Int.J.Curr.Microbiol.App.Sci (2017) *6*(5): 2746-2756

International Journal of Current Microbiology and Applied Sciences
ISSN: 2319-7706 **Volume 6 Number 5 (2017) pp. 2746-2756**
Journal homepage: http://www.ijcmas.com

Original Research Article https://doi.org/10.20546/ijcmas.2017.605.308

Evaluation of Potential DNA Barcoding Loci from Plastid Genome: Intraspecies Discrimination in Rice (*Oryza species*)

Jyoti Singh*, Datta P. Kakade, Mayur R. Wallalwar, Rishiraj Raghuvanshi, Miranda Kongbrailatpam, Satish B. Verulkar and Shubha Banerjee

Department of Plant Molecular Biology and Biotechnology, College of Agriculture, Indira Gandhi Krishi Vishwavidyalaya, Raipur, Chhattisgarh-492012, India
**Corresponding author*

Keywords

Rice, DNA Barcoding, Chloroplast 'DNA, rbcl, 'matK, Species.

Article Info

Accepted:
26 April 2017
Available Online:
10 May 2017

A B S T R A C T

DNA barcoding is a technique that makes use of short sequences from a standardized region of a genome to provide quick and reliable identification of species among all forms of life. The presence of uniqueness and variability required for DNA barcoding is well reported in animal system based on mitochondrial gene CO1.On the other hand, limited information is available on universal barcode for plants. Candidate loci belonging to chloroplast genome (CpG) and nuclear genome have been analyzed in various plants to identify universal barcoding loci capable of inter and intraspecies discrimination. In this study, relative potential of 24 candidate loci (Dong *et al*, 2012) from plastid genome were validated on set of 231 diverse rice genotypes, for selection of suitable barcoding loci for DNA barcoding in rice. Results indicated that only one of the chloroplast CGS primer pair "*psbA-trnH*" showed (100%) amplification efficiency followed by "*rbcL*" (89.61%), "*atpH-atpl*" (68.39%), "*matK*" (66.2%) and "*petA-psbJ*" (62.33%). While 9 primers showed lower amplification efficiency between 5.19% and 52.81%. Based on amplification efficiency, reproducibility and amplicon size (as per Consortium for the Barcode of Life standard) five primers were selected for amplicon sequencing and further study of phylogenetic and phylogeographical relationships among above genotypes.

Introduction

Rice is a global food crop as well as current medium of economic support for millions of peoples and that's why for half of the humanity "rice is life." Two major subspecies of cultivated rice, *indica* and *japonica*, are the products of separate domestication events from the ancestral species, *O. rufipogon*, an assumption initially based on studies of biochemical traits (Second *et al.,* 1982). Geographically or ecologically diverse groups of rice are expected to show greater genetic differentiation as rice is predominantly autogamous and hence, gene flow is limited than would be the case in an outcrossing species. Because of this a greater proportion of diversity is expected to exist in terms of variation between homozygous lines within a heterogeneous landrace in rice (Olufowote *et al.*, 1997).

Chhattisgarh is traditionally rich in rice diversity containing the wild progenitors of cultivated rice. An organized collection of rice germplasm from Madhya Pradesh

including Chhattisgarh during 1972 to 1981 collected a total of 18,541 accessions of rice. Chhattisgarh is prominent for rice diversity and considered as one of the secondary centre of diversity. Further explorations in association with NBPGR, New Delhi were organized and new collections were added to the gene pool which at present has 23,500 accessions including 210 accessions of wild species. This germplasm has only moderately been characterized for various biotic and abiotic stress tolerances (Pandey *et al.,* 2010). In order to understand the genetic variability and study phylogenetic variation in rice germplasm belongs to CG a representative samples of 231 rice genotypes were selected based on their morphological and physiological characters (Table 1).

Genetic diversity serves as an insurance against selection pressure a few crop failures. Earlier studies of variation are based on morphological character however; present studies focus on molecular level that are primarily based on the changes of DNA sequences among populations of a species and higher taxa (Hamby and Zimmer, 1992). A diverse array of molecular techniques is available for studying genetic variability. For future crop improvement conservation and cataloguing of genetic diversity is very essential to explore genetic potential of plants and their wild relatives. Collection and characterization of existing germplasm is not only important for utilizing the appropriate attribute, in breeding programmes, but is also essential for protecting the unique identification of a genotype worldwide. Thus scientific community today is concerned on genetic variability of organisms located at various sites of life. This has advanced greatly in the last decade with the development of the molecular biology techniques (Soltis *et al.,* 1998; Hollingsworth *et al.,* 1999; Wen and Pandey, 2005; Mondini *et al.,* 2009). In accordance to these views and to study the phylogenetic as well as phylogeographic discrimination of the rice genotypes belonging to the geographical area of Chhattisgarh, the DNA barcoding of its unique germplasm is a promising tool. Next generation sequencing is a high throughput technique which is adopted for "DNA barcoding" with aim to develop an inexpensive, fast and standardized method for species identification that is accessible to the other non taxonomists'.

