Determination of molecular diversity based on Internal Transcribed Spacer (ITS) sequences of nrDNA in *Commiphora wightii* (Art.) Bhandari (Guggul)

Joshi Ramesh*, Joshi Vinod, Angel Bennet, Angel Annette and Rao Satyawada Rama
Department of Life Sciences, Sharda School of Basic Sciences and Research, Sharda University, Greater Noida, U.P. 201310, INDIA
*rameshjoshi10@gmail.com; ramesh.joshi1@sharda.ac.in

Abstract

Commiphora wightii has a vast economic value and a wide range of medicinal uses. The presence of guggulsterone in the form of E and Z as a steroidal compound in oleo-gum resin is being used since ancient times for treating various disorders and ailments. The population studies on this plant revealed the drastic decline in its numbers and the imminent threat of extinction, if no concrete conservation plans are implemented. The present investigation was carried out for genetic diversity among populations of Commiphora wightii revealed by Internal Transcribed Spacer (ITS) regions of Nuclear Ribosomal DNA (nrDNA). Plant materials were obtained from various eco-climatic regions of Rajasthan and genomic DNA was extracted using a modified CTAB method. The primer ITS4 and ITS5 amplified the 5.8S-coding region and the internal transcribed spacers (ITS1 and ITS2).

The bands of DNA were eluted from the gel using the Qiagen Min-Elute gel extraction kit followed by sequencing at Desert Medicine Research Centre, ICMR using an automated sequencer, ABI-3730xl. The plant population of Jodhpur (J6), which acted outliner in the analysis showed significant diversity in comparison to rest of the population due to the reasons being entirely different climatic conditions in the area as compared to the rest of populations. Analysis of ITS regions inferred that the plant species Commiphora wightii might have advanced/evolved during their reproductive-isolation where continuous habitat loss and fragmentation of population may be the probable reasons.

Keywords: *Commiphora wightii*, Molecular diversity, Internal transcribed spacer.

Introduction

Genus *Commiphora wightii* (2n=26) in other words Guggul sometimes Myrrh tree is a small tree belonging to Burseraceae family[46]. In ancient as well as modern therapeutics, this plant has contributed as vast economic value and medicinal use. The plant is known as ‘‘Guggul’’ as it contains the E and Z form of steroidal compound called guggulsterone, in the oleo-gum resin. Since ancient times, various ailments and disorders have been treated using this compound (2000 B.C.). The plant exudates are complex in nature, are the mixture of 61% resin, 29.3% gum and 6.1% other chemicals. Till date, the reports estimate that 150 compounds or even more and new compounds continue to be reported[16].

Overexploitation of this species has brought its population under threat and already has been reported in the Data Deficient category of IUCN. The survey report of Rajasthan, India during 2007-2009[41] showed the growth of this species in only 2% of the plots which confirms its rarity. Population studies conducted on the species have verified the drastic decline in its numbers and the imminent threat of extinction, if no concrete conservation plans are implemented. It is expected that this species will be eroded from its natural abundance state if this exploitation continues without proper management of genetic resources. Conservation managers and evolutionary biologists have shown interest in the genetics of threatened species[4,12,23]. It becomes crucial to know the nature of a species population-structure because conserving and maintaining a species without taking into an account the way its population has been arranged, may fail to preserve the entire range of benefits which that species provides[18]. There continues to be a substantial need for research on many aspects of extent and distribution of genetic diversity. Research is especially required on factors such as the distribution of allelic variation within and between populations and the geographic patterns of these variations. These studies related to genetic diversity will help conservationists and breeders to estimate the important traits and their variability also to help them select the parents for conservation as well as hybridization programs.

In the past few decades, traditional methods of evaluating genetic variation have been complemented by a number of efficient molecular techniques[51] proven beneficial in distinguishing different species and identifying genes of interest[29]. The technique of molecular markers useful in detection of polymorphic loci between individuals or linked to major genes is emerging as indispensable tools in crop improvement programs[36]. Very few molecular studies have been conducted on the molecular characterization of *Commiphora* at both international and national level.

Cytoplasmic DNA i.e. mitochondrial (mt) and chloroplast (cp) DNA and nuclear (nr) DNA are the most significant tools for phylogenetic studies. Most chloroplast genes have limited potential to gather mutations[39] but much of the diversity is confined to non-coding regions which accumulate mutations at higher rate than actual genes. These regions provide more phylogenetic information at the intra

and interspecific population level in a variety of angiosperms[36] and have also been used to study the relationship at higher taxonomic level in angiosperms[9] and other groups of plants[19,27].

