BIODIVERSITAS
Volume 25, Number 2, February 2024
Pages: 654-663

ISSN: 1412-033X
E-ISSN: 2085-4722
DOI: 10.13057/biodiv/d250224

High-quality genomic DNA extraction methods of Yellow *Spathoglottis* Blume complex for next-generation sequencing

JASIM HAIDER MAHMOD JASIM[1,3], AHMAD SOFIMAN OTHMAN[1,2,♥], FARAH ALIA NORDIN[1,♥♥], NURUL SHAKINA MOHD TALKAH[1]

[1]School of Biological Sciences, Universiti Sains Malaysia. 1112, Persiaran Sains, 11800 Gelugor, Pulau Pinang, Malaysia. Tel.: +604-653-3181, ♥♥email: farahalianordin@usm.my

[2]Centre of Chemical Biology, Universiti Sains Malaysia. Sains@USM, Blok B No. 10, Persiaran Bukit Jambul 11900, Bayan Lepas, Pulau Pinang, Malaysia. Tel: +604-653-5500, Fax.: +604-653-5514, ♥email: sofiman@usm.my

[3]School of Biological Sciences, Al-Muthanna University. Samawah, 66001, Iraq

Manuscript received: 21 December 2023. Revision accepted: 16 February 2024.

Abstract. *Jasim JHM, Othman AS, Nordin FA, Talkah NSM. 2024. High-quality genomic DNA extraction methods of Yellow* Spathoglottis *Blume complex for next-generation sequencing. Biodiversitas 25: 654-663.* Advancements in genomic research have spurred a growing interest in extracting high-quality genomic DNA, particularly from intricate plant species like the Yellow *Spathoglottis* Blume Complex. The preparation of genomic DNA high-quality from a plant sample is a crucial step for genomic and genetic analysis studies. A variety of genomic DNA extraction methods such as the traditional Hexadecyltrimethylammonium Bromide (CTAB) method and commercially available kits have been reported. Even so, they are either low purity and yield or costly. We develop a good-quality genomic DNA extraction method from fresh and dried leaves of the Yellow *Spathoglottis* Blume complex in Peninsular Malaysia, specifically tailored for Next-Generation Sequencing (NGS) applications. Three DNA extraction methods were compared traditional CTAB, modified CTAB, and the DNeasy Qiagen plant mini kit methods. The yield and quality of extracted DNA were assessed by PCR amplification using nuclear ETS and ITS genes performed to evaluate the success of DNA extraction. Additionally, the suitability of the modified CTAB method for NGS library preparation and sequencing was tested. It demonstrated better PCR amplification results, yielding high-concentration bands for both ETS and ITS primers. Furthermore, the modified CTAB method proved suitable for high molecular weight DNA extraction, without contamination. The extracted DNA from *Spathoglottis aurea* exhibited high molecular weight and passed the quality control assessment for NGS library preparation. Sequencing results demonstrated an average quality score of 36.0, with an accurate base call rate of approximately 99.9%. The developed method enables reliable genomic analysis and sequencing of the Yellow *Spathoglottis* Blume complex, contributing to further research and understanding of this species group.

Keywords: CTAB method, Fasta QC, internal transcribed spacer, molecular studies, next-generation sequencing

Abbreviations: CTAB: Cetyltrimethylammonium Bromide, NGS: Next-Generation Sequencing, ITS: Internal Transcribed Spacer, SSR: Microsatellites, ISSR: Inter-Simple Sequence Repeats, AFLP: Amplified Fragment Length Polymorphism, CIA: Chloroform Isoamyl Alcohol

INTRODUCTION

The genus *Spathoglottis* Blume complex belongs to the Orchidaceae family and comprises 44 subspecies. The yellow *Spathoglottis* Blume complex species are a problematic taxonomy (Nordin et al. 2023). A molecular study among this species in Malaysia was limited, therefore, a genomic approach will be done to solve the taxonomic confusion (Ginibun et al. 2019). Genetic studies have been conducted with molecular markers in *Spathoglottis* Blume complex, such as DNA Barcoding (Tsaballa et al. 2023) and DNA amplification fingerprinting (Nordin et al. 2022), microsatellites (SSR) or Inter-Simple Sequence Repeats (ISSR) (Sari and Aloysius 2021), and Amplified Fragment Length Polymorphism (AFLP) (Ginibun et al. 2018). All previous techniques need high-quality DNA with sufficient amount. High-quality genomic DNA extraction is the cornerstone for various downstream applications in modern molecular biology. The simplicity of the methods described in the literature (Grohme et al. 2018), underscores the practicality and efficiency of these approaches. Ensuring the success of NGS techniques requires large-scale high molecular weight genomic DNA extraction. With the advent of Next-Generation Sequencing (NGS) technologies, the investigation into the genome studies of plant species has been easier. NGS technologies, such as the Illumina HiSeq and TruSeq platforms, allow for the investigation of the entire genome of plants and as such require the input of several micrograms of high-quality DNA. The quality and quantity requirements for plant DNA sequencing by NGS often confine extraction methods to using leaf material, which is problematic due to the accumulation of high amounts of phenolics and polysaccharides. According to (Ambra et al. 2023) polysaccharides are known to obstruct the creation of NGS libraries because of their chemical characteristics, which cause them to coprecipitate with genomic DNA and give solutions a sticky, glue-like look. high-quality DNA is characterized as DNA that is

predominantly high molecular weight with an A260/280 ratio between 1.8 and 2.0 and without contaminating substances, such as polysaccharides or phenolics (Healey et al. 2014; Habibi et al. 2022). Since the prerequisite of NGS starting material is high-quality genomic DNA extract, it is critical to establish a method that delivers such requirements with high reproducibility (Tsaballa et al. 2023). However, the extraction of genomic DNA from *Spathoglottis* Blume Complex individuals is found to be challenging due to the structure of its leaves which are fibrous and thick lignin elaborate physical structure of the leaves makes it hard to extract DNA (Chikmawati 2013), based on the foregoing, there are many ways to extracted DNA from plants, one of the most widely method used for obtaining total DNA from plants uses the ionic detergent Cetyltrimethylammonium Bromide (CTAB) (Dharajiya et al. 2017; Kouakou et al. 2022). Also, several commercial kits are available to extract genomic DNA from plant tissues with sufficient quality, but the yield of DNA produced from commercial kits is often low and the cost can be prohibitive (Grohme et al. 2018). However, these methods produce either insufficient amounts or inconsistent quality of DNA for molecular studies and it is different from plant to plant depending on plant and tissue type (Jabeen et al. 2022). In plants, DNA can be extracted from various tissues, such as leaf, stem, root, fruit endocarp and embryo. These tissues can be used to answer different ecological and microevolutionary questions. Leaf and stem, for example, are often used in population genetics studies (Li et al. 2020). Therefore, these methods need modification and standardization in DNA isolation protocols. Modification is usually done in the concentration of chemicals used during the extraction procedure according to the plant used for DNA extraction (Leza et al. 2017; Pandey et al. 2019) or the modification includes some protocol steps change (Dharajiya et al. 2017).

