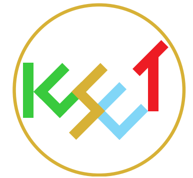
Kathmandu University Journal of Science, Engineering and Technology, Vol. 17, No. 1, June 2023

Kathmandu University
Journal of Science, Engineering and Technology



DNA extraction and PCR optimization for DNA barcode analysis of commercially-grown coffee varieties in Nepal

Shreejan Pokharel[a], Bignya Chandra Khanal[a], Anish Basnet[b], Gyanu Raj Pandey*[b], and Samsher Basnet[c]

[a]National Biotechnology Research Centre, Nepal Agricultural Research Council, Lalitpur, Nepal.
[b]Shubham Biotech Nepal Pvt. Ltd., Chitwan, Nepal.
[c]National Entomology Research Centre, Nepal Agricultural Research Council, Lalitpur, Nepal.

Abstract

The isolation of high-quality genomic DNA is an essential criterion for further molecular analysis. *Coffea* genus is well known for its high amount of polyphenols, polysaccharides, and other secondary metabolites that degrades the quality of the DNA isolation needed for further down streaming processes. The present work was carried out to obtain a simple and efficient DNA isolation protocol generating high-quality amplification for barcoding. The protocol involves modifying the CTAB extraction, incorporating the use of polyvinylpyrrolidone and β-mercaptoethanol yielding quality DNA with a ratio (A260/280) between 1.8–2.0 indicating low contamination. The PCR conditions were optimized for high amplification based on the optimal concentration of $MgCl_2$ (3 mM), primer (0.5 µM), Taq polymerase (0.2 U), 50–60 ng of DNA template, and cycle conditions as initial denaturation of 94°C for 4 min followed by 35 cycles of denaturation at 94°C for 50 sec, annealing (respective of barcodes) for 50 sec and extension at 72°C for 80 sec, followed by a final extension at 72°C for 7 min. The optimal conditions produced highly amplified reproducible data. Thus, the optimized method proposed enabled a simple DNA extraction and PCR amplification for *Coffea* genus and may serve as an efficient tool for further molecular analysis.

Keywords: Internal Transcribed Spacers (ITS); Maturase K (MatK); Ribulose-1 and 5-bisphosphate carboxylase large subunit (RbcL)

1. Introduction

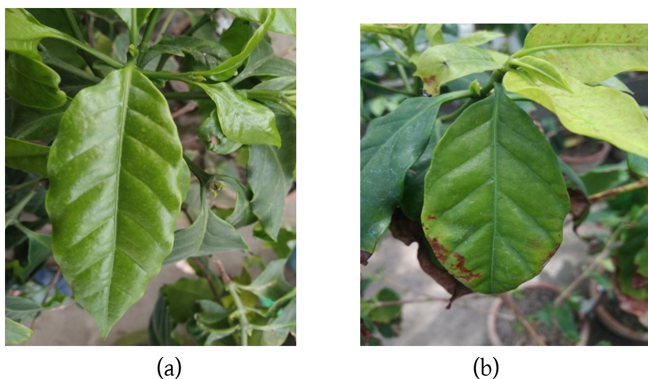
The family Rubiaceae includes the genus *Coffea*. Out of the 124 species in the genus *Coffea* [1], *Coffea arabica* and *Coffea canephora* are the two most widely farmed around the world. Among the popular ones, *Coffea arabica* is the highly cultivated one as it has superior taste, rich aroma, and low caffeine and is responsible for 70% of the world's coffee production. The genus's other species are diploid (2n=2x=22) and genetically incompatible with one another, but *C. arabica* is the sole tertraploid (2n=4x=44) species that is self-fertile [2]. *C. arabica* is created when *C. canephora* and *C. eugenioides* hybridize [3]. With an area of about 11 million ha, coffee is primarily grown in tropical and subtropical climates and is a significant cash crop in over 60 nations in South and Central America, Asia, and Africa [4].

Coffea is not indigenous plant to Nepal; its seed were first brought and planted in Nepal in 1938 A.D. Its popularity has made it one of the developing and valuable cash-generating commodities for Nepal's hill farmers. Of the species, *C. arabica* is mostly cultivated by farmers in Nepal [5] Because Nepali coffee is grown at higher elevations (800-1600m), distant from the typical coffee-growing zone in other parts of the world, it is known for having a particular scent and flavor. As a consequence, coffee cultivated in Nepal is marketed as "Specialty Coffee" in a number of international markets and is also praised for having an organic certification and following fair trading principles. A coffee plantation occupies more than 1000 hectares of land and produces 250 mt of dry cherries overall [6]. The initial coffee variety that was planted and the diversity of *Coffea* species over the past seven decades are not officially documented.

