ACADEMIA
MOLECULAR BIOLOGY &
GENOMICS

Research Article
Published: 2025-10-31
https://doi.org/10.20935/AcadMolBioGen7956

High-quality DNA extraction method enabling long-read sequencing in challenging native plant species

Bruno Cesar Rossini[1,2,*], Romain Guyot[3], Marcelo Augusto Mendes Alcantara[4,5], Alef Carvalho Batista[2], Yann Bourgeois[3], Marcela Aparecida de Moraes Silvestre[6], Patricia Ferreira Alves[6], Selma Maria Bozzite de Moraes[6], Mario Luiz Teixeira de Moraes[6], Celso Luis Marino[2,5]

Academic Editor: Chia-Yi Cheng

Abstract

The extraction of high-quality, high-molecular-weight (HMW) DNA from plant tissue is a critical step for successful long-read sequencing, particularly in species with high secondary metabolite content. *Myracrodruon urundeuva* ("Aroeira", Anacardiaceae), a native tree species from Brazil, presents significant challenges for DNA extraction due to its rich composition of polyphenols and polysaccharides, which can inhibit enzymatic reactions and compromise sequencing efficiency. In this study, we present an optimized DNA extraction protocol specifically tailored for native forest species with high secondary metabolite content. The protocol integrates a sorbitol pre-wash step to reduce contaminants, an improved CTAB-based extraction method, and a subsequent clean-up process using proteinase K and RNAse A, followed by short-fragment elimination to enrich HMW DNA. Our approach significantly improved DNA purity, achieving 260/280 ratios close to 1.8 and 260/230 ratios above 1.9, with DNA fragment sizes averaging over 60 kbp. The extracted DNA was successfully used in Oxford Nanopore Technologies (ONT) long-read sequencing, yielding high read lengths and sequencing coverage. These findings enhance genomic studies of native species, facilitating conservation efforts, genetic diversity assessments, and sustainable management strategies. The proposed protocol offers an efficient and cost-effective alternative to obtain high-quality DNA in species with challenging biochemical compositions, supporting the advancement of third-generation sequencing technologies in plant genomics.

Keywords: *high-molecular-weight DNA, Nanopore sequencing, Myracrodruon urundeuva, plant DNA extraction, genomic studies*

Citation: Rossini BC, Guyot R, Alcantara MAM, Batista AC, Bourgeois Y, de Moraes Silvestre MA, et al. High-quality DNA extraction method enabling long-read sequencing in challenging native plant species. *Academia Molecular Biology and Genomics* 2025;2. https://doi.org/10.20935/AcadMolBioGen7956

1. Introduction

The study of native forest species is significant for biodiversity conservation and sustainable development [1]. These species play essential ecological roles, such as maintaining ecosystem services and promoting environmental resilience. Moreover, the precise scientific identification of these species is necessary to recognize those with economic potential, as well as those that are legally protected, endemic, rare, or at risk of extinction, information that is indispensable for environmental analyses and vegetation management permit processes [2]. In this context, *Myracrodruon urundeuva* Allemão, commonly known as *Aroeira*, *Aroeira-do-sertão*, or *Aroeira-preta*, is a tree species belonging to the Anacardiaceae family that is widely distributed across the Northeast, Southeast, and Central-West regions of Brazil and extends into areas of Paraguay, Bolivia, and Argentina [3]. It is characterized as a medium- to large-sized tree, reaching heights of 15 to 30 m and a diameter of 80 to 100 cm. Its trunk is straight, with a broad canopy and drooping branches, while the bark has a brownish-gray color and scaly texture [4]. The wood is notably dense, ranging from 0.6 g/cm³ to 0.8 g/cm³, giving it remarkable mechanical strength and durability. These properties make it highly valued for various end uses, including civil construction, furniture manufacturing, and the production of utility poles and railroad sleepers [5–7]. Additionally, the species holds significant ecological importance, as it is part of seasonal deciduous forests and contributes to maintaining biodiversity within these ecosystems [8, 9].

Recent studies have investigated the chemical and pharmacological properties of *M. urundeuva*, particularly concerning the volatile compounds present in its bark and leaves. These studies indicate that tannins extracted from the bark exhibit anti-inflammatory and antioxidant properties, suggesting possible therapeutic use in the treatment of inflammatory and neurodegenerative diseases. Such findings indicate potential for the sustainable use of this species, highlighting its value not only in terms

[1]Department of Bioprocesses and Biotechnology, School of Agricultural Sciences, São Paulo State University (UNESP), Botucatu, Brazil.
[2]Institute of Biotechnology (IBTEC), São Paulo State University (UNESP), Botucatu, Brazil.
[3]UMR DIADE, Research Institute for Development (IRD), University of Montpellier, CIRAD, Montpellier, France.
[4]Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC, USA.
[5]Department of Genetics and Micro-Immunology, Institute of Biosciences, São Paulo State University (UNESP), Botucatu, Brazil.
[6]Department of Crop Science, Food Technology and Socioeconomics, School of Engineering, São Paulo State University (UNESP), Ilha Solteira, Brazil.
*email: bruno.rossini@unesp.br

of high-quality wood but also in both traditional and modern medicine [1, 8, 10]. Despite its ecological and economic significance, *M. urundeuva* is threatened by predatory exploitation and the degradation of its natural habitats. Intensive logging, combined with agricultural expansion and urbanization, has led to a decline in natural populations, resulting in the species being classified as vulnerable in several regions [11]. Conservation efforts, including reforestation programs and sustainable management practices, are imperative to ensure the preservation of this species and the ecosystem services it provides [12–14].

The integration of advanced molecular techniques in the study of native forest species not only enhances knowledge of genetic variability and phylogenetic relationships but also supports conservation strategies and the sustainable use of these resources. The application of optimized DNA extraction protocols, combined with modern genomic approaches, strengthens the identification of genes of interest and the understanding of adaptive mechanisms, contributing to the preservation and valuing of native forests. Obtaining high-quality genomic DNA is a fundamental step in conducting genetic studies on native forest species.