The main objective of this study to find out which method produces the highest DNA concentration and purity for molecular studies such as NGS, by using the modified CTAB Method to compare with the traditional CTAB method as described in Doyle and Doyle (1990), and the DNeasy Qiagen mini plant kit method to extract DNA from dried and fresh leaves of various yellow *Spathoglottis* Blume complex.

MATERIALS AND METHODS

Experimental designs and plant materials

Spathoglottis individuals were collected from Peninsular Malaysia and deposited in the USM herbarium (Table 1). The fresh samples were kept in plastic bags with silica gels to dry before proceeding for extraction. In this study, nineteen yellow *Spathoglottis* species and one sample purple *Spathoglottis plicata* as outgroups were used. Three DNA extraction methods were utilized; the first method follows Doyle and Doyle (1990) (Handayani et al. 2016) as the traditional CTAB method. Secondly, the CTAB method was utilized with some modifications and thirdly, the extraction process was done using the commercially available kit DNeasy Qiagen plant mini kit.

Solutions, reagents, and consumables

CTAB buffers required for this study were prepared manually. Noteworthy, the distilled water utilized to prepare the buffers is sterilized to avoid any potential contamination. Additionally, it is critical to ensure all chemical components in the buffer are homogenized by briefly heating and swirling inside the water bath. It is also necessary to adjust the pH for buffers before use as pH 8.0. Once β-mercaptoethanol was added to CTAB free buffer and CTAB extract buffer, the buffers' shelf life is only 2-3 days. Thus, β-mercaptoethanol was only added to the buffers before the extraction process started.

Total genomics DNA extraction methods

Genomic DNA extraction was performed according to the procedure described by Doyle and Doyle (1990) as the traditional method, CTAB with some modifications as modified method and using a commercially available extraction kit (DNeasy Qiagen mini plant kit) following manufacturer protocol. The modifications of the CTAB method included as (i) at the beginning of the protocol, and after grinding leaves tissue with liquid nitrogen, CTAB-free buffer was added it was then kept at -20°C for 10 min, then the sample was centrifuged at maximum speed for 5 min. After that, the supernatant was discarded, this step was repeated three times with dry samples and one time with fresh samples; (ii) the Chloroform Isoamyl Alcohol (CIA) 24:1 was added three instead of two times, and (iii) after adding the ice-cold isopropanol, the mixture was stored at -20°C for 1-2 hours. The main purpose of this modification is to reduce proteins, polysaccharides, and phenol contaminations in which resulting from the breakdown of cell components.

CTAB modified method

The amount of 0.1-0.5 g of fresh and dried leaves were flash frozen with liquid nitrogen and thoroughly ground to fine powders using a mortar and pestle. The powders were then transferred immediately to a sterile and clean 2.0 mL microcentrifuge tube before it thawed. This step should be done quickly to prevent the degradation of DNA by endogenous endonucleases (≥10 minutes). Subsequently 750 µL CTAB free buffer [200 mM Tris-HCl pH 8.0, 50 mM EDTA and 250 mM NaCl and buffer 5 µL β-mercaptoethanol] and 5 µL β-mercaptoethanol were then added before vortexed for few seconds and incubated at -20°C for 10 minutes. Centrifugation at maximum speed for 5 minutes, was done before supernatant was discarded. The step was repeated 1-3 times depending on the type of samples used. Approximately 600 µL of preheated CTAB extraction buffer [2% (w/v) CTAB, 2%, (w/v) Polyvinyl-Pyrrolidone (PVP-40T), 1.4 M sodium chloride (NaCl), 100 mM Tris-HCl, pH 8.0; 20 mM Ethylenediamine Tetra-Acetic Acid (EDTA), and 1% (v/v) β-mercaptoethanol] was added to the sample and, mixed by vortex and the homogenized mixture was then incubated at 65°C in a water bath for 60 minutes. After incubation, the mixture was let to cool at room temperature for 2 minutes and, then 400 µL of chloroform-isoamyl alcohol (24:1) was added to the mixture. The solution was thoroughly mixed into a

single phase by gently inverting the tube and centrifuged at 10,000 rpm for 10 minutes. After that, the supernatant (uppermost layer) was transferred into a new sterile 2.0 mL microcentrifuge tube, while the lower phase was discarded. In the next step, 200 µL of 10% CTAB (w/v) solution and 400 µL of chloroform-isoamyl alcohol (24:1) were added into the uppermost layer. Again, the solution was mixed by gently inverting the tube, followed by centrifuged at 10,000 rpm for 10 minutes. The uppermost layer (excluding the green and white layer) was then transferred to a new 1.5 mL microcentrifuge tube. Then 5 µL RNase was added to the upper phase and incubated in a water bath at 37°C for 30 min, after incubation, the DNA was extracted again by using an equal amount of chloroform-isoamyl alcohol (24: 1) and centrifuged at 10,000 rpm for 10 minutes. The uppermost aqueous layer was then transferred into a new sterilized 1.5 µL microcentrifuge tube before 200 µL of CTAB precipitation [1% CTAB; 10 mM EDTA, pH 8.0; 50 mM Tris-HCl, pH 8.0] and 400 µL of ice-cold isopropanol] 400 µL were added. Gently, the tube was inverted to mix the solution. At this step, the tube should not be shaken vigorously because DNA is very sensitive to fragmentations. Then, to precipitate the DNA, the samples were left to incubate at -20°C for 1-2 hours. The samples were then centrifuged at 10,000 rpm in 10 minutes. The supernatant was discarded carefully without touching the pellet. The pellet which contains DNA was washed using 600 µL of wash buffer I [76% ethanol; 2.5 M sodium acetate (NaAc), pH 5.0] and left at room temperature for 30 minutes. The samples were then centrifuged at 10,000 rpm for 1 minute. The supernatant was discarded. Again, the DNA pellet was washed using 400 µL of wash buffer II [76% ethanol; 1.0 M sodium acetate (NaAc), pH 5.0] and left at room temperature for 30 minutes with another round of centrifugation at 10,000 rpm for 1 minute. The supernatant was discarded, and the DNA pellet was allowed to air dry for 2 hours at room temperature. At the end, the dried DNA was reconstituted with 100 µL of ddH20 or 1X TE Buffer. The DNA extract was stored at -20°C.

