International Journal of Pharmaceutical Science Invention
ISSN (Online): 2319-6718, ISSN (Print): 2319- 670X
www.ijpsi.org Volume 9 Issue 6 ǁ June 2020 ǁ PP. 06-16

DNA Barcoding of Medicinal plant: a Systemic Review

R. Rajphriyadharshini[1], O.V.D.S.J Weerasena[1]
1. Institute of Biochemistry, Molecular Biology and Biotechnology, University of Colombo, Sri Lanka.

ABSTRACT
DNA barcoding technique generates DNA sequencing data from genetic regions in an organism and is used to identify specific species even by the non-specialist. DNA barcodes can be used to monitor the adulteration of herbs Sida is a most prominently used medicinal plant. However, it is being substituted by other plants of the same genus, in trade due to difficulty to discriminate based only on morphology. Such substituting by other plants that are not used for the medical purpose might induce harmful effects. Also, the beneficial effect of Sida may be lost entirely. Additionally, taxonomists find it difficult to identify the authentic Sida plant from processed products. Thus this obstacle can now be overcome by using DNA barcoding. Due to its economical and biological significance developing DNA barcodes may contribute to the trade control, where the endangered species like Sida are protected from export outside of the country and are conserved. This review analyses the existing published articles related to the pros and cons of DNA barcoding in the field of Medicinal plants, specifically related to the Sida plant.

KEYWORDS: *DNA, Sida, Barcoding, Adulteration, Sequencing*

Date of Submission: 08-08-2020 Date of acceptance: 21-08-2020

I. INTRODUCTION

Sida is a genus of flowering plants of Malvaceae that includes about 200 species distributed in tropical and subtropical regions of the world. *Sida* showed a wide range of morphological variability. Nearly 18 species of *Sida* (*Sidae, Sida Lilm., Nelvagae, Cordifoliae, Spinosae, Malachroideae*, and *Stenindae*, etc) were known to occur in Sri Lanka (Aguilar et al. 2003). Even though so far 200 species were recorded under the *Sida* genus, over 1000 species were predicted under the genus, due to uncertain morphological variability of *Sida* plants. In contrast, some studies including Santhosh et al., (2015), Simplice et al., (2007), and Vassou et al., (2015) accepted that *Sida* only includes 98 species in it. Thus the diversity of *Sida* is still unclear which gives rise to a high rate of adulteration of *Sida* plants.

The vernacular name of *Sida* species in Sinhala is '*Babila*'. It is a valuable medicinal plant and has enormous applications in the Ayurvedic system of medicine. Also, it possesses various biological effects including strengthening the central nervous system and treat neurological disorders such as hemiplegia, sciatica, facial paralysis, emaciation, and cervical spondylosis (Mishra *et al*., 2015). *Sida* that is used in traditional medicine is being substituted by other plants of the same genus, in trade due to difficulty to discriminate based only on morphology. Such substituting by another plant that is not used for the medical purpose might induce harmful effects. Also, the beneficial effect of *Sida* may be lost entirely. Additionally, taxonomists find it difficult to identify the authentic *Sida* plant from processed products.

Thus this obstacle can now be overcome by using DNA barcoding. Due to its economical and biological significance developing DNA barcodes may contribute to the trade control, where the endangered species like *Sida* are protected from export outside of the country and are conserved.

DNA barcoding technique generates DNA sequencing data from genetic regions in an organism and is used to identify specific species even by a non-specialist. To employ such an identification method, a suitable DNA barcode should be selected. DNA barcodes are defined as the short segment of the gene sequence that evolves radically to differentiate species. DNA barcode effectively assesses the species admixtures in the raw drug trade (Nithaniyal *et al*., 2017). This shows that DNA barcodes can be used to monitor the adulteration of herbs. Therefore the present study will be carried out to develop DNA barcodes specific for *Sida* species by using several potential exiting barcoding candidates.

There is no single universal DNA barcode marker for plants. The Consortium for the Barcoding of Life (CBOL) has recommended chloroplast genes ribosomal RNA maturase K (matK) coding gene and ribulose 1, 5-bisphosphate carboxylase/oxygenase large subunit coding region (rbcL) as the core markers for DNA barcoding of plants. However, those regions are not 100% efficient for the discrimination of some plant species. Additionally, the non-coding chloroplast psbA-trnH marker was also recommended as a secondary marker by CBOL as it was found to discriminate species in the Malvaceae plant group. Furthermore, the use of an internal transcribed spacer (ITS2) region marker was recommended for large scale DNA barcoding (Vassou et al.,

2015). Hence this review analyses the existing published articles related to the pros and cons of DNA barcoding in the field of Medicinal plants, specifically related to the *Sida* plant.

1. *Sida* Species: an overview

1.1 Genus *Sida*

Domain: Eukaryota

Kingdom: Plantae

Phylum: Spermatophyta

Subphylum: Angiospermae

Class: Dicotyledonae

Order: Malvales

Family: Malvaceae

Genus: *Sida*

Species: *Sida acuta*

Figure 1: Classification of Sida species (Mishra et al., 2015)

Sida is a weed of tropical, semi-arid, and rarely also sub-tropical and warmer temperate regions. It grows in different habits such as open woodlands, waterways pastures, plantations, gardens, crops, roadsides, and waste areas. Commonly found in South Asia, mainly in India and Sri Lanka. Also, it is widely naturalized throughout eastern Asia (China and Taiwan), the northern parts of Australia, and on several Pacific islands (Santhosh et al., 2015).

The Common vernacular names used for *Sida* species are broom grass, broom weed, cheese weed, common fan petals, clock plant, common wire weed, morning mallow, southern *Sida*, spiny-headed *Sida***.** The genus *Sida* is polyphyletic (Aguilar et al. 2003). Table 1 showed the common synonyms used for *Sida* species recorded in Sri Lanka. Taxonomic Keys are used by biologists for identifying unknown organisms. The taxonomic key to *the Sida* plant is bellowed.

Table 1: *Sida* species recorded in Sri Lanka, their local (Sinhala) names, and synonyms (Gunatilaka et al., 1987).

