nature research

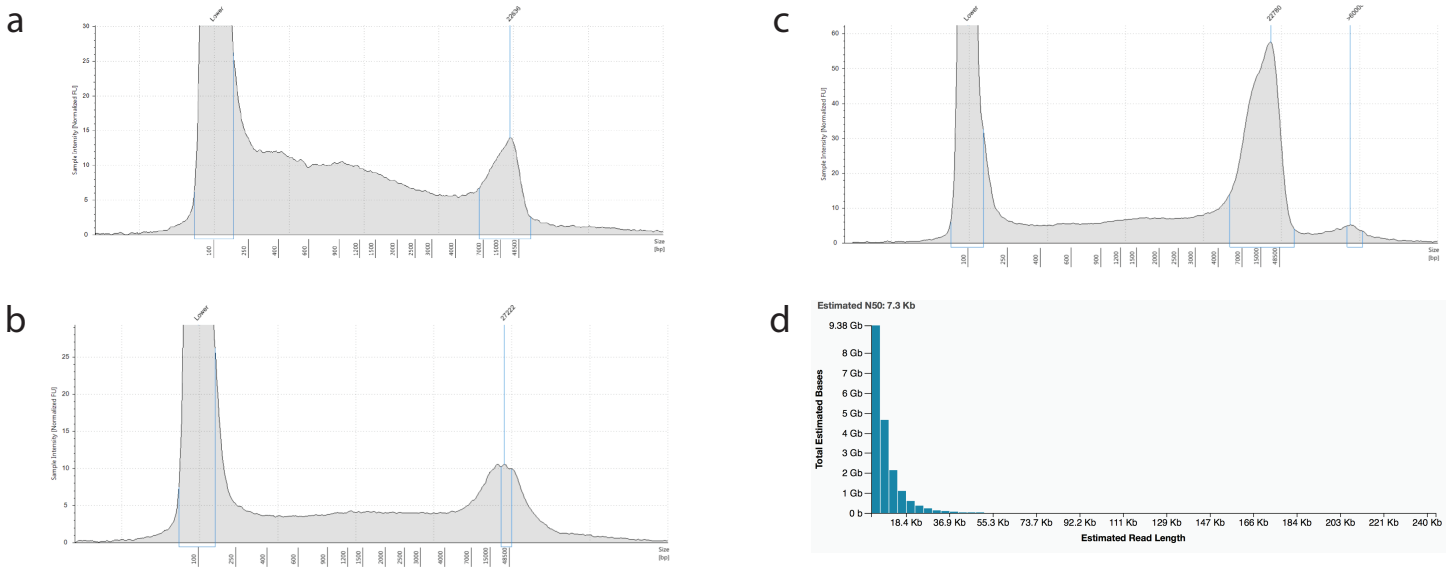
https://doi.org/10.1038/s41596-020-00424-x

Supplementary information

Improved high-molecular-weight DNA extraction, nanopore sequencing and metagenomic assembly from the human gut microbiome

In the format provided by the authors and unedited

Supplementary Figure 1: TapeStation traces for various DNA extraction methods



TapeStation traces of DNA fragment lengths for three samples, where the x-axis represents fragment size, the y-axis represents relative fluorescence at that length, and the peak centered at 100 bp indicates a molecular weight standard. A) A sample extracted with standard bead beating approaches that is not recommended for nanopore sequencing, with no clear peak in DNA fragment length and excessive mass below 2.5 kb. B) A sample extracted with the presented method that could be used directly for sequencing or optionally size selected once more, which has a clear peak above 15 kb but includes mass below 2.5 kb. C) A sample extracted with the presented method that can be used directly for sequencing, with a clear peak above 15 kb and minimal mass below 2.5 kb. D) Read length distribution from sample in panel (C) as displayed in MinKNOW when run on an Oxford Nanopore MinION RevD R9.4 flow cell.

Supplementary Note 1: Abbreviated high molecular weight extraction protocol

Here, we provide an abbreviated protocol for high molecular weight DNA extraction. This protocol substitutes the Qiagen Genomic-tip purification with a MicroSpin S-400 HR column purification, which is a faster column protocol. Additionally, the RNase A and proteinase K digestion steps are incorporated into other incubation steps, and the sample is incubated in SDS for more effective denaturation. Altogether, this protocol can be performed in approximately 4 hours.

Alternative reagents:

- MicroSpin S-400 HR column (Millipore Sigma cat. No. 27-5140-01)

Protocol:

1. Keeping stool sample on dry ice as much as possible to maintain sample integrity, use a biopsy punch to aliquot 150 mg stool into a 2 ml microcentrifuge tube. Suspend the sample in 500 µl PBS and vortex for 3-4 seconds to mix. For lower biomass stool samples, aliquot up to 300 mg stool.
2. Add 5 µL Qiagen lytic enzyme solution and 2 µl MetaPolyzyme to the stool suspension. Mix by inverting six times slowly and gently. Incubate the mixture in a 37°C heat block for 30 min.
3. Add 2 µL RNAse A and invert to mix. Incubate for 30 minutes at 37C. Add 12 µL 20% SDS and 2 µL Proteinase K and invert to mix. Incubate for 30 minutes at 56C.
4. In a fume hood, add 500 µl Phenol/Chloroform pH 8. Add approximately 100 µl of phase-lock gel to the microcentrifuge tube. Alternatively, add approximately 100 µl of phase-

lock gel to the inside cap of the microcentrifuge tube rather than directly into the tube for ease of application.