Table 1. Species list, collection numbers, and collection localities of the yellow *Spathoglottis* species and outgroup used in this present work

Species name	Collection number	Collection areas	Sample type	GPS coordinates (Lat. & Lon.)
Spathoglottis affinis	FAN028	Padang Tok Sheikh, G. Jerai, Kedah	Dried leaves	N05°47'29", E100°26'04"
Spathoglottis aurea	FAN030	Abu Suradi Trail, Fraser's Hill, Pahang	Dried leaves	N03°42'49", E101°44'05"
Spathoglottis aurea	FAN037	G. Brinchang, Cameron Highlands, Pahang	Dried leaves	N04°31'33", E101°23'05"
Spathoglottis aurea	FAN054	Padang Tok Sheikh, G. Jerai, Kedah	Dried leaves	N05°47'29", E100°26'04"
Spathoglottis aurea	FAN057	G. Bunga Buah, Selangor	Dried leaves	N03°22'47", E101°45'44"
Spathoglottis aurea	FAN099	Mile 39, Tanah Rata, Cameron Highlands, Pahang	Dried leaves	N04°10'49", E102°11'37"
Spathoglottis eburnean Gagnep	FAN022	Fang District, Chiang Mai, Thai Myanmar Border	Dried leaves	N20°05'36", E99°06'15"
Spathoglottis gracilis	FAN094	Kg. Liposu, Ranau, Sabah	Dried leaves	N05°57'39", E116°38'06"
Spathoglottis kimballiana Hook.f.	FAN085	Ranau, Sabah	Dried leaves	*
Spathoglottis kimballiana var. *angustifolia Ames*	FAN076	BiduBidu FR, Telupid, Sabah	Dried leaves	N05°46'39", E117°14'45"
Spathoglottis kimballiana var. *kimballiana*	FAN067	Mt. Kinabalu, Ranau, Sabah	Dried leaves	N06°02'03", E100°26'04
Spathoglottis microchilina	FAN083	Kinabalu Park Research Centre, Sabah	Dried leaves	N06°00'15", E116°32'39"
Spathoglottis microchilina	FAN086	Entrance of Mamut Copper Mine, Sabah	Dried leaves	N06°00'38", E116°39'28"
Spathoglottis microchilina	FAN091	Sg. Lohan, Ranau, Sabah	Dried leaves	N06°01'37", E116°40'29"
Spathoglottis aurea	FAN039	G. Ulu Kali, Genting Highlands, Pahang	Dried leaves	N03°26'20", E101°47'02"
Spathoglottis pubescens	FAN068	Fang District, Chiang Mai, North Thailand	Dried leaves	N19°56'60", E99°10'59"
Spathoglottis aurea	UNS001	G. Jerai, Kedah	Fresh leaves	N05°48'55", E100°26'03"
Spathoglottis plicata	*	Taman Rimba, G. Jerai, Kedah	Fresh leaves	N05°48'37", E100°26'03"
Spathoglottis affinis	*	Tok Sheikh, G. Jerai, Kedah	Fresh leaves	N05°54'977", E100°26'04"
Spathoglottis aurea	*	G. Jerai, Kedah	Fresh leaves	*

Note: (*) Collection Numbers and GPS coordinates are not available for locations denoted with an asterisk

PCR amplification of genomic DNA

Internal Transcribed Spacer (ITS) of nuclear ribosomal DNA is one of the most used DNA markers in plant phylogenetic studies. The marker is recommended as a core plant DNA barcode. The ITS primers sequences utilized for this study (Forward-5' GGAAGTAAAAGTCGTAACA AGG-3'; Reverse- 5′-TCCTCCGCTTATTGATATGC-3') acquired from (Pandey et al. 2019). Additionally, another nuclear marker that has been reported to give good resolutions in plant phylogenetic studies is External Transcribed Spacer (ETS). The primers utilized for amplification of ETS sequences are retrieved from (Alonso et al. 2014) ETS-AF: 5′-GACCGTGACGGYACGTGAG-3' 18S-R: 5′-AGACAAGCATATGACTACTGGCAGG-3'). In this study, nuclear DNA primers were used to check the applicability of the extracted DNA in PCR (Nordin et al. 2022; Tsaballa et al. 2023). All primers were synthesized commercially, purchased from Integrated DNA Technologies company in Asia Pacific. In brief, the PCR reaction mixtures (25 µL) contained 0.5 µL (10 ng) of sample DNA, 12.5 µL of 2X master mix and 0.5 µL of each forward and reverse primer (10 µM) then completed of volume to 25 µL with free nuclease water. The reaction was carried out in Biometra thermal cyclers (Analytik Jena, Germany) with initial activation of the DNA polymerase at 94°C for 5 min, followed by 30 cycles of denaturation at 94°C for 60 s, annealing at 50-56°C for 60 s and extension at 72°C for 60 s. The final extension was carried out at 72°C for 7 min. The PCR products were visualized on 1.0% agarose gel and ran at 90 V for 1 h. Gel images were captured on a gel documentation system (Chemidoc MP, BioRad, USA). Ultimately, PCR products amplified with ITS and ETS primers resulted in 748 and 633 bp DNA fragments respectively.

Next-generation sequencing library submission, preparation, and sequencing

High-quality DNA was extracted from *S. aurea* (UNS001) for NGS library preparation and sequencing. To ensure that the DNA was high yield and quality, electrophoresis and NanoDrop spectrophotometer were performed. The electrophoresis conditions 1% agarose gel, 45 mins running at 110 V with 1 µL of DNA loaded. Three replicates with 50 µL of total genomic DNA isolated by the three methods were sent for short-reads Next-Generation Sequencing (NGS). After initial Quality Checks (QC), the DNA was processed for library preparation (150 bp paired-end) and sequenced on an Illumina Novaseq 6000 platform.

Paired-end reads were removed of low-quality reads (below Phred score Q20) and sequence adaptor using Cutadapt Version 3.50 (Freedman et al. 2020) implemented in Trim-galore Version 0.6.7 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). The quality of the clean read was inspected using FastQC Version 0.11.8 and compiled into a single report using MultiQC Version 1.12 (Ewels et al. 2016).

Statistical analyses

To compare the efficiency of the three DNA extraction methods, average DNA yield rates were calculated and compared based on the DNA concentration and purity. DNA yield rate (log-transformed for normal distribution) and absorbance ratio were analyzed using one-way ANOVA followed by T-test with a threshold of $P<0.01$ in SPSS Statistics version 19.0.0 (SPSS Inc, Chicago, IL, USA).

RESULTS AND DISCUSSION

Isolation of plant genomic DNA and checking of the yield and quality

Three methods for DNA extraction were compared in this study. The genomic DNA was extracted using both fresh and dried leaves using 2X Cetyltrimethylammonium Bromide (CTAB) method as explained in Doyle and Doyle (1990), CTAB modified method with minor modifications and DNeasy Qiagen extraction kit. The methods for DNA extraction from yellow *Spathoglottis* Blume complex species were traditional CTAB, modified CTAB, and DNeasy Qiagen plant mini kit methods. Each protocol provided significantly different results concerning DNA concentration, quality, and contaminant carryover, (Figure 2). The mean and median for concentration and purity were higher in modified CTAB than both other methods (traditional CTAB and Qiagen kit methods) see Table 2. Also, the range of DNA concentrations with dried *Spathoglottis* leaves was higher in modified CTAB method (345-1424 ng/µL), while traditional CTAB method recorded 294.5-1112 ng/µL and the low concentration range (22.1-168 ng/µL) recorded by Qiagen kit method. The optimum range DNA purity A260/280 for dried leaves was high in the modified CTAB at 1.78-1.91, then Qiagen kit and traditional CTAB methods (1.64-1.78) and (1.57-1.77) respectively.