Sida species	Local name(s)	Synonym(s)
S.acuta BURM.F	*Gas-bevila* *Kesera- bevila*	*S.carpinifolia L* *S.lanceolata WILLD.RETZ* *S.herbaceae CAV*
S.cordifolia.L.	*Wal-bevila* *Sudu-bevila* *Suluboo-bevila*	*S.althaefolia SWARTZ* *S.rotundifolia CAV*
S.racemosa BURM.F.	*Sirivadi-bevila* *Girivadi-bevilaa*	*S.glutinosa CAV* *S.mysorensis W and A*
,S.rhombifolia L.	*Kotikan-bevila*	*S.canariensis WILLD* *S.compressa WALL.* *S.rhombifolia L.*
S.rhomboidea ROXB		*Var.rhomboidea MASTERS* *S.orientalis CAV.*

1.2 Morphology

Sida plant is a long-lived herbaceous plant or small shrub that usually grows 30-100 cm tall and occasionally reaches up to 1.5 m in height. Most often have a short life span (annual). The stems of *Sida* are branched and either upright or spreading in nature which is sparsely covered with star-shaped hairs. The leaves are elongated in shape and yellowish-green in color. As well as typically have pointed tips and margin. These

leaves can be either hairless or covered sparsely with star-shaped (also known as stellate) hairs (Baracho , George Y, and Maria, 2016).

Its flowers are borne singly or in small clusters in the leaf forks on short stalks. These flowers have five sepals and five yellow petals. *Sida* produces fruit that breaks up into 5-8 wedge-shaped segments when mature. These one-seeded segments are topped with two sharp spines. This species reproduces by seed, which readily attaches to animals, clothing, and other materials. Seeds may also be dispersed in mud and contaminated agricultural produce (Simplice et al., 2007).

1.3 Varieties of *Sida* species

The *Sida* spices have common characters (as mentioned in 1.3) which enable taxonomists to differentiate *Sida* plants from others (Santhosh et al., 2015). However, the variation among the interspecies is very less, which makes it difficult to identify individual species using morphological characters. Most commonly the *Sida* acuta which is widely used in raw drug confused with *Sida spinosa*, *Sida rhombifolia*, *Sida subspicata*, and *Sida cordifolia*. Baracho, George Y, and Maria, (2016) have pointed out certain characters that are unique to each species.

Sida acuta has elongated leaves that are hairless or sparsely hairy on both surfaces and have pointed tips (i.e. acute species). Its flowers are borne singly or in small clusters on short stalks (2-8 mm long) and their sepals are mostly hairless. The fruit is topped with two sharp awns. *Sida spinosa* has elongated leaves that are relatively broad and have densely hairy on their undersides with blunt tips. Its flowers are relatively short stalks (3-15 mm long) and have finely haired sepals. The fruit usually breaks up into 5 wedge-shaped seeds that are topped with two sharp awns (about 1 mm long).

Sida rhombifolia has oval or diamond-shaped leaves with a dense covering of hairs on undersides rounded tips. Its flowers are borne singly and their sepals with sparsely hairy. The fruit usually breaks up into 8-12 wedge-shaped seeds. *Sida subspicata* has relatively narrow leaves, which ardently covered with hairs on both sides. These leaves have pointed tips Its flowers are almost with a few small leaves sometimes interspersed between them. and the fruit breaks up into 4-6 wedge-shaped 'seeds' that usually have rounded tips.

1.4 Pharmacological and biological uses

Sida is one of the most widely used medicinal plants in Ayurvedic medicine. The vernacular name for *S. cordifolia* in Ayurvedic formulary is 'Bala', (meaning strength). Its seeds are called as Bijabanda in Ayurveda. The roots of the *Sida* plant are used for strengthening the central nervous system and treating neurological disorders such as hemiplegia, facial paralysis, sciatica, neuralgia, and neurosis (Koman, 1921). It is also a common ingredient in several important Ayurvedic preparations like Ksirabala, Dhanvantaram, Balaristam, and Rasnadi Kasayam which are used as a nervine tonic. *Sida* is also used to treat urinary tract infection, where the powdered root bark is administered with milk as a treatment for urinary urgency and leucorrhoea. Several studies (Cheung et al., 1993; Fazekas et al., 2008; Montelone et al., 1981) found out that the exterior of seeds of *Sida* can reduce body weight. However, this character is not yet tested on humans. Moreover, Obah, Akerele, and Obasuyi (2007) stated that *Sida* can be used to regulate blood pressure and can cure cardiac arrythamitcs. Even the extract of the whole plant is being used to treat spermatorrhoea. Santhosh et al., (2015) reported that bark of the plant can effective in curing facial paralysis and sciatica. Besides *Sida* plant is used as one of the ingredients in oils that are used topically to the sore joints and sore muscles in rheumatism and arthritis. It is diaphoretic in nature and helps in lowering fever. In Mexico, *Sida* leaves are smoked for its simulative effects. For a similar purpose, it is used in tea in several parts of India.

1.5 Composition and properties of *Sida* species

It has been reported that the roots of the *Sida* plant contain alkaloids including ephedrine, saponine, choline, pseudoephedrine, beta phenethylamine, vaccine, hipaphorine, and related indole alkaloids (sitoindoside, acylsteryglycoside). Whereas the whole plant has alkaloids, hydrocarbons, fatty acids, and ephedrine. Kunmvmiata and Rapp, (2001) reported that due to the presence of different chemicals including alkaloids that arrow leaf *Sida* is not liked by cattle.

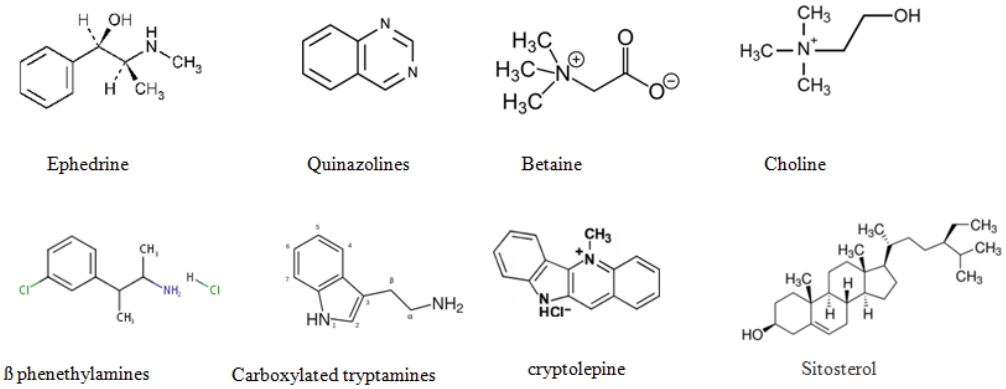


Figure 2: Major chemical constituents in Sida species (Source: Obah et al., 2007)

Prakash et al., (1981) isolated three types of alkaloidal constituents, ß phenethylamines, Gunatilaka et al., (1987) have reported the major alkaloids of some *Sida* species from Sri Lanka, including those of *S. acuta* and *S. rhombifolia*. They found that the major alkaloids of Sida acuta are cryptolepine and ephedrine which have antimicrobial activity against Proteus vulgaris. Besides this, a study carried out by Obah et al., (2007) isolated traces of sitosterol, palmic acid, and stearic acid from this plant. According to Simplice et al., (2007), the main alkaloid present in *Sida* is asparagin. He also reported that *Sida* does not contain any tannin or glycoside.