At present the techniques for studying the molecular phylogeny of plants depends mainly on chloroplast genome sequence data. The reason behind this is that the chloroplast genome has a simple and stable genetic structure, it is haploid there are no (or very rare) recombination, it is generally uniparental transferred. Along with these ease PCR amplification and sequencing of chloroplast genes. The short, variable and standardized DNA sequence can be termed as DNA barcode when it mirrors the distributions of intra-and interspecific variabilities separated by a distance called 'DNA barcoding gap' and characterizes conserved flanking regions for development of universal primers across highly divergent taxa (Kress *et al.,* 2005; Savolainen *et al.,* 2005; Hollingsworth *et al.,* 2009). As *Oryza sativa* is a dominant cultivated rice species its complete Chloroplast genome sequences are available with the availability of existing data our aim was to generate information that allows the identification of most variable chloroplast genome in rice. Due to the high level of conservation, analysis of the chloroplast genome has become a valuable tool for plant phylogenetic studies (Waters *et al.,* 2012; Yang *et al.,* 2013). Entire chloroplast genome analysis provides high-resolution plant phylogenies (Parks *et al.,* 2009). Earlier, only a few chloroplast markers have been applied in studies of plant diversity and evolution (Schroeder *et al.,* 2011)

DNA Barcodig Contains sufficient variation to discriminate between higher plant species based on conserved flanks for universal primers of land plants. For precise application in DNA barcoding the loci must be more than 500bp in length and highly reproducible the technique has been used in degraded samples which make DNA BARCODING rapid, accurate and automatable species identification technique by using standardized DNA sequences as tags (Hebert *et al.,* 2003). The mitochondrial cytochrome *c oxidase1* (*cox1*) gene has been used as a universal barcode in animal. Due to low rate of nucleotide substitution in plant mitochondrial genomes preclude the use of *CO1* as a universal plant barcode (Fazekas *et al.,* 2008). As *CO1* was not useful in plants, many loci have been proposed as plant barcodes, including *ITS* (Chase *et al.,* 2009), *rbcL* (Kress and Erickson, 2007), *psbA-trnH* and *matK* (Chase *et al.,* 2009). The identification of high resolution DNA barcodes at species level is critical. The third International Barcode of Life Conference (CBOL, 2009) concluded with a remark that *matK* and *rbcL* are sequences as the universal barcode sequence, for land plants.

In spite of this useful recommendation, both the identification and the combination of the most appropriate regions for plant DNA barcoding remain debatable (Bruni *et al.,* 2010). Since 24 regions of chloroplast genome like *psbA-trnH*, *rbcL*, *atpH-atpl*, *petA-psbJ*, *ndhA-ndhA, trnK-trnK, petB-petD, ndhC-trnV*, *trnS1-trmG1*, *trnW-psaJ*, *clpP-clpP*, *trnT-psbD, rbcL-accD, accD-psal, ndhF, petN-psbM, psbM-trnD, psbE-petL, Rpl32-trnL, rpoB-trnC, rps16-trnQ, trnH-psbA, trnS2-trnG2* and, *matK* were used in our study as all this used for development of candidate markers in plant DNA barcoding (Dong *et al.,* 2012). Based on the information the aim of present study is to evaluate the performance of different barcoding loci and efficiency for discrimination of the different species and cultivars, also tried to find out highly informative primers designed from chloroplast genomes on the basis of PCR amplification efficiency in *Oryza sativa L*. As a result, such regions resolve phylogenies and for DNA barcoding intraspecies of (*Oryza sativa L.)*.A set of 24 primer pair were validated on set of 231 (table 1) diverse rice genotypes.

Materials and Methods

The experimental materials consisted of 231 diverse rice genotypes including germplsm lines, elite, varieties and wild rice, which were taken from the rice germplasm collection I.G.K.V, Raipur (Table 1.) DNA was extracted from leaf tissue from individual plant from each accessions genomic DNA was extracted using MiniPrep method (Doyle and Doyle, 1987).The concentration and quality of the extracted DNA were determined using gel electrophoresis and a Nano Drop spectrophotometer (Thermo scientific 30304-Ace-600).The isolated genomic DNA was stored at -20^0 C until used. A total volume of 20 µl of PCR reaction mixture contained the following: 2 µl (50 ng /µl) DNA, 2µl 10mM dNTPs mix (Invitrogen), 2µl of 10X PCR buffer with 15mM MgCl2 (Invitrogen), 2µl of 10 pMo primer (1µl of each forward and reverse primer), 0.1µl of Taq DNA poly 5U/µl (Invitrogen) and rest was adjusted with nuclease free water (Sigma Aldrich). The 24 primer pairs (Table 2.) were used for the PCR (Imperial Life Sciences). The PCR was done Veriti 96-Well Thermal Cycler (Applied Biosystems) as follows: 94^0 C for 4 min, followed by 35 cycles of 94^0 C for 30 s, 50^0 C-65^0 C for 30s, and 72^0 C for 1 min, followed by an elongation step at 72^0 C for 7 min. A long (Horizontal electrophoresis unit Max Fill) 1.5% horizontal agarose gel using 1X TAE buffer containing 0.5ul/mL ethidium

bromide was used for resolving PCR. Gel images were documented using a (Gel Doc XR^{+} BIORAD ET9970616AA) UV transilluminator opticom imaging system. The PCR product sizes were determined using a 100-bp ladder. PCR products were purified using (Thermo Scientific Gene JET Gel Extraction Kit) as per manufacturing instruction.