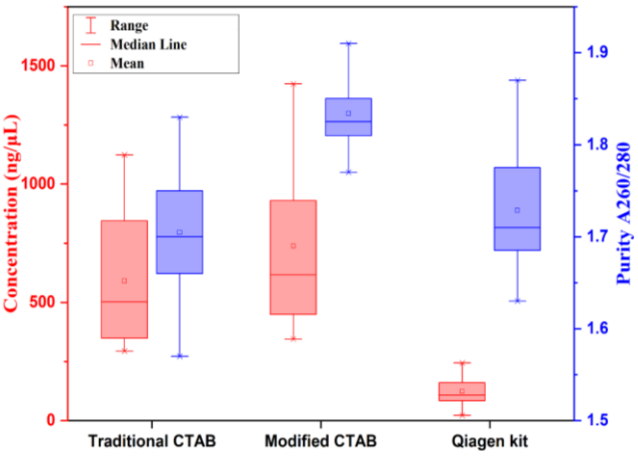

Figure 1. The range, median, and mean of DNA concentration and purity for three DNA extraction methods: Traditional CTAB method, Modified CTAB method, and Qiagen kit method. Data were drawn by origin software 10.5.124 version

Table 2. Statistics of DNA concentration and purity extracted from yellow *Spathoglottis* species using three methods. Data analysis by SPSS Statistics version 19.0.0 (SPSS Inc, Chicago, IL, USA) was shown for one-way ANOVA, followed by a t-test a posteriori test with statistically significant P values. SD: Standard Deviation, SE: Standard Error, and *Significantly higher (t-test p<0.01)

	Traditional CTAB method		Modified CTAB method		Qiagen kit method	
	DNA concentration (ng/µL)	DNA purity A260/280 (OD)	DNA concentration (ng/µL)	DNA purity A260/280 (OD)	DNA concentration (ng/µL)	DNA purity A260/280 (OD)
Samples no.	20		20		20	
Mean	573.4	1.704	737.5	1.853	123.3	1.74
SD	260.6	0.071	341.5*	0.046*	61.36	0.090
Minimum	294.5	1.57	345.4	1.78	22.1	1.63
Median	502.1	1.7	616	1.84*	108.2	1.72
Maximum	1124	1.83	1424	1.91	243.4	1.87
SE of mean	58.28	0.015	76.36	0.010	13.72	0.020

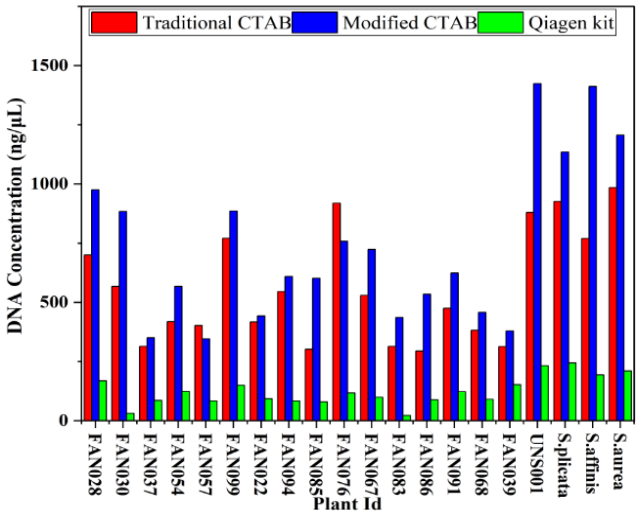

Figure 2. DNA concentration for traditional CTAB, modified CTAB, and DNeasy Qiagen plant mini kit methods. (FAN028 FAN093) referred to plant Id for dried Yellow *Spathoglottis* leaves and fresh leaves (UNS001-*S. aurea*) used

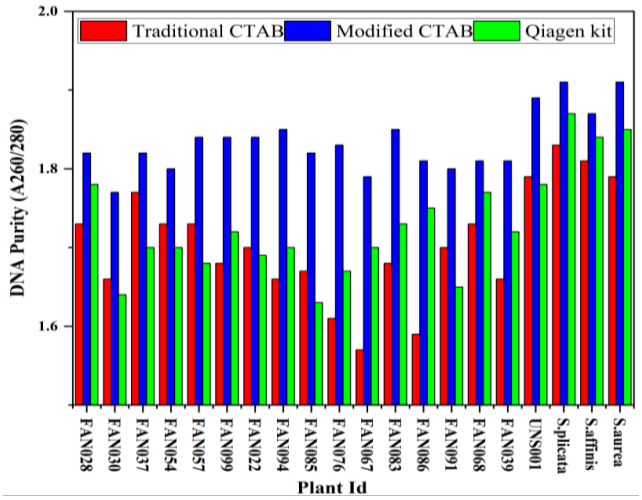

Figure 3. DNA purity (A260/280) for traditional CTAB, modified CTAB, and DNeasy Qiagen plant mini kit methods. (FAN028-FAN093) referred to plant Id for dried Yellow *Spathoglottis* leaves and fresh leaves (UNS001, *S.plicata*, *S affinis* and *S. aurea*) used

DNA extraction from dried and fresh leaves

To compare DNA extraction methods from dried and fresh leaves, we used fresh leaves for four yellow *Spathoglottis* species and dried leaves for 16 species collected from the Malaysian Peninsulas see Table 1. Generally, the amount of DNA obtained from modified CTAB method by used fresh leaves was higher yield than the DNA obtained from dried leaves as show in Figure 4. The highest concentrations were recorded with fresh leaves of *S. aurea* (UNS001) at 1424.1 ng/µL, while the dried samples, were recorded with *Spathoglottis.affinis* (FAN028) at 975.2 ng/µL were used.

On the other hand, the purity of DNA extracted by modified CTAB at wavelength A260/A280 was better in fresh leaves than with dried leaves as shown in Figure 5, where it was high in *S. plicata* and *S. aurea* recorded the same absorption at 1.91, while *S. aurea* "FAN038" was recorded at 1.85 by using dried leaves.

The modified CTAB method gave us a positive DNA extraction in both fresh and Dried samples with good and clear bands see Figures 5C and 5D. While in the traditional CTAB method, all samples were positive, but the bands were smeared more, and degraded DNA was found as shown in Figures 5A and 5B. In the Qiagen kit method, six species were negative bands and positive DNA bands have very low concentrations, unclear and with some smear (see Figure 5).

PCR amplification

An important component of any DNA extraction method is its use in downstream applications. Therefore, we compared PCR amplification of yellow *Spathoglottis* species with two nuclear primers; ITS and ETS (see Table 3).

A representative agarose gel containing 633 bp fragment of the ETS gene and a ~748 bp fragment of the ITS gene that was successfully amplified from DNA samples extracted by three methods used in this study. Good band and amplification occurred within DNA extracted by the modified CTAB method with all samples (see Figures 6 and 7).

