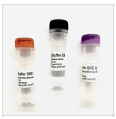


Apr 09, 2020

DNA clean-up and size selection for long-read sequencing V.3

Ashley Jones[1], Neeraj Purushotham[2], Jamila Nasim[1], Benjamin Schwessinger[1]

[1]Research School of Biology, Australian National University, Canberra, ACT, Australia, [2]Institute for Agriculture and the Environment, Research and Innovation Division, University of Southern Queensland, Toowoomba, QLD, Australia

1 *Works for me* dx.doi.org/10.17504/protocols.io.betdjei6

High molecular weight DNA extraction from all kingdoms
Tech. support email: **See@each.protocol**


Ashley Jones
Australian National University

ABSTRACT

DNA extractions often contain impurities which limit the output of long-read sequencing technologies. Here a protocol is provided which removes impurities and size selects for longer fragments. To remove residual RNA and protein, an additional RNAse A and Proteinase K treatment is performed. A clean-up with chloroform: isoamyl alcohol (24:1) removes these proteins and other hydrophobic organics such as lipids. A low volume ethanol precipitation and wash is used to concentrate the DNA, hopefully also reducing polysaccharides. An optional needle shearing is described which can help create a more uniform DNA length to maximise sequencing output. Finally, a Short-Read Eliminator (SRE) kit by Circulomics is utilised for size selection, which also appears to clean the DNA. This was trialled for the sorghum rot fungus *Macrophomina phaseolina*, providing highly promising results with an Oxford Nanopore MinION. One strain yielded 13.71 Gbases with an N50 of 21.75 kb, another strain yielded 9.72 Gbases with an N50 of 43.50 kb. Similar results have been obtained with other fungi, plants, reptiles and insects.

GUIDELINES

This protocol is optimsed for high-molecular weight DNA. For best results, avoid column based DNA extractions (which shear the DNA) and avoid using phenol (interferes with sequencing, can irreversibly bind to DNA).

MATERIALS TEXT

Chloroform: isoamyl alcohol (24:1 v/v)
1.5 mL Eppendorf DNA LoBind tubes
Ethanol (100% and 70%)
Needle (29 or 26 gauge) with syringe (optional shearing)
Proteinase K (20 mg/mL)
RNAse A (20 mg/mL)
Short Read Eliminator kit (Circulomics)
3 M Sodium acetate pH 5.2
10 mM Tris-HCl pH 8 (nuclease-free)
Water (nuclease-free)

Citation: Ashley Jones, Neeraj Purushotham, Jamila Nasim, Benjamin Schwessinger (04/09/2020). DNA clean-up and size selection for long-read sequencing. https://dx.doi.org/10.17504/protocols.io.betdjei6

RNA AND PROTEIN REMOVAL

1 Aliquot 10-30 µg of DNA into a 1.5 mL eppendorf tube. Increase volume to 200 µL with 10 mM Tris-HCl pH 8.

- *DNA quantification must be based on Qubit Fluorometer (Thermo Fisher Scientific), or similar device. A high quantity of RNA may be present.*
- *Volume can exceed 200 µL, maximum of 600 µL due to tube capacity at later steps (ethanol precipitation).*

2 Add RNAse A and Proteinase K

Enzyme	Stock	Quantity
100 µg/mL RNAse A	20 mg/mL	1 µL
100 µg/mL Proteinase K	20 mg/mL	1 µL

3 Incubate the samples at 50-60°C for 20 min, shaking at 400-900 rpm if possible.

CHLOROFORM CLEAN

4 Increase the volume to 600 µL with 10 mM Tris-HCl pH 8 (add 400 µL).

5 Add an equal volume of chloroform: isoamyl alcohol (24:1, v/v) (600 µL) and mix by inverting 10-15 times. Ensure the organic and aqueous phases become mixed at least temporarily.

6 Separate the phases by centrifuging at 16,000 rcf for 1 min at 20°C.

7 Transfer the upper aqueous phase to a new 1.5 mL Eppendorf tube.

8 Repeat the chloroform: isoamyl alcohol clean (equal volume).

DNA PRECIPITATION

9 Add 1.5x volume of 100% ethanol (~900 µL) and 0.1x volume of 3 M sodium acetate pH 5.2 (~60 µL). Incubate on ice for 1 min.

Sample volume should be a maximum 600 µL after cleaning with chloroform: isoamyl alcohol.

10 Centrifuge at 16,000 rcf for 1 min. Carefully decant the supernatant as soon as possible, without disturbing the pellet.

High-molecular weight DNA should pellet easily, thus could pulse spin for 10 s at maximum speed. If no pellet can be seen, centrifuge for longer. Bringing ethanol concentration to 2x volume is also an option.