Many PCR Based molecular marker techniques have already been used extensively for genetic diversity analyses, genotype identification in gene bank management, molecular phylogenetic studies, and the development of species diagnostic protocols. These techniques include those based on Random Amplified Polymorphic DNA (RAPD), microsatellites or Simple Sequence Repeats (SSRs), Inter-Simple Sequence Repeats (ISSRs), Amplified Fragment Length Polymorphism (AFLPs), and Inter Transcribed Spacers (ITS) sequences [7]. Coffea species easily cross pollinate and create comparatively fertile hybrids [8,9]. Morphological characteristics have been used to create infrageneric categories [10,11]. However, grouping criteria are currently seen as having little utility since they have grown to be quite complicated and rather muddled [10]. Therefore, further research is needed to understand the evolutionary connections between these species [8].

Quality Nucleic acid isolation is essential for further down streaming application (PCR amplification, Restriction Digestion, Cloning, Genotyping, and Sequencing). Coffee, like many other kinds of tree plants, includes significant levels of polysaccharides, polyphenols, and secondary metabolites such alkaloids, flavonoids, and phenols that often obstruct DNA isolation [12]. When isolating DNA, polysaccharides, flavonoids, and phenols securely attach to nucleic acids and obstruct DNA replication. With ensuing responses[13,14]. The main issue that was found during coffee DNA separation was the co-isolation of very sticky polysaccharides. For reducing the procedures and expense of DNA extraction, several techniques have been developed [12,15]. One isolation approach could not provide for the best DNA yields in various taxa, there-

*Corresponding author. Email: gyanupandey9@gmail.com



(a) (b)

Figure 1: a) Coffee leaf (Sanramon) with fungicide treatment; b) Coffee leaf without fungicide treatment.

fore an effective DNA isolation and PCR optimization are essential. To get DNA of the appropriate quality, researchers frequently alter a technique or combine two or more distinct processes [16].

Here, we report a simple Genomic DNA extraction protocol [17] with some modification for high quality DNA yield and optimization of PCR conditions is carried out by the means of ITS, MatK and Rbcl sequences for the further use in identifying phylogenetic relationship of *Coffea* species.

2. Materials and methods

2.1. Plant materials

For DNA extraction, 23 samples (Table 1) of coffee were collected from Coffee Research Center, Gulmi, Nepal (Latitude: 27°56'17.85" N to 27°5'44.87" N; Longitude: 83° 25' 29.2"1 E to 83°25'30.20" E) ; Coffee Development Center, Aapchaur-05; Horticulture Research Centre, Malepatan which were maintained ex-situ in Nepal Agricultural Research Council, National Biotechnology Research Centre, Khumaltar, Nepal. To protect plants from fungal diseases like coffee leaf rust 1% of Bordeaux mixture treatment on leaf surface is allowed in organic coffee farming. The difference on treatment of 1% Bordeaux mixture on the leaf is seen in field in Fig. 1.

2.2. Solutions

An extraction buffer consisting of 2% CTAB (w/v), Tris HCL pH 8.0 (1M); EDTA pH 8.0 (0.5 M); NaCl (5 M); 2 %PVP (w/v); 1 % β-mercaptoethanol (added just before Pre warming) was prepared. Other additional solutions needed were Tween 20, Chloroform: Isoamylalcohol (24:1 v/v), Ethanol (70% and 100%), and TE buffer (Tris-HCl 10 mM, EDTA 1 mM ,pH 8.0).

2.3. DNA extraction

DNA extraction technique (Doyle and Doyle, 1987) was used with some modifications to optimize the extraction process.

2.3.1. DNA extraction protocol

Freshly harvested young leaf sample (1g) was taken and surfaced cleaned with Tween 20 followed by periodic cleaning with Distilled water. Then the sample was ground in liquid Nitrogen using mortar and pestle. The pulverized leaves were immediately transferred to tube and a freshly prepared pre-warmed (65°C) extraction buffer (700 µl) was kept. The tubes were incubated at 65°C in water bath for 1 hr with inverting the tubes every 10 mins. Then an equal volume of Chloroform: Isoamylalcohol (24:1) was added and mixed properly by inversion for couple of minutes and centrifuged at 15,000 rpm for 15 min at Room temperature. After the phase separation the supernatant was carefully pipetted out and transferred into a new tube. To the supernatant, repeated process of keeping Chloroform: Isoamylalcohol (24:1) of 700µl was done so that complete removal of pigments and debris can be obtained. The tube was centrifuged at 15000 rpm for 7 min. Then the supernatant was transferred to a new tube and precipitated by adding equal volume cold absolute alcohol/isopropanol and gently mixed by inverting and was stored at -20°C overnight. The samples were centrifuged at 10,000 rpm for 10 min. The pellets were washed with 70% ethanol twice. The pellet was air dried and re-suspended in TE buffer (1X).

2.3.2. DNA quantification and purity

DNA quantification and quality was assessed by using Nano drop (Quawell Q-5000). The purity of the DNA was assessed by the absorbance ($A_{260/280}$) ratio and running the DNA samples on 0.8% agarose comparing it with 1kb DNA ladder.