High molecular weight DNA extraction for next-generation sequence

The suitability of using modified CTAB method for NGS platforms was evaluated based the purity and concentration of extracted genomic DNA. *S. aurea*

(UMS001) was high molecular wight at 1424.1 ng/ml and 1.89 absorbance peak at A260/280. The gel electrophoresis image performed by Macrogen referred to high molecular weight DNA band with no RNA or polysaccharide contamination. Based on the Quality Control (QC) report of NGS library, the modified protocol was suitable for the isolation of DNA for whole genome NGS library preparation and sequencing (Table 4). The other samples which were extracted by traditional CTAB and Qiagen kits were not passed for library preparation and sequenced on an Illumina Novaseq 6000 platform, because there was more contamination and less concentration.

Read quality distributions based upon FASTAQC quality scores, ranged between 34.0-37.0 (mean-36.0). A base call accuracy of approximately 99.9%, favorably desirable for big data sequencing applications, was obtained (Figures 10A and 10B).

Table 3. The quality of PCR product for two used primers (ETS and ITS genes) with *Spathoglottis* species

Plant ID	Traditional CTAB method		Modified CTAB method		Qiagen kit method	
	ETS	ITS	ETS	ITS	ETS	ITS
FAN028	+++SD	++SD	++	+++	+D	+++
FAN030	-	-	++	+++	++	+++
FAN037	++	-	+++	++	-	-
FAN054	+++SD	+++SD	++	+++	-	-
FAN057	+++SD	+++SD	+++	++	+	++
FAN099	++	+++SD	+++	+++	+	+
FAN022	++	+D	+++	++	++	+++
FAN094	+++SD	+++SD	+++	+++	-	-
FAN085	+++SD	+++SD	+++	+++	+	+
FAN076	+++SD	+D	+++	++	+	+++
FAN067	+++SD	+++SD	+++	++	+	+
FAN083	+++SD	++D	+++	+++	-	-
FAN086	-	-	+++	++	-	-
FAN091	+	+++SD	++	+++	-	-
FAN068	+++SD	+++SD	++	+++	-	-
FAN039	+++SD	+++SD	+++	+++	-	-
UMS001	+++SD	+++SD	+++	+++	++	+++
S. plicata	+++SD	+++SD	+++	+++	++	+++D
S. affinis	+++SD	+++SD	+++	+++	++	+++
S. aurea	+++SD	+++SD	+++	+++	++	+++

Note: +++: High, ++: Medium, and +: Low concentration band, -: Negative sample, S: Smeared, D: Degraded

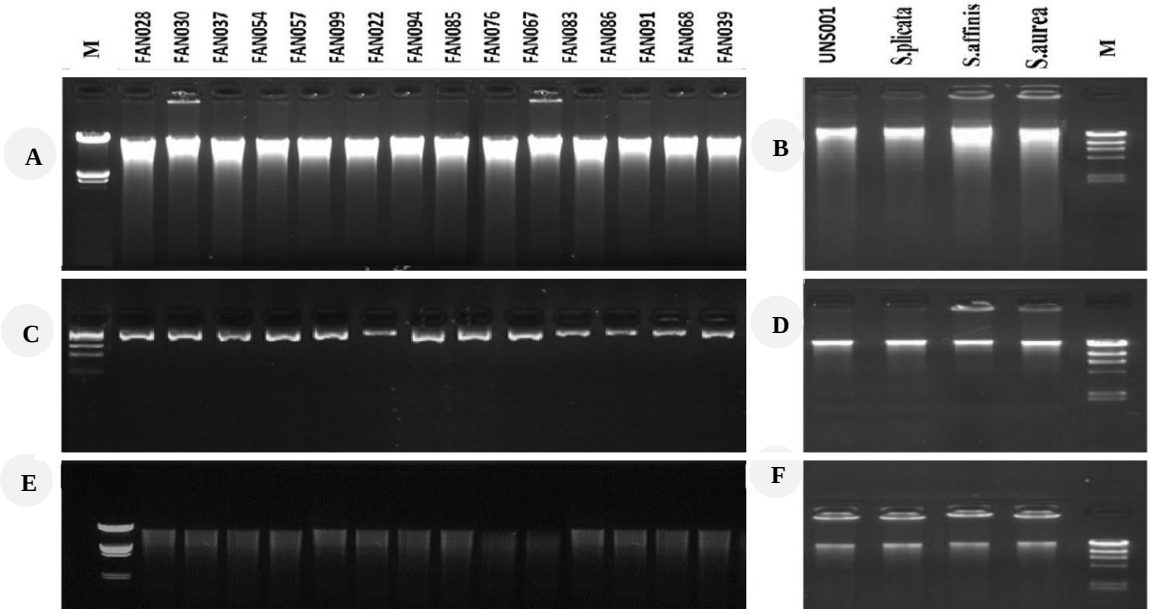

Figure 4. Agarose gel (0.8%) electrophoresis of total DNA extracted from dry (A,C,E) and fresh leaves (B,D,F) of *Spathoglottis* species by using three commonly used methods. A and B represent genomic DNA isolation by the traditional CTAB method, C and D refer to total DNA extraction by the modified CTAB method, then E and F refer to DNA isolated by DNeasy Qiagen plant mini kit method. Line M lambda/HindIII DNA. Line from FNA028 to FNA039 represents dried samples of yellow *Spathoglottis* species. Line UNS001-S4 refers to the fresh leaves of the *Spathoglottis* species

Table 4. NGS library QC for *Spathoglottis aurea* (UMS001)

Library name	Library type	Concentration (ng/uL)	Concentration (nM)	Size (bp)	Total amount (ug)	Result
UMS001	TruSeq Nano DNA (350)	13.30	43.53	612	1.441	Pass

