Automated phenol-chloroform extraction of high molecular weight genomic DNA for use in long-read single-molecule sequencing [version 1; peer review: 1 approved]

Andrew W. Liu1, Alejandro Villar-Briones2, Nicholas M. Luscombe1, Charles Plessy1

1Genomics and Regulatory Systems Unit, Okinawa Institute of Science and Technology, Onna-son, Okinawa, 904-0495, Japan
2Instrument Analysis Section, Okinawa Institute of Science and Technology, Onna-son, Okinawa, 904-0495, Japan

Abstract

Background: Automation has increasingly become more commonplace in the research laboratory workspace. The introduction of articulated robotic arms allows the researcher more flexibility in the tasks a single piece of automated machinery can perform. We set out to incorporate automation in processing of genomic DNA organic extractions to increase throughput and limit researchers to the exposure of organic solvents.

Methods: In order to automate the genome sequencing pipeline in our laboratory, we programmed a dual-arm anthropomorphic robot, the Robotic Biology Institute's Maholo LabDroid, to perform organic solvent-based genomic DNA extraction from cell lysates. To the best of our knowledge, this is the first time that automation of phenol-chloroform extraction has been reported.

Results: We achieved routine extraction of high molecular weight genomic DNA (>100 kb) from diverse biological samples including algae cultured in sea water, bacteria, whole insects, and human cell lines. The results of pulse-field electrophoresis size analysis and the N50 sequencing metrics of reads obtained from Nanopore MinION runs verified the presence of intact DNA suitable for direct sequencing.

Conclusions: We present the workflow that can be used to program similar robots and discuss the problems and solutions we encountered in developing the workflow. The protocol can be adapted to analogous methods such as RNA extraction, and there is ongoing work to incorporate further post-extraction steps such as library construction. This work shows the potential for automated robotic workflows to free molecular biological researchers from manual interventions in routine experimental work. A time-lapse movie of the entire automated run is included in this report.
Keywords
Automation, Robotics, Organic Extraction, DNA extraction, Genome Sequencing

This article is included in the Cell & Molecular Biology gateway.
Introduction

DNA extraction is a routine, yet critical step in genomics research. Across diverse settings and applications, the purity of isolated DNA affects the sensitivity of downstream processes such as polymerase chain reaction (PCR) and sequencing. In addition, recent developments in sequencing technologies have increased the emphasis of extracting intact or high-molecular weight genomic DNA from a wide variety of biological materials. The Oxford Nanopore Technologies (Kasianowicz et al., 1996; Jain et al., 2018) and PacBio Single Molecule Real-Time (Eid et al., 2009; Chin et al., 2013) platforms enable users to obtain long sequence reads – a major limiting factor being the availability of long DNA molecules. Sequencing data from such platforms are routinely used to generate whole genome sequence assemblies.

Organic, or phenol-chloroform, extraction is one of the oldest and, for many years, the most widely used method for DNA isolation. A phenol-chloroform mixture is added to lysed or homogenised biological material. When centrifuged, the unwanted proteins and cellular debris are separated away in the organic phase, and DNA molecules in the aqueous phase can be cleanly isolated for analysis. Among the many applications, organic extraction was historically used to isolate nucleic acids from viruses, which had proved challenging owing to the chemically resistant protein coats surrounding their genomes (Sinsheimer, 1959; Thomas & Berns, 1961; Saito & Miura, 1963). Despite the development of modern extraction approaches, the phenol-chloroform method continues to be relevant because it works reliably for many biological samples and consistently gives high yields of high molecular weight DNA (Ghaheri et al., 2016; Bouso & Planet, 2019; Torii et al., 2021). However, it is also time-consuming, involves the use of hazardous organic solvents, and requires samples to be transferred between multiple tubes, increasing the risk of error or contamination.

Here, to establish consistent isolation of high molecular weight genomic DNA and reduce the amount of manual work needed for a routine protocol involving toxic reagents, we programmed the Robotic Biology Institute's Maholo LabDroid, a dual-arm anthropomorphic robot, to perform organic genomic DNA extraction from lysed cells and homogenized tissues.