11 Add approximately 700 µL of freshly prepared 70% ethanol, enough to cover the pellet. Let the pellet soak for 1 min at room temperature to dissolve excess salts.

Citation: Ashley Jones, Neeraj Purushotham, Jamila Nasim, Benjamin Schwessinger (04/09/2020). DNA clean-up and size selection for long-read sequencing. https://dx.doi.org/10.17504/protocols.io.betdjei6

12 Centrifuge at 16,000 rcf for 1 min at 4°C. Carefully decant the supernatant as soon as possible, without disturbing the pellet.

13 Repeat previous steps for a second 70% ethanol wash.

14 Air-dry the pellet by placing the tube upside-down on tissue paper for 5-15 min (until all ethanol has evaporated). Be careful of water touching or dislodging the pellet.

15 Dissolve DNA with 60 µL of 10 mM Tris-HCl pH 8.

16 Quantify the DNA on a Nanodrop and a Qubit fluorometer (dsDNA broad-range assay).

NEEDLE SHEARING (OPTIONAL)

17 Needle shearing is optional and helps provide a more uniform DNA length. This can maximise long-read sequencing output at the expense length. For instance, repetitive polyploid plant DNA over 25 kb can block nanopores, stopping sequencing prematurely.

Choose a needle and a syringe; increasing needle gauge (smaller diameter), causes more DNA shearing.

Needle	Passes	Estimated N50
29 gauge	5	10-20 kb
26 gauge	5	20-35 kb

- *No difference was found between 5, 10, 15 and 20 passes using the 29 gauge needle with plant DNA. The 26 gauge needle has not been extensivley tested.*
- *The chloroform clean-up and DNA precipitation from previous sections has already caused some DNA shearing.*

18 Aliquot 3-9 µg of DNA into a 1.5 mL Eppendorf DNA LoBind tube. Adjust the volume to exactly 60 µL using 10 mM Tris-HCl (pH 8).

19 Connect the needle to a syringe and perform 5 passes with DNA (aspirate and dispense 5 times). Proceed directly to short-read elimination.

SHORT-READ ELIMINATION

20 Choose a Short-Read Eliminator (SRE) product supplied by Circulomics. Recovery estimated based on high-molecular weight DNA from non-model species. Expect higher loss if DNA is significantly sheared.

Product	Catalogue number	Size selection	Advertised recovery	Realistic recovery
SRE XS	SKU SS-100-121-01	≥ 10 kb	50-90%	50-75%
SRE	SKU SS-100-101-01	≥ 25 kb	50-70%	30-60%
SRE XL	SKU SS-100-111-01	≥ 40 kb	40-50%	< 45%

Citation: Ashley Jones, Neeraj Purushotham, Jamila Nasim, Benjamin Schwessinger (04/09/2020). DNA clean-up and size selection for long-read sequencing. https://dx.doi.org/10.17504/protocols.io.betdjei6

21 Aliquot 3-9 µg of DNA into a 1.5 mL Eppendorf DNA LoBind tube. Adjust the volume to exactly 60 µL using 10 mM Tris-HCl (pH 8) or Buffer EB (supplied by Circulomics).

- *DNA mass must be measured by Qubit Fluorometer (Thermo Fisher Scientific) or equivalent.*
- *DNA must not contain high levels of salts, polyphenols, polysaccharides or other contaminants.*

22 Add 60 µL of Buffer SRE to the sample (supplied by Circulomics). Mix thoroughly by gently tapping the tube or by gently pipetting up and down.

23 Centrifuge at 10,000 rcf for 30 min at room temperature. Note where the DNA will pellet, by marking the tube or placing the tubes in a consistent orientation within the rotor.

Recovery will be impacted if centrifugation is performed at low temperature (e.g. 4°C).

24 Carefully remove the supernatant with a pipette as soon as possible, without disturbing the pellet. Care must be taken as the pellet is fragile, being easily dislodged.

The DNA pellet may not be visible. Be careful of accidentally losing the pellet; mark the expected spot.

25 Add 200 µL of freshly prepared 70% ethanol, enough to cover the pellet. Do not tap or mix after adding ethanol.

26 Centrifuge at 10,000 rcf for 2 min at room temperature. Carefully remove the supernatant with a pipette as soon as possible, without disturbing the pellet.

27 Repeat previous steps for a second 70% ethanol wash.

28 Air-dry the pellet by placing the tube upside-down on tissue paper for 5-15 min (until all ethanol has evaporated). Be careful of water touching or dislodging the pellet.

29 Add 50 µL of Buffer EB (supplied by Circulomics). Mix by gently tapping the tube. Pellet should hydrate at room temperature within 10 min. If not, incubate at 50°C for 10 min.

Original Circulomics protocol recommends 50°C for 1 h (excessive). DNA concentration increased < 8%. Limit DNA shearing by avoiding high temperatures. Pellet will hydrate further when stored at 4°C.