Efficient DNA extraction is challenged by the presence of secondary metabolites in plants, such as polysaccharides and phenols, which can interfere with the purity and integrity of genetic material [15, 16]. CTAB (cetyltrimethylammonium bromide)-based methods have been widely employed to overcome these difficulties, providing DNA of sufficient quality for subsequent molecular analyses [17]. However, based on our experience with Brazilian native species, which are rich in secondary metabolites, commercial kits and generalized protocols are often ineffective, potentially leading to failures in third-generation sequencing, such as Oxford Nanopore. While many studies focus on obtaining high-molecular-weight DNA [18, 19], field samples from plants under intense selection and competition, such as those found in Brazil, present additional challenges. Therefore, protocol optimization is essential to ensure successful genetic analyses for projects involving complete genome sequencing and assembly specifically directed at third-generation long reads, where large amounts of pure native DNA are required. These protocols also support the analysis of epigenetic modifications associated with sequencing of methylated DNA regions. Nucleic acid extraction protocols designed for PCR-based sequencing or for use in population genetics studies, such as Short Sequence Repeats (SSR) or Inter Simple Sequence Repeat (ISSR) markers, may not adequately support these genomic studies due to the presence of secondary compound contaminants in the sample. As an example, a study on individuals of another genus of the same family, *Schinus terebinthifolius*, showed that the composition of essential oils varied among different populations, reflecting both environmental adaptability and distinct chemical profiles across regions [20]. These factors represent potential contaminants for genomic analyses that are based on third-generation applications, such as Oxford Nanopore long reads.

Therefore, the choice of a DNA extraction protocol must consider the specific characteristics of each species and the type of plant tissue used. In this context, methods such as that of Mogg and Bond [21] offer practicality and improved DNA quality while having a lower impact compared to the traditional CTAB method [22]. Additionally, adapting protocols to the specific characteristics of native species is necessary to ensure DNA is obtained efficiently with high purity and sufficient concentration, as suggested by Faleiro et al. [23] and Silva et al. [24]. However, the application of these methods without prior washing or subsequent fragment selection may also compromise the results of Oxford Nanopore sequencing.

Through the optimization of various DNA extraction methods, we developed a protocol for high-molecular-weight DNA extraction for Brazilian native forest species for use with third-generation Oxford Nanopore long-read sequencing. This protocol includes preliminary washing steps and can be easily applied to process a small number of samples, as suggested for genome projects where no reference genome is available. The efficiency of DNA extraction from plant tissue depends on selecting the appropriate protocol, considering the specific nature of the tissue and the presence of interfering compounds. Moreover, the standardization of these methods is needed to ensure the reproducibility and reliability of results in studies addressing the genetic diversity of native forest species.

2. Materials and methods

Young leaves were selected in the field, separated into portions of approximately two grams, and wrapped in aluminum foil immediately after collection. Another option for collection is to select branches with young leaves without detaching the leaves. To avoid desiccation, these branches can be wrapped and kept in moistened cotton bags until processing in the laboratory. Samples frozen at -80 °C can also be used long after collection (over a year). The strategy of collecting young leaves for some deciduous species such as Aroeira is important, as the tree loses all its leaves during flowering, and therefore, the best collection time is at the beginning of the rainy season. Late collection in summer is not recommended, considering that early senescence occurs. Old leaves did not result in sufficient DNA extraction due to the presence of large amounts of secondary metabolites, as indicated by low ratios in Nanodrop (Thermo Fisher Scientific, Waltham, USA) measurements. Permission to access genetic heritage is registered in the SisGen system (National Genetic Heritage and Associated Traditional Knowledge Management System), number AEF203A.

2.1. DNA extraction

For this protocol, two commercial kits were tested: the Wizard® HMW DNA Extraction kit (Promega Corporation) and the Nanobind® plant nuclei kit (PacBio-Pacific Biosciences of California, Inc.). In addition, we modified protocols to isolate HMW DNA described in the literature, including those proposed by Rai [25] and Jones et al. [26], as both use magnetic beads to purify DNA. We optimized the steps to avoid their use, because the cost can be high and prohibitive in some situations, such as in poorly equipped field laboratories.

As a first step, maceration of the material must be carried out in liquid nitrogen. The mortar and pestle must be cooled beforehand, the leaves weighed (2 g), and the central veins removed with scissors. The leaf tissue must be constantly immersed in liquid nitrogen (about 3–5 l are needed). Maceration must be carried out for at least 20 min, as proposed in the Nanobind® plant nuclei kit (PacBio), until a fine floury powder is obtained. Maceration for shorter periods did not result in sufficient quantity for subsequent cleaning, purification, fragment selection, and library preparation for sequencing steps. The macerated tissue (still in liquid nitrogen) is transferred into two 50 ml Falcon tubes that have been cooled in liquid nitrogen (about 1 g for each tube). After nitrogen evaporation, 10 ml of sorbitol wash buffer is quickly added to help reduce polysaccharides and polyphenols (**Figure 1**).