The LabDroid is an example of a versatile modular system that has demonstrated a practical application in a laboratory environment (Ochiai et al., 2020). Maholo’s design allows users to design a workspace/equipment layout within the reach of the robot (Figure 1). Automating the procedure was expected to enhance reproducibility, increase throughput, and reduce the possibility of DNA shearing into shorter fragments by minimizing excessive handling. Devices for automated DNA isolation are available, but to the best of our knowledge, this is the first report that a phenol-chloroform extraction has been automated.

Methods

Ethical considerations

The human cell lines, laboratory animals and microbial stocks used in this study do not require ethics approval in accordance with Okinawa Institute of Science & Technology Graduate University (OIST)’s Animal Care and Use Committee and Japan Society for the Promotion of Science (JSPS) guidelines. All procedures regarding the welfare and handling of organisms meet ARRIVE guidelines for the care and use of experimental animals. In particular, Drosophila were anaesthetized with carbon dioxide (CO₂) prior to mechanical cell disruption.

Sample preparation

Samples were manually homogenized either by grinding using a simple tube-based mortar and pestle (single whole eclosed fruit fly) or freeze-thawing five times (cultured cyanobacteria and algae). The human cell culture was not subjected to mechanical disruption.

Three newly eclosed single X7 Drosophila melanogaster (RRID:BDSC_5478) were isolated in vials containing food but no yeast for 18 hours, put to sleep using CO₂, and then individually ground gently in 200 μl ice-cold extraction buffer (50mM Tris-Cl pH 8.0, 100 mM NaCl, 2.5 mM EDTA, 0.5% SDS, 0.2 mg/ml Proteinase-K, 1% 2-mercaptoethanol) with a tube-based mortar & pestle on ice. The homogenized extract was incubated overnight for 18 hours at 55°C. The cell lysate was centrifuged at 11,000g (Tomy, MX-302) for 10 minutes at 4°C and 200 μl of the cleared lysate used for robotic extraction.

Cultured human embryonic kidney cells, HEK293 (RRID: CVCL_45), were dissociated from a confluent tissue culture flask with 0.25% Trypsin-EDTA for 3 minutes at 37°C (Gibco, 25200056), counted on a hemocytometer, centrifuged at 800g for 5 minutes and washed with 10 ml phosphate buffered saline (PBS). 1 × 10⁶ cells were drawn off and spun again at 800g for 5 minutes. The PBS was aspirated, and the cell pellet was resuspended in extraction buffer 50 mM Tris-Cl pH 8.0, 100 mM NaCl, 10 mM ethylenediaminetetraacetic acid (EDTA), 0.5% SDS, 0.2 mg/ml Proteinase-K, 1% 2-mercaptoethanol and incubated for 2 hours at 55°C. The cell lysate was processed immediately by centrifugation at 11,000g for 10 minutes at 4°C and 200 μl of the cleared lysate used for robotic extraction.
The cyanobacterium Synechococcus sp. was centrifuged at 2,500g for 5 minutes from a one-month-old culture; the resulting pellet was washed with filtered autoclaved seawater three times and gently resuspended in 250 μl of extraction buffer (50 mM Tris-Cl pH 8.0, 100 mM NaCl, 20 mM EDTA, 1% SDS, 2% 2-mercaptoethanol) and frozen overnight at –80°C. The next morning the cyanobacteria were thawed on ice, an equal volume of 5% SDS, 50 mM EDTA & 50 mM Tris-Cl pH 8 was mixed by inversion and subjected to four more freeze/thaw cycles. Mixed algal cultures of 6 × 10^6 Chaetoceros calcitrans, 1 × 10^7 Isochrysis sp. and 1 × 10^6 Rhinomonas reticulata were centrifuged similarly from a one-month-old culture; the resulting pellet was washed with filtered autoclaved seawater three times and gently resuspended in 250 μl of ice-cold extraction buffer (100 mM Tris-Cl pH 8.0, 1.4 M NaCl, 20 mM EDTA, 2% cetyltrimethylammonium bromide, 0.1% polyvinylpolypyrrolidone, 2% 2-mercaptoethanol) and then frozen overnight at –80°C. The next morning, the algal extractions were thawed on ice and subjected to four more freeze/thaw cycles. All extractions were centrifuged at 4°C for 10 min at 15,000g and 200 μl of the cleared lysate used for robotic extraction.

