DNA of other types, such as mitochondrial DNA, viral DNA, etc. Furthermore, by using Nanopore sequencing, our method can reconstruct full-length eccDNA sequences and reveal their genomic origins, including distinguishing between the self-ligation of a single genomic fragment and multiple ligations of fragments from different genomic regions[2]. Our computational method can calculate the number of full passes, which measures the reoccurrence of the same structure (concatemer) for rolling-circle-amplified circular DNA. A minimum of two full passes ensures the circularity of the original DNA because linear DNA would not form concatemers comprising two or more identical units during RCA. To avoid artificial concatemers generated from branching and template switching during phi29 amplification, we keep only eccDNAs with all fragments in the same locus mapped to the same genomic strand, since artificial concatemers generated from branching and template switching would map to different genomic strands. Thus, our purification and computational analysis method together ensure the high reliability of the identified eccDNAs.

Well-trained researchers in biochemistry and molecular biology can easily implement 3SEP. However, if the final eccDNA yield is <10 ng, we strongly suggest confirmation of the eccDNA purity by imaging using atomic force microscopy (AFM) or electron microscopy. In this case, relevant training in sample preparation and microscopy or assistance from relevant facilities are needed. To obtain the full-length sequence of eccDNA by Nanopore long-read sequencing, basic computational skills or training provided by Oxford Nanopore are required for setting up the sequencing platform in a laboratory. In addition, basic Linux and Python skills are needed for processing the long reads and constructing full-length sequences of eccDNA by running the tools.

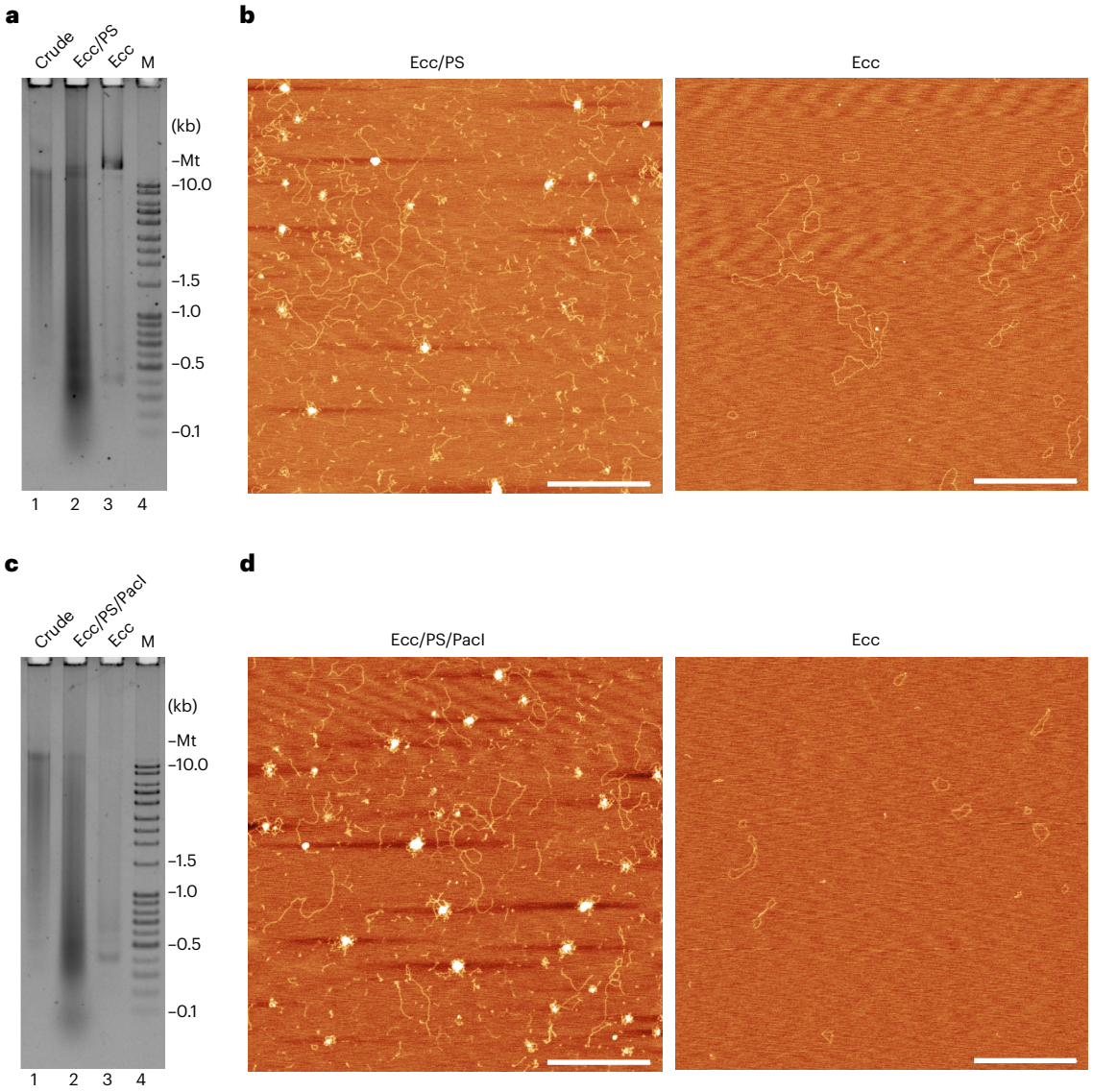

Fig. 2 | Agarose gel and AFM images of eccDNAs at different purification steps. a–d, HeLa cells were cultured in DMEM supplemented with 10% FBS. Cells were cultured for another 48 h after reaching 100% confluence. Both floating and adherent cells were collected, washed with PBS three times, resuspended in PBS and counted. Cells were fixed by adding absolute methanol to a final concentration of 95% (vol/vol) and put on ice for at least 10 min. Then, 3SEP procedures were performed without PacI (**a**,**b**) or with PacI linearization of mtDNA (**c**,**d**). **a**, Agarose gel image of DNA from different steps of 3SEP without PacI linearization of mtDNA. Lane 1: crude circular DNAs (Crude) product after alkaline lysis; lane 2: eccDNAs after PS DNase treatment (Ecc/PS); lane 3: eccDNAs (Ecc) after Solution A purification; lane 4: linear DNA ladder; mtDNA (Mt) indicates the mitochondrial DNA band in lane 3. Note: to avoid over-loading the gel, DNAs were not loaded proportionally. **b**, AFM images of the DNAs in lane 2 (left, Ecc/PS) and lane 3 (right, eccDNA) in **a**. The two big circular DNAs on the right are mtDNAs purified by 3SEP in the absence of PacI. **c**, Agarose gel image of DNA from different steps of 3SEP with PacI linearization of mtDNA. All the information is the same as in **a**, except that the 3SEP included PacI treatment (lane 2), which linearizes mtDNA and results in its disappearance from the eccDNA product, as indicated by the missing corresponding mtDNA band in lane 3 of the gel (compared with lane 3 in **a**. Mt, mitochondrial DNA. **d**, AFM images of the DNA in lane 2 (left, Ecc/P.S/PacI) and lane 3 (right, eccDNA) in **c**. Scale bars, 500 nm.