Results and Discussion

The search for appropriate DNA barcoding locus for plants are most important issue for practical use of the technique in modern years, hence studies on evolution /comparison of DNA barcodes are extremely important. On the basis of recent development, it is admitted that the barcode databases will grow rapidly. Consequently, the International Nucleotide Sequences Database (INSD: GenBank European Molecular Biology Laboratory (EMBL) and DNA data bank of japan (DDBJ) has adopted a unique keyword identifier (BARCODE) to recognize standard barcode sequences specified by the scientific community. Mainly plants posses three genomes i.e. nucleus, chloroplast and mitochondria, Chloroplast DNA (cpDNA) possesses the most ideal DNA sequence for phylogenetic analysis. The reason behind this is they are relatively easy to purify, characterize, clone and sequence (Cleg *et al.,* 1990) also endemic to plants. Thus Chloroplast DNA barcodes avoid the DNA contamination from other organisms without chloroplasts, such as animals and fungi. The chloroplast genome sequence of rice Nipponbare (*O. sativa L.ssp. japonica*) was reported to have a length of 134,525 bp (Hiratsuka *et al.,* 1989). Chloroplasts restrain both highly conserved genes important to plant life and more variable regions, which have been informative over broad time scales. Relative studies of the genomic structural design showed that the order of genes and the contents of essential genes are highly conserved among most chloroplast genomes (De Las Rivas *et al.,* 2002). Nevertheless, variations between different and closely related genomes have occurred during evolution (Tang *et al.,* 2004).

Hollingsworth *et al,* 2011 proposed the seven candidate plastid region *rpoB, rpoC, matK, rbcl, atpF-atpH and psbI and trnH-psbA* for groups of land plants. Similarly Chase *et al,* 2007 proposed to make universal barcodes with combination of *matK+rpoc1+rpoB* and *matK+rpoc1+trnH& psbA* out of which combination of *rbcl+matK* has been suggested for the terrestrial plants as the main barcode (CBOL, 2009), although (Dong *et al.,* 2012) scanned entire chloroplast genomes of 12 genera to explore for extremely variable region. In view of that the suggested primers by various scientists on basis of their study, we also tried to amplify and sequencing of highly variable loci in our study in order to find out and validate most variable loci in rice (*Oryza sativa L.*)

A set of 24 primers were used for PCR amplification of 231 genotypes to find out the highly informative primers to validate specific region of the chloroplast genome of rice for barcoding. Primers from 24 selected region on 231 diverse rice genotypes including germplasm lines, elite, varieties and wild rice (Table 1). Our results indicate that (*psbA-trnH*) showed 100% amplification in 231 genotypes. Kress *et al.,* 2010 also recommended the (trnH-psbA) plastid intergenic spacer region could become and appropriate candidate as universal barcode for land plants which seems to be ideal confirmation as the primer pairs validate 100% amplification efficiency in rice (*Oryza sativa L*). Followed by (rbcL) with 89.61% amplification efficiency amplified in 219 rice genotypes. The rbcl gene among various loci of plastids reported as most well characterized

gene and is sufficiently reported for the recovery of bidirectional sequences of high quality.

The rbcl gene that codes for "RUBISCO", ribulose 1, 5-biophosphate-carboxylase/ oxygenase a free enzyme present in stroma in the single copy region of chloroplast genome and the coding region is separated by intergenic spacer (600-800) nucleotide (Savolainen *et al.,* 2000) also considered as integral component for species discrimination (Janzen *et al.,* 2009). The rbcl based DNA barcoding also seems to be efficient to resolve the issues on taxonomic confusion on the familia and higher levels and also on lower (inter/Intra generic) levels lived in cupressaceae, Cornaceae, Ericaceae, Graniaceae (Gille *et al.,* 1994). While 10 chloroplast genome specific primer pairs showed efficiency ranges from 5.19% to 68.39 %. Primer pair showed *rbcL-accD* 5.19%, *trnT-psbD* 9.09%, *clpP-clpP* 19.48%, *trnW-psaJ* 32.03%, *trnS1-trmG1* 33.33%, *ndhC-trnV* 33.33%, *petB-petD* 34.60%, *trnK-trnK* 41.90%, *ndhA-ndhA* 52.81%, *petA-psbJ* 62.33%, *matK* 66.2%, *atpH-atpl* 68.39%, amplification efficiency.

Like *rbcl, matK* is another widely used barcode for plants is another cpDNA gene region which codes for maturase of higher plants while the *matK* exon being located within the trnk intron (Ems *et al.,* 1999). Among the most preferred choice *matK* is also included for systemic studies for higher plants as contains greater number of non-synonymous mutations, indels (insertions and deletions) and nucleotide substation (Olmstead *et al.,* 1994 and Hilu *et al.,* 1997). The other ten primer pairs of our panel did not amplified at all in any of the rice genotypes shows no amplification efficiency they are *accD-psal, ndhF, petN-psbM, psbM-trnD, psbE-petL, Rpl32-trnL, rpoB-trnC, rps16-trnQ, trnH-psbA, trnS2-trnG2* (Table 3).