2.3.3. Gel electrophoresis

The quality of extracted DNA was also assessed by using 1.5% agarose gel electrophoresis (Cleaver scientific, UK) in 1XTAE (50X TAE; 242gm Tris-base, 57.1 ml acetic acid (or 100% glacial acid) and 100 ml of 0.5 M EDTA (pH-8.0) at 70V for 1 hr. PCR amplification products were analyzed by using 2% agarose gel at 70V for 2.5-3 hr. using the same buffer system. The gel was stained with ethidiumbromide and photographed using Gel Documentation system (VWR®Genosmart 2, UK).

2.4. Primer sequence

Primer set for barcoding analysis is shown in Table 2.

2.5. PCR reaction

The reactions were carried out in Mygene L series thermo cycler (LongGene Scientific Instrument Co. Ltd). Reactions without DNA were used as Negative Controls. The conditions were optimized by varying parameters as directed in Table 3. The reaction contains about 50-60 ng of template DNA, 2x master mix (Promega Corporation, USA), 0.5 µm of single Primer (Macrogene Inc. , South Korea), with additional 25 mM $MgCl_2$ (Himedia laboratories Pvt. Ltd, India), 0.2mM dNTP Mix (Promega Corporation, USA),0.2U of *Taq* polymerase (Promega Corporation, USA), 1 mg/ml BSA. The thermo cycler was programmed for an initial denaturation step of 4 min at 94°C, followed by 35 cycles of 50s annealing at 54°C (for ITS, MatK and rbcl marker), extension was carried out at 72°C for 80 sec and final extension at 72°C for 7 min and hold temperature of 4°C at the end. PCR products were electrophoresed on 2% (w/v) agarose gel stained with ethidium bromide, at 70V for 3 hrs. Gels were visualized and photographed by using Gel documentation system (Genosmart v2, VWR, UK). 1kb and 100 bp ladder (Thermo scientific), was used as molecular marker for the size comparison of the visible fragments.

3. Results and discussion

3.1. DNA extraction and quality

The time of grounding of the samples in liquid Nitrogen to placement of vials into water bath plays a vital role in DNA quality as high polyphenolic contents are present in *Coffea* species which oxidizes quickly giving brownish coloration which masks the reading of $A_{260/280}$. Leaves samples should be young and freshly picked; leaves stored in -80°C or liquid nitrogen several weeks prior can

Table 1: Coffee sample varieties collected from different sources (collected from Coffee Research Program (CRP), Nepal).

Code	Variety	Source (Location)	Year of Introduction (B.S.)	Young leaf Color	Fruit Color
C1	Gulmi local	CDC, Gulmi	2071	Green	Red (G to R)
C2	Chhetradeep	HRS, Malepatan, Pokhara	2071	Green	Red (G to R)
C3	Selection-10	HRS, Malepatan, Pokhara	2071	Green	Maroon red (G to R)
C4	Selection-12 (Cavery)	Thanapathi VDC, Thorga, Gulmi	2073	Green	Red (G to Y to R)
C5	Yellow Catura	HRS, Malepatan, Pokhara	2071	Green	Yellow (G to Y)
C6	Catimor (Red)	HRS, Malepatan, Pokhara	2071	Green	Red (G to R)
C7	Pacamara	HRS, Malepatan, Pokhara	2071	Green	Red (G to R)
C8	Tekisic	-	-		-
C9	Pacas	HRS, Malepatan, Pokhara	2071	Green	Red (G to R)
C10	Robusta Coffee	CDC, Gulmi	2073	Green	-
C11	Catuai Amarillo (Brazillian)	HRS, Malepatan, Pokhara	2071	Green	Yellow (G to Y)
C12	Sanramon	Indian	2057	Light Green	Dark Red (G to R)
C13	Catisic	-	-	-	-
C14	Bourbon Amarillo	HRS, Malepatan, Pokhara	2071	Green	Red (G to Y to R)
C15	Mundo Novo (Brazillian)	HRS, Malepatan, Pokhara	2071	Green	Red (G to Y to R)
C16	Caturra Amarillo		2059	Light Green	Yellow (G to Y)
C17	Syangja Special	Highland coffee nursery, Syangja	2071	Green	Red (G to R)
C18	Catuai Vermello	-	-	-	-
C19	Argakhachi local	-	2072	Green	Red (G to R)
C20	Kaski local	HRS, Malepatan, Pokhara	2071	Green	-
C21	Indonesia	DCPA, Kaski	2071	Green	Maroon red (G to R)
C22	Bourbon Vermelo	HRS, Malepatan, Pokhara	2071	Light Green	Red (G to R)
C23	Hawaii Kona	DCPA, Kaski	2071	-	-

G: Green; Y: yellow; and R: Red.

Information source: list of variety, sample type, source (location) year of introduction were provided by CRP, Gulmi. Young leaf colour and fruit colour were collected by visual observation and coding by research team during sample collection at research blocks and field gene bank of CRP, Gulmi [18].

Table 2: Primer set for barcoding analysis.