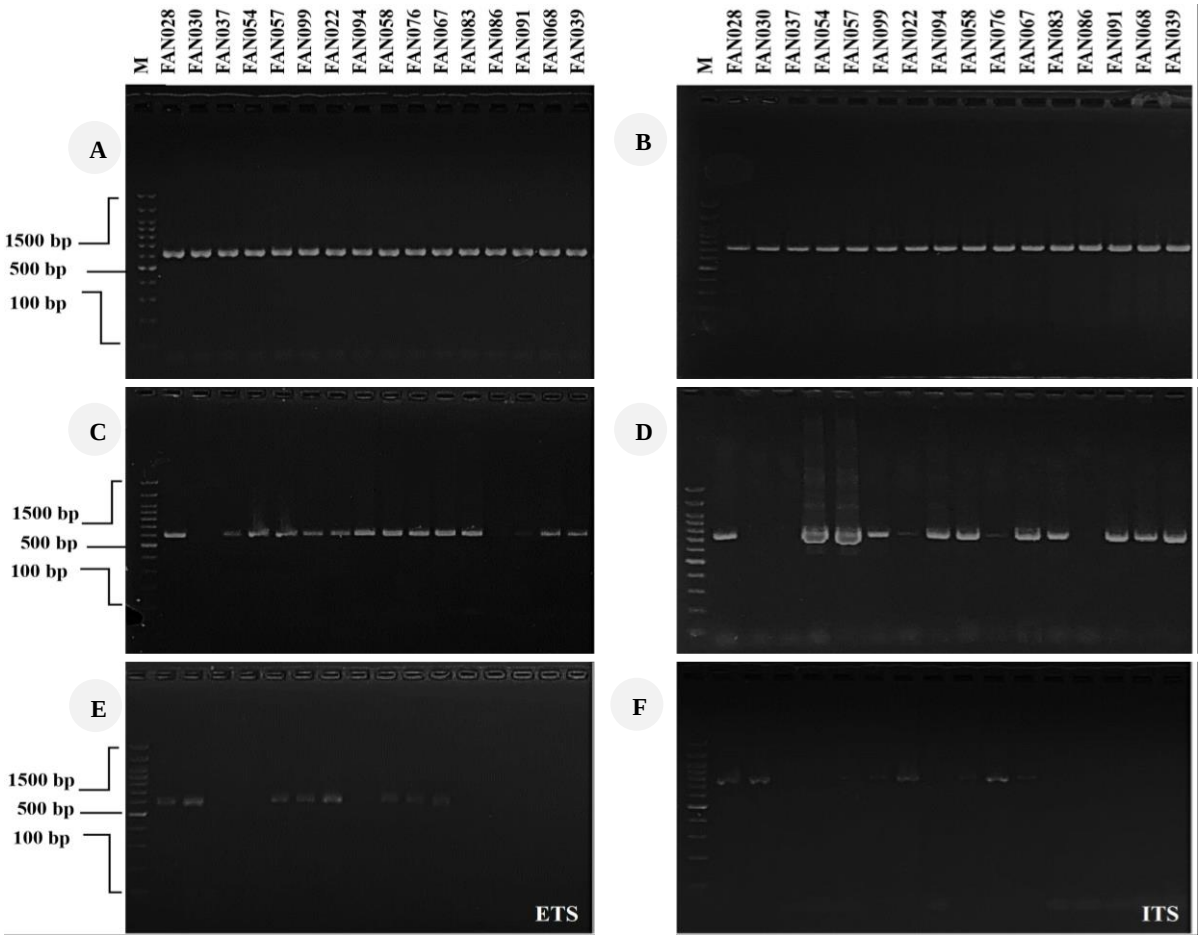

Figure 5. Comparison of PCR amplification using genomic DNA extracted with dried leaves of *Spathoglottis* species. Genomic DNA was used to amplify a 633-bp and 748-bp fragment of a *Spathoglottis* species with ETS (A, C and E) and ITS (B, D and F) genes. The PCR products were separated in a 1% agarose gel. Images A and B represent PCR products amplified with DNA template isolation by modified CTAB method, C and D refer to PCR products amplified with total DNA extraction by the traditional CTAB method, E and F refer to PCR products amplified with DNA isolated by DNeasy Qiagen plant mini kit method. Line from FNA028 to FNA039 represents spices ID for *Spathoglottis* species. Line M referred to 1500 bp DNA marker

Discussion

Since the advent of the CTAB-based extraction method from plant leaves by Doyle and Doyle in (1990) (Handayani et al. 2016) many different iterations have been published, each with modifications to contend with the co-extractives of polyphenolics and polysaccharides present in the leaves of many plant species (Tamari et al. 2013; Drábková 2014; Dharajiya et al. 2017; Kouakou et al. 2022).

Our study highlights the focus of the research on developing an efficient and reliable method for extracting genomic DNA from the yellow *Spathoglottis* Blume complex. The objective is to obtain high-quality DNA suitable for genomic studies, especially Next-Generation Sequencing (NGS) applications. For next-generation sequencing has revolutionized genomic research by enabling rapid and cost-effective analysis. However, the success of NGS largely depends on the quality and integrity of the input DNA. Therefore, the development of an optimized DNA extraction method is crucial to ensure accurate and reliable sequencing results. (Healey et al. 2014; Habibi et al. 2022). We made three critical modifications to the CTAB method. First, prewash with free CTAB buffer to reduce proteins and polysaccharides (Pandey et al. 2019). This step can be repeated from one to three times, depending on the clarity and viscosity of the supernatant. When adding this step, we noticed an increase the DNA purity.

The second improvement, the chloroform extraction step was added three times instead of two times after the CTAB extraction buffer. The step chloroform extraction significantly improved the DNA yield and quality, without

this step, we were barely able to recover 200 µL of the aqueous phase out of 750 µL extraction buffer added to dry tissue. By adding a chloroform extraction step, we could easily recover over 600 µL of the aqueous phase. In a high throughput format, we routinely took 500 µL of the aqueous phase from each sample with a micropipette without the risk of contamination from the organic phase. In addition, chloroform extraction removes polysaccharides, lipids, and other nonpolar substances from the aqueous phase, resulting in cleaner DNA.

The third improvement was the time decrease of DNA precipitation, which led to a decrease in the precipitation of contamination and thus leads to an increase in the purity of the DNA. Most CTAB DNA extraction methods reported in the literature (Tamari et al. 2013; Lucas et al. 2019) are variations of the method reported by Doyle and Doyle 1990, in which DNA in a CTAB extraction buffer is precipitated by adding 0.5 volume isopropanol then incubated for two days. In our experiment, we reduced the storage time after adding the precipitation buffer to two hours, therefore we got sufficient DNA yield with high quality than other methods (Table 3 and Figure 2), these supported by (Leza et al. 2017). The results demonstrate that the modified CTAB method consistently outperforms the traditional CTAB and DNeasy Qiagen plant mini kit methods in terms of DNA concentration, purity, and amplification quality. The modified CTAB method yields higher DNA concentrations, providing more starting material for NGS library preparation. These findings suggest that the modified CTAB method effectively removes contaminants and optimizes DNA extraction from the yellow *Spathoglottis* Blume complex. Furthermore, the study verifies the suitability of the modified CTAB method for NGS by assessing the high molecular weight and absence of RNA or polysaccharide contamination in the extracted DNA. The successful PCR amplification of target genes using the extracted DNA confirms its quality and suitability for downstream NGS applications The traditional CTAB method gave quality DNA yield with fresh samples then dried samples, this suitable for PCR but not for NGS. Additionally, the purity of extracted DNA was bad at A260/A280 ratio (Tables 4), suggesting that the preparations were not sufficiently free of proteins and polyphenolics/polysaccharide compounds. These results are at par with the findings of (do Nascimento Silva et al. 2021).

