Pak. J. Biochem. Mol. Biol., 2022, 55(3):169-178

Research paper

Comparison of Amplification Ability of Mat-K and RbcL DNA Barcodes in the Identification of Wild Grasses of the Quetta Ecosystem

Aemad Tabasum[1], Shahjahan Shabbir Ahmed Rana[1]*, Nazeer Ahmed[1], Imran Ali Sani[2], Dawood Shahid[1], Shazia Irfan[2], Syed Moeezullah[1] and Muhammad Mushtaq[1]

[1]Department of Biotechnology, Faculty of Life Sciences, Balochistan University of Information Technology Engineering and Management Sciences (BUITEMS), Quetta, Pakistan.

[2]Sardar Bahadur Khan Women University (SBK), Quetta, Pakistan.

*corresponding author: Dr. Shahjahan Shabbir Ahmed Rana

(Shahjahan.shabbir@buitms.edu.pk).

ABSTRACT

DNA barcoding is an innovative method that utilizes a short-length DNA region registered as a DNA barcode for species investigation, by comparing the particular barcoding regions with the sequences already placed within the reference library. The study has been designed to decipher DNA barcode applications in the identification of wild grasses of the Quetta ecosystem. For this purpose, DNA extraction of multiple grass samples was performed using an inorganic CTAB DNA extraction protocol. All the samples were electrophoresed on 1% agarose gel and run for PCR analysis. Samples were undergone for amplification and sequencing of chloroplast DNA regions i.e., Mat-k and RbcL. The results showed that the RbcL and Mat-k have the potential as bar code markers. 7 samples showed positive results for the RbcL and only 4 were positive when Mat-k was used as the primer. Of these two primers, RbcL has more potential to be used as the barcode marker as it showed more positive results than the Mat-k. Using these markers, we identified the samples up to the generic level. All the samples were identified, and vouchers were marked and labeled by the seasoned taxonomist. In this study, we concluded that RbcL and Mat-k are the potential markers for barcodes in grasses and they resolved the grasses to the generic and species level. The finding of this study suggested that RbcL is an effective barcode at the generic level as compared to Mat-k.

KEYWORDS DNA barcoding, Grasses, RbcL, Mat-k, BLAST, Quetta Ecosystem

INTRODUCTION

DNA Barcoding is an innovative blend of taxonomy, genetics, and computer science that automates the process of obtaining expert species identifications [1]. DNA barcoding relies on finding different conserved regions in divergent species to produce a large-scale reference genome library [2]. It is reported earlier that the primers designed on nuclear and mitochondrial regions could map the route of species identification. Moreover, these specific sequences are termed barcodes that are later used in the phylogenetic analysis, genetic diversity, and species discrimination in different organisms. This system not only helps to classify the organisms but also reveals genetic information for species ancestral

inheritance and flagging of new species [2,3,4]. Things become trickier to identify effective conserved regions coupled with complex genomic diversity in the case of plants. Chloroplasts are active metabolic machinery in green plants to convert light energy to carbohydrates. High-throughput sequencing technology has resulted sequencing of about 800 chloroplast genomes from different plants [5]. Two conserved regions from plastid (chloroplast) genome maturase-K (Mat-k) and ribulose-1, 5-bisphosphate carboxylase/oxygenase (RbcL) were proposed as barcode primers to discriminate large group of angiosperms [2].

Ram, Jain, Mishra, Mandal, and Abdin [6], suggest that DNA barcoding is a multidisciplinary field. It involves a unique amalgamation of taxonomy, computer sciences and genetics that implies the process of obtaining expert species identifications. An earlier report by [7] speculate that the grasses are prevalent in all continents of the world except the arctic regions of Antarctica.