Dong *et al.,* (2012) reported in his work that while testing the twenty-three most variable regions in chloroplast genomes of 12 genera with two or more species. Genus consists of *Acorus, Aethionema, Calycanthus, Chimonanthus, Eucalyptus, Gossypium, Nicotiana, Oenothera, Oryza, Paeonia, Populus, Solanum* primer *accD-psaI* shows no fragment length, π value (nucleotide diversity per site) and number of indels and inversions are also not obtained. Most accepted reason behind these will be rapidly evolving regions of the chloroplast genome; evolutionary events that occur include the formation of secondary structures, multiple-hit sites, and intra-molecular recombination actions. These troubles seem less serious in phylogenetic analyses of closely related species. However, aim is to accurately solve phylogenetic relationships by using the loci identified by various study may not always be achieved because of other problems. Some authors (Borsch and Quandt, 2009) have speculated that intraspecific inversions might be problematic for barcoding, but did not test this assumption with empirical data. Prior to this paper, intraspecific inversions have rarely been reported but are not unknown. In accordance to the result obtained in present study, (Kress *et al.,* 2007) compared regions *atpB-rbcL, ITS, psbM-trnD, trnC-ycf6, trnH-psbA, trnL-F,trnk-rps16, trnV-atpE, rpL36-rps8, ycf6-psbM* sampling strategy applied by them is they used 19 individuals,19 species from 7 angiosperm families they reported the universality percentage success *trnH-psbA, rpl136-rpf8,,trnL-F=100%, trnC-ycf6, ycf6-psbM=90%*. Other regions shows 73-80% sequence divergence *ITS* (2.81%), *trnH-psbA* (1.24) thus for barcode region recommendation by them was *ITS* and *trnH-psbA.*

Moreover chloroplast genome phylogenetic analysis revealed that the *Oryza nivara* is closed to *O. sativa L*. spp. indica and the *O.*

sativa L. spp. japonica is closed to *Oryza rufipogon* in Asian cultivated and wild rice (Brozynska *et al.*, 2014) and the African rice (*Oryza glaberrima* and *Oryza barthii*) were cluster together but in separate group with the Asian rice (Wambugu *et al.*, 2015).

In the present scenario it become possible to overcome from conventional sequencing of plant chloroplast genomes to next generation sequencing (NGS), it has become progressively more feasible to examine the entire genome of the chloroplast, rather than targeting individual regions (Nock *et al.*, 2011; Straub *et al.*, 2012). However, the chloroplast genome only represents the maternal evolutionary history. In addition, it also cannot be fully applied to rapidly diverging taxa, as the chloroplast has a slow rate of evolution (Parks *et al.*, 2009).

Table.1 Detail of 231 accessions used for validations

Wild rice	Variety	Landraces	Advance breeding lines	Germplasm
WR3, WR41, WR99,	Annada,ARB8,Abhaya,ARB6,Bamleshwari,CT9993,IR36,MTU1010, Punjab Bas3,IR64,Kranti,Mahamaya,Samleshwari,Swarna,Swarna sub1,Vandana, IBD-1,Danteshwari,Poornima, ,Badshahbhog, Aganni, Karma masuri, Safri 17, Dubraj, , Jitpiti, Durgeshwari, Shymala, Rajeshwari, Chandrahasini, Indira sugndhit dhan-1, Elayachi, Jeeradhan, Nagina-22, Tarunbhog, CHIR-8, CGZR-1, Basmati 370, Basmati 1, IR64, Swarna, IBD1	Buddha,Bakal,Bhataphool,Batro, Bhatajhooli,Deshi lal Dhan, DeshiNo.17,DagadDeshi, Lalmati,Laloo14,BotkiGurmatia (2728), PRATAO, Chuva Dau 130, DJOGOLON-DJOGOL , Azucena Bhansapanchi, Banda, Bada gada khuta, Reg-695, GP-145-40, RKVY-104, RKVY -211, Dular, BAM 1292, BAM 5446, BAM 5926, Moroberekan, BAM 5997, Kalanamak, GP-145-37, SL 62, GP-145-41, , CHAU DAU,Karigilas, Azucina, Azucina , Mikhudeb, Moshur, Moshur, Binuhangin, Dangar, Dhala Shaita, Gul Murali, Jabor Sail, Moyna Moti, Uri, ARC 10376, Dharia Boalia, Aus 257,Chengri 2, Juma, Koi Murali,Ramjiyawan, Shennong-89366, E-1701, E-1702, E-1703, E-1827, E-2010, E-2312, E-2367, M:4628, E-1857, E-2526, M-114, M-184, M-1051, M-1433, Sehra dabri, chitrakot, Reg-1035, Reg-1038, IR74371-70-1-1,IR 83381-B-B-55-4, GP-145-66,GP-145-66, RKVY-77, GP-145-103, GP-145-78, GP-145-43, G1, G5, G8, G21, GP-145-43, GP-145-59, GP-145-136, GP-145-50, GP-145-65, GP-145-114, , GP-145-11, G108, GP-145-20, G114, GP-145-34, G127, GP-145-5,GP-145-138	IR 62266, IC-267982,IR42253, IR 84984-17-83-48-1-BSahabhagi Dhan, IR84984-83-15-862-B, IR 90019-17-159-B, IR 90019-22-28-2-B, B-6, R-RF-78, IR 55419-04, IR 86931-B-400, IR 86918-B-305, IR 87728-75-B-B, IR 87728-367-B-B,IR 84984-83-15-110-B, CR 5272, EPAGRI-2, PINKAEO, RYT 3275, PR 122, SLO-16, Kalia, AVT-1-IME-3, R1570, AVT-2 ASG-5,BPT 204(Improved),BPT5204(Improved),AVT-2-IME-10, AVT-2-E-TP-6, AVT-1-ASG, R-RHZ-LI-23, R-RHZ-IB-13, R-RHZ-SM-14, R-RHZ-MI-30, R-56, RR-100, A-GM-AS-45, GP-145-42, G21, G23, G42, G47, G69, G93,G100,G102,Cross116,RGMATN47,IR55419,KALOKUCHI,G132,G134,Kalamkati,G136,G158,G173,G186, G194, G196, G198, G200, G203, R-RF-69, ARC10955,R-RF-75, RR-152, RR -137, RR-149, RR-8 M011, G104,	Gurmatia(2676),Gurmatia (3053),Bangla Gurmatia (2711), Sultu Gurmatia (2788), Bisni-I, CHAPTI GURMATIA,Chepti Gurmatia (3011),JhunkiGurmatia (2739), Kalam Nunki Gurmatia (2784), Sultu Gurmatia (2788), Srikamal, Jhilli IET 23829, Kadamphool