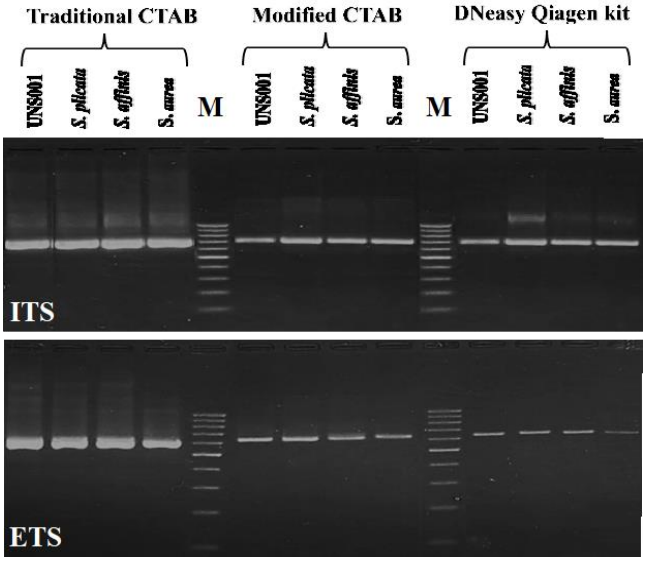

Figure 6. Comparison of PCR amplification using genomic DNA extracted with fresh leaves of yellow *Spathoglottis* species. The genomic DNA was used to amplify a 633-bp and 748-bp fragment of a yellow *Spathoglottis* Blume complex with two genes (ITS above image and ETS below image). The PCR products were separated in a 1% agarose gel. Line UNS001, *S. plicatia,S. affinis* and *S. aurea* represent spices Id. Line M referred to 100 pb DNA marker

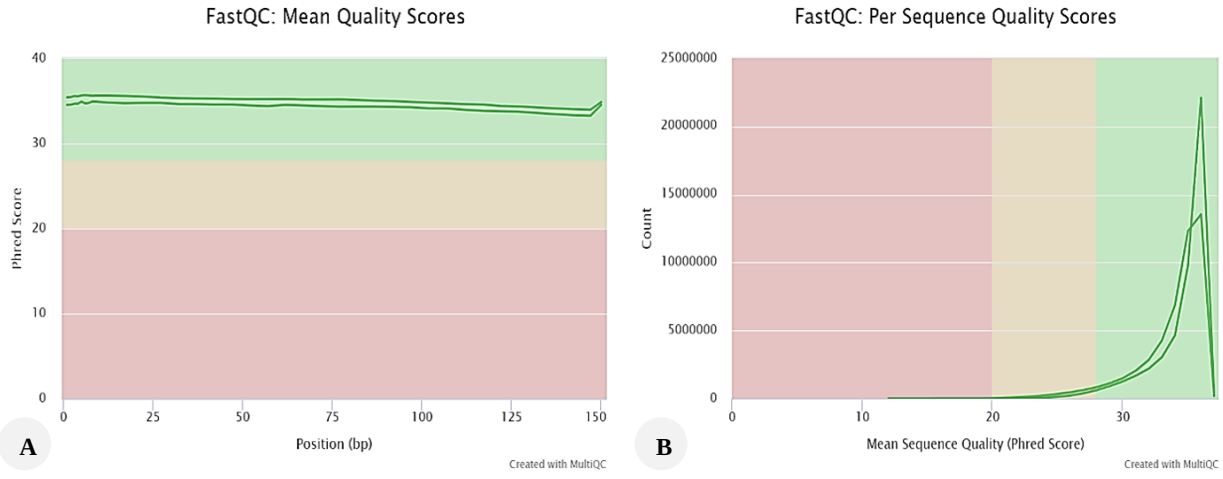

Figure 7. DNA sequencing quality details. FastQC report illustrating the average quality scores across all bases of the paired-end datasets for the sample (UNS001). A. Phred quality scores for each nucleotide position are represented as a box and whisker plot. The green line represents the mean quality. The yellow box represents the inter-quartile range (25-75%). The upper and lower whiskers represent the 10 and 40 % points. Image B. Average quality per sequence Quality Scores. The report was Produced by FastQC Version 0.11.8

In terms of DNA yield, the modified CTAB method showed higher concentrations of DNA when fresh leaves were used compared to dried leaves. The highest concentration of DNA was obtained from fresh leaves of *S. aurea* (UNS001) at 1424.1 ng/µL (Table 5). Among the dried samples, *Spathoglottis affinis* (FAN028) had the highest DNA concentration at 975.2 ng/µL. Regarding DNA purity, as measured by the A260/280 ratio, the modified CTAB method exhibited better results with fresh leaves compared to dried leaves. *S. plicata* and *S. aurea* showed the same absorption ratio of 1.91 when fresh leaves were used.

The visualization of the extracted DNA bands on an agarose gel (0.8%) showed that the modified CTAB method consistently produced clear and distinct bands for both fresh and dried samples (Figure 5C and 5D). In contrast, the traditional CTAB method resulted in smeared bands and degraded DNA (Figure 5A and 5B). The DNeasy Qiagen plant mini kit method had some limitations, with six species showing negative bands and low concentrations of positive DNA bands that were unclear and smeared (Figures 5E and 5F).

The results regarding PCR amplification of yellow *Spathoglottis* species using two nuclear primers (ETS and ITS) are presented in Table 4. The amplification results were compared for the DNA extracted using the traditional CTAB method, modified CTAB method, and DNeasy Qiagen plant mini kit method. From the table, it can be observed that the modified CTAB method consistently provided good PCR amplification results for both ETS and ITS genes, as indicated by the presence of high or medium concentration bands. In contrast, the traditional CTAB and DNeasy Qiagen plant mini kit methods showed more varied results, with some samples exhibiting negative amplification or smearing. The gel image confirms that the modified CTAB method resulted in successful amplification for all samples. In comparison, the traditional CTAB method and the DNeasy Qiagen plant mini kit method showed less consistent amplification. Additionally, the suitability of the modified CTAB method for high molecular weight DNA extraction for next-generation sequencing (NGS) platforms was assessed. The results showed that the DNA extracted from *S. aurea* (UM0001) using the modified CTAB method had high molecular weight and exhibited no contamination. Gel electrophoresis confirmed the presence of high molecular weight DNA bands without RNA or polysaccharide contamination. In terms of sequencing quality, the FastQC analysis demonstrated an average quality score ranging between 34.0 and 37.0, with a mean of 36.0. The base call accuracy was approximately 99.9%, which is desirable for reliable big data sequencing applications.

In conclusion, the improved CTAB approach is a significant advancement over previous techniques for improving genomic DNA extractions from leaf tissue of yellow *Spathoglottis* Blume complex species. High-quality DNA was extracted from yellow *Spathoglottis* Blume complex species using the procedure, and these species passed the stringent criteria for submission to the Illumina NGS library. Additionally, by evaluating the high molecular weight and absence of RNA or polysaccharide contamination in the extracted DNA, the study confirms the suitability of the modified CTAB approach for NGS. The extracted DNA's quality and appropriateness for subsequent NGS applications are validated by the successful PCR amplification of target genes using it.

ACKNOWLEDGEMENTS

We thank the School of Biology, Universiti Sains Malaysia for technical assistance with flow experiments, also we thank Herbarium in USM for providing us dried samples. The whole genome generated from *Spathoglottis aurea* is available online in GenBank under Accession OQ411079.

REFERENCES

Alonso A, Bull RD, Acedo C, Gillespie LJ. 2014. Design of plant-specific PCR primers for the ETS region with enhanced specificity for Tribe Bromeae and their application to other grasses (Poaceae). Botany 92 (10): 693-699. DOI: 10.1139/cjb-2014-0062.