Limitations

Although circular DNA larger than 16 kb (such as mitochondrial DNA) could be efficiently purified using our method, the Nanopore sequencing method we describe here can mainly process eccDNAs shorter than 10 kb. For eccDNAs longer than 10 kb, please refer to other long sequencing methods[26].

Experimental design

In general, eccDNA abundance in biological samples is low and sensitive to cellular status. Approximately 10 ng eccDNAs were obtained from 600 million adherent mouse embryonic stem cells (mESCs) of exponential growth. However, dozens to hundreds times more eccDNA can be produced

if all input mESCs commit to death induced by apoptosis inducers such as etoposide, staurosporine and UV-C[2]. We speculate eccDNA biogenesis is sensitive to cellular health status; thus, shielding samples from undesired stress is crucial for consistent results and their correct interpretation. Therefore, the required amount of input material may need to be determined by pilot tests if starting with new samples.

In our experience, eccDNA contents vary a lot among biological samples. We recommend beginners learn 3SEP using HeLa cells. To reduce the burden of HeLa cell culture and reagent cost for eccDNA purification, over-confluent growth conditions can be used to increase eccDNA yield. However, to avoid artifacts, researchers must keep in mind that stress conditions, particularly those that can introduce DNA damage and double-strand breaks, may dramatically increase eccDNA contents[2].

Although ≥0.5 ng eccDNA, obtained by this protocol, is enough to be amplified by RCA for Nanopore library preparation, we recommend validating eccDNA purity by direct microscopy imaging if the total eccDNA yield from one preparation is <10 ng from a sample. We do not suggest using inverse PCR to validate eccDNA purity because products of both tandem ligation and self-circularization of a DNA fragment can be amplified with a pair of primers across the ligation site by inverse PCR.

Materials

Biological materials

- HeLa S3 (CVCL_0058; American Type Culture Collection, cat. no. CCL-2.2; https://www.atcc.org/products/ccl-2.2) **! CAUTION** The cell lines used in your research should be in excellent health and should be regularly checked to ensure they are not infected with mycoplasma.

Reagents

- Methanol, high-performance liquid chromatography grade (EMD Millipore, cat. no.MX0475-1) **! CAUTION** Methanol is highly flammable and volatile and is toxic upon inhalation or contact. Keep away from ignition sources, wear protective gloves, and avoid inhalation, swallowing and contact with skin.
- Pyrrolidine, 99% (Sigma-Aldrich, cat. no. P73803-100ML) **! CAUTION** Pyrrolidine is highly flammable with a strong ammonia-like odor, and irritates skin and eyes upon contact. Keep away from ignition sources, wear protective gloves, and avoid inhalation, swallowing and contact with skin.
- Sodium dodecyl sulfate (SDS), 20% solution (Fisher Chemical, cat. no. BP1311200)
- 2-Mercaptoethanol ≥99.0% (Sigma-Aldrich, cat. no. M7522)
- QIAGEN Plasmid Plus Midi Kit (25) (Qiagen, cat. no. 12943)
- ATP-Dependent PS DNase (Lucigen, cat. no. E3110k)
- Solution A from cccDNA purification from either Bingene (cat. no. 220501-50; http://www.bingene.com/product/46028.html) or Biofargo (cat. no. 220501; https://biofargo.com/products/100ml-cccdna-purification-kit-222501-100?variant=42043981037749) **! CAUTION** Solution A contains phenol, which is corrosive and an irritant; work under a hood.
- Qubit 1× dsDNA HS Assay Kit (Thermo Scientific, cat. no. Q33231)
- Phase Lock Gel, QuantaBio, Phase Lock Gel Heavy (VWR, cat. no. 10847-802)
- Phenol/chloroform/isoamyl (PCI) alcohol (25:24:1 (vol/vol/vol) mixture, pH 8.0) (Thermo Fisher Scientific, cat. no. BP1752I-400) **! CAUTION** Phenol is corrosive and an irritant; work under a hood.
- Ethyl alcohol, pure, 200 proof, for molecular biology (Sigma-Aldrich,cat. no. E7023) **! CAUTION** Absolute ethanol is highly flammable.
- 1 M Tris pH 7.0 (Fisher Scientific, cat. no. AM9851)
- Nickel(II) chloride (Sigma-Aldrich, cat. no. 451193)
- UltraPure DNase/RNase-free distilled water (Thermo Fisher Scientific, cat. no. 10977-023)
- SYBR Gold Nucleic Acid Gel Stain (10,000× concentrate in dimethyl sulfoxide) (Fisher Scientific, cat. no. S11494)
- Glycogen, molecular biology grade 20 MG (Sigma-Aldrich, cat. no. 10901393001)
- Dynabeads MyOne Silane (Thermo Fisher Scientific, cat. no. 37002D)
- PacI (New England Biolabs, cat. no. R0547)
- UltraPure agarose (Thermo Fisher Scientific, cat. no. 16-500-100)

- Ethylenediaminetetraacetic acid (EDTA) (0.5 M), pH 8.0 (Thermo Fisher Scientific, cat. no. AM9260G)
- 50× Tris–acetate–EDTA (TAE) buffer (Thermo Fisher Scientific, cat. no. B49)
- Magnesium acetate tetrahydrate (Sigma-Aldrich, cat. no. M5661)
- 5 M NaCl (Thermo Fisher Scientific, cat. no. AM9760G)
- Sodium acetate (3 M), pH 5.5 (Thermo Fisher Scientific, cat. no. AM9740)
- 1× TE buffer (10 mM Tris–HCl, 1 mM EDTA, pH 8.0) (Thermo Fisher Scientific, cat. no. AM9858)
- Thermo Scientific Exo Resistant Random Primer (Thermo Fisher Scientific, cat. no. SO181)
- phi29 DNA polymerase (New England Biolabs (NEB), cat. no. M0269)
- Deoxynucleotide Solution Set (NEB, cat. no. N0446)
- Pyrophosphatase, inorganic (NEB, cat. no. M2403)
- BSA, molecular biology grade (NEB, cat. no. B9000S)
- SPRIselect (Beckman Coulter, cat. no. B23317)
- T7I (NEB, cat. no. M0302)
- Ligation Sequencing Kit (Oxford Nanopore, cat. no. SQK-LSK109)
- FBS (Sigma-Aldrich, cat. no. F6178)
- Dulbecco's modified Eagle medium (DMEM), high glucose (Thermo Fisher Scientific, cat. no. 11995065)
- Penicillin–streptomycin (Thermo Fisher Scientific, cat. no. 15140122)
- Glacial acetic acid (Sigma-Aldrich, cat. no. 1000661000) **! CAUTION** Always keep glacial acetic acid away from heat and ignition sources; wear protective gloves and avoid inhalation and contact with skin; handle acetic acid under a hood or where it is well ventilated.
- Sodium acetate (Sigma-Aldrich, cat. no. S2889)