1.6 Adulteration and Substitution

The annual demand for *Sida* plants in South Asia is estimated to be about 50- 100 million kg/year. The heavy demand for *S. cordifolia* and *S. rhombifolia* is often beyond the available supply, this could predispose the adulteration of *Sida* products in the market (Fazekas et al., 2008).

Adulteration usually refers to mixing other matters of an inferior which may occasionally harmful to the quality of food, drink, or chemical product that are intended to be sold. Adulteration of herbal products includes admixtures, product substitution, and contamination which could cause life-threatening health implications. Thus it is important to develop an accurate method that confirms the authenticity of raw material, which is a fundamental component of quality assurance in the field of the food supply industry. Currently, there were many chemical fingerprints developed to unveil the authenticity of raw material and to overcome species adulteration. In herbal plants, drug adulteration usually refers to mixing completely different plants with genuine plant materials. Substitution refers to replace one genuine plant by other plants of the same genus. Hence DNA barcoding help to establish the correct authentication of the plants that undergo both adulteration and substitution.

Generally, the identity of the medicinal plant is determined by chemical and physical methods such as mass spectrometry, High-performance liquid chromatography, Thin-layer chromatography, Fourier transform infrared (FT-NIR) spectroscopy and NMR spectroscopy.

However, there are countless obstacles and pitfalls with these methods. Most often the metabolic profiles of the same species can be altered due to the influence of storage conditions and external factors (light). Thus multitudinous researchers are interested in DNA based adulteration detection methods due to its reliability. The goal of DNA based authentication is to create an enormous online digital library of DNA barcode that serves as a standard that enables authentication of unidentified samples by matching their DNA sequence to the standard.

Even though several chemical fingerprints were developed to unravel species adulteration. As these chemical markers are prone to vary with age and the geographic population of the plants, it is not recommended as a robust method to overcome issues related to species adulteration. Thus in recent years, many researchers attempt to use DNA barcoding in exhibiting the adulteration of herbal products.

Newmaster et al. (2013) used DNA barcoding techniques to detect contamination as well as product substitution in natural health products. Their study showed that 59% of the products tested contained species that are not listed on the product labels. Equally, Srirama et al. (2010) demonstrated the competence of DNA barcoding in identifying species admixtures in *Phyllanthus L*. (Euphorbiaceae). Their study revealed that up to 24% of the market samples were adulterated with other phenotypically similar *Phyllanthus* species.

In the case of *Sida* plants, the root of this plant is the important source used in raw drug '*Sida*'. Since the morphological characters of *Sida* roots are not well defined, its identity is controversial leading to adulteration. Santhosh et al., (2015) reported in India, the Ayurvedic physicians of Kerala use *S. alnifolia* as

'*Sida*' whereas in northern India *S. cordifolia* is used. Various species of *Sida* such as *S. rhombifolia, S. rhomboidea, S. Scabrida, S. acuta, S. cordifolia, S. fryxellii* are used as '*Sida*'. A few species of allied genera such as Abutilon, Pavonia, Urena are also used as '*Sida*' in the raw drug market either as a substitute or as an adulterant. This includes species of *Abutilon* such as *A. indicum, A. asiaticum,* and *A. graveolens;* species of *Urena* such as *U. lobata* and *U. sinuata:* species of *Pavonio* such as *P. odorata* and *P. zeylanica* and one or two species of *Cirewia.* Vassou et al., (2015) revealed that *Triumfetta rhomhoidea Jacq* is the common adulterant of *Sida*.

II. DNA BARCODING OF MEDICINAL PLANT

2.1 DNA barcode definition and primary objectives

A DNA barcode is a relatively short gene sequence present in the genome that is unique to species (Jian, 2014). DNA barcoding is a helpful tool for taxonomic classification and identification of species by sequencing a very short standardized DNA sequence in a well-defined gene. Identification of the species is carried out by amplifying highly variable region (DNA barcode region of the nuclear, chloroplast or mitochondrial genome) using Polymerase Chain Reaction (PCR). Region widely used for DNA barcoding include nuclear DNA (e.g. *ITS*), chloroplast DNA (e.g. *rbcL, trnL-F, matK, psbA, trnH, psbK*) and mitochondrial DNA (e.g. *COI*). DNA barcoding is used for a wide range of purposes such as the authenticity of labeling by confirming identity or purity to support ownership or intellectual property rights to reveal cryptic species, in forensics to link biological samples to crime scenes, to support food safety and in ecological and environmental genomic studies.

Paul Herbert and colleagues in 2003 first proposed the use of short DNA sequences as a method of identifying species, with the aim of rapid species-level identifications across all life forms. Hebert *et al.* (2003), proposed to use the mitochondrial gene Cytochrome c oxidase subunit l gene (CO1) as the standard barcode for all animals. This was readily adopted by the scientific community, and assessments have since shown that CO1 can be used to distinguish over 90% of species in most animal groups. In recent years the barcoding movement has grown substantially, and worldwide efforts coordinated by CBOL (the Consortium for the Barcode of Life) are now being put into retrieving barcode sequences from all organisms (CBOL, 2009).

Jaakola and his colleagues (2007) in their study on DNA barcoding of berry species, developed novel approaches to overcome adulteration of bilberries in the market. They designed high resolution melting analysis of ITS and plastid DNA of wild berries. The method is rapid and easily adapted to other food materials. The study showed it can be performed within a short period as it requires single DNA isolation and PCR steps.

Chan Jian and his colleague (2012) on their study on "A successful case of DNA barcoding used in international trade" identified contamination in the ultra-fine powder of roasted barley tea. The research problem rose as when this product was exported from China but returned due to adulteration. They amplified four DNA barcodes including rbcl, ITS-2, PsbA-trnH, and matK. The result of the study showed 77% of the product was contaminated by other plant substances. The study found there is a difficulty in amplifying matK as it is only present in a single copy in the cell, thus the researcher developed an alternative protocol for amplification of matK region.

Therodoridis et al (2014) have employed DNA barcoding in the native plants of Lamiaceae family (lavender, mint, oregano, thyme, sage) of Greece and Turkey. They used three barcoding markers including rbcl, matK, and psbA-trnH. The result of the study showed that psbA-trnH and matK can discriminate species of Lamiacea family and contribute to the trade control as well as conservation of plant resources.