Barcode	Primer	Primer sequences(5'-3')	References
ITS (ITS1-5.8s-ITS7)	ITSL ITSR	TCGTAACAAGGTTTCCGTAGGTG TATGCTTAAAYTCAGCGGG-3'	[19]
matK	3F_Kim f 1R_Kim r	CGTACAGTACTTTTGTGTTTACGAG ACCCAGTCCATCTGGAAATCTTGGTTC	[20]
rbcl	rbclF rbclR	ATGTCACCACAAACAGAGACTAAAGC GAAACGGTCTCTCCAACGCAT	[20]

ITS: internal transcribed Spacers

matK: maturase K

rbcl: ribulose-1,5-bisphosphate carboxylase large subunit.

Table 3: Optimization of PCR reaction parameters for *Coffea* species.

PCR Parameter	Tested Range	Optimum Condition	Observations
DNA Concentration(ng)	10,20,30,40,50,60,80,100,150 and 200	50-60ng	High Amplification observed from within range. No amplification on lower concentration and presence of smear in higher concentration.
$MgCl_2$ Concentration (mM)	1 , 1.5 , 2 ,2.5 , 3 , 3.5, 4	3 mM	High /low concentration increases the non-specificity and no yield of product.
dNTPs Concentration (mM)	0.12 , 0.13 , 0.14 , 0.15 , 0.16, 0.2	0.16 – 0.2mM	Increase in concentration reduces the intensity and amplification.
Primer Concentration (μM)	0.1 , 0.5, 1 , 1.5 , 2 , 2.5 , 3, 3.5	0.5 μM	Intensity of amplified bands were same from 1-3 μM concentration. No amplification above 3.5 μM and faint amplification below 0.5 μM
Taq Polymerase (units)	0.1 , 0.2 , 0.3 , 0.4 , 0.5	0.2 U	Lower concentration showed faint bands and high concentration decreased the specificity.
Initial Denaturation time(min) at 94 °C	2 ,3 , 4 and 5	94 °C ,4 min	Unreproducible amplification in higher/lower time intervals, loss of Taq activity leading to poor amplification.
Annealing Temperature (°C) Time Interval (s)	45, 48 ,50, 52, 53, 55 and 62 40,45,50,55,60 and 65	ITS 54 °C for 50sec	Low/no amplification in higher and lower temperature and time interval than optimum range.
		matK 54 °C for 50 sec	
		RbcL 54 °C for 50 sec	
Reaction volume (μl)	10 , 12,13 ,15	13 μl	Determine cost of the PCR ingredients
Number of Cycles	25, 30 ,35 and 40	35	Low/no amplification in higher and lower cycles from optimum.
DMSO (Dimethyl sulfoxide)	3 , 5 , 7, 10 %	7%	High intensified bands in optimum conditions. Faint bands/ no bands on high/ low concentration than optimum.

also be used. Consequently, high quality DNA was obtained from presented protocol with low-polysaccharide genomic DNA.

Many plant species include significant oxidizing agents called polyphenols, which can impair yield and purity by binding covalently to extracted DNA, rendering it worthless for the majority of research uses [21-23]. Polyphenols bind to DNA and interfere with downstream processing by co-precipitating with the nucleic acid. Poor DNA quality was produced as a result of the traditional extraction approach [24][17]. High purity is indicated by a ratio of absorbance (A260/280) in the range of 1.8 to 2.0 [25][26]. By oxidizing phenolic chemicals, the traditional extraction method (Doyle and Doyle) produced unsatisfactory absorbance ratio (A260/280) values below 1.5 and above 2.2, which caused the DNA pellets to become brown. The Doyle and Doyle technique was changed to create the protocol. In these modifications, CTAB and its components are used at greater concentrations than in the original Doyle and Doyle procedure (2% CTAB (w/v), Tris HCL pH 8.0 (1 M), EDTA pH 8.0 (0.5 M), NaCl (5 M), and the use of PVP and β-mercaptoethanol. In biological material extractions, the use of CTAB, a potent ionic detergent, has made it easier to separate proteins from nucleic acids [27]. A high concentration of NaCI facilitates the removal of polysaccharides, and the high ionic strength of CTAB forms complexes with protein and the bulk of the acidic polysaccharides [28]. PVP and β-mercaptoethanol, on the other hand, stop the oxidation of secondary metabolites, such phenols, during cell lysis. When phenolic compounds are centrifuged in the presence of chloroform, PVP and phenolic compounds create hydrogen bonds that cause the phenolic compounds to collect in the interphase, which is between the organic and aqueous phases [28]. β-Mercaptoethanol prevents the polymerization of tannins, which hinder the isolation process in a manner similar to polysaccharides [14].

With all the modification and optimization in the modified protocol, increased the DNA yield and quality, in particular in the $A_{260/230}$ ratio ranging from 1.8-2.0 indicating high purity (data not shown). The excess of salts in the pellets gets removed by washing in 70% ethanol leading to quality DNA. The DNA yield obtained ranged from 500-2000 ng/μl. The presence of DNA was assessed by running the genomic DNA (Only first 17 samples were tested, extracted on same batch of 23 samples) (Table 1) in 0.8% agarose (Fig. 2) and DNA quantification was assessed by using Nanodrop (Quawell Q-5000).