Despite some early attempts of molecular inventories in higher plants quite meager knowledge is available about grasses identification [8]. Poaceae formerly known as Graminae members are some of the most economically and environmentally important temperate native and introduced pasture and turf grasses, including *Lolium* L. (Ryegrasses), *Poa* L. (Tussock-grasses), and *Puccinellia* Parl. (Salt-grasses) [8,9]. Plethora of cereals and bamboos cultivated lawn are vital members of this family. Vegetation of grasses in numerous ecosystems makes them inevitable for numerous residents of different species of both animals and plant, they are vital in numerous territories, including wetlands, timberlands, and tundra. Grasses as forage and game reserves to enhance the aesthetic beauty of the sites/parks and golf clubs are inevitable, almost all the game reserves rely on the growth and progress of the grasses in their lawns [10]. Implication of grasses as forage and turf could be ambiguous if their properties are not classified and distinguished at species level. DNA barcoding is the first hand choice to chalk out differences of the species exist in many wild grasses [11]. *Bromus tectorum*, *Cynodon dactylon*, *Lollium perrenne*, *Aegilops touschii Coss*, are well known grass species for their aesthetic, forages, and dietary uses [12]. To exhibit a successful DNA barcoding venture, differences in sequences between two species must be higher enough to discriminate both. In contrary, intra and interspecific variation of the sequences must be lower enough to distinguish the variations [2].

Unanimous decision taken by the Consortium for the Barcode of Life Plant Working Group that seven chloroplast genomic regions are qualified as barcodes and hence, a combination of Mat-k with RbcL was recommended as universal barcode for the identification and authentication of flowering plants [13]. The universal primer of RbcL possesses high universality but low-resolution power whereas Mat-k deem low universality with high resolution power among different plant species. It is therefore suggested to use a combination of both markers can be helpful to discriminate variety of species on a broader range. However, in order to gain maximum discrimination power between two closely related species, a combination of Internal Transcribed Spacer (ITS) with Mat-k and RBCK was recommended by China Plant BOL Group [14].

The efficiency and discrimination power of mentioned markers used in DNA over RNA is better hypothesized just because of the stability of the molecule and therefore, suggested to use in most of the terrestrial plants like grasses. Present study design in view of questing the following points.

MATERIALS AND METHODS

Present study design in view of questing the following points. (a) Amplification and sequencing of Mat-k and RbcL regions in

nine different grass species collected from the Quetta ecosystem (b) Determining functional annotation and homology modeling of both sequences among nine species by the help of Basic Local Alignment Search Tool (BLAST) (c) Comparative genomic analysis to depict sequences in detail comparative genomic analysis in order to study these sequences in detail among different grass species (d) To conduct phylogenetic analysis among these nine grass species collected from Quetta ecosystem. We believe that the output of the study would further open the horizon of molecular identification, authentication of the grasses dwelling in the northwestern parts of Balochistan and other allied regions.

Plant Material and Sample Collection

The samples of wild grasses were collected across different regions of Quetta, Balochistan ecosystem and data of these samples were recorded that can be seen in Table 1. Samples were collected randomly on the basis of quadrat method Samples were preserved in polythene zipper bags with silica gel Taxonomic identification carried out and voucher number were assigned by the credible taxonomist of Sardar Bahadur Khan Women (SBK), University, Quetta, Pakistan.

DNA Extraction and Estimation

DNA was extracted from leaves by CTAB (cetyltrimethylammonium bromide detergent) method with minor modifications.

Amplification PCR Products

The PCR reaction was performed for each sample in a volume of 20 µl solution using commercial master mix. Each PCR tube containing 10µl PCR Master mix, 1µl template DNA, 1µl both reverse and forward primers (10pM) and 8µl PCR water. The reaction was carried out in a single run using temperature gradient. The software was designed in such a way that annealing temperature was set at 51°C for RbcL and 52 °C for Mat-K to anneal the primers on specific region. The detail about each primer is shown in table 2. PCR cycling conditions was followed by initial denaturation, 94°C for 5 min for both RbcL and Mat-k. 30-35 cycles of denaturation, annealing and extension was followed as show in Figure 1 (a) and Figure 1(b). After completion of cycles the final extension was set at 72°C for 4 min.