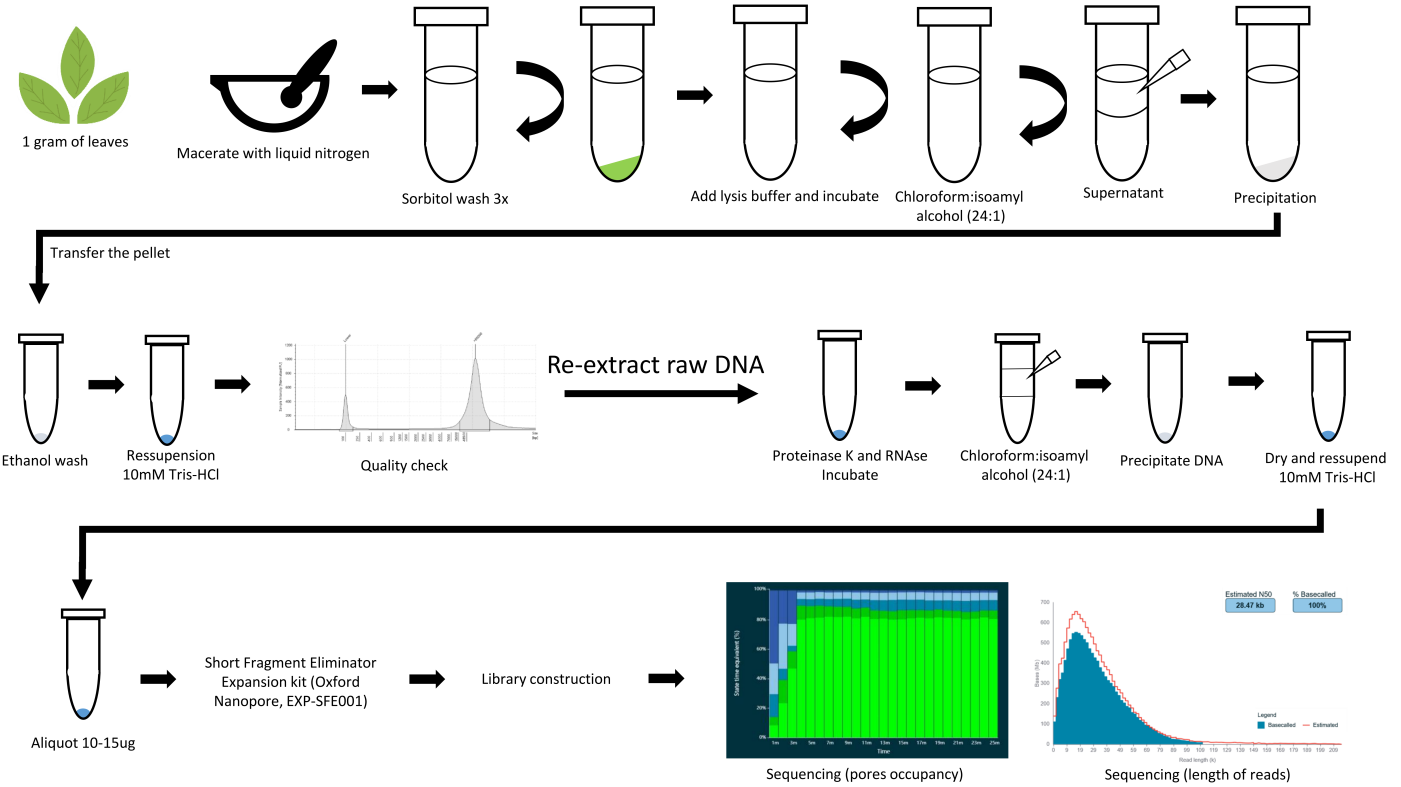

Figure 1 • Protocol overview for isolating a quantity of high-quality DNA from *M. urundeuva*.

The sorbitol wash buffer is composed of 0.35 M sorbitol, 5 mM EDTA pH 8.0, 100 mM Tris-HCl pH 8.0, and Polyvinylpyrrolidone 1% (*w*/*v*) (average molecular weight 360,000; PVP-360; modified from Inglis et al. [27]), with the addition of 2-mercaptoethanol 1% immediately before use. Using PVP-360 instead of PVP-40 resulted in better DNA metrics for several other native species (data not shown). The buffer must be kept at 4 °C and used within six months. At this stage, vortexing does affect DNA fragmentation. Thus, mix by vortexing for at least 30 s or until there are no clumps, followed by centrifugation at 5000 rcf for 5 min at 4 °C. Discard the supernatant and repeat the process for at least three washes. After washing, the supernatant should appear clear and not brown/dark green as it was at the beginning. If a dark color persists, repeat the wash sequence until obtaining a lighter-colored supernatant.

After the last wash, immediately add 20 ml of the lysis buffer proposed by Rai [25], preheated to 60 °C (Tris-HCl 100 mM, EDTA 20 mM, CTAB (*w*/*v*) 4%, NaCl 1.4 M, PVP-360 (*w*/*v*) 1%, β-mercaptoethanol 2% (added at the time of use)). Continue with the following steps:

-Vortex to avoid the formation of clumps. Adjust the volume to 20 times the initial weight to maximize the extraction of as much DNA as possible.

-Place in a dry bath at 60 °C for 1 h at 300 rcf. After 20 min of initial incubation, add 200 µl of Proteinase-K 20 mg/ml (incubation of 40 min in Proteinase-K solution). Gently invert the tubes about 10 times to mix the proteinase solution.

-After incubation, allow the tubes to cool to room temperature under slow agitation in a Roto-shake Genie (Scientific Industries) or slow manual inversion for 10 min. Centrifuge for five minutes at 3000 rcf at room temperature and transfer the supernatant to another tube.

-Add an equal volume of chloroform/isoamyl alcohol (24:1) and invert the tubes for 5 min in a Roto-shake or manually until completely mixed (milky appearance). Centrifuge for 10 min at 3000 rcf at room temperature and transfer the upper phase (supernatant) to another tube without touching the interphase using a wide-bore tip or a 10 ml Pasteur pipette.

-Add twice the volume of isopropanol at room temperature, invert ~20 times to ensure complete mixing, and incubate at room temperature for at least 1 h. Centrifuge at 3000 rcf for five minutes at room temperature. At this point, a viscous white or almost clear pellet should be seen at the bottom of the tube.

-Carefully remove and discard the entire supernatant using a pipette and resuspend the pellet in 1 ml of freshly prepared 70% ethanol. Use a wide-bore tip to transfer the resuspended pellet into a 2 ml LoBind tube (Eppendorf). Rinse the tube with another 1 ml of 70% ethanol and add to the 2 ml LoBind tube.

-Incubate the tubes on a Roto-shake at approximately 9–10 rpm for five minutes. Centrifuge at 13,000 rcf for 10 min, discard the supernatant, and repeat the wash with 70% ethanol.

-Finally, centrifuge under the previous conditions and wash the pellet with 2 ml of 100% ethanol. Discard the supernatant and dry for five minutes in a fume hood.

-Add 75 μl of prewarmed 10 mM Tris-HCl solution at 37 °C. If the pellet is large, add another 75 μl and homogenize with caution using a wide-bore pipette. Add 1 μl of RNaseA solution (20 mg/ml) for each 75 μl of elution buffer and incubate at 37 °C for 30 min.