5. Place tubes into the multi-position vortexer and vortex for 5 seconds at minimum speed. Centrifuge the tube for 5 min at 10,000*g* at room temperature. Decant the aqueous phase into a new 2 mL microcentrifuge tube.
6. Add 90 µl 3M sodium acetate and 500 µl isopropanol. Invert the tube thrice slowly to mix. Incubate the mixture at room temperature for 10 min.
7. Spin the tube for 10 min at 10,000*g* at room temperature, making sure that the hinge is facing the outside edge. While being very careful not to disrupt the pellet, remove and discard the supernatant and allow the pellet to air dry. Resuspend pellet in 110 µl nuclease free water.
8. Break bottom off of MicroSpin S-400 HR column and loosen cap by a quarter turn. Spin at 1 min at 500*g*.
9. Place the column in a fresh microcentrifuge tube and apply all of the sample to the column. Spin for 2 min at 500*g*.
10. Prepare beads in a custom buffer as has been described. Add 0.8 volumes (80 µL) of the custom bead suspension to the tube and gently flick to mix. Incubate the tube for 10 min on a Hula mixer at room temperature.

    CRITICAL STEP Bead suspension to sample ratio will vary with each preparation of the custom buffer. Test the selection stringency of each bead preparation with a non-precious sample to ensure proper selection.

11. Spin the tube down briefly and place the tube on a magnetic rack to pellet beads. Wait for approximately 3 min, or until the solution has become clear. Carefully remove the supernatant with a pipette. Wash pelleted beads with 200 µL freshly prepared 80%

ethanol, then pipette off ethanol. Repeat the wash step once more. Remove the tube from the magnetic rack, spin it down quickly, place the tube back on the magnetic rack, and pipette off any residual ethanol. Air dry the beads for 30 seconds.

CRITICAL STEP Do not overdry the beads, as this may negatively impact DNA recovery.

12. Remove the tube from the magnetic rack and resuspend beads in 50 µL nuclease-free water. Optionally resuspend in 15 µL nuclease-free water if proceeding with the Rapid Sequencing library preparation protocol. Incubate the suspension for 10 min at 37°C. Pellet the beads on the magnetic rack and transfer the eluent to a fresh microcentrifuge tube.

    PAUSE POINT The extracted DNA can be stored at 4°C for several months

13. Quantify the DNA concentration using a Qubit. The suggested minimum concentration is 20 ng/µL.
14. Quantify the DNA purity using a nanodrop. The suggested purity is A260/A230 > 2, A260/A280 > 1.8.
15. Quantify the DNA size distribution with a TapeStation. The suggested size distribution is a major peak mean greater than 15 kilobases, with minimal mass (<50 fluorescence units) below 2.5 kilobases.

    PAUSE POINT The extracted DNA can be stored at 4°C for several months

Supplementary Note 2. Lathe workflow output directory structure

Output directory of the Lathe workflow includes folders with output files from basecalling, assembly, polishing, and circularization. Subfolders are indicated in blue. Files commonly used for downstream applications include the basecalled reads (sample.fq), the circularized genomes (3.circular_sequences), and the final assembly (sample_final.fa), which are indicated with italics.

0.basecall
-- *sample.fq*
-- data_links
-- nanoplots
1.assemble
-- assemble_100m (if specified)
-- assemble_250m (if specified)
-- sample_merged.fasta
-- sample_raw_assembly.fa
-- sample_raw_assembly.fa.amb
-- sample_raw_assembly.fa.ann
-- sample_raw_assembly.fa.bwt
-- sample_raw_assembly.fa.fai
-- sample_raw_assembly.fa.pac
-- sample_raw_assembly.fa.sa
2.polish
-- sample_polished.fasta
-- sample_polished.fasta.bam
-- sample_polished.fasta.bam.bai
-- sample_polished.fasta.fai
-- pilon (if specified)
-- racon (if specified)
-- medaka (if specified)
3.circularization
-- 1.candidate_genomes
-- 2.circularization
-- *3.circular_sequences*
-- 4.sample_circularized.corrected.fasta
-- 4.sample_circularized.fasta
-- 4.sample_circularized.fasta.bam
-- 4.sample_circularized.fasta.bam.bai
-- 4.sample_circularized.fasta.fai
-- 4.sample_circularized.fasta.misassemblies.tsv
5.final
-- *sample_final.fa*