Equipment

- (Optional) DNA LoBind Tubes 1.5 mL, PCR clean (Fisher Scientific, cat. no. 13-698-791)
- Ultra-0.5 Centrifugal Filter Unit with Ultracel-10 membrane (Amicon, cat. no. UFC5010)
- QIAvac 24 Plus (Qiagen, cat. no. 19413)
- Thermal cycler
- Nanodrop (Thermo Scientific)
- Magnet (compatible with 1.5 ml tubes)
- Qubit fluorometer (Fisher scientific)
- Syringe filter unit, 0.22 μm (EMD Millipore, cat. no. SLGVM33RS)
- Fluorescence plate reader
- Black 384-well plate (Greiner Bio-One, cat. no. 781906)
- Benchtop centrifuge
- Mini-PROTEAN Tetra Vertical Electrophoresis Cell for Mini Precast Gels, 4-gel (Biorad, cat. no. 1658004)
- Mini-PROTEAN Tetra Cell Casting Module (15-well, 1.5 mm gel) (Biorad, cat. no. 1658022)
- (Optional) Highest Grade V1 AFM Mica Discs (Fisher Scientific, cat. no. NC1535937)
- (Optional) Dust-Off Disposable Air Duster (Falcon, DPSJB-12)
- Flow Cell (R9.4.1) (Oxford Nanopore, FLO-MIN106D)
- Sequencer (Oxford Nanopore)
- Computational hardware: a computer/server with at least 32 GB memory and eight-core central processing unit (CPU) running 64-bit Linux or macOS system. If the analysis starts from Nanopore raw data, a computer/server with NVIDIA graphics processing unit (GPU) (NVIDIA compute version of 6.1 or higher) would substantially reduce the base-calling time

Software

- Guppy: base-calling from Nanopore raw signals; registration is required to download. (https://community.nanoporetech.com/downloads)
- eccDNA calling toolkit (https://github.com/YiZhang-lab/eccDNA_RCA_nanopore)
- minimap2 (ref. [27]) (https://github.com/lh3/minimap2)
- Porechop (https://github.com/rrwick/Porechop)
- Snakemake[28] (https://snakemake.github.io)
- Git (https://git-scm.com/downloads)
- Python3 (https://www.python.org/downloads/)

- R (https://cran.r-project.org)
- Python packages:
  - pyfaidx (https://pypi.org/project/pyfaidx/)
  - pyfastx (https://pypi.org/project/pyfastx/)
  - Biopython (https://biopython.org)
- R packages
  - caTools (https://cran.r-project.org/web/packages/caTools/)
  - data.table (https://cran.r-project.org/web/packages/data.table/)
  - digest (https://cran.r-project.org/web/packages/digest/)
  - dplyr (https://cran.r-project.org/web/packages/dplyr/)
  - DT (https://cran.r-project.org/web/packages/DT/)
  - emojifont (https://cran.r-project.org/web/packages/emojifont/)
  - extrafont (https://cran.r-project.org/web/packages/extrafont/)
  - fastmatch (https://cran.r-project.org/web/packages/fastmatch/)
  - flexdashboard (https://cran.r-project.org/web/packages/flexdashboard/)
  - futile.logger (https://cran.r-project.org/web/packages/futile.logger/)
  - ggplot2 (https://cran.r-project.org/web/packages/ggplot2/)
  - ggExtra (https://cran.r-project.org/web/packages/ggExtra/)
  - ggridges (https://cran.r-project.org/web/packages/ggridges/)
  - knitr (https://cran.r-project.org/web/packages/knitr/)
  - optparse (https://cran.r-project.org/web/packages/optparse/)
  - plyr (https://cran.r-project.org/web/packages/plyr/)
  - RColorBrewer (https://cran.r-project.org/web/packages/RColorBrewer/)
  - readr (https://cran.r-project.org/web/packages/readr/)
  - reshape2 (https://cran.r-project.org/web/packages/reshape2/)
  - scales (https://cran.r-project.org/web/packages/scales/)
  - tufte (https://cran.r-project.org/web/packages/tufte/)
  - viridis (https://cran.r-project.org/web/packages/viridis/)
  - yaml (https://cran.r-project.org/web/packages/yaml/)

Reagent setup

Suspension buffer (100 ml)

Add 2 ml 0.5 M EDTA pH 8.0, 3 ml 5 M NaCl and 1 ml glycerol to 94 ml diH_2O to a final concentration of 10 mM EDTA pH 8.0, 150 mM NaCl, 1% (vol/vol) glycerol. Then add 100 µl Lysis blue (Qiagen Kit) and 100 µl RNase A (Qiagen Kit, 110 mg/ml). The buffer is stable for 6 months at 4 °C. Freshly add 1/500 volume 2-mercaptoethanol before use.

Pyr buffer (100 ml)

Add 4.17 ml pyrrolidine, 4 ml 0.5 M EDTA pH 8.0 and 5 ml 20% (wt/vol) SDS to 80 ml diH_2O; adjust pH to 11.80 with 2 M sodium acetate pH 4.0 (to prepare 100 ml 2 M sodium acetate pH 4.0, weigh and dissove 16.46 g sodium acetate with UltraPure water, adjust pH to 4.0 with glacial acetic acid, then bring the volume to 100 ml with water), then supplement diH_2O to 100 ml to a final concentration of 0.5 M pyrrolidine, 20 mM EDTA and 1% SDS, pH 11.80. The Pyr buffer is stable at room temperature (RT, 20–25 °C) for 2 months. Freshly add 1/500 volume 2-mercaptoethanol before use.

(Optional) 10× AFM imaging buffer (10 ml)

Add 129.6 mg nickel(II) chloride ($NiCl_2$) and 1 ml Tris–HCl, pH 8.0, to 9.0 ml UltraPure water to make the final concentration: 100 mM $NiCl_2$; 100 mM Tris–HCl, pH 8.0. Filtrate the buffer through a 0.22 µm syringe filter. The buffer is stable at RT for >1 year.

(Optional) 500× AFM sampling washing buffer

Weigh 2.145 g magnesium acetate tetrahydrate and add UltraPure water to 10.0 ml to make 1 M magnesium acetate. Filtrate the buffer through a 0.22 µm syringe filter. The buffer is stable at RT.

Equipment setup

Software setup

The link for each software listed above provides installation instructions. If the software is installed in a user-specific path, this path should be added to the `PATH` environment variable with the following shell command line:

```
export PATH=/path/to/software/bin:$PATH.
```

To set up Python packages, use:

```
pip install package_name
```

To install R packages, use the following code inside R:

```
install.packages(package_name)
```

Alternatively, software installation can be done by using the conda package management system. Instructions to set up conda are available at https://bioconda.github.io/user/install.html#install-conda. After setting up conda, install software and packages with:

```
conda install package_name
```

Example data

We provide an example dataset of raw sequencing signals in fast5 files generated by Nanopore sequencing of rolling-circle-amplified eccDNA purified from mESC. These fast5 files are the raw input for the whole eccDNA calling pipeline. The dataset was subsampled from a whole run of MinION R9.4.1 Flow Cell (Oxford Nanopore, FLO-MIN106D) with sequencing kit (Oxford Nanopore, SQK-LSK109) in the original data in ref. [2]. The example dataset can be downloaded at https://figshare.com/articles/dataset/Nanopore_reads_of_eccDNA/17046158.