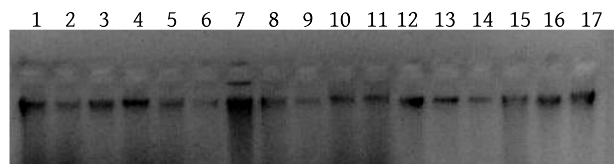


Figure 2: Lanes 1-17 represents the agarose gel (0.8%) of Genomic DNA isolated from first 17 species of coffee viz 1. Gulmi local, 2. Chhetradeep, 3. Selection-10, 4. Selection-12, 5. Yellow Cataura, 6. Catimor, 7. Pacamara, 8. Tekisic , 9. Pacas, 10. Robusta, 11. Catuai Amarillo, 12. Sanaramon, 13. Catisic , 14. Bourbon Amarillo, 15. Mundo Novo, 16. Caturra Amarillo, 17. Syangja Special. (Only 17 species out of 23 were ran on gel for validity of the extraction process done on the same batch).

3.2. PCR Reaction

Further downstreaming process, includes PCR reaction prior to sequencing and other molecular analysis. With all the components in PCR reaction, $MgCl_2$ concentration and T*aq* DNA polymerase plays the most vital role on getting good PCR product. Magnesium acts as cofactor to the thermostable DNA polymerases during

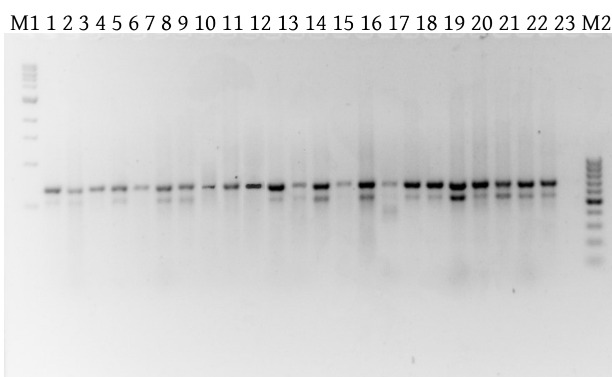


Figure 3: M1 and M2 represents DNA ladder of 1KB and 100 bp respectively. PCR Amplification products were fractioned in 2% agarose gel. Lane 1-23 are the 23 coffee samples from Table 1 where two amplified bands were seen from genomic DNA using ITS region (ITS1-5.8s-ITS7) ITSL (5'-TCGTAACAAGGTTTCCGTAGGTG-3') and ITSR (5'-TATGCTTAAAYTCAGCGGG-3';[19].

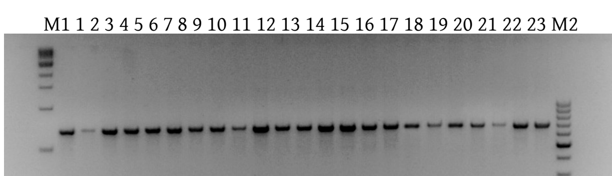


Figure 4: M1 and M2 represents DNA ladder of 1KB and 100 bp respectively. PCR Amplification products were fractioned in 2% agarose gel. Lane 1-23 are the 23 coffee samples from Table 1. Amplified from genomic DNA using ITS region (ITS1-5.8s-ITS7) was amplified at 700bp with primers ITSL (5'-TCGTAACAAGGTTTCCGTAGGTG-3') and ITSR (5'-TATGCTTAAAYTCAGCGGG-3' [19].

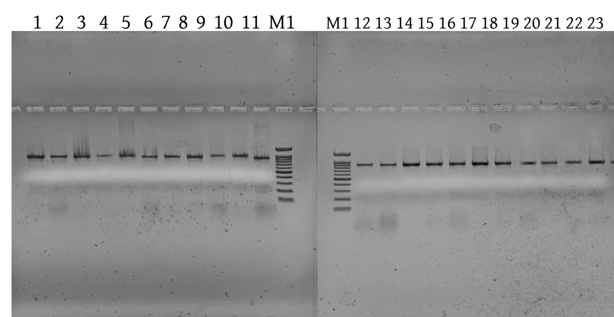


Figure 5: PCR Amplification products were fractioned in 2% agarose gel. Lane 1-23 are the 23 coffee samples from Table 1. Gel amplified by Oligonucleotide primer from matK region (Macrogen .Inc. , South Korea)matK F-5'-CGTACAGTACTTTTGTGTTTACGAG3'; matK R-5'-ACCCAGTCCATCTGGAAATCTTGGTTC-3' [19].