The power of DNA barcoding to identify species has been demonstrated in several studies. Bruni et al. (2010) have employed DNA barcoding for the rapid identification of poisonous plant materials. Moreoer Mattia et al. (2011) have demonstrated that a DNA barcoding approach can be used for the species identification in processed plant materials of commercial kitchen spices. Cornara et al. (2013) have also used DNA barcoding in the identification of ingredients in commercial plant mixtures. However, only a few studies such as; Kool et al. 2012; Wallace et al. 2012; Newmaster et al. 2013, showed the successful approach of DNA barcoding for species adulteration of natural health products and medicinal plant raw drugs.

2.2 The need for a special DNA barcode for plants

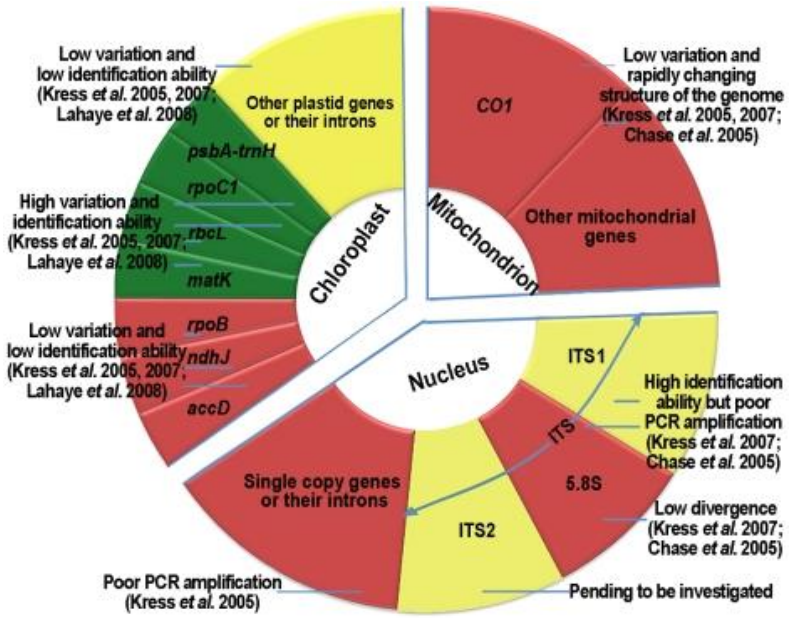
Barcoding of plants, however, has developed at a markedly slower pace. Early on, it became clear that the mitochondrial genome evolves far too slowly in plants to allow it to distinguish between species (Lee, 2016). According to Kress et al (2005), the main reason that barcodes are not applied to plants was that plant mitochondrial genes are poor candidates for species-level discrimination due to their low divergence. Moreover, the mitochondrial genome structure of plants is changed rapidly avoiding the existence of universal intergenic spacers that would be unique identifiers at the species level (Kress et al, 2005).

According to Kress and Erickson, (2008) a gene region must satisfy three criteria to be practical as a DNA barcode. They are; the gene region that should contain a sufficient variation to discriminate between

species, consist of conserved flanks to develop universal PCR primers and have a short sequence length to facilitate current capabilities of DNA extraction and amplification. A single barcoding locus combining these traits has not been found, and a combination of two or more, be required to approach the level of species discrimination and universality for plants (Jian,2014). The reliable candidates for plant DNA barcodes are several chloroplast gene regions nucleotide coding loci or noncoding spacers and coding genes (Rubinoff et al., 2006). Chloroplast genome is appropriate due to the high copy number, conserved structure, and diversity of substitution rates across genes, introns, and intergenic spacers. Figure 1.6 summarizes the candidates for DNA barcodes. Examples of them arc ribosomal RNA maturase K (matK) coding gene and Ribulose 1, 5-bisphosphate carboxylase/oxygenase large subunit coding region (rbcL) which considered as core barcodes.

The CBOL recently recommended the two-locus combination of *matK+rbcL* as the best plant barcode with a discriminatory efficiency of only 72% (CBOL, 2009). Taxonomists have suggested that a multi-locus method may be necessary to discriminate plant species. The combination of the spacer region tRNA-His and photosystem II protein D1 (psbA-trnH spacer) and ITS is also used as a plant barcode for the majority of plant families in addition to core barcode markers (Kress and Lrickson 2008h; Chen et al, 2010).

Locations of these DNA barcodes in plastid genome are presented in Figure 1.7. According to CBOL, matK and rbcL genes are needed to be supplemented by additional loci discriminate among closely related species. The psbA-trnH region remains as the source of additional data (Vijayan and Tsou, 2010; Barbara et al. 2010). The other plastid loci sequenced in plant systematics for phylogenetic purposes are trnL-F, rpoCl, rpoB2, 911, etc, with different degrees of success (Kress et al, 2005).



igure 3: Genes from three genomes in plants that are considered as candidate barcodes. Green represents potential barcodes, red represents poor candidates and yellow represent barcode regions yet to be investigated (Source: Chen et al., 2010).

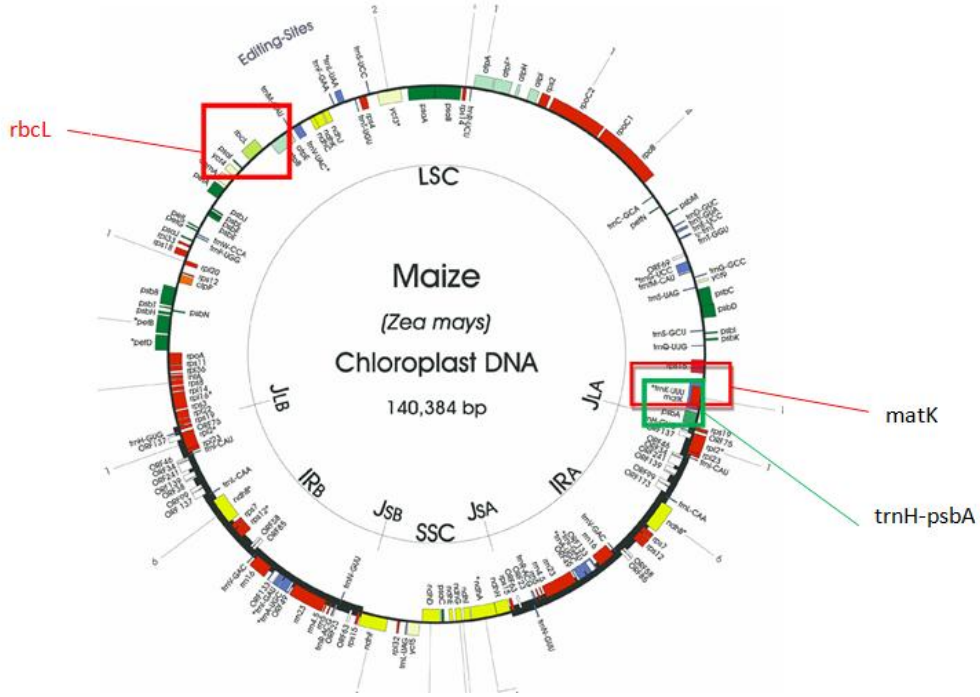


Figure 4: Locations of three potential candidate plant DNA barcodes in chloroplast genome (Source: Kress et al, 2005)