Confirmation of Amplified PCR products on Agarose gel

The amplified PCR products were resolved on 2% agarose gel. The gel was prepared in 1x TBE buffer (Tris-base, boric acid, EDTA) added with 0.5µl ethidium bromide for visualization of bands. Gel electrophoresis was performed at 110 voltages for 45 minutes in 1x TBE buffer and the length of PCR products were compared with 100bp DNA ladder. The DNA bands were then visualized under UV light at 260nm in a gel documentation system.

Nucleotide Sequencing

The PCR amplified products were sent for Sanger DNA sequencing service from Beijing Genome Institute (BGI), Shenzhen, China.

RESULTS AND DISCUSSION

RbcL and Mat-k markers have the potential for the identification and authentication of flowering plants. In our study we extracted DNA of wild grasses collected across different regions of Quetta (Figure 2), the amplified products of conserved regions of both primers were required to qualify as barcodes for the species discrimination. Our results with RbcL and Mat-K primers indicated vibrant PCR products of nine and four different bands, respectively using templates of DNA (Figure 3&4). Earlier report of [15] about the barcode amplification in angiosperms, gymnosperms and liverwort are also coincided with our results.

Homology of nucleotide sequences were analyzed through BLAST [16]. The average sequence length of RbcL was ranging between 550 to 600 nucleotides of ten species with *Cynodon dactylon, Cynodon dactylon, Lolium perenee* (L.), *Aegilopus touschii Coss, Austrostipa ramosissima, Stipa lessingiana, Hordium murinium, Lolium perenee* (L.),

Cynodon dactylon and *Bromus tectorum* (Table 4). Similarly, with Mat-k, the sequence length of four species was ranging between 800 to 900 nucleotides with *Bromus tectorum*, *Cynodon dactylon, Hordium murinium* and *Austrostipa ramosissima* (Table 3).

With Mat-k, the sequence homology of *Austrostipa ramosissima* was 93% with *Secale strictum* and *Elymus tauri*. The sequence homology of *Hordeum murinum* was 94% with *Hordeum murinum* and *Hordeum pubiflorum* with maximum cover was 1242. The sequence of *Cynodon dactylon* was 99% with *Cynodon transvaalensis* and 99% with *Cynodon radiatus*. The sequence of *Bromus tectorum* was 95% similar with *Cynodon transvaalensis* as shown in Table 3 and Mat-k phylogenetic tree is shown in Figure 5a.

With RbcL, the sequence homology of *Cynodon dactylon* was 99% with *Eleocharis uniglumis*. The sequence homology of *Cynodon datylon* was 99% with *Cynodon dactylon*. The sequence homology of *Lolium perenne* was 100% with *Lolium perenne.* The sequence homology of *Aegilopus touschii Coss* was 99% with *Aegilops umbellulata*. The sequence homology of *Austrostipa ramosissima* was 99% with *Austrostipa ramosissima.* The sequence homology of *Stipa lessingiana* was 99% with *Stipa lessingiana.* The sequence homology of *Hordium murinium* was 99% with *Hordium murinium.* The sequence homology of *Lolium perenee* was 99% with *Lolium rigidum.* The sequence homology of *Cynodon dactylon* was 99% with *Cynodon radiatus.* The sequence homology of *Bromus tectorum* was 95% with *Bromus boissiera* and 98% as shown in Table 4 and its phylogenetic tree is shown in Figure 5b.

The quality of extracted DNA plays a pivotal role in the discrimination power of barcode primers used especially that hampers by the presences of metabolites in most of the cases [17]. Application of DNA barcode markers as universal loci primers are broadly evident to discriminate species of wide range of terrestrial plants [18-20]. The present study shows amplification and nucleotide sequencing of fourteen wild grasses' two conserved regions of chloroplast. These grasses are the representative of major chunk of natural grasses exist in the southwestern regions of the Balochistan consider as forage resource and main stay of the livestock. It is the first report about the implementation of DNA barcodes in the species discrimination of forage grasses of the natural ecosystem of the Quetta. Our findings suggest that the application of RbcL and Mat-K as DNA barcoding is quite successful for generating genus and species discrimination power in the wild grasses of the natural ecosystems and pastures. Although the amplification success rates were 70% in the case of RbcL and less than 50% in case of Mat-k, but their BLAST analysis revealed their efficiency and reliability more than 85%. Some previous studies also coincided with the findings of current study on the basis of increased levels of metabolites [21].