Procedure

Part 1: purify eccDNAs by 3SEP ● Timing 1–2 d

Isolate crude circular DNA by modified alkaline lysis ● Timing 1.5–2 h

▲ **CRITICAL** Over-confluent HeLa cells are prepared by continued culturing for another 48–72 h in the complete medium after reaching 100% confluence. Two to three 10 cm dishes of cultures are enough to obtain 35 million over-confluent cells to be used for the 3SEP procedures below. Of note, researchers need to determine the cell numbers required for eccDNA purification when applying 3SEP to other biological samples.

1 Harvest cells. Collect both floating cells and adherent cells (by trypsinization). Wash cells with PBS three times and count the number of cells. Fix cells in 95% (vol/vol) methanol by resuspending cells with 0.5 ml PBS and mixing with 9.5 ml absolute methanol, then put them on ice for 10 min. When 3SEP cannot be immediately performed, keep cells in 95% methanol at −20 °C. Although 3SEP works very well with fresh cells, we recommend fixing cells with methanol as we found methanol fixation reduces nonspecific adherence of magnetic beads to the tube in Steps 22–32, and promotes beads settling down on the side of the magnet.
■ **PAUSE POINT** Fixed cells can be stored at −20 °C for 1 month.

2 Collect 35 million HeLa cells in a 50 ml tube, then centrifuge for 10 min at 2,000*g* at 4 °C.

3 Resuspend cells in suspension buffer. Suspend cells with 10 ml suspension buffer after supplementing 20 μl 2-mercaptoethanol (for more cells, proportionally scale up the volume of buffer).

4 Add 10 ml Pyr buffer (equal volume to Step 3) after supplementing with 20 μl 2-mercaptoethanol and gently mix by inverting the tube five to ten times. The lysate will become blue; keep for 5 min at RT.
! **CAUTION** Vigorous mixing at this step can damage both chromosomal DNA and eccDNA. Breakage of chromosomal DNA will increase linear DNA contamination and the difficulty in downstream linear DNA depletion, while damage to eccDNA will decrease final eccDNA yield.
! **CAUTION** Pyr buffer contains pyrrolidine, which smells of ammonia, and we recommend to handle it in a fume hood or under a benchtop exhaust snorkel.

5 Add 10 ml Buffer S3 (from QIAGEN Plasmid Plus Midi Kit, equal volume to Step 3) and gently invert the tube until the solution turns white.
! CAUTION It is critical to avoid vigorous mixing.
6 Centrifuge the lysate for 10 min at 4,500*g* at RT.
? TROUBLESHOOTING
7 Transfer the clear lysate to a QIAfilter Cartridge (QIAGEN Plasmid Plus Midi Kit), and filter the cell lysate into a 50 ml tube.
8 Estimate the volume of filtrated lysate and add one-third of the estimated volume Buffer BB (QIAGEN Plasmid Plus Midi Kit). Mix by inverting the tube four to eight times.
9 Transfer the lysate to a spin column on the Qiavac 24 plus and apply vacuum until all liquid is drawn through the column according to Qiagen's instructions.
10 Wash the column with 0.7 ml ETR buffer (QIAGEN Plasmid Plus Midi Kit) and apply vacuum until all liquid is drawn through the column. Repeat the wash with 0.7 ml PE buffer.
11 Centrifuge the column at 10,000*g* for 2 min at RT and transfer the column to a clean 1.5 ml tube.
12 Elute crude circular DNA by adding 100–200 µl 0.1× EB buffer (10× dilution of EB buffer from QIAGEN Plasmid Plus Midi Kit) to the column, wait >2 min and centrifuge at 10,000*g* for 1 min at RT.
13 Measure the DNA concentration by either Nanodrop or Qubit 1× dsDNA HS Assay Kit.
■ PAUSE POINT DNA can be kept at 4 °C overnight or −20 °C for long-term storage.

Digest linear DNA with PS DNase ● Timing 3.5 h to overnight

14 Linearize mtDNA by PacI and digest linear DNA with ATP-dependent PS DNase. First, combine the reagents as outlined below (proportionally scale up the reaction according to the DNA amount in Step 13. Typically, >100 µg DNA may be obtained in Step 13 and may require >1.5 ml digestion reaction)

10× ATP-dependent PS DNase buffer	5 µl
ATP (25 mM)	2.0 µl
ATP-dependent PS DNase	1.0 µl
Crude circular DNA (Step 13)	1–3 µg
PacI	0.5 µl
RNase A/T1	0.1 µl
H_2O	To 50 µl

Incubate the reaction at 37 °C for 2 h to overnight.
▲ CRITICAL STEP We use PacI to linearize mtDNA (both human and mouse mtDNA have three PacI recognition sites) here because it is a rare cutter (8 bp recognition site) and is also fully active in the PS DNase buffer. Thus, linearizing mtDNA and removing linear DNA could be performed in a single step. However, if PacI digestion is not desired, this enzyme could be omitted without any additional changes.
▲ CRITICAL STEP PacI could be replaced by any other rare cutter, or even clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 with single guide RNA targeting mtDNA or any unwanted circular DNA. However, if the new rare cutters are not active in the PS DNase buffer, the linearization must be performed before PS DNase digestion, and proper buffer exchange (e.g., follow Steps 15–18) must be done before proceeding to PS DNase treatment.
15 (Optional) Concentrate solution. Filter the reaction solution from Step 14 using an Ultra-0.5 Centrifugal Filter Unit (10 kDa) to 250–360 µl according to the manufacturer's instruction.
16 PCI alcohol extraction. If the volume from Step 15 is <250 µl, bring it to 250 µl with UltraPure water. Transfer the solution from Step 14 or concentrate from Step 15 to a Phase Lock Gel tube, add an equal volume of PCI alcohol (25:24:1 mixture, pH 8.0), shake the tube by hand thoroughly and centrifuge at 14,000*g* for 7.5 min at 4 °C.
! CAUTION Phenol should be handled in a fume hood or under a benchtop exhaust snorkel.
17 Precipitate DNA. Transfer the aqueous phase to a new tube, add 1/10 volume of sodium acetate (3 M, pH 5.5), add 1 µl glycogen and three volumes of 200 proof ethanol, mix and put at −80 °C for at least 30 min.
■ PAUSE POINT DNA can be kept at −80 °C for overnight or longer.