2.3. The matk region

Ribosomal RNA maturase K (matk) is a rapidly evolving coding gene (Hollingsworth et al, 2009). Such a region with 800-1500 base pairs (bp) is located within the intron of the chloroplast gene trnK (Figure 1.7). The studied aspects of this region are rate and types of nucleotide substitutions in the gene and the use of the sequence variation in constructing phylogenies from the tribe to the division level (Lee, 2016).

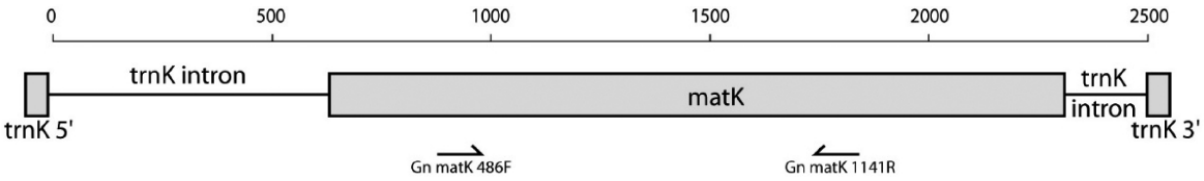


Figure 5: Relative position of matK region. Boxes represent coding regions and connecting show spacer regions (Source: Kress et al, 2005)

According to Windelspecht (2007), MatK has a high evolutionary rate, suitable length, and obvious interspecific divergence as well as a low transition/ transversion rate. Unfortunately, *matK* is difficult to amplify universally using currently available primer sets.

The CBOL Plant Working Group (2009) revealed a nearly 90% success rate in amplifying angiosperm DNA using a single primer pair. However, the success was limited in gymnosperms (83%) and much worse in cryptogams (10%) even with multiple primer sets.

Different primer pairs were necessary for different taxonomic groups (Chase *et al.*, 2007; Hollingsworth, 2008). Lahaye *et al.* (2008) used specific primers to amplify the matK gene of 1667 angiosperm plant samples and achieved a success rate of 100%. A further challenge is the discriminating rates of different taxonomic groups. Matk can discriminate more than 90% of species in the Orchidaceae (Kress & Erickson, 2007) but less than 49% in the nutmeg family. Fazekas *et al.* (2008) attempted the identification of 92 species from 32 genera using the matK barcode but only achieved a success rate of 56%. These findings demonstrate that generally, the matK barcode alone is not a suitable universal barcode.

2. 4 The rbcL region

Ribulose 1, 5- bisphosphate carboxylase/ oxygenase large subunit coding region (rbcL) which is located in the plastid genome has a length of 600-750 bp. The rbcl, region is a candidate for plant barcoding as it can detect generic level evolutionary relationships Windelspecht (2007).

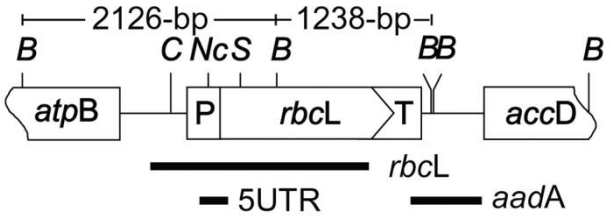


Figure 6 : **The tobacco plastome in rbcL region** (Source: Chen et al., 2010).

The rbcL can be easy to amplify as well as, sequenced and aligned in most land plants. Thus it is a good DNA barcoding region for plants of both at the family and genus levels. Kress *et al.*, (2005) showed that since rbcL sequences evolve slowly, it has a low divergence of plastid genes in flowering plants. Hence it is not much suitable for discriminating at the species level. Moreover, the length is also a drawback where four primers are required for its double-stranded sequencing. However, despite its limitation rbcL is still being wildly used for plant barcoding due to a large amount of easily accessible data and straightforward recovery of the entire gene sequence. Although rbcL by itself does not meet the desired qualities of a barcoding locus, rbcL in combination with various plastid or nuclear loci can make accurate identifications (Gielly&Taberlet, 1994; Renner, 1999; Salazar *et al.*, 2003).

2.5 The trnH-psbA region

Located between tRNA-His and photosystem II protein D 1, trnH-psbA non-coding spacer on in most of the plants shows many features considered as desirable in a barcode (Figure1.9). The trnH-psbA gene with 800-1500 base pairs (bp) is located within the intron of the chloroplast DNA. Since the region has a highly conserved coding sequence, enable in designing feasible universal primers, where a single primer can be used to amplify almost all the angiosperms. The noncoding intergenic region has a high rate of insertion and deletion making trnH psbA a suitable candidate for plant barcoding (Shaw *et al.*, 2007).

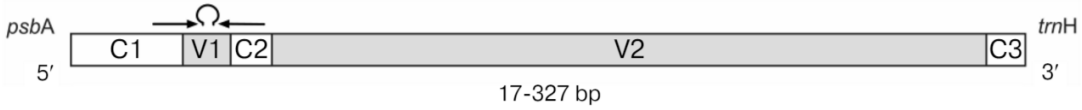


Figure 7: General scheme of the organization of the psbA-trnH spacer, Conserved (C1, C2, and C3) and variable (V1, V2) regions are marked (Shaw *et al.*, 2007).

However, trnH-psbA may undergo frequent inversion in certain plant linkages that may cause incorrect phylogenetic assignment. It some cases, where taxon-specific internal primers are not available, the mononucleotide repeats can have premature termination of sequencing (Lee, 2016). This may lead in difficulty to retrieve longer trnH-psbA regions. Therefore trnH-psbA can be used as a three locus barcode system for adequate identification.