The nuclear ribosomal transcription unit; NRTU has 18S, 5.8S and 28S genes in it along with two internal transcribed spacers (ITS-1 and ITS-2) and one intergenic spacer (IGS). Mature rRNAs are the key components of cytoplasmic ribosomes which are produced after transcription by processing the NRTU. For the last one-decade, internal transcribed spacers (ITS) sequences of NRTUs are being globally used in inferring phylogenetic-relationships, to resolve transformation in a large range of plants and genetic diversity[11,18]. In the present investigation, nucleotide-sequences of ITS1-5.8S-ITS2 of the nr DNA were utilized to evaluate molecular differentiation/variations in *C. wightii* population and their relationship with the geo-graphical parameters of the areas they inhabit and also to provide genetic information needed to frame conservation strategies of *C. wightii.*

Material and Methods

Germplasm Collection: Survey was conducted to analyze the gender of plant in each population. Plant materials were obtained from various eco-climatic regions of Rajasthan where *Commiphora wightii* thrives in their natural habitat. The sample size for each area was varying in this study because of sample being collected from their natural habitats and the growth of the species is different in each area.

Fresh young leaves were collected from plant which has natural growth and belonged to twelve populations from Rajasthan, India (Table 1). 0.4 g of leaves sample were used to obtain genomic DNA using modified CTAB methods. DNA quality was assessed using agarose (0.8%) gel electrophoresis and spectrophotometer. Primer ITS4 and ITS5 amplified the 5.8S-coding region and the internal transcribed spacers (ITS1 and ITS2). The products that were amplified underwent electrophoresis with 1.4% agarose gels that were produced using a 1X TAE (Tris-acetate-EDTA) buffer system. The gels were run at a voltage of 5V/cm and stained with 0.5 µg/L ethidium bromide solution.

A 100 bp DNA ladder was used (Hyperladder IV, Bioline Ltd., London, UK), to estimate the amplified product's molecular weights. After a quick examination of the stained gels, bands were removed and the gel was eluted using the Qiagen Min-Elute gel extraction kit. To get accurate and dependable sequences, the samples were directly sequenced at Desert Medicine Research Centre, ICMR using an automated sequencer, ABI-3730xl. For every amplified sample chosen for the investigation, two sequences which overlapped for sense strand and antisense strand were obtained by sequencing the samples using flanking (ITS5 and ITS4) and internal (ITS1 and ITS2) primers.

Sequencer 4.8 (Gene Codes Corporation, MI, USA) was used to build complementary strands of the ITS region and the NCBI BLAST tool (http://www.ncbi.nlm.nih.gov/BLAST/) was used to determine homology. ClustalX was used to perform multiple alignments of nucleotides[32]. Polymorphism of nucleotide was measured by θw[33]. Diversity measured by π[34] was calculated by DnaSP v4.5[35]. Sequence alignments were obtained and adjusted manually. Indels were single characters when there was confidence for positional-homology. There were several gaps which were considered as missing data among phylogenetic analyses. Nucleotide multiple alignments were calculated with MUSCLE (Mega 6).

Table 1
Location of the sample population of *C. wightii* with terrains and average climatic conditions

S.N.	Area	Terrain	Population Code	Sample size	Sample code	Climatic Conditions
Ajmer						
1.	Ajmer	Hilly	A	7	A3	Hot, Semi-Arid
2.	Pushkar	Hilly	PB	5	PB1	Hot, Semi-Arid
3.	Srinagar	Rocky	SN	5	SN1	Hot, Semi-Arid
4.	Mangaliyawas	Hilly	M	10	M423	Hot, Semi-Arid
Jaipur						
5.	Jobner	Plains	JOB	6	JOB6	Hot, Semi-Arid
Jodhpur						
6.	Kailana	Rocky	J	7	J6	Hot, Arid Climate
Rajsamand						
7.	Bheem	Rocky	BH	1	BH	Sub-tropical dry climate
8.	Gomti Chouraha	Rocky	H	6	H2	Sub-tropical dry climate
Udaipur						
9.		Hilly	KG	4	KG2	Tropical climate
10.	Kavita	Hilly	KVT	2	KVT2	Tropical climate
11.	Neemach Mata	Hilly	NM	2	NM1	Tropical climate
12.	Thoor	Hilly	TH	5	TH2	Tropical climate

Phylogenetic and molecular evolutionary analyses and Maximum Parsimony (MP) tree were generated using the Close-Neighbor-Interchange algorithm in MEGA (Molecular Evolutionary Genetics Analysis**)** version 6 [48].