ACKNOWLEDGEMENTS

The authors acknowledge the office of the research innovation and commercialization (ORIC), BUITEMS for the financial support.

CONFLICT OF INTEREST

The authors declare there is no conflict of interest in the designed study.

Table 1: Samples with voucher codes

No.	Code	Name	Quadrates	Elevation feet (ft)
1	**CPC**	***Cynodon dactylon***	**N= 30°16'19.1 E= 066°56'20.1**	**5354**
2	**CPD**	***Cynodon dactylon***	**N= 30°16'16.8 E= 066°56'18.3**	**5285**
3	**GS1**	***Lolium perenee (L.)***	**N= 30°16'12.6 E= 066°56'15.8**	**5202**
4	**GS2**	***Aegilopus touschii Coss***	**N= 30°16'19.4 E= 066°56'17.8**	**5197**
5	**GS3**	***Stipa lessingiana***	**N= 30°16'15.9 E= 066°56'21.2**	**5201**
6	**AZ1**	***Austrostipa ramosissima***	**N= 30°16'09.7 E= 066°56'31.3**	**5211**
7	**BU3**	***Hordium murinium***	**N= 30°16'09.6 E= 066°56'31.3**	**5211**
8	**LF1**	***Lolium perenee (L.)***	**N= 30°15'59.0 E= 066°56'41.8**	**5211**
9	**BZ1**	***Cynodon dactylon***	**N=40°16'59.0 E= 069°58'41.8**	**6233**
10	**BU1**	***Bromus tectorum***	**N= 30°17'59.0 E= 056°56'41.8**	**5210**

Table 2 Characteristics of RbcL (Ribulose-1,5 Bisphosphate Carboxylase) and Mat-K (Megakaryocyte-Associated Tyrosine Kinase) primers used for PCR amplification.

Primer Name	**PCR product size**	**5'-sequence-3'**	**Tm℃**	**GC%**
RbcL-α-F	650bp	ATGTCACCACAAACAGAGACTAAAGC	64.7	42.31
RbcL-α-R		GTAAAATCAAGTCCACCRCG	56.4	45.00
Mat-K-390 -F	850bp	CGATCTATTCATTCAATATTTC	52.8	27.27
Mat-K -1326-R		GTAAAATCAAGTCCACCRCG	60.03	45.45