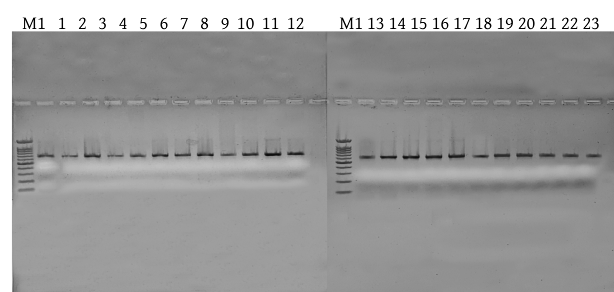


Figure 6: PCR Amplification products were fractioned in 2% agarose gel. Lane 1-23 are the 23 coffee samples from Table 1. Oligonucleotide primer from rbcl region (Macrogen .Inc. , South Korea) were used for optimization of rbcl reactions. RbclF- 5'-ATGTCACCACAAACAGAGACTAAAGC-3'; RbclR- 5'-GAAACGGTCTCTCCAACGCAT-3' [20].

the reaction process and changing the magnesium concentration is one of the easiest reagents to manipulate with perhaps the greatest impact on the stringency of PCR [29]. In general, the PCR product yield will increase with the addition of greater concentrations of Mg^{2+}. However, increased concentrations of Mg^{2+} will also decrease the specificity and fidelity of the DNA polymerase [29]. Also, BSA (bovine serum albumin) and DMSO (dimethyl sulfoxide) acts as a PCR enhancer for high amplification of the PCR product. DMSO is thought to reduce secondary structure and is particularly useful for GC rich templates. A number of investigators have accordingly proposed that organic additives like DMSO enhance PCR by hydrogen bonding to the major and minor grooves of template DNA and destabilizing the double-helix [30]; and references therein). DMSO at 2-10% may be necessary for amplification of some templates; however 10% DMSO can reduce Taq polymerase activity by up to 50% so it should not be used routinely. BSA has proven particularly useful when attempting to amplify ancient DNA or templates which contain PCR inhibitors such as melanin. BSA has also been reported to reduce the adhesion of the DNA polymerase to glass capillaries used in certain real-time PCR instruments. This is important for reducing the loss of reagents through adsorption to tube walls. BSA concentration of 0.01μg/μl to 0.1μg/ μl can be used [31].

3.3. Effect of leaf sample

The result showed two amplified bands: one near 700 bp and other around 600 b.p (Fig. 3). The full ITS region in fungi has an average length of 500 and 600 base pairs (b.p) for ascomycetes and basidiomycetes, respectively, and an average length of 600 bp across all fungal lineages [30]. About 80% of vascular plants are thought to host fungus, according to some estimation cited by the same authors, and co-evolution between plants and fungi has also been proposed [32]. Many types of "endophytic" fungus thrive as symptomless parasite fungi inside plant tissues without obviously harming the host plant [33]. The primers ITSL and ITSR used in this study were designed specifically for plants but unsatisfactory results were seen (Fig. 3) with non-specific amplification. The universal primers for the ITS region [34] were primarily developed as from comparisons across sequences of fungus. Because they amplify the whole ITS region (ITS1-5.8s-ITS7), it is most likely that they can amplify even fungal sequences if any fungal DNA is present. It follows that their great sensitivity to these species, which may potentially be used in clinical analyses to detect fungus, is comprehensible [35]. Since the untreated fungal leaf samples (Fig. 1) were utilized for DNA isolation, the non-specific bands shown in our results were thought to represent an amplified fungal sequence. Sterilization of the tissue prior to DNA extraction is one method to prevent fungus proliferation, as shown by various researches [36] . It can be easy yet crucial to properly clean the host tissue's surface (with ethanol, detergents, and reducing agents) in order to get rid of the phylloplane fungus and other organisms that are present there [36]. With that in mind, treatment of the coffee plants with fungicide was done to remove potential fungal infection (Fig. 1) and during the extraction process, surface cleaning was done by detergent Tween 20. With these implications, the result was improved with amplified bands only at 700b.p which is beyond the fungal range [30] (Fig. 4).

With improvised extraction protocol focused on fungal contamination removal from leaf surface optimization was done for ITS-PCR reaction by varying different parameters like: template DNA, primer concentration, dNTPs, Magnesium chloride, T*aq* Polymerase, temperature and time intervals during denaturation, annealing and extension were optimized for high amplification and reproducible data. The optimized conditions for ITS protocol are given in Table 3 with band fragment size of ~700 bp (Fig. 4). Higher or lower concentration of tested constituents beyond the opti-

mized concentration resulting into lack of reproducibility. The modified DNA extraction protocol and the optimization of PCR conditions help in generating quality and reproducible data for further analysis. With less extraction process and reagents used in the modified protocol can be considered a simple and fast, giving good quality DNA.