-Leave at room temperature overnight. Then, assess DNA quality using a NanoDrop (Thermo Fisher Scientific), a Qubit (Thermo Fisher Scientific), and a TapeStation instrument (Genomic DNA ScreenTape, Agilent Technologies). Pulse Field Gel Electrophoresis is also an option. Store DNA at 4 °C. Do not freeze.

The raw DNA is then further purified using the method described by Jones et al. [26] with modified enzyme volumes. To the approximately 15 μg of extracted DNA (around 150 ng/μl), we added 10 mM Tris-HCl pH 8 to a total volume of 200 μl.

-Add 4 μl of Proteinase-K solution (20 mg/ml) and 4 ul of RNAseA solution (20 mg/ml) and incubate the samples at 56 °C for 20 min at 400 rpm (if possible, using a dry bath with shaking or, alternatively, leaving under gentle agitation). Adjust the volume to 600 ul with 10 mM Tris-HCl pH 8, add an equal volume of chloroform/isoamyl alcohol (24:1, *v*/*v*) and invert the tubes constantly for five minutes. Centrifuge the resulting emulsion at 16,000 rcf for one minute at room temperature.

-Transfer the supernatant to a new tube and repeat the chloroform/isoamyl alcohol step. After centrifugation, transfer the supernatant to a new tube and add 1.5 volumes of 100% ethanol and 0.1 volumes of 3 M sodium acetate pH 5.2. Incubate on ice for one minute.

-Centrifuge at 16,000 rcf for one minute at room temperature and discard supernatant. Wash the pellet twice with 700 ul of 70% ethanol and leave to soak for one minute. Centrifuge at 16,000 rcf for one minute at 4 °C.

-After the second wash, allow the pellet to dry in a fume hood for 5–10 min and resuspend it in 50 μl of 10 mM Tris-HCl pH 8. Quantify by Qubit and use 10 ug. To eliminate short reads, use the Short Fragment Eliminator Expansion kit (Oxford Nanopore Technologies, EXP-SFE001) following the manufacturer's instructions. Quantify in Nanodrop, Qubit, and if possible, TapeStation. Set aside around 6 μg of DNA to prepare your libraries.

2.2. Library preparation

To prepare the libraries for sequencing in a MinION R10 4.1 flow cell, use the Ligation Sequencing Kit V14 (SQK-LSK114, Oxford Nanopore Technologies). Modify the manufacturer's instructions as follows: Use around 6 μg of DNA instead of the recommended 1 μg, as losses occur during subsequent washes with beads in the protocol. Expect around 1.5 μg of final library, which is enough for at least two injections into a MinION R10.4.1 flow cell. Less than 1 μg is enough for one injection. During library preparation, increase the incubation times in the repair and end-prep step to 20 min at 20 °C, followed by 20 min at 65 °C. For bead purification, increase incubation from 5 to 10 min. During adapter ligation, increase the incubation time from 10 to 20 min and use 0.5X Ampure XP beads (Beckman Coulter, included in the SQK-LSK114 Oxford Nanopore Technologies) in the next stage (save these beads for another purification if necessary). For washing, use 250 ul of Short Fragment Buffer (SFB, included in SQK-LSK114 Oxford Nanopore Technologies), and in the second wash use 250 ul of Long Fragment Buffer (LFB, included in SQK-LSK114 Oxford Nanopore Technologies). After final quantification, expect between 90 and 100 ng/μl in ~12 ul of the final library. If the library is below this concentration, perform a second purification with the beads from the first step. Add another 40 μl of Ampure XP beads (Beckman Coulter) and repeat the steps above with SFB and LFB (you can use the same SFB and LFB recovered from previous washes). Elute in 12 μl and quantify.

3. Results

The DNA extraction protocol presented here resulted in good readings during sequencing. The washing step with sorbitol buffer significantly reduced contamination from secondary compounds. In plants with many secondary compounds, such as Aroeira and other members of the Anacardiaceae family, implementing prewashing is essential. Without this, the pellet was generally brownish at the end of elution, and quality data obtained on the Nanodrop showed 260/230 ratios of less than 1.00 and 260/280 ratios of less than 1.4. The omission of magnetic beads for purification increases the number of handling steps. However, the inclusion of a clean-up step after extraction, again using proteinase, RNAse, and chloroform/isoamyl alcohol, with consequent use of the Short Fragment Eliminator Expansion kit, resulted in excellent metrics for quality and quantity of DNA (**Table 1**). The total amount recovered after using the Short Fragment Eliminator kit varied between 3.17 ug and 13.15 ug. TapeStation data indicated that the material had an average of 95% of fragments above 60,000 bp (Supplementary material S1, Figure S1; quality of sequencing reads Figure S2). As for commercial kits applied to Aroeira samples, the Wizard® HMW DNA Extraction Kit (Promega) resulted in very fragmented DNA and the Nanobind® plant nuclei kit (PacBio) resulted in a low concentration, with an average of 4.81 ng/μl, which is insufficient for library preparation. These methods were disregarded.

Changing the incubation times and increasing the initial input well above the values recommended by the kit manufacturer proved to be essential for the success of library preparation. The successive losses of material during bead washing reduce the final yield of DNA, which may result in an insufficient amount of final library for flow cell injection. For the MinION R10.4.1 flow cell, our studies indicate that a total amount of 750 ng is sufficient to maximize pores for long-read sequencing; however, amounts of ~450 ng also resulted in good reads. For example, a library prepared under the conditions above with an injection of 750 ng and a second injection of 450 ng resulted in ~14 Gb of data, with a read length of N50 28.47 Kb and more than 900,000 reads (**Figure 2** and **Figure 3**).

Table 1 • Summary of quantification DNA ratio values obtained in Nanodrop before and after DNA purification and Qubit final quantification after Short Fragment Eliminator kit (Oxford Nanopore).