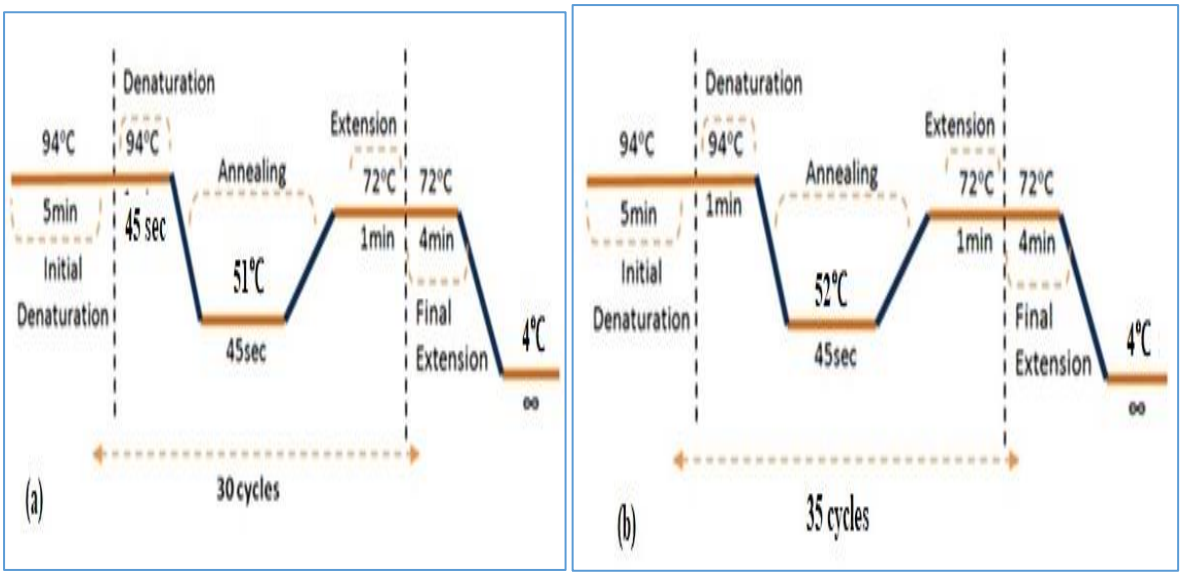

Figure 1(a): PCR cyclic condition for RBCL.

Figure 1(b): PCR cyclic condition for Mat-K

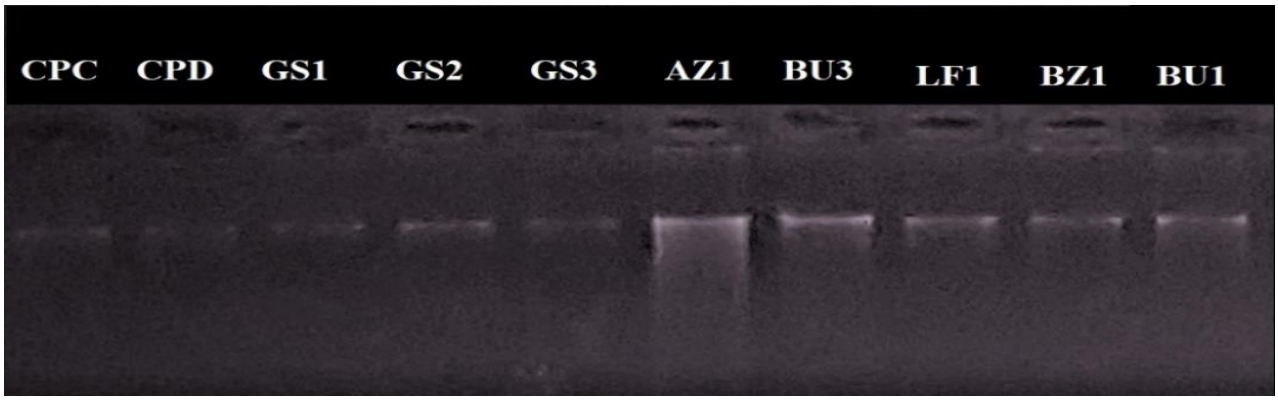

Figure 2: Total extracted DNA of samples used in this study.