With the optimization of the ITS region, the parameter Optimization of Matk and rbcl gene was done with respect to ITS protocol. Good amplifications of the bands were seen for MatK-Rbcl PCR at 950 b.p and 680 b.p respectively (Fig. 5 and 6). The range for Matk-Rbcl amplification was 900-1500 b.p and 650-700 b.p respectively, so the amplification was at desired range. In plant biology, DNA barcoding has been frequently utilized to clarify the phylogenies of related species. Even closely related species might be distinguished from one another using nucleotide variation in these standard locations. However, in other instances, it was observed that the standard areas' nucleotide variability was insufficient to discriminate between closely related species [37]. MatK exhibits fast development and a strong capacity for interspecific recognition, although the primer is not all-encompassing. RbcL, on the other hand, offers great generality, ease of amplification, and comparability, but its species-level selective efficiency is ineffective. Because of this, other chloroplast areas such trnH-psbA, trnL, and trnL-F as well as the nuclear ribosomal Internal Transcribed Spacer (ITS) region are frequently utilized as supplemental barcodes in addition to matK and rbcL [38]. As compared to morphological and biochemical markers, molecular markers are more effective, accurate, and trustworthy at differentiating between closely related species and cultivars. Additionally, it has been shown that certain areas of the chloroplast genome may be employed as DNA barcodes in a variety of plant species [39]. According to these claims, effective DNA extraction and PCR optimization are essential for amplifying DNA bands, and because a single barcode cannot distinguish, the usage of many additional barcodes can further validate the distinction.

In conclusion, reliable DNA isolation protocol from coffee leaves and PCR optimization for 3 barcoding genes protocol can be scaled up or down depending on the samples and requirements for particular analysis. In consideration to the DNA isolation and optimization of the PCR, can be further used for barcoding and other down streaming applications.

Conflict of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

The authors thankfully acknowledge National Biotechnology Research Centre, NARC for funding the research. They are also grateful to Shubham Biotech Nepal Pvt. Ltd. for their guidance and support during the research work.

References

[1] Davis A P, Psilanthus mannii, the type species of Psilanthus, transferred to Coffea, *Nord. J. Bot.*, 29 4 (2011) 471.

[2] Clarindo W R & Carvalho C R, First Coffea arabica karyogram showing that this species is a true allotetraploid, *Plant Syst. Evol.*, (2008).

[3] Consortium T A C G, Towards a Better Understanding of the Coffea Arabica Genome Structure, *Int. Conf. Coffee Sci.*, November (2014) 42.

[4] Geleta M, Herrera I, Monzn A & Bryngelsson T, Genetic diversity of arabica coffee (Coffea arabica L.) in Nicaragua as estimated by simple sequence repeat markers, *Sci. World J.*, 2012 (2012).

[5] Gotame T P, Research needs on coffee enterprise development in Nepal, March (2016).

[6] NTCDB, Analysis of Habitat Suitability of Coffee (Coffea arabica) in Nepal, *Natl. Tea Coffee Dev. Board*, (2018) 1.

[7] Milbourne D, Meyer R, Bradshaw J E, Baird E, Bonar N, Provan J, Powell W & Waugh R, Comparison of PCR-based marker systems for the analysis of genetic relationships in cultivated potato, *Mol. Breed.*, 3 2 (1997) 127.

[8] Charrier A & Berthaud J, *Botanical Classification of Coffee*, In: Coffee, (1985).

[9] Ky C L, Barre P, Lorieux M, Trouslot P, Akaffou S, Louarn J, Charrier A, Hamon S & Noirot M, Interspecific genetic linkage map, segregation distortion and genetic conversion in coffee (Coffea sp.), *Theor. Appl. Genet.*, 101 4 (2000) 669.

[10] Lashermes P, Combes M C, Trouslot P & Charrier A, Phylogenetic relationships of coffee-tree species (Coffea L.) as inferred from ITS sequences of nuclear ribosomal DNA, *Theor. Appl. Genet.*, 94 6–7 (1997) 947.

[11] Anthony F, Bertrand B, Quiros O, Wilches A, Lashermes P, Berthaud J & Charrier A, Genetic diversity of wild coffee (Coffea arabica L.) using molecular markers, *Euphytica*, 118 1 (2001) 53.

[12] Mishra M K, Rani N S, Ram A S, Sreenath H L & Jayarama, A simple method of DNA extraction from coffee seeds suitable for PCR analysis, *African J. Biotechnol.*, 7 4 (2008) 409.

[13] Demeke T & Jenkins G R, Influence of DNA extraction methods, PCR inhibitors and quantification methods on real-time PCR assay of biotechnology-derived traits, *Anal. Bioanal. Chem.*, 396 6 (2010) 1977.

[14] Sultana S, Islam M N & Hoque M E, DNA fingerprinting and molecular diversity analysis for the improvement of brinjal (Solanum melongena l.) cultivars, *J. Adv. Biotechnol. Exp. Ther.*, 1 1 (2018) 01.

[15] Shrestha S, Sijapati J, Rana N, Malla D, Regmi P & Raskoti B, Optimization of RAPD-PCR conditions for the study of genetic diversity in Nepal's *Swertia chirayita* (Roxb. Ex Fleming) H. Karst, *Himal. J. Sci.*, 6 8 (2011) 35.

[16] Thomas H, Turner L, Bat T & Jones N, Markers and mapping revisited : finding your gene Related papers Tansley review Markers and mapping revisited : finding.