18 Centrifuge the DNA for 30 min at >16,400*g* at 4 °C, wash the pellet once with 1 ml 80% (vol/vol) ethanol. After pipetting out the ethanol, spin down the tube for 30 s at >16,400*g* at 4 °C, pipette out residual ethanol, dry the pellet by leaving the tube lid open for 1–2 min, then resuspend the pellet with 50 µl 2 mM Tris–HCl 7.0 before the pellet totally dries out.

Selectively recover eccDNA with Solution A ● Timing 1–1.5 h

19 After equilibrating Solution A at RT for 30 min, add 700 µl Solution A to the sample from Step 18, mix thoroughly by pipetting up and down ten times, and let the tube stand for 5 min at RT.
! CAUTION Solution A contains phenol and should be handled in a fume hood or under a benchtop exhaust snorkel.

20 Thoroughly vortex to resuspend the Dynabeads MyOne Silane beads, take 10 µl to a new tube and place the tube on an appropriate magnet for 2 min. Remove the liquid and resuspend the beads in 20 µl Solution A.
! CAUTION Solution A contains phenol and should be handled in a fume hood or under a benchtop exhaust snorkel.

21 Combine the suspended beads with the solution from Step 19 by pipetting up and down ten times, then incubate at RT for 5 min.

22 Place the tube on a magnet and allow the beads to settle, then remove and discard the solution (circular DNAs are on the beads now).

23 Quickly spin down the tube using a benchtop centrifuge for 10 s and put it on magnet again, then remove the residual liquid with a 10 µl micro pipette tip.
▲ CRITICAL STEP The less free liquid left, the purer the eccDNA will be, but it is important to take care not to remove the beads to which circular DNAs are bound.

24 Take the tube off the magnet, resuspend the beads in 300 µl Solution A by pipetting and incubate the tube for 2 min.

25 Place the tube on the magnet and allow the beads settle on the magnet, then remove and discard the solution.

26 Quickly spin down the tube using benchtop centrifuge for 10 s and put it on a magnet again, then remove the residual liquid with a 10 µl pipette tip.
▲ CRITICAL STEP Remove as much free liquid as possible, but take care not to remove any beads.

27 Repeat Steps 24–26 once more.

28 Keeping the tube on the magnet, add 700 µl 3.5 M NaCl and wait for 1 min without disturbing the beads, then remove the solution. Repeat once.

29 Keeping the tube on the magnet, add 800 µl freshly prepared 80% (vol/vol) ethanol, wait for 1 min without disturbing the beads and remove the ethanol. Repeat once.

30 Quickly spin down the tube using a benchtop centrifuge for 30 s, put the tube on the magnet and remove all residual liquid with a 10 µl pipette tip, taking care not to remove beads.

31 Take the tube off the magnet, thoroughly resuspend the beads with 100 µl 0.1× EB buffer before beads are completely dried out and rotate the tube slowly on a rotator for >3 min.

32 Place the tube on the magnet to allow the beads to settle. Transfer the eluate to a new DNA LoBind Tube.

Measure the eccDNA concentration ● Timing <0.5 h

▲ CRITICAL At least 10 ng pure eccDNA could be obtained from 35 million over-confluent HeLa cells.

▲ CRITICAL EccDNA size affects its accurate quantification. Since the amount of recovered eccDNA is generally low, highly sensitive quantification methods, such as fluorescence-based approaches, are required for accurate quantification. However, we found commercial fluorescence dye-based kits, including Qubit 1× dsDNA HS Assay Kit, underestimate the DNA concentration by one- to fourfold for circular DNA of hundreds bp in a size-dependent manner.

33 For eccDNA ≥1 kb, the Qubit 1× dsDNA HS Assay Kit works for accurate eccDNA quantification. Perform the quantification according to the manufacturer's instruction using eccDNA from Step 32.

34 To accurately measure the concentration of eccDNA that is hundreds of base pairs in size, we use a SYBR gold-based quantification method that could accurately quantify both linear dsDNA and circular dsDNA with curve generated by linear standards from the Qubit 1× dsDNA HS Assay Kit. First, prepare the working solution for DNA quantification. Dilute SYBR gold dye with 1× TE buffer (10 mM Tris 8.0 and 1 mM EDTA) at a ratio of 1:10,000.

35 Generate a DNA standard curve for quantification. Prepare 50 ng, 10 ng, 2 ng and 0.4 ng linear DNA standard from the Qubit 1× dsDNA HS Assay Kit by making 100 µl serial dilutions with working solution from the previous step. Duplicates or triplicates are recommended.

36 Make the eccDNA sample mixtures. Mix 1–10 µl eccDNA from Step 32 with the working solution to a final volume of 100 µl. Duplicates or triplicates are recommended.

37 Quantify eccDNA with a fluorescence plate reader. Pipette the standard mixtures (Step 35) and eccDNA sample mixtures (Step 36) to wells in a black 384-well plate, 100 µl in each well. To evaluate the background fluorescence, in another three wells, add 100 µl working solution without DNA (blank wells) to each. Then, read the fluorescence intensity at excitation/emission = 495 ∓ 8/536 ∓ 11 nm after 30 s of preshaking. Subtract the readout of the blank wells from the fluorescence intensity of both the standard and eccDNA wells, then calculate DNA concentrations according to the curve generated with the linear standards DNA by applying linear regression.

? TROUBLESHOOTING

Part 2: vertical agarose gel electrophoresis ● Timing 2 h

38 Mix 1.0 g UltraPure agarose with 100 ml 1× TAE in a microwave flask and microwave for 3–5 min until the agarose is completely dissolved.
▲CRITICAL STEP Agarose concentration may vary within the range of 1–2% (wt/vol). 1% agarose gel works well for eccDNA samples (a few hundred base pairs) that also retain intact mitochondrial DNA (e.g., when eccDNAs are purified from mitochondria-containing samples and mtDNA has not been linearized, for example, because PacI was omitted from Step 14). 2% agarose gel can provide good resolution for eccDNAs of a few hundred base pairs without intact mitochondrial DNA.

39 Pour the agarose into the preassembled glass plates immediately after the agarose is dissolved without allowing the solution to cool down. Insert the comb and let the gel cool down for at least 20 min.
! CAUTION Wear heat-resistant gloves when handling the flask. Be careful during stirring as the boiling solution may splash.

40 Remove the comb after dismounting the gel from the casting chamber, and, optionally, remove residual gel pieces within the wells using sharp tweezers.

41 Assemble the gel in the apparatus according to the manufacturer's instructions, then fill the chamber with 1× TAE.

42 Carefully pipette samples from Step 32 (>1 ng eccDNA could be visualized) into wells after mixing with loading buffer, and load a suitable amount of DNA ladder in an adjacent well.

43 Separate DNA at 80 V for 35 min. Running time may be different depending on the concentration of agarose in the gel.

44 Turn off the power, open the glass plates and transfer the gel to a 15-cm-diameter plastic dish, then add 50–70 ml 1× TAE.

45 Add 5 µl 10,000× Sybr Gold and shake the gel for >15 min in the dark.

46 Visualize the gel using any device that has a blue light as excitation source (e.g., Axygen Gel Documentation System-BL).