	Raw DNA efore clean-up			After clean-up and short fragment eliminator kit			Qubit quantification
Sample repetitions	**260/280**	**260/230**	**ng/µl**	**260/280**	**260/230**	**ng/µl**	**ng/µl**
#2	1.50	0.67	2216.7	1.80	1.94	528.4	218
#3	1.05	0.41	517.7	1.78	1.69	44.1	63.4
#4	1.58	0.79	2651.9	1.84	1.88	1078.9	251
#5	1.57	0.57	922.5	1.87	1.99	946.2	259
#6	1.85	0.65	1290.6	1.87	1.86	1242	263

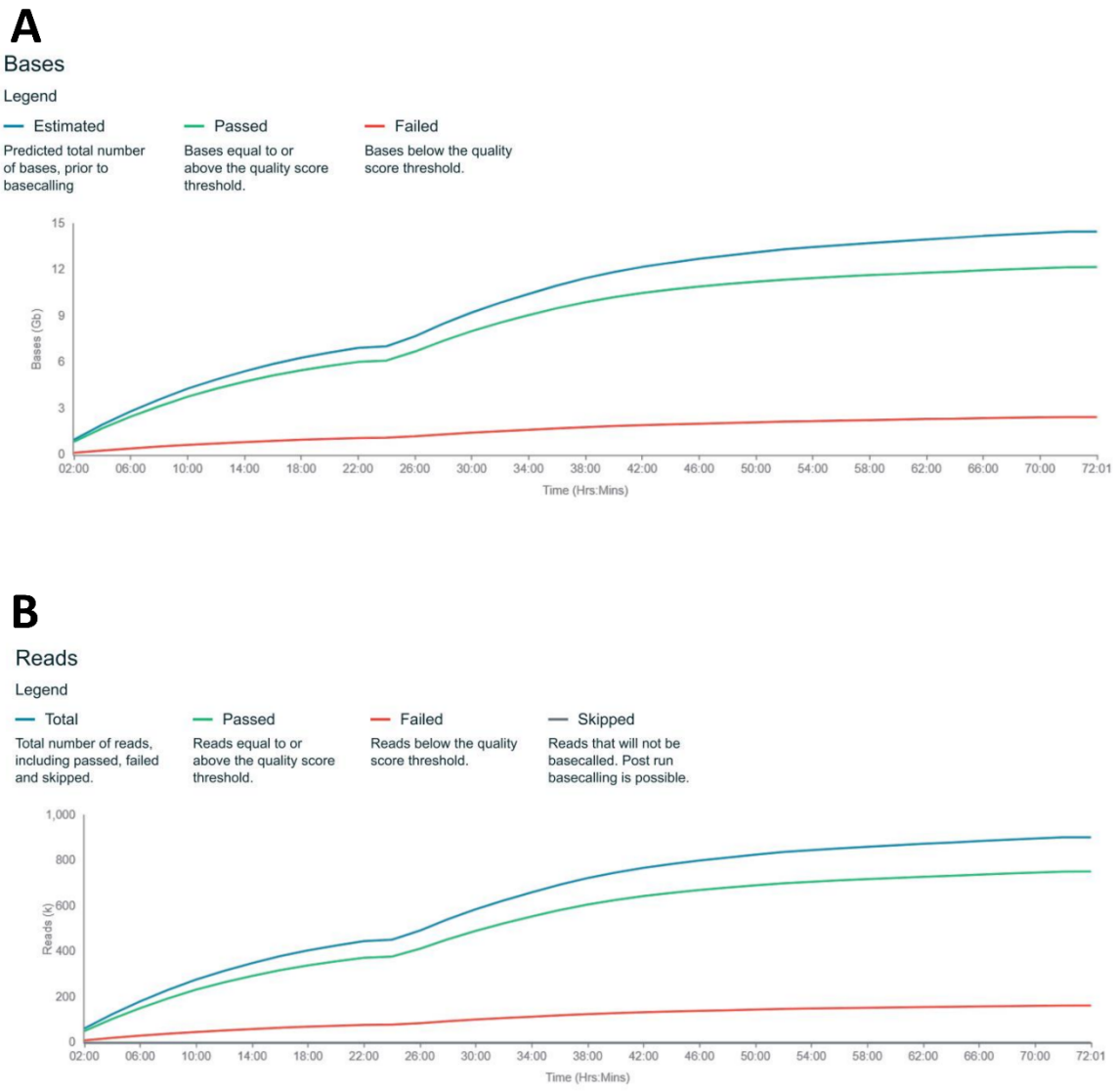

Figure 2 • Graphics of yield in bases (Gb) of sequenced data (**A**) and number of reads (**B**) in a MinION R10.4.1 flow cell.

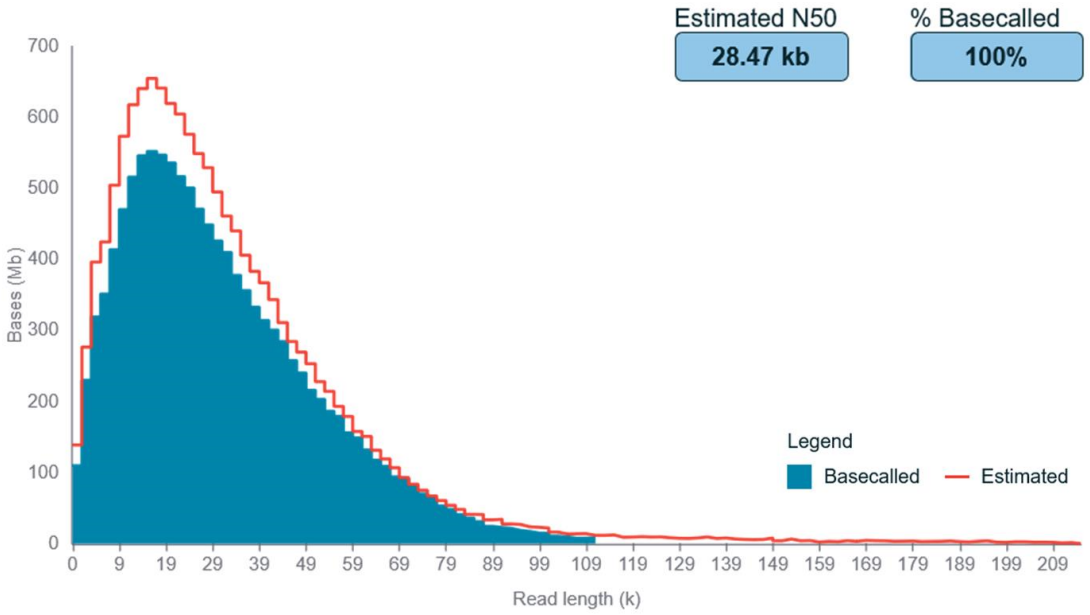

Figure 3 • Length distribution of reads of a library prepared with the protocol described here in a MinION R10.4.1 flow cell.