2.6 The ITS region

ITS (Internal transcribed spacer) region is located between the 16S and 23S rRNA genes. It is recommended only as a supplementary locus by CBOL. ITS region is not permitted to being a core barcode marker. Firstly the ITS region is different to amplify and sequence. Secondly, it has an incomplete concerted evolution. Thirdly, the ITS region is not specific for plants as it is also present in fungus (fungal contamination), thus ITS sequence amplification can be confused with fungi sequence (Mishra et al., 2015).

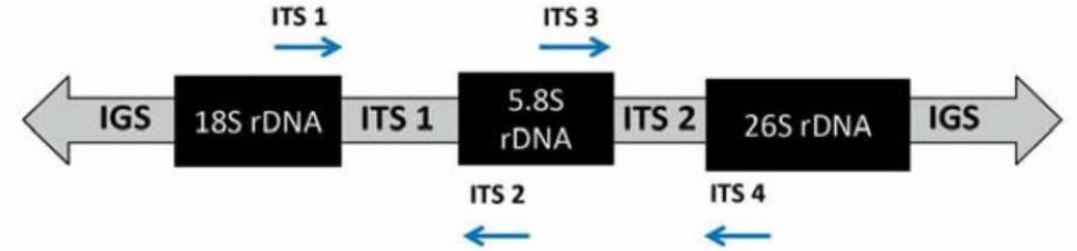


Figure 8: Relative position of ITS region. Boxes represent coding regions and connecting show spacer regions (Source: Kress et al, 2005)

In contrast to the above China Plant, BOL Group presented a different view as the ITS loci can be used as the core barcodes. They suggested to overcome difficulties in sequencing the entire ITS, ITS 2 can be used as back up. Since ITS 2 have a conserved sequence, it can reduce the problem related to the amplification and sequencing of the entire ITS region. Later on, ITS 2 was accepted as a novel universal barcode for a broad range of plant taxa. Mishra et al., (2015) revealed ITS 2 can also be used for herbarium specimens with degraded DNA. Zhen L. (2010) Found ITS 2 usage can increase PCR efficiency.

2.6 The suitable combination of Barcodes for plant identification

Despite extensive efforts to identify a universal plant barcode comparable to CO1 in animals, the task has proved difficult due to the lack of adequate variation within single loci. CBOL suggested that a multi-locus method will be required to obtain adequate species discrimination. Several factors must be weighed in selecting a standard plant DNA barcode. Erickson et al, (2008) and Kress and Erickson, (2008) suggest five factors that should be considered to select a standard land plant DNA barcode. They are universal PCR amplification, power of species differentiation, range of taxonomic diversity and application, complementation among loci and bioinformatic analysis, and application.

Universal PCR amplification is the highest relative rate of recovery of a barcode region by PCR (Erickson et al, 2008). There are several aspects of the power of species differentiation. There are a fraction of species groups that cannot be resolved with any suggested DNA barcode marker, but recovery can be improved through marker selection. When addressing the range of taxonomic diversity and application, many barcode markers are limited in their application of non-angiosperms (Erickson et al, 2008). In the complementation among loci, the agreement supports a plant barcode having two or more marker loci (Kress and Erickson, 2007; Chase et al, 2007; Rubinoff et al, 2006). The complementation of loci has shown a great improvement in the results (Kress and Erickson, 2007). Bioinformatics is indispensable because DNA barcodes work by matching sequence data from a query to a reference sequence from a voucher specimen (Montelone et al., 1981). Hence, an open-access on-line barcode library serves as a standard to the DNA barcode of an unidentified sample that can be matched to a known sequence in the database.

matk +rbcl were recommended as universal barcode combination by CBOL as the rbcl region can be directly recovered and matK sequence has high discriminating power. Even though matk +rbcl combination produces higher identification efficiency than other combination such combination fails to acquire the original purpose of universal DNA barcode. Firstly Matk +rbcl combination has very low PCR efficiency. Secondly, discrimination power is lower than COI in animals.

Even though combined barcodes are recommended as it has numerous beneficial outcomes, the opposite argument on its usage is also being raised. Chen et al., (2010) stated that compared to the single-locus marker, the combined barcode causes increase analytical difficulty. Moreover, CBOL also demonstrated in one study that the use of seven different candidate loci combinations did not improve species level discriminatory ability compared to rbcl+ matK. Some authors stated inconsistency between the plastid gene tree and species boundaries could cause the failure of multiple-locus barcodes. They showed that such multiple-locus barcodes cannot eliminate the inherent deficiencies of the current DNA barcoding of plants (Dubey et al., 2016).

III. DNA BARCODES OF *SIDA* SPECIES

Medicinal plants are utilized for the preparation of herbal health supplements in Ayurveda and indigenous medicine. Plants used as drugs, dietary supplements, and herbal medicines are identified at the species level. DNA-based methods are developed for medicinal plant identification. Phytochemical and genetic data are correlated, but only genetic data allow for medicinal plant differentiation at the species level (Sucher and Caries, 2000).

Ashfaq et al (2013) studied DNA barcoding efficiency on the cotton of the Malvaceae family. They tested the effectiveness of three barcoding markers including matK, rbcL, and ITS2 in discriminating 20 diploid and five tetraploid species of cotton. Based on the 'best match' analysis, they found the combination of matK+ITS2 was best. Moreover, based on all species barcodes analysis, ITS2 gave the maximum percentage of correct species identifications. Although all three barcodes separated the species concerning their genome type, the study concluded that no single combination of barcodes able to differentiate all the *Gossypium* species.

Poovitha et al. (2016) analyzed the discriminatory power of rbcL and matK. trnH-psbA and ITS2 on 16 species of *Hibiscu*s, of Malvaceae family distributed around southern peninsular India. Differentiation of all the species of *Hibiscus* was possible with the matK DNA barcode marker. Two marker combinations that include matK gave 100% species differentiation as well as better species resolution. The research concluded that matK is more suitable when compared to rbcL, trnH-psbA, and ITS2 for species identification of *Hibiscus.*

In the study on "DNA barcoding to assess species adulteration in the raw drug trade of "Bala" (genus: *Sida L.*) Herbal products in South India", Kumar et al (2015) investigate species adulteration in the raw herbal trade of *Sida* in southern India. ITS, matK, and psbA-trnH were used as DNA barcode markers in the study. Whereas psbA-trnH and ITS region were found to have discriminating power for the species of *Sida* according to inter specific distance. The study indicated that species adulteration in the market samples is uncontrolled especially in the case of *Sida cordifolia*, where all the market samples analyzed were *Sida acuta*.