? TROUBLESHOOTING

Part 3: (optional but strongly recommended) AFM imaging ● Timing ~2–4 h

47 To 4.5 µl eccDNA (from Step 32) with a concentration of 0.6–1.0 ng/µl, add 0.5 µl (1/10 volume) 10× imaging buffer and mix by pipetting.

48 Cleave mica surface by double-sided tape, spread the DNA mixture on the fresh mica surface and incubate for 2 min. The mica disc comprises many very thin but dense layers, and each layer can serve as a stage for sample scanning. Double-sided tape can remove the top layer of the mica disc and expose a fresh layer below to allow spreading a new sample.

49 Blow off the liquid with compressed gas and rinse the mica with 30 µl of 2 mM magnesium acetate. Repeat once. Blow off the washing buffer between and after washing with compressed gas.

50 Acquire image by using tip C of an SNL-10 probe and scanning in Air mode and processing with Gwyddion (e.g., we scanned samples with Veeco MultiMode atomic-force microscope with a Nanoscope V Controller in 'ScanAsyst in Air mode').

Part 4: rolling-circle amplification and Nanopore sequencing library construction ● Timing 2–3 d

51 Assemble the RCA as follows, using eccDNA from Step 32.

10× phi29 DNA Polymerase Reaction Buffer	2 µl
25 mM dNTPs (pool of equal volumes of each 100 mM dNTP, NEB)	2 µl
Exo Resistant Random Primer (Thermo Fisher Scientific)	1 µl
eccDNA	>200 pg
H_2O	To 17.6 µl

Incubate at 95 °C for 5 min, then ramp to 30 °C at 1% ramp rate on a thermocycler (~30 min). Remove the tube from the thermocycler and sequentially supplement the reaction with the indicated amount of enzymes and BSA in the table below.

phi29 DNA Polymerase (NEB)	1 µl
Pyrophosphatase, inorganic (NEB)	1 µl
BSA, molecular biology grade	0.4 µl

Mix the reaction thoroughly by pipetting up and down at least ten times. Incubate at 30 °C for 14 h, then at 65 °C for 10 min to inactivate the phi29 DNA polymerase.

52 Dilute the RCA product with 180 µl UltraPure water and recover the high molecular weight (HMW) DNA with 80 µl SPRIselect beads (0.4× volume) according to the manufacturer's instructions.

53 Debranch RCA products. First, assemble the debranching reaction as shown below, then incubate at 37 °C for 15 min.

10× NEBuffer 2 (NEB)	15 µl
T7I	7.5 µl
(HMW) DNA from Step 52	3 µg
H_2O	To 150 µl

54 Add 210 µl elution buffer (Qiagen), immediately extract the debranched products with 360 µl PCI alcohol (25:24:1 mixture, pH 8.0), then precipitate DNA as in Step 17.
! CAUTION Phenol should be handled in a fume hood or under a benchtop exhaust snorkel.

55 Resuspend DNA from Step 54 with 200 µl as described in Step 18.

56 Select the HMW DNA with 80 µl (0.4× V) SPRIselect beads as in Step 52.

57 Construct the Nanopore sequencing library with the Ligation Sequencing Kit (Oxford Nanopore) according to the manufacturer's instructions.

Part 5: nanopore sequencing ● Timing 2 d

58 Perform Nanopore long-read sequencing with an appropriate flow cell (e.g., FLO-MIN106 for R9.4.1 flow cell) according to the manufacturer's instructions.

? TROUBLESHOOTING

Part 6: data processing and analysis ● Timing 1 d to 1 week

Data preparation

59 Download the reference genome sequences of the studied species in fasta format from NCBI, EMBL-EBI, UCSC genome browser or Ensembl. For example, mouse reference genome (GRCm38/mm10) can be obtained by:

```
wget -O GRCm38.fa.gz https://ftp.ebi.ac.uk/pub/databases/gencode/Gencode_mouse/release_M25/GRCm38.primary_assembly.genome.fa.gz

gzip -d GRCm38.fa.gz
```

60 Build a minimap2 index of the reference genome.

```
minimap2 -x map-ont -d GRCm38.mmi GRCm38.fa
```

61 Create a working folder and download the eccDNA analysis scripts.

```
mkdir eccDNA

cd eccDNA

git clone https://github.com/YiZhang-lab/eccDNA_RCA_nanopore.git
```

To run the example data, use the following commands to download the data:

```
wget -O nanopore_fast5.zip https://figshare.com/ndownloader/files/31526759
unzip nanopore_fast5.zip
rm nanopore_fast5.zip
```

Base calling and mapping

62 The sequencing data processing and mapping can be done using a single Snakemake based script with a configuration file (option A) or by running step-by-step commands (option B). The output of an Oxford Nanopore sequencing run is the raw signals stored in multiple fast5 files. `guppy` will be used to perform base calling, which would convert the raw signals to DNA sequences stored in fastq format.

! CAUTION `guppy` base calling supports running in either CPU or GPU. However, running in CPU would be very time consuming. Using a CUDA-enabled GPU would greatly reduce the base calling time.

(A) Base calling and mapping using a single script

(i) The `eccDNA_RCA_nanopore/mapping/Snakefile` is the Snakemake-based script to perform base calling and read mapping. Before running it, the configuration file `config.json` should be properly configured, using any text editor.

```
cd eccDNA_RCA_nanopore/mapping

vi config.json
```

Edit the parameters in the `config.json` according to the guidance below.

Parameter	Description
Sample	Sample name
Fast5	The absolute path to the folder of fast5 files generated by the Nanopore sequencer
Flowcell	The Nanopore flow cell used when sequencing (e.g., FLO-MIN106 for R9.4.1 flow cell)
Kit	The library construction kit used for sequencing (e.g., SQK-LSK109)
Threads	Number of threads used in the run
Reference	Path to the minimpa2 indexed reference genome ending with `.mmi`
Workdir	Absolute path to the working folder
GPU	Whether GPU is available for base calling (Y: yes; N: no)

(ii) Start the script with the following command. The `-j` option specifies the number of threads used in this run, which should be adjusted on the basis of available CPU/GPU resources.

```
snakemake -j 8 -p
```

When the script finishes running, the base calling results will be stored in the *`sample_name.fastq.gz`* file, and the mapped results in the *`sample_name.paf`* file. The quality-control reports for the Nanopore sequencing run are in the *`sample_name_qc`* folder.

(B) Step-by-step base calling and mapping

(i) Use `guppy` to generated read sequences from raw signals:

```
guppy_basecaller --input_path <path/to/fast5/> --save_path <path/to/output/folder> --flowcell <flow_cell_name> --kit <library_kit_name> --calib_detect --num_callers <threads> --trim_barcodes --trim_strategy dna --disable_pings --device auto
```

▲ **CRITICAL STEP** The above command is for running with GPU. If GPU is not available, remove the `--device auto` option to run with CPU.

! **CAUTION** When using fastq files as input, skip Step 62B(i) and start from Step 62B(ii).