4. Discussion

Obtaining HMW DNA is a fundamental step in the successful sequencing of long reads using the MinION flow cell. Furthermore, the purity of the sample is crucial to the success of the experiment. With the optimization and compilation of previously described protocols, this study demonstrates that it is possible to obtain HMW DNA for a plant species with a high concentration of secondary compounds, as reported for several Anacardiaceae species [28–30]. The quantity and purity of DNA obtained in the protocol presented here are similar to previous studies, but without the use of magnetic beads for purification [18]. Our results are also comparable to those obtained for cultivated species in terms of purity and quantity, such as cotton, blackgrass, and strawberry [31]. Considering that Aroeira is a native species that is not cultivated, our results are satisfactory and can be used for research into Anacardiaceae species or other families with a high content of secondary compounds. The presence of these compounds can not only result in low purity and integrity of the extracted DNA but also inhibit enzymes in the library preparation protocol, such as the repair and end-prep, which negatively impact Oxford Nanopore sequencing performance [32]. For the generation of genomic data, previous studies have shown that species such as brown algae contain polyphenols that are capable of blocking nanopores, leading to significantly shorter read lengths compared to controls [33]. Similarly, in our initial attempts with DNA extracted using protocols optimized for short-read sequencing or PCR, we observed pore blockage and short fragments, indicating the need to adapt protocols for native DNA extraction from non-model species. This is particularly important for samples obtained from natural habitats and diverse biomes, where secondary metabolites or polysaccharides can compromise long-read sequencing performance. DNA extraction methods for Aroeira have been proposed for PCR-based applications and molecular markers, but as observed in the electrophoretic patterns reported in these studies, all isolated samples exhibited smearing [34, 35]. Since there are no capillary electrophoresis profiles nor the application of sequencing technologies associated with these studies, we consider that our protocol generally outperforms previous methods, including the significant presence of fragments larger than 60,000 bp, as indicated by TapeStation (Agilent Technologies) results, and OD 260/230 ratios better than those obtained in previous studies. Indeed, using extraction protocols without the re-extraction step proposed herein resulted in sequencing failure, as noted above.

In general, studies involving DNA extraction for use with new sequencing technologies, such as third-generation Oxford Nanopore, include several steps to ensure sufficient quality and quantity for subsequent applications. In this sense, the use of liquid nitrogen and extended maceration (20–25 min) favors the release and isolation of nuclei in later stages, ensuring good results for many species. Furthermore, all methods require several steps to achieve this purification, including the use of magnetic beads, which can be costly for some laboratories [18, 31, 36]. As examples of challenges posed by high concentrations of secondary metabolites, which highlight the need for prior adaptations like those implemented in this protocol to overcome species-specific difficulties, can be found in studies reporting high biological and chemical variability associated with Native American plants. This variability has been observed in Anacardiaceae species, such as cashew (*Anacardium occidentale*), which presents a high level of contaminants even when tested with different protocols [37], and mango (*Mangifera indica*), where a glucose buffer was used to reduce the presence of contaminants [38].

For the protocol to be successful, some specific considerations are necessary. For example, the age of the sampled leaves, storage conditions, and maceration time are all important. If obtaining fresh leaf material is not possible, one alternative is the immediate immersion in liquid nitrogen for sample processing after field collection. Furthermore, if wide-bore tips are not available, we recommend cutting the tip of a 1000 ul pipette tip to increase the diameter. This not only facilitates pipetting, but it also avoids DNA fragmentation at all stages of the isolation procedure. Use the same precautions when preparing the library.

5. Conclusions

With the continuous development of new sequencing methodologies, DNA purity and integrity are becoming increasingly necessary. Therefore, protocols adapted to native species that have a high content of secondary compounds are a challenge for Nanopore sequencing, as such compounds can inhibit repair and end-prep enzymes in library preparation. Our protocol for HMW DNA extraction works well without using magnetic beads for purification, and the results obtained are comparable to those from cultivated species in terms of purity and integrity. High yields of final DNA were obtained (more than 10 ug per gram of tissue), and the method can be scaled up to increase yield even further. Finally, we suggest some changes to the Nanopore library preparation protocol in order to improve sequencing output.

Acknowledgments

We thank the Laboratory of Population Genetics and Silviculture, UNESP—Ilha Solteira, for support during sample collection.

Funding

This research was sponsored by the São Paulo Research Foundation (FAPESP) under grant (22/03320-5) and the National Council for Scientific and Technological Development—CNPq under grant (441084/2023-0).

Author contributions

Conceptualization, B.C.R.; methodology, B.C.R., R.G. and Y.B.; software, B.C.R.; validation, B.C.R., M.A.M.A., P.F.A., S.M.B.d.M. and A.C.B.; formal analysis, B.C.R.; investigation, B.C.R., M.A.M.A., A.C.B., M.A.d.M.S., P.F.A. and S.M.B.d.M.; resources, B.C.R., C.L.M.; data curation, B.C.R.; writing—original draft preparation, B.C.R.; writing—review and editing, B.C.R., R.G., Y.B., M.L.T.d.M. and C.L.M.; visualization, B.C.R.; supervision, B.C.R.; project administration, B.C.R. and C.L.M.; funding acquisition, B.C.R. and C.L.M. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare that they have no competing interests.

Data availability statement

The data supporting the findings of this publication has been made available within a publicly accessible repository at NCBI (National Center for Biotechnology Information) (accession number PRJNA1263011; https://www.ncbi.nlm.nih.gov/sra/?term=SRR33570753 [accessed on 2025 Oct 22]).