Table.2 Primer used for amplifying and /or sequencing 24 highly informative loci (source; Dong *et al.*, 2012 and Holligsworth *et al.*, 2011)

	Forward primer			Reverse primer	
	Locus	Name	Sequence 5'to 3'	Name	Sequence 5'to 3'
1	*rbcL-accD*	rbcL-f	tagctgctgcttgtgaggtatgga	accD-r	aaatactaggcccactaaagg
2	*accD-psaI*	accD-f	ggtaaaagagtaattgaacaaac	psaI-r	ggaaatactaagcccactaaaggcaca
3	*atpH-atpI*	atpH-f	aacaaaaggattcgcaaataaaag	atpI-r	agttgttgttcttgtttctttagt
4	*clpP*	clpP-f	gcttgggcttctcttgctgacat	clpP-r	tcctaatcaaccgactttatcgag
5	*ndhA*	ndhA-f	tcaactatatcaactgtacttgaac	ndhA-r	cgagctgctgctcaatcgat
6	*ndhC-trnV*	ndhC-f	agaccattccaatgccccctttcgcc	trnV-r	gttcgagtccgtatagcccta
7	*ndhF*	ndhF-f	acaccaacgccattcgtaatgccatc	ndhF-r	aagatgaaattcttaatgatagttgg
8	*petA-psbJ*	petA-f	ggatttggtcagggagatgc	psbJ-r	atggccgatactactggaagg
9	*petN-psbM*	petN-f	atggatatagtaagtctcgcttgg	psbM-r	atggaagtaaatattcttgcat
10	*psbM-trnD*	psbM-f	tttgactgactgtttttacgta	trnD-r	cagagcaccgccctgtcaag
11	*petB-petD*	petB-f	caatccactttgactcgtttt	petD-r	ggttcaccaatcattgatggttc
12	*psbE-petL*	psbE-f	atctactaaattcatcgagttgttcc	petL-r	tatcttgctcagaccaataaataga
13	*rpl32-trnL*	rpl32-f	gcgtattcgtaaaaatatttggaa	trnL-r	ttcctaagagcagcgtgtctacc
14	*rpoB-trnC*	rpoB-f	acaaaatccttcaaattgtatctga	trnC-r	tttgttaatcaggcgacacccgg
15	*rps16-trnQ*	rps16-f	tttatcggatcataaaaacccact	trnQ-r	tggggcgtggccaagcggt
16	*trnT-psbD*	trnT-f	gccctttaactcagtggtagag	psbD-r	ccaaataggaactggccaatc
17	*trnH-psbA*	trnH-f	cgcgcatggtggattcacaaatc	psbA-r	tgcatggttccttggtaacttc
18	*trnK*	trnK-f	gggactcgaacccggaacta	trnK-r	agtactcggcttttaagtgcg
19	*trnW-psaJ*	trnW-f	tctaccgaactgaactaagagcgc	psaJ-r	cgattaatctctatcaatagacctgc
20	$trnS^{GCU}$-$trnG^{GCC}$	trnS1-f	aacggattagcaatccgacgcttta	trnG1-r	cttttaccactaaactatacccgc
21	$trnS^{UGA}$-$trnG^{UCC}$	trnS2-f	cggttttcaagaccggagctatcaa	trnG2-r	cataaccttgaggtcacgggttcaaat
22	*rbcL*	rbcL-f	atgtcaccacaaacagaaac	rbcl-r	tcgcatgtacctgcagtagc
23	*matK*	matK-f	cgatctattcattcaatatttc	matK-r	tctagcacacgaaagtcgaagt
24	*psbA-trnH*	*psbA-trnHF*	Gttatgcatgaacgtaatgctc	*psbA-trnHR*	cgcgcatggtggattcacaattc

Fig.1 Amplification profiles of the chloroplast genomic loci; (a) psbA-trnHF (100%) (b) primer atpH-atpL (68.39%) amplification profile; (c) primer trnW-psaJ (32.03%)

.