30 After incubation, gently tap the tube to ensure that the DNA is properly resuspended and mixed.

31 Quantify the DNA on a Nanodrop and a Qubit fluorometer (dsDNA broad-range assay).

Citation: Ashley Jones, Neeraj Purushotham, Jamila Nasim, Benjamin Schwessinger (04/09/2020). DNA clean-up and size selection for long-read sequencing. https://dx.doi.org/10.17504/protocols.io.betdjei6

32 Store DNA at 4°C to prevent cycles of freeze-thawing that shear the DNA.

No effects on DNA integrity have been noticed for samples stored at 4°C for extended periods.

33 **EXPECTED RESULTS**

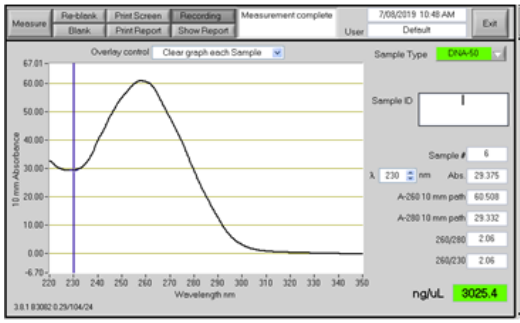
Fresh mycelia (approximately 2 g) was used to extract crude high-molecular weight DNA using a CTAB method previously described. See Jones *et al.* (2019) High-molecular weight DNA extraction from challenging fungi using CTAB and gel purification, Protocols.io. The DNA was cleaned and then size selected with a Short-Read Eliminator kit (SRE), as described in this protocol. Initially the Qubit:Nanodrop ratio had a large discrepancy (Table 1 and Figure 1). This is due to the presence of RNA, given the 260/280 is above 1.8. Cleaning removed this discrepancy, particularly with the removal of all RNA. With RNA not present, the 260/230 ratio is more representative, despite appearing worse.
For sequencing, we used a MinION sequencer from Oxford Nanopore Technologies, creating a native genomic DNA library (SQK-LSK109) following the manufacturer's instructions. Although the DNA was not perfect, a high sequencing output was achieved (Table 2). Size selection with SRE resulted in a long-reads and a high N50 (Figure 2).

Citation: Ashley Jones, Neeraj Purushotham, Jamila Nasim, Benjamin Schwessinger (04/09/2020). DNA clean-up and size selection for long-read sequencing. https://dx.doi.org/10.17504/protocols.io.betdjei6

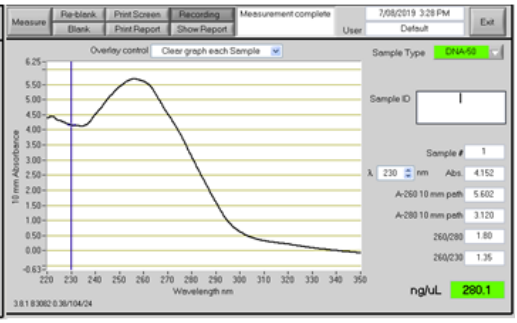
Table 1: Quantification of the crude DNA, DNA after cleaning and after size selection.

Sample	DNA input µg	SRE kit	Qubit ng/µL	Nano ng/µL	Qubit: Nano	260/ 280	260/ 230	Vol µL	Yield µg (Qubit)
Mp strain 1 crude DNA	NA	NA	64.30	526.80	1:8.19	2.05	2.09	300	19.29
Mp strain 1 clean DNA	19.00	NA	271	431.80	1:1.59	1.92	2.16	50	13.55
Mp strain 1 clean + SRE	9.00	25 kb	74.10	97.60	1:1.31	1.88	1.67	50	3.71
Mp strain 2 crude DNA	NA	NA	258	3,025	1:11.72	2.06	2.06	300	77.40
Mp strain 2 clean DNA	19.00	NA	272	280.10	1:1.03	1.80	1.35	50	13.60
Mp strain 2 clean + SRE	9.00	25 kb	94	146.10	1:1.55	1.82	1.29	50	4.70