**Robot-assisted organic extraction**

A workflow based on a modified organic solvent based back-extraction method (Giles et al., 1980) was programmed into the LabDroid with Robotic Biology Institute’s BioPortal Protocol Maker graphical user interface (Liu and Plessy, 2022b). Though the robot is capable of processing 14 samples in a run, the development was carried out with three samples. The procedural schematic and robotic workflow are outlined in Figure 2A and B, respectively. To start, 200 μl of cleared lysate was transferred to a new 2 ml tube (Eppendorf 0030 120.094) and used for the robot-assisted protocol. 200 μl of Phenol, chloroform and isoamyl alcohol (25:24:1, Sigma-Aldrich P3803) was used for the first two robot “hand-mixed” organic extractions. The closed-capped tubes were mixed by the robot’s arm/manipulators by inversion at a rate of one every 3 seconds, for a total of 5 minutes. These were later centrifuged at 15,000g for 1 minute at 4°C. The first extraction aqueous phase was transferred to a new 2 ml Eppendorf tube awaiting at 4°C. Remaining DNA left behind was back-extracted by the same means after addition of fresh extraction buffer to the phenol phase and the second aqueous phase was pooled in same 2 ml tube with the first aqueous phase. 400 μl of chloroform (Wako 3034-02603) was saturated with
10 mM Tris-Cl pH 8.5 1 mM EDTA and then used to extract the pooled aqueous phases. These were “hand-mixed” by inversion with the robotic arm and centrifuged similarly as the phenol extractions. The 325 μl of the aqueous phase from the chloroform extraction was then transferred to a new 1.5 ml centrifuge tube (Eppendorf 0030 125.215). 650 μl of 100% ethanol and 33 μl of 3 M sodium acetate pH 5.2 (Sigma Aldrich, S7899) was added, “hand-mixed” for 5 minutes and centrifuged at 15,000g for 30 minutes at 4°C. 50 μg glycogen (Thermo, R0561) was added to the final 1.5 ml tubes before the start of the robotic run to help visualize pellets after precipitation and centrifugation. After centrifugation, the tubes were placed into a 4°C block awaiting sample recovery. An entire run was video-captured, and the steps are highlighted at key points described in this manuscript as a visual demonstration of the robot’s movement (below).
Pulsed field electrophoresis and long-read sequencing
Subsequently, the ethanol supernatant from the last automated step was removed from the pelleted DNA by manual aspiration. The pellets were carefully washed with 70% ethanol, removed by manual aspiration and air dried for 5 minutes at the bench. The washed DNA pellets were gently resuspended in 15 μl 10 mM Tris-Cl and 1 mM EDTA and quantitated with Qubit high-sensitivity double-stranded DNA reagents and fluorometer (Thermo Fisher, Q32851). 10 ng of the sample was used for pulsed-field (Agilent, FP-1002-0275) or standard capillary electrophoresis Tape Station (Agilent, 5067-5365). 1000 ng of gDNA was used for adapter ligation (Oxford Nanopore Technologies, ONT LSK-109). Between 40–500 ng of adapter ligated DNA was loaded onto a MinION R9.4 flow cell (ONT FLO-MIN106D) and allowed to run for 18-24 hours. N50 metrics were determined with the MinKNOW software (ONT).

Results
Movie 1 shows an entire run was of LabDroid’s extraction process, and the steps are highlighted at key points described in this manuscript as a visual demonstration of the robot’s movement.

Table 1. Representative Nanopore MinION N50 raw-reads and fragment length analysis.

<table>
<thead>
<tr>
<th>Organism</th>
<th>Sample material</th>
<th>N50 (kb)</th>
<th>Peak fragment size (kb)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Synechococcus</em> sp.</td>
<td>Bacterial culture</td>
<td>19.13</td>
<td>152.6</td>
</tr>
<tr>
<td><em>Chaetoceros calcitrans</em></td>
<td>Algal culture</td>
<td>35.55</td>
<td>&gt;165</td>
</tr>
<tr>
<td><em>Isochrysis</em> sp.</td>
<td>Algal culture</td>
<td>39.67</td>
<td>&gt;165</td>
</tr>
<tr>
<td><em>Rhinomonas reticulata</em></td>
<td>Algal culture</td>
<td>29.54</td>
<td>152.2</td>
</tr>
<tr>
<td><em>Drosophila melanogaster</em></td>
<td>Whole flies</td>
<td>16.43</td>
<td>&gt;60 (Tape Station)</td>
</tr>
<tr>
<td><em>Homo sapiens</em> (HEK293)</td>
<td>Tissue culture</td>
<td>ND</td>
<td>120.5</td>
</tr>
</tbody>
</table>
Figure 3. Representative Femto Pulse fragment length results from: A, cyanobacteria (*Synechococcus* sp.) and B, microalgae (*Rhinomonas reticulata*).