(ii) Trim adaptors from reads using `porechop`:

```
porechop --extra_end_trim 0 --discard_middle -i <guppy/output/pass/> | gzip > sample_name.fastq.gz
```

(iii) Map reads with `minimap2`:

```
minimap2 -x map-ont -c --secondary=no -t <threads> <path/to/reference.mmi> sample_name.fastq.gz > sample_name.paf
```

eccDNA calling

63 Perform consensus eccDNA calling and full-length sequence reconstruction with eccDNA_RCA_nanopore.py using *sample_name*.fastq.gz and *sample_name*.paf generated in the last step as input:

```
./eccDNA_RCA_nanopore.py --fastq mapping/sample_name.fastq.gz --paf mapping/sample_name.paf --info <info.tsv> --seq <seq.fa> --var <var.tsv> --reference <path/to/reference.fa> --verbose | tee <out.log>
```

Additional parameters that can be adjusted can be found by running the command: `./eccDNA_RCA_nanopore.py -h`.

Anticipated results

By omitting or including PacI (Step 14 of 3SEP), a restriction enzyme with 8 bp recognition site that can make three cuts to linearize human mtDNA, mtDNA was either retained (Fig. 2a,b) or depleted (Fig. 2c,d) from eccDNAs obtained from over-confluent HeLa cells. Agarose gel electrophoresis revealed the stepwise DNA pattern changes when eccDNAs were enriched by 3SEP (Fig. 2a,c). In general, more than 95% of contaminating DNA present in the crude extrachromosomal circular DNA produced by alkaline lysis (Crude in lane 1 of Fig. 2a,c) can be degraded by overnight PS DNase treatment. However, the remaining 5% contaminating linear DNAs that escaped the PS DNase treatment (lane 2 of Fig. 2a,c) still dominate the eccDNA preparation when viewed by AFM imaging (Fig. 2b,d left). Nevertheless, upon further purification in part 3 of 3SEP, eccDNAs of high purity were obtained and linear DNAs were barely detected (Fig. 2b,d right). These results support the efficacy and robustness of 3SEP in eccDNA enrichment.

The purified eccDNAs are then subjected to rolling-circle amplification, library construction and Nanopore sequencing. After successfully running the eccDNA calling tool (Flec), four output files are generated. The reconstructed eccDNA information is stored in the output `info.tsv` file. Table 1 explains the format of the eccDNA calling results. Example outputs of consensus eccDNA calling are provided in Table 2. eccDNAs with a full-pass number of at least 2 are high-confidence eccDNAs and can be used for further analysis. It is possible that reads with a single pass could also derive from eccDNAs. However, they cannot be discriminated from amplification products of linear DNAs. Thus, to avoid false positives, we suggest using a minimum of two full passes to make sure that the called eccDNAs are real circular DNA. The full-length sequence of each eccDNA is available in the `seq.fa` output file in fasta format. `var.tsv` contains information about the identified genomic variants including single nucleotide variants, insertions and deletions, with six columns representing chromosome, position, reference nucleotides, alternative nucleotides, supportive coverage and total coverage, respectively. `out.log` stores the detailed reconstruction and identification process for each eccDNA (examples provided in Box 1).

Table 1 | Description of eccDNA calling results in the `info.tsv` file

Field	Description
readname	The name (id) of each read generated by Nanopore
Nfullpass	Number of full passes for this eccDNA covered by this read, when the Nfullpass is 0, it means no eccDNA was identified for this Nanopore read
Nfragment	Number of fragment(s) that form this eccDNA
refLength	The length of the reference locus that this eccDNA was mapped to
seqLength	Actual sequence length of this eccDNA
fragments	The genomic location(s) for each fragment composing this eccDNA. The coordinates of fragments are 1-based and inclusive. Multiple fragments are separated by '|'

Table 2 | Example outputs of eccDNA calling results

readname	Nfullpass	Nfragment	refLength	seqLength	fragments
1f99ee19-a07e-468b-aa76-a85c2d42e922	0	0	0	0	
de6f925d-010a-4ab5-a2ef-699d815d9904	5	1	700	700	chr1:45731400-45732099(+)
45ad109f-34e8-48a7-bd3b-8bc07adb96f3	4	1	519	520	chr2:19276179-19276697(-)
7573a6ef-1930-4d65-ad73-a03edf6c862c	4	2	955	954	chr5:99484102-99484687(-)|chr19:16574470-16574838(+)
882d8765-0a1a-4292-ac1e-d2a7d3d58f29	5	2	399	399	chr11:60300254-60300409(+)|chr7:29267126-29267368(+)

Troubleshooting

Troubleshooting advice can be found in Table 3.

Table 3 | Troubleshooting table

Step	Problem	Possible reason	Possible solution
6	Incomplete precipitation of the white SDS–potassium–DNA complex	The ratio of input cells (or other biological samples) to buffer is too high	Use more buffer to suspend the sample in Step 3, and proportionally do so in Steps 4 and 5
33, 37, 46	No eccDNA detected by either Qubit or SYBR Gold Dye	The amount of obtained eccDNA is below the detection limit of Qubit or SYBR gold	Increase input eccDNA for quantification DNA degradation, use fresh samples Increase input sample for eccDNA purification The obtained eccDNA products can be concentrated and resolved with agarose gel and SYBR Gold staining as described in Steps 38–46 to determine whether eccDNAs have been purified
40	Removing the comb draws out the agarose gel from the glass plates	This is normal	Place the glass plates and gel sandwich on the bench to let the gel move back to its original position If the gel does not automatically slip back to its original position by placing the glass plates and gel sandwich on bench, add some 1× TAE buffer to the top of the gel, wet the contact surfaces of gel and glass plates, then gently slap the bottom of the sandwich on the bench to accelerate the gel moving back Do not introduce air bubbles between the gel and glass plates, as these bubbles may interfere with electrophoresis After moving the gel back to its original position, lock the glass plates and gel sandwich into the clamping frame and proceed to load the sample and start electrophoresis. The gel cannot move any more after being locked by the clamping frame, and we have not observed sample leakage after locking in frame
58	Nanopore sequencing terminates in hours	Nanopores of the flow cell are blocked	Optimize debranching reaction in Step 53 by increasing reaction time or T7I

Timing

Steps	Tasks	Time
1–13	Isolate crude circular DNA products by buffered alkaline lysis	1.5–2 h
14–18	Digest crude circular DNA products by PS DNase	3.5 h to overnight
19–32	Selectively recover eccDNA with Solution A	1–1.5 h
33–37	Quantify eccDNA concentration	~0.5 h
38–46	Vertical agarose gel electrophoresis (optional but strongly recommended)	2 h
47–50	AFM imaging (optional but strongly recommended)	2–4 h
51–57	Rolling-circle amplification and construct Nanopore sequencing library	2–3 d
58	Nanopore sequencing	48 h
69–63	Nanopore data analyzing	1 d to 1 week

Time estimated based on the handling of one sample. Time may vary depending on the sample size and sample processing time before alkaline lysis.