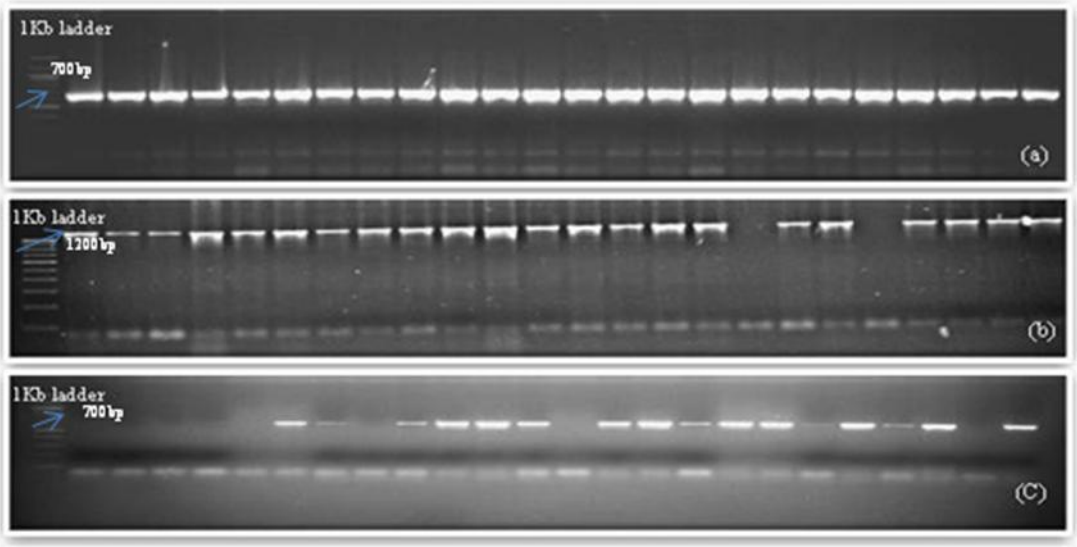
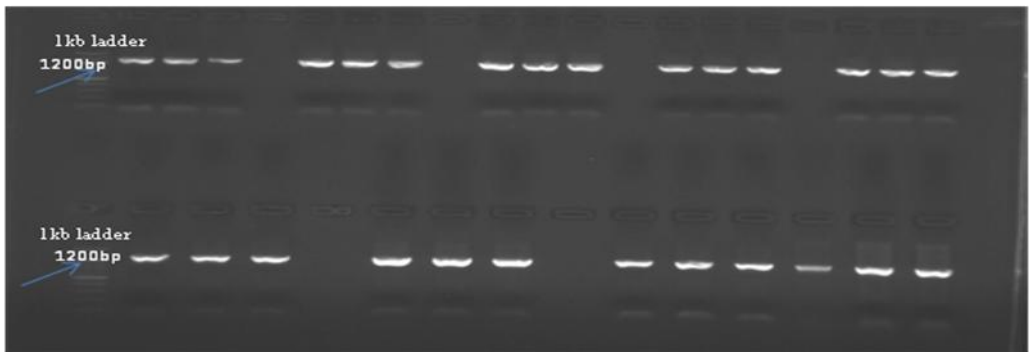


Table.3 Amplification efficiency of 24 chloroplast specific marker in 231 rice genotypes

Sr. No.	Primer	Monomorphic	Polymorphic	No amplification	Amplification efficiency (%)	Amplicon size (bp)
1	rbcL-f/accD-r	12	0	219	5.19	800
2	accD-f/psal-r	0	0	231	NA	0
3	atpH-f/atpl-r	158	0	73	68.39	1200
4	clpP-f/clpP-r	0	45	186	19.48	800
5	ndhA-f/ndhA-r	122	0	109	52.81	1200
6	ndhC-f/trnV-r	77	0	154	33.33	1200
7	ndhF-f/ndhF-r	0	0	231	NA	0
8	petA-f/psbJ-r	144	0	87	62.33	1200
9	petN-f/psbM-r	0	0	231	NA	0
10	psbM-f/trnD-r	0	0	231	NA	0
11	petB-f/petD-r	80	0	151	34.6	1200
12	psbE-f/petL-r	0	0	NA	NA	0
13	Rpl32-f/trnL-r	0	0	NA	NA	0
14	rpoB-f/trnC-r	0	0	NA	NA	0
15	rps16-f/trnQ-r	0	0	NA	NA	0
16	trnT-f/psbD-r	21	0	217	9.09	700
17	trnH-f/psbA-r	0	0	NA	NA	0
18	trnK-f/trnK-r	46	51	134	41.9	1200
19	trnW-f/psaJ-r	74	0	157	32.03	1200
20	trnS1-f/trmG1-r	64	13	154	33.33	800
21	trnS2-f/trnG2-r	0	0	NA	NA	0
22	rbcl-f/rbcl-r	207	0	24	89.6	1200
23	matK-f/matK-r	153	0	78	66.2	800
24	psbAtrnHF/psbA-trnHR	231	0	231	100	700