Table 2. Summary of problems encountered and solutions.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Organic solvent deformation of centrifuge tube</td>
<td>Limit time of tube exposure to organic solvents</td>
</tr>
<tr>
<td>Tubes containing organics dropped by the LabDroid</td>
<td>Grip tubes more tightly</td>
</tr>
<tr>
<td>Irreproducible position of the tubes’ hinge after centrifugation</td>
<td>Increase reproducibility of organic solvent dispensing (see below)</td>
</tr>
<tr>
<td>Liquid-solvent interface unpredictable because of variable volume of aqueous phase after organic extraction</td>
<td>Water-saturate organic solvents</td>
</tr>
<tr>
<td>Liquid-solvent interface unpredictable because of variable volume of organic phase</td>
<td>Dispense solvents with new unwetted tip for each sample</td>
</tr>
<tr>
<td>Liquid-solvent interface disturbed by pipetting</td>
<td>Equip Maholo with pipettes capable of slower aspiration speed</td>
</tr>
<tr>
<td>Liquid-solvent interface disturbed by pipetting</td>
<td>Position tip-end further from interface and extract the remaining water phase again</td>
</tr>
</tbody>
</table>
precision of the robot’s pipetting capabilities relative to the phase interface. It also impaired the robot’s handling of the centrifugation step since the slight imbalance caused individual tubes to rotate about their vertical axes as they balanced the mass distribution in the rotor. This changed the position of the hinge at the end of the run which prevented the robot from removing the rotated tube from the rotor. To ensure volume reproducibility, chloroform was saturated with 10 mM Tris-Cl pH 7.5 and 1 mM EDTA (TE) the night before (1:1 phenol and chloroform was purchased pre-equilibrated with TE pH 8) to prevent any aqueous phase volume loss due to the solvent’s slight water solubility. Finally, we also programmed the LabDroid to change the pipette tips after each aliquot of solvent was added, as the actual volume of solvent transferred by the tips was slightly greater after the tip had been pre-wetted.

The greatest hurdle was the collection of aqueous phases after extraction and centrifugal phase separation; this occurs three times in the protocol. The LabDroid initially was fitted with a brand of Bluetooth actuated liquid handling micro-pipettors that could not aspirate the aqueous phase without disturbing the organic/aqueous interface. As a result, the aqueous phases were contaminated with solvent which caused purity problems downstream. This was solved when new electronic pipettors (Viaflow, Integra Biosciences), capable of slower rates of aspiration, were outfitted to replace the earlier model. The aspiration offset height (the distance between the end of the pipette’s tip to the bottom of the tube) was also increased by 1 mm to avoid disturbing the organic/aqueous interface. As this increased the dead volume, we added a back-extraction step to collect the leftover aqueous phase.

**Attempts at extending the protocol beyond organic extraction steps**

Currently, the automated protocol ends with the pelleted DNA in salted ethanol solution, which is suitable for long-term DNA storage. We tried to incorporate a last step to precipitate the pellet in 70% ethanol and resuspend it in TE. We were unsuccessful as the pellet dislodged after the wash centrifugation. While a skilled technician can carefully remove supernatant in the presence of a floating pellet, a robot with no visual sensors cannot. To overcome this, we tried partial aspiration of the 70% ethanol wash to prevent loss of the pellet and drying 50-70 μl of remaining supernatant with a vacuum centrifuge for 40 minutes at room temperature at 0.02 bar, but we found this method partially degraded preparations. As a result, we decided to leave out the 70% ethanol wash from the automated protocol.

**Discussion**

Our main challenges in developing the automated protocol were problems that an experienced human investigator tends to resolve instinctively, but which prevents a deterministically programmed robot from continuing to the next action, thus aborting the run.