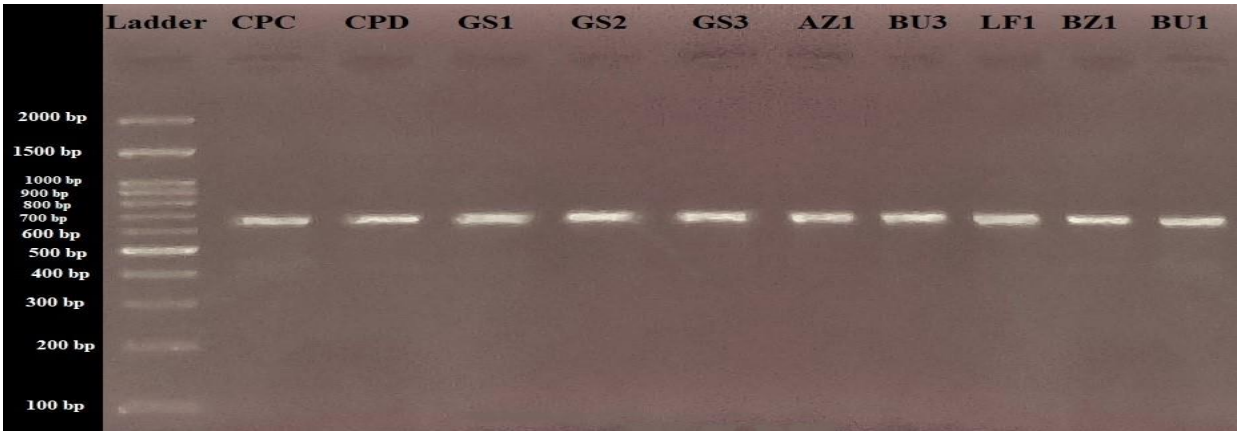

Figure 3: Amplified PCR Products of RbcLprimers and Standard 100 bp DNA ladder (MCLAB # bp DNA 250) on 2% Agarose gel.

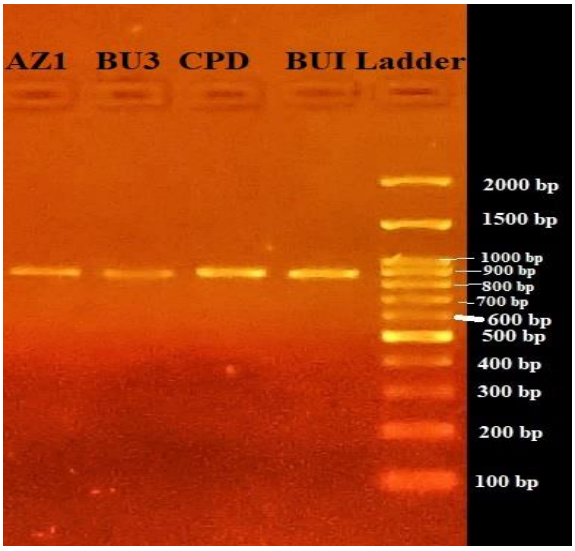

Figure 4: Four amplified PCR products of Mat-k primers and Standard 100 bp DNA ladder (MCLAB #bpDNA 250), on 2% agarose gel