Sample availability

Samples used for this research are available upon request and under the terms of the Brazilian biodiversity law 13123/2015 (National System for the Management of Genetic Heritage and Associated Traditional Knowledge—SisGen).

Supplementary materials

The Supplementary materials are available at https://10.20935/AcadMolBioGen7956.

Additional information

Received: 2025-07-10

Accepted: 2025-09-29

Published: 2025-10-31

Academia Molecular Biology and Genomics papers should be cited as *Academia Molecular Biology and Genomics 2025*, ISSN 3064-9765, https://doi.org/10.20935/AcadMolBioGen7956 . The journal's official abbreviation is *Acad. Mol. Biol. Genom.*

Publisher's note

Copyright

References

1. Lucena RFP, Farias DC, Carvalho TKN, Lucena CM, Vasconcelos Neto CFA, Albuquerque UP. Use and knowledge of aroeira (*Myracrodruon urundeuva*) by traditional communities in the semiarid region of Brazil. Sitientibus Ser Cien Biol. 2011;11(2):255–64. doi: 10.13102/scb109

2. Wäldchen J, Wittich HC, Rzanny M, Fritz A, Mäder P. Towards more effective identification keys: a study of people identifying plant species characters. People Nat. 2022;4(6):1603–15. doi: 10.1002/pan3.10405

3. Moraes MLT, Kageyama PY, Sebbenn AM. Diversity and spatial genetic structure in two populations of *Myracrodruon urundeuva* Fr. All. under different anthropogenic conditions. Revista Árvore. 2005;29(2):281–9. doi: 10.1590/S0100-67622005000200011

4. Salomao AN. *Myracrodruon urundeuva*: aroeira. In: Vieira RF, Camillo J, Coradin L, editors. Native species of Brazilian flora with current or potential economic value: plants for the future: Central-West Region. Brasília: MMA;

2018. p. 835–40. [accessed on 2025 Jul 10]. Available from: https://www.embrapa.br/busca-de-publicacoes/-/publicacao/1073295/especies-nativas-da-flora-brasileira-de-valor-economico-atual-ou-potencial-plantas-para-o-futuro-regiao-centro-oeste

5. Florsheim SMB, Tomazello Filho M. Variação da densidade básica da madeira de aroeira *Myracrodruon urundeuva* F.F. & M.F. Allemão (Anacardiaceae). Rev Inst Florestal. 1998;10(1):53–61. [accessed on 2025 Jul 10]. Available from: https://smastr16.blob.core.windows.net/iflorestal/ifref/RIF10-1/RIF10-1_53-61.pdf

6. Guimarães Junior JB, Cangussu JT, Lisboa FJN, Guimarães ÍL. Caracterização tecnológica da madeira de *Myracrodruon urundeuva* em diferentes classes diamétricas. Enciclopédia Biosfera. 2016;13(24):250–61. doi: 10.18677/EnciBio_2016B_022

7. Tung ESC, Freitas MLM, Flörsheim SMB, Lima IL, Longui EL, Moraes MLT, et al. Variação, divergência e correlações genéticas entre caracteres silviculturais e densidade básica da madeira em progênies de *Myracrodruon urundeuva* (Engler) Fr. Allem. Rev Inst Florest. 2011;23(1):1–12. doi: 10.24278/2178-5031.2011231281

8. Domingos FR, Silva MAP. Use, knowledge and conservation of Myracrodruon urundeuva: a systematic review. Res Soc Dev. 2020;9(11):e2329118851. doi: 10.33448/rsd-v9i11.8851

9. Nunes YRF, Fagundes M, Almeida HS, Veloso MMD. Ecological aspects of aroeira (*Myracrodruon urundeuva* Allemão-Anacardiceae): phenology and seed germination. Rev Árvore. 2008;32(2):233–43. doi: 10.1590/S0100-67622008000200006

10. Figueiredo YG, Corrêa EA, de Oliveira AH Jr, Mazzinghy ACdC, Mendonça HdOP, Lobo YJG, et al. Profile of *Myracrodruon urundeuva* volatile compounds ease of extraction and biodegradability and in silico evaluation of their interactions with COX-1 and iNOS. Molecules. 2022;27(5):1633. doi: 10.3390/molecules27051633

11. Souza AS, Monteiro JM, Santos SS, Ferreira EC, Lucena RFP. Population structure and level of exploitation of medicinal plants by the Fulni-ô Indians in Northeast Brazil. Rev Bras Geogr Física. 2025;18(1):424–41. doi: 10.26848/rbgf.v18.1.p424-441

12. Freitas MLM, Sebbenn AM, Zanatto ACS, Moraes E. Graping of seeds by moves from the selection inside the teste of progênies of *Myracrodruon urudeuva* Fr. All. Rev Inst Florestal. 2007;19(2):65–72. doi: 10.24278/2178-5031.2007192349

13. Sebbenn AM, Ettori LC. Ex situ genetic conservation of Esenbeckia leiocarpa, Myracrodruon urundeuva E Peltophorum dubium in misto progênies teste. Rev Inst Florestal. 2001;13(2):201–11. doi: 10.24278/2178-5031.2001132642

14. Viegas MP, Silva CLSP, Moreira JP, Cardin LT, Azevedo VCR, Ciampi AY, et al. Genetic diversity and effective size in two *Myracrodruon urundeuva* Fr. All. population under ex situ conservation. Rev Árvore. 2011;35(4):769–79. doi: 10.1590/S0100-67622011000500002

15. Porebski S, Bailey LG, Baum BR. Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. Plant Mol Biol Rep. 1997;15:8–15. doi: 10.1007/BF02772108

16. Aboul-Maaty NAF, Oraby HAS. Extraction of high-quality genomic DNA from different plant orders applying a modified CTAB-based method. Bull Natl Res Cent. 2019;43:25. doi: 10.1186/s42269-019-0066-1

17. Molinari HB, Crochemore ML. Extração de DNA genômico de *Passiflora* spp. para análises PCR-RAPD. Rev Bras Frutic. 2001;23(2):447–50. [accessed on 2025 Jul 10]. Available from: https://www.scielo.br/j/rbf/a/RtJfK893z97FQLqD3b78mKS/?format=pdf&lang=pt