Fig.2 Gel image of fragments (atpH-atpL) primer pairs purified and sent for sequencing



As a result, chloroplast-based evolutionary studies must sometimes be complemented by nuclear genomic information. Closer evolutionary relationships between *indica* and *aus* strains were observed using both nuclear and chloroplast genome data, as well as among the tropical *japonica*, temperate *japonica*, and aromatic groups (Garris *et al.*, 2005). The *indica* subpopulation was shown to contain the highest degree of chloroplast

diversity (Garris *et al.,* 2005). Kim *et al.,* 2014 evaluated 67 improved varieties and 13 landraces from the Democratic People's Republic of Korea (DPRK) at both nuclear and chloroplast levels, and they found a temperate japonica subgroup that was less diverse than the *indica* ancestor group at the nuclear level but more diverse at the chloroplast level (Kim *et al.,* 2014).

The outcome of our study indicated that more standardization of universal primers is required to improve amplification efficiency and to get of higher number of informative loci. Further designing of new primers from the specific site of rice chloroplast genome will help in precise amplification of reproducible chloroplast genome specific loci. As more loci will be identified and validated using sequencing informative data for analyzing intra species variation in rice will be achievable. This will further strengthen barcoding of local rice genotype of various regions of India like Chhattisgarh and across the world. The normally used method for classifying DNA sequence is likely to be based on distance. Primers which show amplification efficiency were sequenced and analysis for species discrimination is ongoing.

References

Altschul, S., Madden, T., Schaffer, A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. 1997. Gapped BLAST and PSIBLAST: A new generation of protein database search programs. *Nucleic Acids*, 25: 3389-3402.

Borsch, T. and Quandt, D. 2009. Mutational dynamics and phylogenetic utility of noncoding chloroplast DNA. *Plant Systematics and Evolution*, 282: 169-199.

Brozynska, M., Furtado, A. and Henry, R.J. 2014. Direct chloroplast sequencing: comparison of sequencing platforms and analysis tools for whole chloroplast barcoding. *PLoS ONE,* 9: e110387.

Bruni, I., De Mattia, F., Galimberti, A., Galasso, G., Banfi, E., Casiraghi, M. and Labra, M. 2010. Identification of poisonous plants by DNA barcoding approach. *Int. J. Legal Med.,* 124: 595-603.

Chase, M.W. and Fay, M.F. 2009. Barcoding of plants and fungi. *Sci.,* 325: 682-683.

Clegg, M.T., Learn, G.H. and Golenberg, E.M. 1991. Molecular evolution of chloroplast DNA. Evolution at the molecular level/edited by Robert K. Selander, Andrew G. Clark, and Thomas S. Whittman.

De Las Rivas, J., Lozano, J.J. and Ortiz, A.R. 2002. Comparative analysis of chloroplast genomes: functional annotation, genome-based phylogeny, and deduced evolutionary patterns. *Genome Res.,* 12: 567-583.

Dong, W., Liu, J., Yu, J., Wang, L. and Zhou, S. 2012. Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding. *PLoS ONE,* 7: e35071.

Doyle, J.J. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.,* 19: 11-15.

Ems, S.C., Morden, C.W., Dixon, C.K., Wolfe, K.H., de Pamphili,s C W and Palmer J D. 1995. Transcription, splicing and editing of plastid RNAs in the nonphotosynthetic plant Epifagus virginiana. *Plant Molecular Biol.,* 29: 721-733.

Fazekas, A.J., Kesanakurti, P.R., Burgess, K.S., Percy, D.M., Graham, S.W., Barrett, S.C., Newmaster, S.G, Hajibabaei M and Husband, B.C. 2009. Are plant species inherently harder to discriminate than animal species using DNA barcoding markers? *Molecular Ecol. Res.,* 9: 130-139.

Garris, A.J., McCouch, R. and Kresovich, S. 2003. Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the xa5 locus of rice (Oryza sativa L.). *Genetics,* 165: 759-769.

Gielly, L. and Taberlet, P. 1994. The use of chloroplast DNA to resolve plant

phylogenies: noncoding versus rbcL sequences. *Molecular Biol. Evol.*, 11: 769-777.

Group, C.P.W., Hollingsworth, P.M., Forrest, L.L, Spouge, J.L., Hajibabaei, M., Ratnasingham, S., van der Bank M, Chase M.W, Cowan, R.S. and Erickson, D.L. 2009. A DNA barcode for land plants. *Proceedings of the National Academy of Sciences*, 106: 12794-12797.

Hamby, R.K. and Zimmer, E.A. 1992 Ribosomal RNA as a phylogenetic tool in plant systematics. In Molecular systematics of plants. pp 50-91. Springer.

Hebert, P.D., Cywinska, A. and Ball, S.L. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London B: Biological Sciences* 270: 313-321.

Hilu, K. and Liang, H. 1997. The matK gene: sequence variation and application in plant systematics. *American J. Botany*, 84: 830-830.

Hiratsuka, J., Shimada, H., Whittier, R., Ishibashi, T., Sakamoto, M., Mori M, Kondo C, Honji Y, Sun C-R and Meng B-Y. 1989. The complete sequence of the rice (Oryza sativa) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Mol. General Genetics MGG*, 217: 185-194.