(A) *M. phaseolina* before clean



(B) *M. phaseolina* after chloroform cleaning



(B) *M. phaseolina* after clean and short read eliminator

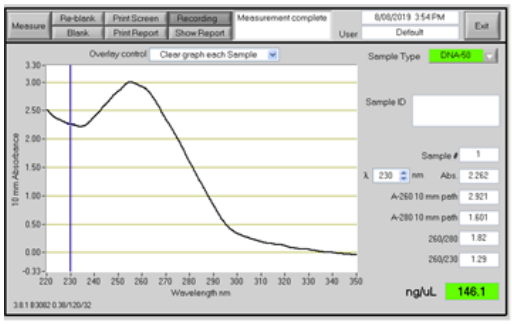


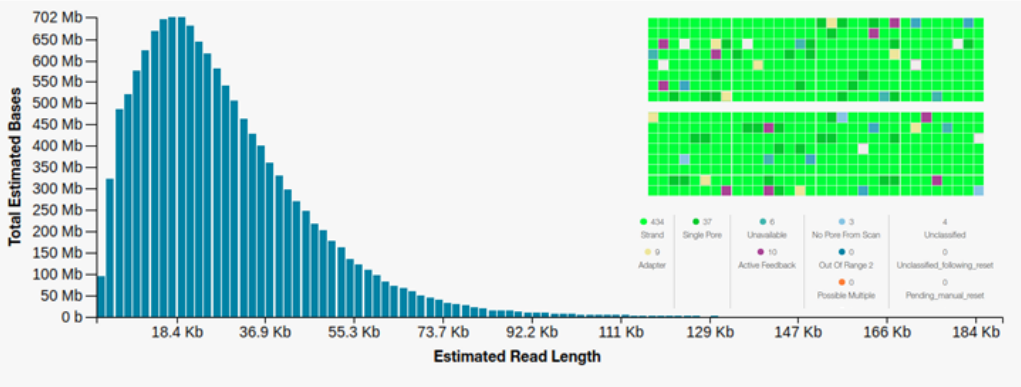
Figure 1: Spectrophotometer results of DNA at multiple stages of the protocol. Strain 2 is shown as an example. (A) Initial DNA before cleaning. (B) DNA after an initial clean with chloroform and ethanol precipitation. (C) DNA after cleaning with chloroform and size selecting with a short read eliminator kit (Circulomics). Readings taken using 1 µL on a Thermo Scientific Nanodrop 1000.

Citation: Ashley Jones, Neeraj Purushotham, Jamila Nasim, Benjamin Schwessinger (04/09/2020). DNA clean-up and size selection for long-read sequencing. https://dx.doi.org/10.17504/protocols.io.betdjei6

Table 2: Sequencing results with a single MinION flow cell per sample (FLO-MIN 106 R9.4.1 revD).

Sample	SRE ≥	Library input µg	Loaded µg	Library preparation	Pores	Output Gb	Base call Gb	N50 kb
Mp strain 1	25 kb	3.00	1.00	Ligation SQK-LSK109	1,557	13.71	13.29	21.75
Mp strain 2	25 kb	3.00	0.94	Ligation SQK-LSK109	1,362	9.72	9.57	43.50

(A) Sorghum rot *Macrophomina phaseolina* strain 1; short read eliminator 25 kb, 13.71 Gb, N50 21.75 kb.



(B) Sorghum rot *Macrophomina phaseolina* strain 2; short read eliminator 25 kb, 9.72 Gb, N50 43.50 kb.

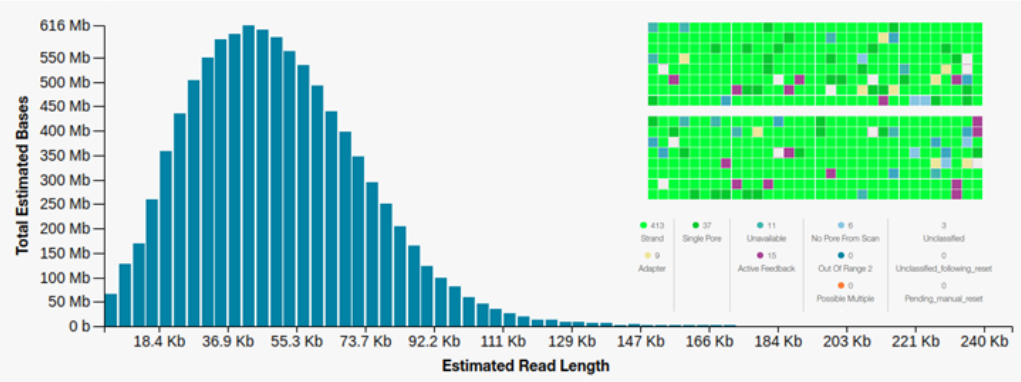


Figure 2: Expected read length histograms on MinKNOW. Both sorghum rot *Macrophomina phaseolina* strains were size selected with a short read eliminator kit (Circulomics) for 25 kb and above. DNA libraries were then prepared with an end ligation kit (Oxford Nanopore SQK-LSK109). Inserts show pore usage, light green indicates pore is active and sequencing is occurring.

Citation: Ashley Jones, Neeraj Purushotham, Jamila Nasim, Benjamin Schwessinger (04/09/2020). DNA clean-up and size selection for long-read sequencing. https://dx.doi.org/10.17504/protocols.io.betdjei6