Data availability

The example dataset of Nanopore sequencing of RCA eccDNA can be downloaded at https://figshare.com/articles/dataset/Nanopore_reads_of_eccDNA/17046158 or https://doi.org/10.6084/m9.figshare.17046158.v1. This example dataset is a subset of the data in Gene Expression Omnibus with accession GSM5058061.

Code availability

All scripts used for data analysis described in this paper are available at https://github.com/YiZhang-lab/eccDNA_RCA_nanopore or https://doi.org/10.5281/zenodo.7045025.

References

1. Hotta, Y. & Bassel, A. Molecular size and circularity of DNA in cells of mammals and higher plants. *Proc. Natl Acad. Sci. USA* **53**, 356–362 (1965).
2. Wang, Y. et al. eccDNAs are apoptotic products with high innate immunostimulatory activity. *Nature* **599**, 308–314 (2021).
3. Paulsen, T., Kumar, P., Koseoglu, M. M. & Dutta, A. Discoveries of extrachromosomal circles of DNA in normal and tumor cells. *Trends Genet.* **34**, 270–278 (2018).
4. Noer, J. B., Horsdal, O. K., Xiang, X., Luo, Y. & Regenberg, B. Extrachromosomal circular DNA in cancer: history, current knowledge, and methods. *Trends Genet.* **38**, 766–781 (2022).
5. Hirt, B. Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* **26**, 365–369 (1967).
6. Smith, C. A. & Vinograd, J. Small polydisperse circular DNA of HeLa cells. *J. Mol. Biol.* **69**, 163–178 (1972).
7. Goebel, W. & Schrempf, H. Isolation and characterization of supercoiled circular deoxyribonucleic acid from beta-hemolytic strains of *Escherichia coli*. *J. Bacteriol.* **106**, 311–317 (1971).
8. Mukai, T., Matsubara, K. & Takagi, Y. Isolation of circular DNA molecules from whole cellular DNA by use of ATP-dependent deoxyribonuclease. *Proc. Natl Acad. Sci. USA* **70**, 2884–2887 (1973).
9. Yamagishi, H. et al. Purification of small polydisperse circular DNA of eukaryotic cells by use of ATP-dependent deoxyribonuclease. *Gene* **26**, 317–321 (1983).
10. Shoura, M. J. et al. Intricate and cell type-specific populations of endogenous circular DNA (eccDNA) in *Caenorhabditis elegans* and *Homo sapiens*. *G3* **7**, 3295–3303 (2017).
11. Radloff, R., Bauer, W. & Vinograd, J. A dye-buoyant-density method for the detection and isolation of closed circular duplex DNA: the closed circular DNA in HeLa cells. *Proc. Natl Acad. Sci. USA* **57**, 1514–1521 (1967).
12. Hardman, N. Formation of rings from segments of HeLa-cell nuclear deoxyribonucleic acid. *Biochem. J.* **143**, 521–534 (1974).
13. Shibata, Y. et al. Extrachromosomal microDNAs and chromosomal microdeletions in normal tissues. *Science* **336**, 82–86 (2012).
14. Moller, H. D., Parsons, L., Jorgensen, T. S., Botstein, D. & Regenberg, B. Extrachromosomal circular DNA is common in yeast. *Proc. Natl Acad. Sci. USA* **112**, E3114–E3122 (2015).
15. Moller, H. D. et al. Genome-wide purification of extrachromosomal circular DNA from eukaryotic cells. *J. Vis. Exp.* https://doi.org/10.3791/54239 (2016).
16. Moller, H. D. Circle-seq: isolation and sequencing of chromosome-derived circular DNA elements in cells. *Methods Mol. Biol.* **2119**, 165–181 (2020).

17. Henssen, A., MacArthur, I., Koche, R. & Dorado-García, H. Purification and sequencing of large circular DNA from human cells. *Protocol Exchange* https://doi.org/10.1038/protex.2019.006 (2019).
18. Guo, L. H. & Wu, R. New rapid methods for DNA sequencing based in exonuclease III digestion followed by repair synthesis. *Nucleic Acids Res.* **10**, 2065–2084 (1982).
19. Machwe, A., Ganunis, R., Bohr, V. A. & Orren, D. K. Selective blockage of the 3′—>5′ exonuclease activity of WRN protein by certain oxidative modifications and bulky lesions in DNA. *Nucleic Acids Res.* **28**, 2762–2770 (2000).
20. Pizzolato, J., Mukherjee, S., Scharer, O. D. & Jiricny, J. FANCD2-associated nuclease 1, but not exonuclease 1 or flap endonuclease 1, is able to unhook DNA interstrand cross-links in vitro. *J. Biol. Chem.* **290**, 22602–22611 (2015).
21. Coté, A.G. In vivo analysis of cruciform extrusion and resolution of DNA palindromes in eukaryotes. PhD thesis, Univ. Toronto (2009).
22. Dillon, L. W. et al. Production of extrachromosomal microDNAs is linked to mismatch repair pathways and transcriptional activity. *Cell Rep.* **11**, 1749–1759 (2015).
23. Ali, N., Rampazzo, R. C. P., Costa, A. D. T. & Krieger, M. A. Current nucleic acid extraction methods and their implications to point-of-care diagnostics. *Biomed. Res. Int.* **2017**, 9306564 (2017).
24. Cloninger, C., Felton, M., Paul, B., Hirakawa, Y. & Metzenberg, S. Control of pH during plasmid preparation by alkaline lysis of *Escherichia coli*. *Anal. Biochem.* **378**, 224–225 (2008).
25. Moller, H. D. et al. Circular DNA elements of chromosomal origin are common in healthy human somatic tissue. *Nat. Commun.* **9**, 1069 (2018).
26. Deshpande, V. et al. Exploring the landscape of focal amplifications in cancer using AmpliconArchitect. *Nat. Commun.* **10**, 392 (2019).
27. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
28. Koster, J. & Rahmann, S. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics* **28**, 2520–2522 (2012).

Acknowledgements

We thank P. Yin for providing the AFM instrument. This work was supported by the Howard Hughes Medical Institute. Y.Z. is an Investigator at the Howard Hughes Medical Institute.

Author contributions

Y.Z. conceived the project; Y.W. and Y.Z. designed the experiments; Y.W. performed the experiments; M.W. developed the bioinformatics analysis tools. Y.W., M.W. and Y.Z. wrote the manuscript.

Competing interests

A patent application covering the eccDNA purification method and eccDNA's application as an immunostimulant and diagnostic marker has been filed by the Boston Children's Medical Center Corporation.

Additional information

Correspondence and requests for materials should be addressed to Yi Zhang.

Peer review information *Nature Protocols* thanks Paul Mischel, Chia-Lin Wei and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Received: 22 December 2021; Accepted: 23 September 2022;
Published online: 14 December 2022

Related links
Key reference using this protocol
Wang, Y. et al. *Nature* **599**, 308–314 (2021): https://doi.org/10.1038/s41586-021-04009-w