Hollingsworth, M.L., Andra Clark A., Forrest, L.L., Richardson, J., Pennington, R., Long D.G., Cowan R., Chase. M.W, Gaudeul M and Hollingsworth, P.M. 2009. Selecting barcoding loci for plants: evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants. *Molecular Ecol. Res.*, 9: 439-457.

Janzen, D.H., Hallwachs, W., Blandin, P., Burns, J.M., Cadiou, J., Chacon, I., Dapkey T, Deans A R, Epstein M E and Espinoza B. 2009. Integration of DNA barcoding into an ongoing inventory of complex tropical biodiversity. *Molecular Ecol. Res.*, 9: 1-26.

Kim, H.M., Oh, S.H., Bhandari, G.S., Kim, C.S. and Park, C.W. 2014. DNA barcoding of Orchidaceae in Korea. *Molecular Ecol. Res.*, 14: 499-507.

Kress, W.J. and Erickson, D.L. 2007. A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. *PLoS ONE*, 2: e508.

Kress, W.J. and Erickson, D.L. 2007. A two-locus global DNA barcode for land plants: the coding rbcL gene complements the non-coding trnH-psbA spacer region. *PLoS ONE*, 2: e508.

Kress, W.J., Erickson, D.L., Swenson, N.G., Thompson, J, Uriarte M and Zimmerman J K. 2010. Advances in the use of DNA barcodes to build a community phylogeny for tropical trees in a Puerto Rican forest dynamics plot. *PLoS ONE*, 5: e15409.

Mondini, L., Noorani, A. and Pagnotta, M.A. 2009. Assessing plant genetic diversity by molecular tools. *Diversity*, 1: 19-35.

Nock, C.J., Waters, D.L., Edwards, M..A, Bowen, S.G., Rice, N, Cordeiro G M and Henry, R.J. 2011. Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnol. J.*, 9: 328-333.

Olmstead, R.G. and Palmer, J.D. 1994. Chloroplast DNA systematics: a review of methods and data analysis. *American J. Botany*, 1205-1224.

Olufowote, J.O., Xu, Y., Chen, X., Goto, M., McCouch, S.R., Park, W.D., Beachel,l H .M and Dilday, R.H. 1997. Comparative evaluation of within-cultivar variation of rice (Oryza sativa L.) using microsatellite and RFLP markers. *Genome*, 40: 370-378.

Pandey, M., Verulkar, S. and Sarawgi, A. 2010. Status paper on rice for Chhattisgarh. *Rice Knowledge Management Portal*, 13-14.

Parks, M., Cronn, R. and Liston, A. 2009. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC*

Biol., 7: 84.

Parks, M., Cronn, R. and Liston, A. 2009. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biol.,* 7: 84.

Savolainen, V. and Hollingsworth, P. 2000. Molecular Systematics and Plant Evolution. (The Systematics Association, Special Vol. 57). JSTOR.

Savolainen, V., Cowan, R.S., Vogler, A.P., Roderick, G.K. and Lane, R. 2005. Towards writing the encyclopaedia of life: an introduction to DNA barcoding. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 360: 1805-1811.

Schroeder, H., Höltken, A. and Fladung, M. 2011 Chloroplast SNP-marker as powerful tool for differentiation of Populus species in reliable poplar breeding and barcoding approaches. In *BMC Proc*., p. P56.

Second, G. 1982. Origin of the genic diversity of cultivated rice (Oryza spp.): study of the polymorphism scored at 40 isozyme loci. *The Japanese J. Genetics*, 57: 25-57.

Straub, S.C., Parks, M., Weitemier, K., Fishbein, M., Cronn, R.C. and Liston, A. 2012. Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. *American J. Botany,* 99: 349-364.

Tang, J., Xia, H.a., Cao, M., Zhang, X., Zeng, W., Hu, S., Tong W, Wang J, Wang J and Yu, J. 2004. A comparison of rice chloroplast genomes. *Plant Physiol.,* 135: 412-420.

Waters, D.L., Nock, C.J., Ishikawa, R., Rice, N. and Henry, R.J. 2012. Chloroplast genome sequence confirms distinctness of Australian and Asian wild rice. *Ecol. Evol.,* 2: 211-217.

Wen, J. and Pandey, A. 2005. Initiating DNA molecular systematic studies in a developing country. Plant Taxonomy: Advances and Relevance. CBS Publishers & Distributors, New Delhi, India: 31-43.

Yang, Y., Li, Y. and Wu, C. 2013. Genomic resources for functional analyses of the rice genome. *Curr. Opinion in Plant Biol.*, 16: 157-163.

How to cite this article:

Jyoti Singh, Datta P. Kakade, Mayur R. Wallalwar, Rishiraj Raghuvanshi, Miranda Kongbrailatpam, Satish B. Verulkar and Shubha Banerjee. 2017. Evaluation of Potential DNA Barcoding Loci from Plastid Genome: Intraspecies Discrimination in Rice (*Oryza species*). *Int.J.Curr.Microbiol.App.Sci.* 6(5): 2746-2756. doi: https://doi.org/10.20546/ijcmas.2017.605.308