Ambra R, Marcelli M, D'Orso F. 2023. Development of a sensitive, easy and high-throughput compliant protocol for maize and soybean DNA extraction and quantitation using a plant-specific universal Taqman Minor Groove binder probe. Genes 14 (9): 1797. DOI: 10.3390/genes14091797.

Chikmawati T. 2013. Anatomical and cytological features of *Spathoglottis plicata* from Java Island. J Trop Life Sci 3 (2): 87-90. DOI: 10.11594/jtls.03.02.03.

Dharajiya DT, Khadia SM, Pagi NK, Khatrani TJ, Jasani HV, Khunt AD, Ravindrababu Y. 2017. Modified method of high quality genomic DNA extraction from Mungbean [*Vigna radiata* (L.) Wilczek] suitable for PCR based amplification. Indian J Sci Technol 10 (20): 1-7. DOI: 10.17485/ijst/2017/v10i19/95082.

do Nascimento Silva J, Viana JPG, Costa MF, de Sá GH, da Costa Gomes MF, de Lima Feitoza L, Dos Santos Valente SE. 2021. A simple and cost-effective method for DNA extraction suitable for PCR in "*Sucupira branca*". Biosci J 37: e37092. DOI: 10.14393/BJ-V37N0A2021-54133.

Doyle JJ, Doyle JL. 1990. Isolation of plant DNA from fresh tissue. Focus 12: 13-15.

Drábková LZ. 2014. DNA extraction from herbarium specimens. Methods Mol Biol 1115: 69-84. DOI: 10.1007/978-1-62703-767-9_4.

Ewels P, Magnusson M, Lundin S, Käller M. 2016. MultiQC: Summarize analysis results for multiple tools and samples in a single report. Bioinformatics 32: 3047-3048. DOI: 10.1093/bioinformatics/btw354.

Freedman AH, Gaspar JM, Sackton TB. 2020. Short paired-end reads trump long single-end reads for expression analysis. BMC Bioinformatics 21 (1): 149. DOI: 10.1186/s12859-020-3484-z.

Ginibun FC, Arens P, Vosman B, Bhassu S, Khalid N, Othman RY. 2018. Genetic diversity of endangered terrestrial orchids *Spathoglottis plicata* in Peninsular Malaysia based on AFLP markers. Plant OMICS 11 (3): 135-44. DOI: 10.21475/poj.11.03.18.p1227.

Ginibun FC, Jaafar MA, Khalid N, Bhassu S, Othman RY. 2019. Intra- and interspecies variation of terrestrial orchid *Spathoglottis* species in Peninsular Malaysia using chloroplast DNA. Acta Hortic 1262: 195-204. DOI: 10.17660/ActaHortic.2019.1262.26.

Grohme MA, Vila-Farré M, Rink JC. 2018. Small- and large-scale high molecular weight genomic DNA extraction from Planarians. Methods Mol Biol 1774: 267-275. DOI: 10.1007/978-1-4939-7802-1_7.

Habibi N, Al Salameen F, Rahman M, Shajan A, Zakir F, Abdulrazzack N. 2022. Comparison and optimization of DNA isolation protocols for high throughput genomic studies of *Acacia pachyceras* Schwartz. MethodsX 9: 101799. DOI: 10.1016/j.mex.2022.101799.

Handayani F, Wulandari RA, Murti RH. 2016. Genomic DNA extraction method from mature leaf of Lai (*Durio kutejensis* Becc.). Agrivita 38 (1): 73-79. DOI: 10.17503/agrivita.v38i1.659.

Healey A, Furtado A, Cooper T, Henry RJ. 2014. Protocol: A simple method for extracting next-generation sequencing quality genomic DNA from recalcitrant plant species. Plant Methods 10: 21. DOI: 10.1186/1746-4811-10-21.

Jabeen R, Habiba U, Khalid S, Shirazi JH. 2022. Extraction and comparison of quality and purity of DNA of medicinal plants by CTAB and kit method. J Contemp Pharm 6 (1): 19-22. DOI: 10.56770/jcp2022613.

Kouakou J-L, Gonedelé-Bi S, Assamoi J-B, N'Guetta S-PA. 2022. Optimization of the Cetyltrimethylammonium Bromide (CTAB) DNA extraction protocol using forest Elephant dung samples. MethodsX 9: 101867. DOI: 10.1016/j.mex.2022.101867.

Leza AB, Hajare ST, Chauhan NM. 2017. Comparative analysis of DNA extraction methods in two popular varieties of Finger millet (*Eleusine coracana*) from Ethiopia. Biotechnol J Intl 20 (2): 1-7. DOI: 10.9734/bji/2017/33145.

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

Lucas MS, da Silva Carvalho C, Hypolito GB, Côrtes MC. 2019. Optimized protocol to isolate high quality genomic DNA from different tissues of a Palm species. Hoehnea 46 (2): e942018. DOI: 10.1590/2236-8906-94/2018.

Nordin FA, Raffi A, Go R, Yong CSY, Saibeh K, Othman AS. 2023. Morphological Systematics of *Spathoglottis* Blume (Orchidaceae: Collabieae) in Peninsular Malaysia and Borneo. Forests 14 (5): 940. DOI: 10.3390/f14050940.

Nordin FA, Saibeh K, Go R, Mangsor KNA, Othman AS. 2022. Molecular phylogenetics of the orchid Genus *Spathoglottis* (Orchidaceae: Collabieae) in Peninsular Malaysia and Borneo. Forests 13 (12): 2079. DOI: 10.3390/f13122079.

Pandey LK, Tantuway SP, Bano Y, Swamy N, Singh R, Bhuvaneswari S. 2019. Developing a protocol for the isolation of high quality DNA from carbohydrate rich medicinal plant for molecular identification. Intl J Mol Biol Biochem 1: 4-8. DOI: 10.33545/26646501.2019.v1.i1a.2.

Sari DR, Aloysius S. 2021. Genetic diversity of *Spathoglottis plicata* Blume orchid variant based on Inter-Simple Sequence Repeat (ISSR) molecular marker. In: Striełkowski W, Black JM, Butterfield SA, Chang C-C, Cheng J, Dumanig FP, Al-Mabuk R, Urban M, Webb S (eds). Proceedings of the 7th International Conference on Research, Implementation, and Education of Mathematics and Sciences (ICRIEMS 2020). Universitas Negeri Yogyakarta, Yogyakarta, 25-26 September 2020. DOI: 10.2991/assehr.k.210305.010.

Tamari F, Hinkley CS, Ramprashad N. 2013. A comparison of DNA extraction methods using Petunia hybrida tissues. J Biomol Technol 24 (3): 113-118. DOI: 10.7171/jbt.13-2403-001.

Tsaballa A, Kelesidis G, Krigas N, Sarropoulou V, Bagatzounis P, Grigoriadou K. 2023. Taxonomic identification and molecular DNA barcoding of collected wild-growing orchids used traditionally for Salep production. Plants 12 (17): 3038. DOI: 10.3390/plants12173038.