[17] Doyle J, Genomic Plant DNA Preparation from Fresh Tissue-CTAB Method, *Phytochemical Bulletin*, Vol. 19 (1987) p. 11.

[18] Annual report 2077/2078 (2020/2021), Coffee Research Program, Baletaksar-1, Bhandaridanda Gulmi, (2021).

[19] Hsaio C, Chatterton N J, Asay K H & Jensen K B, Phylogenetic relationships of 10 grass species: An assessment of phylogenetic utility of the internal transcribed spacer region in nuclear ribosomal DNA in monocots, *Genome*, 37 1 (1994) 112.

[20] Jeanson M L, Labat J N & Little D P, DNA barcoding: A new tool for palm taxonomists?, *Ann. Bot.*, 108 8 (2011) 1445.

[21] Katterman F R H & Shattuck V I, An Effective Method of DNA Isolation from the Mature Leaves of Gossypium Species That Contain Large Amounts of Phenolic Terpenoids and Tannins, *Prep. Biochem.*, 13 4 (1983) 347.

[22] Peterson D G, Boehm K S & Stack S M, Isolation of milligram quantities of nuclear DNA from tomato (Lycopersicon esculentum), A plant containing high levels of polyphenolic compounds, *Plant Mol. Biol. Report.*, 15 2 (1997) 148.

[23] Porebski S, Bailey L G & Baum B R, Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components, *Plant Mol. Biol. Report.*, 15 1 (1997) 8.

[24] Pašakinskienė I & Paplauskienė V, Floral meristems as a source of enhanced yield and quality of DNA in grasses, *Plant Cell Rep.*, 18 6 (1999) 490.

[25] Benti T, Gebre E, Tesfaye K, Berecha G, Lashermes P, Kyallo M & Kouadio Yao N, Genetic diversity among commercial arabica coffee (Coffea arabica L.) varieties in Ethiopia using simple sequence repeat markers, *J. Crop Improv.*, 35 2 (2021) 147.

[26] Hollingsworth M, DNA fingerprinting in plants: principles, methods, and applications, 2nd edn. Weising K, Nybom H, Wolff K, Kahl G. 2005. Boca Raton: CRC Press. $99·95 (paperback). 472 pp., *Ann. Bot.*, 97 3 (2006) 476.

[27] Abu-Romman S, Comparison of methods for isolating high quality dna from sage (salvia officinalis), *J. Med. Plants Res.*, 5 6 (2011) 938.

[28] Aljanabi S, Forget L & Dookun a, An improved and rapid protocol for the isolation of polysaccharide-and polyphenol-free sugarcane DNA, *Plant Mol. Biol. Report.*, 17 (1999) 1.

[29] Lorenz T C, Polymerase Chain Reaction: Basic Protocol Plus Troubleshooting and Optimization Strategies, *J. Vis. Exp.*, 63 (2012).

[30] Chakrabarti R & Schutt C E, The enhancement of PCR amplification by low molecular-weight sulfones, *Gene*, 274 1–2 (2001) 293.

[31] Gene Link & Link G, PCR Additives & Enhancers, *Gene Link Tech. Note*, (2014) 12.

[32] Porter T M & Brian Golding G, Are similarity- or phylogeny-based methods more appropriate for classifying internal transcribed spacer (ITS) metagenomic amplicons?, *New Phytol.*, 192 3 (2011) 775.

[33] C. J. Alexopoulos C W M and M B, *Introductory Mycology*, 4th Editio., John Wiley & Sons Inc. New York, 1996.

[34] Miranda V F O de, Martins V G, Furlan A & Bacci Jr. M, Plant or fungal sequences? An alternative optimized PCR protocol to avoid ITS (nrDNA) misamplification, *Brazilian Arch. Biol. Technol.*, 53 1 (2010) 141.

[35] White T J, Bruns T, Lee S & Taylor J, *Amplification and direct sequencing of fungal ribosomal rna genes for phylogenetics*, In: PCR Protocols, Elsevier (1990) 315–322.

[36] Ferrer C, Colom F, Frasés S, Mulet E, Abad J L & Alió J L, Detection and Identification of Fungal Pathogens by PCR and by ITS2 and 5.8S Ribosomal DNA Typing in Ocular Infections, *J. Clin. Microbiol.*, 39 8 (2001) 2873.

[37] Li Y, Feng Y, Wang X-Y, Liu B & Lv G-H, Failure of DNA barcoding in discriminating Calligonum species, *Nord. J. Bot.*, 32 4 (2014) 511.

[38] Wang J, Yan Z, Zhong P, Shen Z, Yang G & Ma L, Screening of universal DNA barcodes for identifying grass species of Gramineae, *Front. Plant Sci.*, 13 (2022).

[39] Mishra M K, Jingade P & Huded A K C, DNA barcoding analysis and phylogenetic relationships of Indian wild coffee species, *Turk. J. Botany*, 46 2 (2022) 109.