Nithaniyal et al. (2017) reported a case study in which the species identification of the market samples of *Sida cordifolia* was done by DNA barcoding. rbcL, matK, psbA-trnH, and ITS2 regions were used as bar cording markers. Based on the intra-species and inter-species divergence, psbA-trnH and ITS2 were found to be the best two marker combinations for species identification. The study showed that none of the market samples belonged to the authentic species, *S. Cordifolia*, where the majority of the market samples owned to other species of *Sida i*ncluding Sida *acuta, S. alnifolia, S. Spinosa, S. scabrida* and *S. ravii*. As well as other genera such as *Abutilon sp., Ixonanthes sp., Terminalia sp., Fagonia sp.,* and *Tephrosia sp.* This observation was in contrast to the belief that medicinal plants are generally substituted or adulterated with closely related species.

IV. CONCLUSION

The findings of this study indicate that currently used primers of psbA-trnH, ITS-2, rbcL, and matK in DNA barcoding, can amplify the regions of *Sida* DNA. DNA sequencing was successful for all four regions. Blastn search aligned *Sida* species with a plant in the same genus, whereas matK region was not accurately aligned. The results highlight that among the four regions used in this study, ITS-2 and rbcL barcode regions could serve as the best DNA barcodes for the *Sida* species of Sri Lanka. The regions for these species were made available at the Genbank database. These DNA barcodes will contribute to the identification, conservation, and quality control of valuable Sri Lankan medicinal plant species of *Sida*.

In conclusion, as the advantages and limitations of DNA barcoding become evident, it is clear that taxonomic processes incorporating DNA sequencing will achieve maximum efficiency at species identification. DNA barcoding enables us to speed up the work of taxonomists in species identification. Despite some drawbacks of using DNA barcoding, the study reported here showed the success of using the barcoding region in distinguishing *Sida* species of Sri Lanka and resolved species adulteration of Babila in the natural drug market. The study recommends that in the future, efforts should be taken to develop nuclear barcodes to complement the barcoding regions that are currently in use for even more accurate barcoding results.

REFERENCE

[1]. Alaeddini R, Walsh S. and Abbas A. Forensic implications of genetic analyses from degraded DNA-A review. Forensic Science International: Genetics. 2010; 4:148-157.

[2]. Ashfaq M, Asif M, Anjum Z. and Zafar Y. Evaluating the capacity of plant DNA barcodes to discriminate species of cotton (Gossypium: Malvaceae). Molecular Ecology Resources. 2013; 13:573-582.

[3]. Bacich D, Sobek K, Cummings J, Atwood A. and O'Keefe D. False negative results from using common PCR reagents. BMC Research Notes.2014; 4:457-467.

[4]. Baracho S, George Y, and Maria A. A New Record Of Sida Glutinosa (Malvaceae), A Rare Species Of The Caatinga In Northeastern Brazil, With Lectotypifications And Taxonomic Notes On The Allied Sida Glabra. Phytotaxa 2016: 37.

[5]. Bharati, K.. Identification of Indian Sida through mericarp. Pharmacognosy Journal. 2016;8:490-496.

[6]. Chen S, Yao H, Han J, Liu C, Song J, Shi L, Zhu Y, Ma X, Gao T, Pang X, Luo K, Li Y, Li X, Jia X, Lin Y. and Leon C. Validation of the ITS2 Region as a Novel DNA Barcode for Identifying Medicinal Plant Species. PLoS ONE. 2010; 5: 614-618.

[7]. Cheung W, Hubert N. and Landry B. A simple and rapid DNA microextraction method for plant, animal, and insect suitable for RAPD and other PCR analyses. Genome Research. 1993; 3:69-70.

[8]. Costion C, Lowe A, Rossetto M, Kooyman R, Breed M, Ford A. and Crayn D. Building a Plant DNA Barcode Reference Library for a Diverse Tropical Flora: An Example from Queensland, Australia. Diversity. 2016; 8:5-10.

[9]. Coyne K, Hare C, Popels L, Hutchins D. and Cary S. Distribution of Pfiesteria piscicida cyst populations in sediments of the Delaware Inland Bays, USA. Harmful Algae. 2006;5:363-373.

[10]. Dubey S, Tripathi A, Upadhyay B. and Kumar A. Development of conventional and real time PCR assay for detection and quantification of Rhizoctonia solani infecting pulse crops. Biologia. 2016; 71:89-95.

[11]. Fazekas A, Burgess K, Kesanakurti P, Graham S, Newmaster S, Husband B, Percy D, Hajibabaei M. and Barrett S. Multiple Multilocus DNA Barcodes from the Plastid Genome Discriminate Plant Species Equally Well. PLoS ONE. 2008; 3:178-184.

[12]. Gunatilaka, A., Arseculeratne, S., and Panabokke, R. Studies on medicinal plants of Sri Lanka. part 14: toxicity of some traditional medicinal herbs. Journal of Ethnopharmacology, 1985; 13:323-335.

[13]. Hebert P, Cywinska A, Ball S. and deWaard J. Biological identifications through DNA barcodes. Proceedings of the Royal Society B: Biological Sciences. 2003; 270:313-321.

[14]. Hilu K. and Liang H. The matK Gene: Sequence Variation and Application in Plant Systematics. American Journal of Botany. 1997; 84:830-836.

[15]. Jian. "A Successful Case Of DNA Barcoding Used In An International Trade Dispute". DNA Barcodes 2.1. 2014.

[16]. Kane N, Sveinsson S, Dempewolf H, Yang J, Zhang D, Engels J. and Cronk Q. Ultra-barcoding in cacao (Theobroma spp.; Malvaceae) using whole chloroplast genomes and nuclear ribosomal DNA. American Journal of Botany. 2012; 99:320-329.

[17]. Kazi T, Hussain N., Bremner P., Slater A. and Howard C. The application of a DNA-based identification technique to over-the-counter herbal medicines. Fitoterapia. 2013; 87:27-30.