18. Russo A, Mayjonade B, Frei D, Potente G, Kellenberger RT, Frachon L, et al. Low-input high-molecular-weight dna extraction for long-read sequencing from plants of diverse families. Front Plant Sci. 2022;13:883897. doi: 10.3389/fpls.2022.883897

19. Lu D, Liu C, Ji W, Xia R, Li S, Liu Y, et al. Nanopore ultra-long sequencing and adaptive sampling spur plant complete telomere-to-telomere genome assembly. Mol Plant. 2024;17(11):1773–86. doi: 10.1016/j.molp.2024.10.008

20. Velasques J, Crispim BA, Vasconcelos AA, Bajay MM, Cardoso CAL, Barufatti A, et al. Genetic and chemodiversity in native populations of *Schinus terebinthifolia* Raddi along the Brazilian Atlantic forest. Sci Rep. 2021;11:20487. doi: 10.1038/s41598-021-00015-0

21. Mogg RJ, Bond JM. A cheap, reliable and rapid method of extracting high-quality DNA from plants. Mol Ecol Notes. 2003;3(4):666–8. doi: 10.1046/j.1471-8286.2003.00548.x

22. Doyle JJ, Doyle JL. Isolation of plant DNA from fresh tissue. Focus. 1990;12:13–5.

23. Faleiro FG, Faleiro ASG, Cordeiro MCR, Karia CT. Metodologia para operacionalizar a extração de DNA de espécies nativas do cerrado visando a análises moleculares. Planaltina (DF): Embrapa Cerrados; 2003. (Comunicado Técnico; 92). 5 p. [accessed on 2025 Jul 10]. Available from: https://www.infoteca.cnptia.embrapa.br/bitstream/doc/559088/1/comtec92.pdf

24. Silva JN, Viana JPG, Costa MF, Sá GH, Gomes MFC, Feitoza LL, et al. A simple and cost-effective method for DNA extraction suitable for PCR in "sucupira branca". Biosci J. 2021;37:e37092. doi: 10.14393/BJ-v37n0a2021-54133

25. Rai S. High molecular weight total DNA extraction from plant tissues for long read sequencing. Protocols.io. 2022. doi: 10.17504/protocols.io.14egn7yy6v5d/v1

26. Jones A, Torkel C, Stanley D, Nasim J, Borevitz J, Schwessinger B. High-molecular weight DNA extraction, clean-up and size selection for long-read sequencing. PLoS ONE. 2021;16(7):e0253830. doi: 10.1371/journal.pone.0253830

27. Inglis PW, Pappas MCR, Resende LV, Grattapaglia D. Fast and inexpensive protocols for consistent extraction of high quality DNA and RNA from challenging plant and fungal samples for high-throughput SNP genotyping and sequencing applications. PLoS ONE. 2018;13(10):e0206085. doi: 10.1371/journal.pone.0206085

28. Schulze-Kaysers N, Feuereisen MM, Schieber A. Phenolic compounds in edible species of the Anacardiaceae family—a review. RSC Adv. 2015;5(89):73301–14. doi: 10.1039/C5RA11746A

29. Neves AM, Fontenelle ROS, Lopes FS, Mendes JFS, Rodrigues ALM, Marinho MM, et al. Phenolic profile, antioxidant and antifungal activity of extracts from four medicinal plants of the Anacardiaceae family. Res Soc Dev. 2021;10(8):e17410817421. doi: 10.33448/rsd-v10i8.17421

30. Sruthi P, Roopavathi C, Naidu MM. Profiling of phenolics in cashew nut (*Anacardium occidentale* L.) testa and evaluation of their antioxidant and antimicrobial properties. Food Biosci. 2023;51:102246. doi: 10.1016/j.fbio.2022.102246

31. Li Z, Parris S, Saski CA. A simple plant high-molecular-weight DNA extraction method suitable for single-molecule technologies. Plant Methods. 2020;16:38. doi: 10.1186/s13007-020-00579-4

32. Sawicki J, Krawczyk K, Paukszto Ł, Maździarz M, Kurzyński M, Szablińska-Piernik J, et al. Nanopore sequencing technology as an emerging tool for diversity studies of plant organellar genomes. Diversity. 2024;16(3):173. doi: 10.3390/d16030173

33. Pearman WS, Arranz V, Carvajal JI, Whibley A, Liau Y, Johnson K, et al. A cry for kelp: evidence for polyphenolic inhibition of Oxford Nanopore sequencing of brown algae. J Phycol. 2024;60(6):1601–10. doi: 10.1111/jpy.13513

34. Silva JN, Linhares ACS, Palmieri DA, Costa MF, Sá GH, Gomes MFC, et al. Optimized method for DNA extraction and PCR amplification in Aroeira tree. Biosci J. 2023;39:e39001. doi: 10.14393/BJ-v39n0a2023-62577

35. Silva CLS, Moraes MLT, Viegas MP, Campos APS, Lourenção JC, Freitas MLM, et al. Avaliação de protocolos para extração de DNA de aroeira (*Myracrodruon urundeuva*). Cultura Agronômica. 2010;19(1):12–22. [accessed on 2025 Jul 10]. Available from: https://whoamiojs.unesp.br/index.php/rculturaagronomica/article/view/2163/1577

36. Kang M, Chanderbali A, Lee S, Soltis DE, Soltis PS, Kim S. High-molecular-weight DNA extraction for long-read sequencing of plant genomes: An optimization of standard methods. Appl Plant Sci. 2023;11(3):e11528. doi: 10.1002/aps3.11528

37. Castro EBL, Souza Filho MSM, Bordallo PN. Non-destructive method for extracting DNA from cashew seeds. Rev Ciênc Agron. 2025;56:e202392001. doi: 10.5935/1806-6690.20250008

38. Uddin MS, Sun W, He X, Awan FM, Peng L, Qin Q, et al. An improved method to extract DNA from mango *Mangifera indica*. Biologia. 2014;69(1):133–8. doi: 10.2478/s11756-013-0311-2