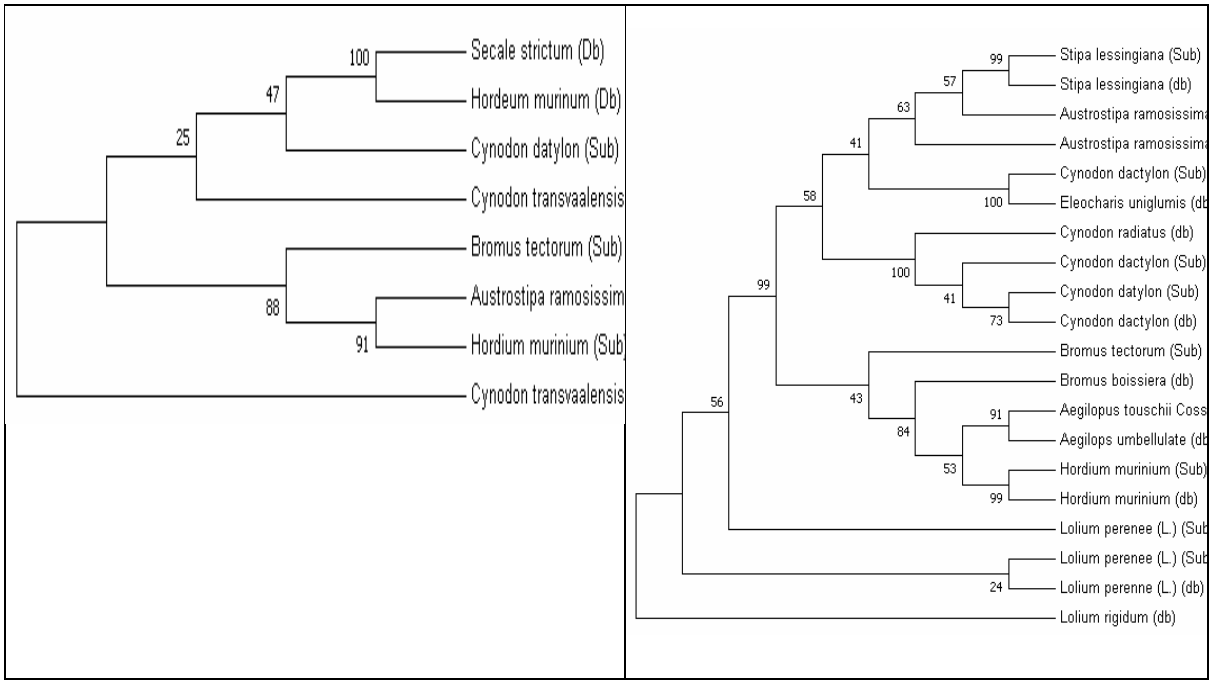

Figure 5: Phylogenetic relationship of wild grasses as identified using Mat-k primers (5.a) and RbcL primers (5.b)

Table 3: Results of BLAST using Mat-k primers for identification of wild grasses

Name	Code	Matched Specie in BLAST	Max score	Query Cover	E value	Ident	Accession
Austrostipa ramosissima	AZ1	*Secale strictum*	1192	100%	0.0	93%	KY636138.1
Hordium murinium	BU3	*Hordium murinium*	1236	99%	0.0	94%	KY636107.1
Cynodon datylon	CPD	*Cynodon transvaalensis*	1480	100%	0.0	99%	HE591380.1
Bromus tectorum	BU1	*Cynodon transvaalensis*	1254	93%	0.0	95%	HE591380.1

Table 4: Results of BLAST using *RbcL* primers for identification of wild grasses

Name	Code	Matched species on BLAST	Max score	Query Cover	E value
Cynodon dactylon	CPC	*Eleocharis uniglumis*	987	92%	0.0
Cynodon datylon	CPD	*Cynodon dactylon*	987	95%	0.0
Lolium perenee (L.)	GS1	*Lolium perenne (L.)*	852	100%	0.0
Aegilopus touschii Coss	GS2	*Aegilops umbellulate*	987	95%	0.0
Stipa lessingiana	GS3	*Stipa lessingiana*	968	100%	0.0
Austrostipa ramosissima	AZ1	*Austrostipa ramosissima*	955	99%	0.0
Hordium murinium	BU3	*Hordium murinium*	987	100%	0.0
Lolium perenee (L.)	LF1	*Lolium rigidum*	983	100%	0.0
Cynodon dactylon	BZ1	*Cynodon radiatus*	987	100%	0.0
Bromus tectorum	BU1	*Bromus boissiera*	935	94%	0.0

REFERENCES

1.Chase, M. W., & Fay, M. F. (2009). Barcoding of plants and fungi. *Science, 325*(5941), 682-683.

2.Wattoo, J. I., Saleem, M. Z., Shahzad, M. S., Arif, A., Hameed, A., & Saleem, M. A. (2016). DNA Barcoding: Amplification and sequence analysis of RbcL and matK genome regions in three divergent plant species. *Advancements in Life Sciences, 4*(1), 03-07.

3. Hebert, P. D., Penton, E. H., Burns, J. M., Janzen, D. H., & Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper

butterfly Astraptes fulgerator. *Proceedings of the National Academy of Sciences, 101*(41), 14812-14817.

4.Son, Y.-O., Kim, J., Lim, J.-C., Chung, Y., Chung, G.-H., & Lee, J.-C. (2003). Ripe fruits of Solanum nigrum L. inhibits cell growth and induces apoptosis in MCF-7 cells. *Food and Chemical Toxicology, 41*(10), 1421-1428.

5.Kelchner, S. A. (2000). The evolution of non-coding chloroplast DNA and its application in plant systematics. *Annals of the Missouri Botanical Garden*, 482-498.

6.Ram, M., Jain, D., Mishra, H., Mandal, S., & Abdin, M. (2014). Recent Advances to Enhance Yield of Artemisinin: A Novel Antimalarial Compound, in Artemisia annua L. Plants. In *Artemisia annua-Pharmacology and Biotechnology* (pp. 173-204): Springer.