We did not include the tissue homogenization and digestive steps in the automation because they involve long incubation times during which the robot would be idle and potentially hinder scheduling of other users of the robot during these longer time slots. More notably, different tissue samples and organisms require different treatments to free the DNA from cells. Nevertheless, it is conceivable to use the platform to determine the optimal incubation for groups of samples of equivalent properties, in the same spirit as the work of Kanda et al. (2020) which reported AI-based optimization of iPS cell-culture conditions with Maholo. Further improvements of the method, such as the use of wide-bore tips to reduce DNA shearing, require more extensive reprogramming of the LabDroid’s movements as it currently relies on a fixed shape for all its tools and consumables.

The LabDroid’s gentle mixing of the organic extractions supplied the ideal amount of agitation necessary to create an organic-aqueous emulsion that resulted in a clean long-read DNA preparation. The results obtained have given us confidence in obtaining intact genomic DNA for routine genome sequencing and assembly, and we have begun using the robot in our sample preparation-to-sequence analysis workflow thus shortening benchwork and reducing exposure to organic solvents of the researcher.

With only minor changes, such as preparative equilibration of the phenol component to pH 4.5 with a small volume of 3M sodium acetate, the protocol can be adapted to total RNA with no further modifications to the robotic workflow (Wallace, 1987). With further development, we hope to program the LabDroid to perform the entire workflow from extraction to loading the adapter-ligated DNA on Nanopore MinION flow-cells.

**Data availability**

**Underlying data**

This project contains the following files:

- 20201207_0559_Dros_X7-2.pdf  MinKNOW Sequencing Run Metrics
- 20210127_0419_Syn.pdf  MinKNOW Sequencing Run Metrics
- 20210203_0018_Chaeto2.pdf  MinKNOW Sequencing Run Metrics
- 20210205_0639_Rhino2.pdf  MinKNOW Sequencing Run Metrics
- 20210208_0638_Iso2.pdf  MinKNOW Sequencing Run Metrics
- gDNA - 2021-08-24-10.59.28.pdf  Drosophila Tape Station Report
- QC26-1.pdf  Algal Femto-Pulse Report
- QC26-2.pdf  Algal Femto-Pulse Report
- QC27.pdf  HEK293 Femto-Pulse Report


Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Extended data

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Acknowledgments
We would like to thank Dr. Takakazu Yokokura (Formation and Regulation of Neuronal Connectivity Research Unit, Okinawa Institute of Science and Technology) for supplying flies, Dr. Narae Kim (Nucleic Acid Chemistry and Engineering Unit, OIST) for providing HEK293 cells, Mayumi Kawamitsu (Sequencing Section, OIST) for providing Femto Pulse results, and Drs. Motohisa Kamei & Shigeji Tasaka (RBI) for input during manuscript drafts. An earlier pre-print version is available on bioRxiv: https://doi.org/10.1101/2022.01.26.477939.

References
Liu AW, Plessy C: Job files for automated phenol-chloroform extraction (RBI LabDroid Maholo). Zenodo. 2022a. Publisher Full Text
Liu AW, Plessy C: Job files for automated phenol-chloroform extraction (RBI LabDroid Maholo). Zenodo. 2022b. Publisher Full Text
Open Peer Review

Current Peer Review Status: ✔️

Version 1

Reviewer Report 08 March 2022

https://doi.org/10.5256/f1000research.120730.r125733

© 2022 Kanda G. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Genki N. Kanda
RIKEN Center for Biosystems Dynamics Research, Osaka, Japan

This paper demonstrates the implementation of the phenol-chloroform DNA extraction procedure, one of the most commonly used experimental techniques in molecular biology, on a dual-armed humanoid robot. The scope of the paper is quite simple, and it clearly and concisely describes not only the results of the implementation, but also the limitations and solutions during development. In the field of laboratory automation, it is greatly useful to share the difficulties encountered during development, and the authors' description should be of sufficient help to the readers.

In the introduction, the authors state that "Automating the procedure was expected to enhance reproducibility, increase throughput, and reduce the possibility of DNA shearing into shorter fragments by minimizing excessive handling." It would be more helpful to the readers if the conclusions about reproducibility, throughput, and quality were described in the discussion.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Laboratory Automation, Molecular and Cellular Biology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com