[18]. Koolman J. and Roehm K. Color atlas of biochemistry. 1st ed. Stuttgart: Thieme. 2013.
[19]. Kress W. and Erickson D. A Two-Locus Global DNA Barcode for Land Plants: The Coding rbcL Gene Complements the Non-Coding trnH-psbA Spacer Region. PLoS ONE. 2007; 2:508-511.
[20]. Kumar D. and R. Kundapur R.Biomedical Applications of Natural Proteins. 1st ed. New Delhi: Springer India.2015.
[21]. Kumar, K., Ramachandra Setty, S. and Narsu, L. Pharmacognostic and Phytochemical Investigations of Stems of Hibiscus micranthus Linn. Pharmacognosy Journal, 2010; 2:21-30.
[22]. Lahaye R, van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, Maurin O, Duthoit, S, Barraclough, T. and Savolainen, V. DNA barcoding the floras of biodiversity hotspots. Proceedings of the National Academy of Sciences, 2008;. 105:2923-2928.
[23]. Lee S, Mahat M, Nazre M. and Mohamed, R. DNA Barcoding of the Endangered Aquilaria (Thymelaeaceae) and Its Application in Species Authentication of Agarwood Products Traded in the Market. PLOS ONE. 2016;11:154-171.
[24]. Lee. DNA Barcoding Of The Endangered Aquilaria (Thymelaeaceae) And Its Application In Species Authentication Of Agarwood Products Traded In The Market. PLOS ONE.2016;87-90.
[25]. Mishra P, Kumar A, Nagireddy A, Mani D, Shukla A, Tiwari R. and Sundaresan V. DNA barcoding: an efficient tool to overcome authentication challenges in the herbal market. Plant Biotechnology Journal.2015; 14: 8-21.
[26]. Montelone B, Prakash S. and Prakash L. Recombination and mutagenesis in rad6 mutants of Saccharomyces cerevisiae: Evidence for multiple functions of the RAD6 gene. MGG Molecular & General Genetics. 1981; 184:410-415.
[27]. Newmaster S, Grguric M, Shanmughanandhan D, Ramalingam S. and Ragupathy S. DNA barcoding detects contamination and substitution in North American herbal products. BMC Medicine. 2013; 11:78-82.
[28]. Nithaniyal S, Vassou S, Poovitha S, Raju B, Parani M. and Cristescu M. Identification of species adulteration in traded medicinal plant raw drugs using DNA barcoding. Genome.2017; 60: 139-146.
[29]. Obah, J O Akerele and O Obasuyi. Antimicrobial Activity Of The Ethanol Extract Of The Aerial Parts Of Sida Acuta Burm.F. (Malvaceae). Tropical Journal of Pharmaceutical Research. 2007;7:451-455.
[30]. Ojeda D. I. DNA Barcodes Successfully Identified Macaronesian Lotus (Leguminosae) Species Within Early Diverged Lineages Of Cape Verde And Mainland Africa. AOB PLANTS.2014;47:784-794.
[31]. Palhares R, Drummond M, Brasil, B, Cosenza G, Brandao M. and Oliveira G. Medicinal Plants Recommended by the World Health Organization: DNA Barcode Identification Associated with Chemical Analyses Guarantees Their Quality. PLOS ONE.2015; 10:116-124.
[32]. Parveen I, Gafner S, Techen N, Murch S. and Khan I. DNA Barcoding for the Identification of Botanicals in Herbal Medicine and Dietary Supplements: Strengths and Limitations. Planta Medica.2016; 82:1225-1235.
[33]. Poovitha S, Stalin N, Balaji R, Parani M. and Bank M. Multi-locus DNA barcoding identifies matK as a suitable marker for species identification in Hibiscus L. Genome. 2016; 59:1150-1156.
[34]. Pusch C. Repair of degraded duplex DNA from prehistoric samples using Escherichia coli DNA polymerase I and T4 DNA ligase. Nucleic Acids Research. 1998;26:857-859.
[35]. Renner S. and Chanderbali A. What Is the Relationship among Hernandiaceae, Lauraceae, and Monimiaceae, and Why Is This Question So Difficult to Answer? International Journal of Plant Sciences. 2000; 161:109-119.
[36]. Sambrook J, Fritsch E. and , Maniatis T. Molecular cloning. 1st ed. Cold Spring Harbor Laboratory Press.1982.
[37]. Santhosh Kumar J, Krishna V, Seethapathy G, Senthilkumar U, Ragupathy S, Ganeshaiah K, Ganesan R, Newmaster S, Ravikanth G. and Uma Shaanker R. DNA barcoding to assess species adulteration in raw drug trade of "Bala" (genus: Sida L.) herbal products in South India. Biochemical Systematics and Ecology.2015;61: 501-509.
[38]. Särkinen T, Staats M, Richardson J, Cowan R. and Bakker F. How to Open the Treasure Chest? Optimising DNA Extraction from Herbarium Specimens. PLoS ONE. 2012;78:80-88.
[39]. Shaw J, Lickey E, Schilling E. and Small R. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. American Journal of Botany. 2007; 94:275-288.
[40]. Simplice D, Wendyam M, Denise P, Djeneba O, Messanvi G, Comlan D. and Jacques, S. Sida acuta Burm medicinal plant with numerous potencies. African Journal of Biotechnology.2007; 6:2953-2959.
[41]. Srirama R, Senthilkumar U, Sreejayan N, Ravikanth G, Gurumurthy B, Shivanna M, Sanjappa M, Ganeshaiah K. and Uma Shaanker R. Assessing species admixtures in raw drug trade of Phyllanthus, a hepato-protective plant using molecular tools. Journal of Ethnopharmacology. 2010; 130:208-215.
[42]. Tamura K, Peterson D, Peterson N, Stecher G, Nei M. and Kumar, S. MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Molecular Biology and Evolution. 2011; 28:10:2731-2739.
[43]. Techen N, Parveen I, Pan Z. and Khan I. DNA barcoding of medicinal plant material for identification. Current Opinion in Biotechnology. 2014; 25:103-110.
[44]. The use of chloroplast DNA to resolve plant phylogenies: noncoding versus rbcL sequences. Molecular Biology and Evolution. 1994.
[45]. Vassou S, Kusuma G. and Parani M. DNA barcoding for species identification from dried and powdered plant parts: A case study with authentication of the raw drug market samples of *Sida* cordifolia. Gene. 2015; 559:86-93.
[46]. Verdcourt B. The Variation of Sida rhombifolia L. (Malvaceae) in East Africa. Kew Bulletin. 2004; 59:2:233-245.
[47]. Wicke S, Costa A, Muñoz J. and Quandt D. Restless 5S: The re-arrangement(s) and evolution of the nuclear ribosomal DNA in land plants. Molecular Phylogenetics and Evolution. 2011; 61:321-332.
[48]. Windelspecht M. (2007). Genetics *101*. 1st ed. Santa Barbara: ABC-CLIO.
[49]. Woodward D. and Woodward V. (1978). Concepts of molecular genetics. 1st ed. New Delhi: Tata McGraw-Hill.
[50]. Zhen L. (2010). DNA barcoding in medicinal plants Caprifoliaceae. China Journal of Chinese Materia Medica.

R. Rajphriyadharshini." DNA Barcoding of Medicinal plant: a Systemic Review." *International Journal of Pharmaceutical Science Invention*, vol. 09(6), 2020, pp 06-16. Journal DOI- 10.35629/6718