7. Henkel, M. (2015). *21st century homestead: Sustainable agriculture III: Agricultural practices*: Lulu. com.

8.Birch, J. L., Walsh, N. G., Cantrill, D. J., Holmes, G. D., & Murphy, D. J. (2017). Testing efficacy of distance and tree-based methods for DNA barcoding of grasses (Poaceae tribe Poeae) in Australia. *PLoS One, 12*(10), e0186259. doi:10.1371/journal.pone.0186259

9.Wilson, A. (2009). Flora of Australia, Volume 44A: Poaceae 2. *Flora of Australia, Volume 44A: Poaceae 2*.

10. Milesi, C., Running, S. W., Elvidge, C. D., Dietz, J. B., Tuttle, B. T., & Nemani, R. R. (2005). Mapping and modeling the biogeochemical cycling of turf grasses in the United States. *Environ Manage, 36*(3), 426-438. doi:10.1007/s00267-004-0316-2

11. Dong, Z., Guo, X., Zhang, X., Qiu, F., Sun, L., Gong, H., & Zhang, F. (2007). Dyadobacter beijingensis sp. nov., isolated from the rhizosphere of turf grasses in China. *Int J Syst Evol Microbiol, 57*(Pt 4), 862-865. doi:10.1099/ijs.0.64754-0

12.Longo, C., Lickwar, C., Hu, Q., Nelson-Vasilchik, K., Viola, D., Hague, J., . . . Kausch, A. P. (2006). Turf Grasses. *Methods Mol Biol, 344*, 83-95. doi:10.1385/1-59745-131-2:83

13. Jinbo, U., Kato, T., & Ito, M. (2011). Current progress in DNA barcoding and future implications for entomology. *Entomological Science, 14*(2), 107-124.

14.Li, F.-W., Kuo, L.-Y., Rothfels, C. J., Ebihara, A., Chiou, W.-L., Windham, M. D., & Pryer, K. M. (2011). RbcL and matK earn two thumbs up as the core DNA barcode for ferns. *PLoS One, 6*(10).

15. Group, C. P. W., Hollingsworth, P. M., Forrest, L. L., Spouge, J. L., Hajibabaei, M., Ratnasingham, S., . . . Erickson, D. L. (2009). A DNA barcode for land plants. *Proceedings of the National Academy of Sciences, 106*(31), 12794-12797.

16.Singh, J., Gupta, S. K., Devanna, B. N., Singh, S., Upadhyay, A., & Sharma, T. R. (2020). Blast resistance gene Pi54 over-

expressed in rice to understand its cellular and sub-cellular localization and response to different pathogens. *Sci Rep, 10*(1), 5243. doi:10.1038/s41598-020-59027-x

17.Khanuja, S. P., Shasany, A. K., Darokar, M. P., & Kumar, S. (1999). Rapid isolation of DNA from dry and fresh samples of plants producing large amounts of secondary metabolites and essential oils. *Plant molecular biology Reporter, 17*(1), 74-74.

18.Kumar, P., Gupta, V., Misra, A., Modi, D., & Pandey, B. (2009). Potential of molecular markers in plant biotechnology. *Plant Omics, 2*(4), 141.

19.Waugh, J. (2007). DNA barcoding in animal species: progress, potential and pitfalls. *BioEssays, 29*(2), 188-197.

20.Wang, D., & Karamyshev, A. L. (2020). Correction to: Next Generation Sequencing (NGS) Application in Multiparameter Gene Expression Analysis. *Methods Mol Biol, 2102*, C1. doi:10.1007/978-1-0716-0223-2_31

21.Sun, L., Zhu, B., Zhang, X., Wang, H., Yan, A., Zhang, G., . . . Xu, H. (2020). The accumulation profiles of terpene metabolites in three Muscat table grape cultivars through HS-SPME-GCMS. *Sci Data, 7*(1), 5. doi:10.1038/s41597-019-0321-1