History of evolution was generated by using a number of methods like the Maximum-Likelihood method, the Neighbor-Joining method [43], the Minimum Evolution method [42], the UPGMA (Unweighted Pair Group Method with Arithmetic Mean) method [45] and the Maximum Parsimony method. These were used for analysis of both the ITS1-5.8S and ITS2-5.8S sequences of populations under study and were procured from the NCBI GenBank database. The Maximum Composite Likelihood Estimate of the pattern of Nucleotide Substitution and Tajima's Neutrality Test [47] were also used for analysing these sequences.

Results

Purification of amplified samples of ITS1-5.8S and ITS2-5.8S regions of *C. wightii* nrDNA was followed by amplification using ddNTPs with individual forward and reverse primers. The ABI sequences were visualized in Chromas LITE 2.1.1 with peaks of individual nucleotides. A comparative analysis of the total number of bases and GC content of all the sequenced samples of *C. wightii* is shown in table 2 and 3.

Out of the 12 samples sequences for the ITS1-5.8S region, sequences for five samples namely A3, KG2, KVT2, NM1 and PB1 could not be generated. Similarly sequences for four samples namely H2, KG2, KVT2 and PB1 for ITS2-5.8S region could also be not obtained, therefore these samples were not included in further sequence analysis. ABI sequences of ITS1-5.8S and ITS2-5.8S of all the samples showed more than 60% GC content. Base number in 'sense' and 'anti-sense' strands of ITS1 region varied from 292 to 540 (sense) and 292 to 492 (anti-sense strands). The GC content ranged between 60 to 68.60%. The substitution probabilities involved nucleotide sequences of 7 populations of *C. wightii* (Table 4).

Table 2
Nucleotide count and GC content (%) of ABI sequences of ITS1-5.8S region of nr DNA of *C. wightii*

Sample	Sense strand of ITS1-5.8S						Anti-sense strand of ITS1-5.8S					
	No. of nucleotides	A	T	G	C	GC (%) Content	No. of nucleotides	A	T	G	C	GC (%) content
A3	-	-	-	-	-	-	-	-	-	-	-	-
JOB6	532	122	95	158	157	68.60	458	91	91	132	144	60.00
J6	320	66	44	103	107	65.62	481	77	96	169	139	64.03
BH	314	63	41	109	101	66.87	292	39	57	96	100	67.12
H2	292	46	61	92	93	63.35	292	42	56	100	94	66.43
KG2	-	-	-	-	-	-	-	-	-	-	-	-
KVT2	-	-	-	-	-	-	-	-	-	-	-	-
NM1	-	-	-	-	-	-	-	-	-	-	-	-
TH2	540	118	71	182	169	65.00	492	81	89	163	159	65.44
SN2	488	105	66	166	151	64.95	456	80	85	147	144	63.81
423M	454	104	73	130	147	61.01	489	76	91	161	161	65.84
PB1	-	-	-	-	-	-	-	-	-	-	-	-

Table 3
Nucleotide count and GC content (%) of ABI sequences of ITS2-5.8S region in the nr DNA of *C. wightii*

Sample	Sense strand of ITS2-5.8S						Anti-sense strand of ITS2-5.8S					
	Total No. of Nucleotides	A	T	G	C	GC (%) content	Total No. of Nucleotides	A	T	G	C	GC (%) content
A3	383	70	60	128	125	66.05	382	65	66	129	122	65.70
JOB6	383	70	54	126	133	67.62	390	74	64	132	120	64.61
J6	380	67	64	118	131	65.52	393	74	67	133	119	64.12
BH	383	69	56	126	132	67.36	394	73	68	135	118	64.21
H2	-	-	-	-	-	-	-	-	-	-	-	-
KG2	-	-	-	-	-	-	-	-	-	-	-	-
KVT2	-	-	-	-	-	-	-	-	-	-	-	-
NM1	381	75	66	119	121	62.99	391	64	71	135	121	65.47
TH2	385	71	58	124	132	66.49	387	67	67	134	119	65.37
SN2	382	68	57	124	133	67.27	390	66	65	139	120	66.41
423M	384	66	56	129	133	68.22	392	80	65	131		
PB1	-	-	-	-	-	-	-	-	-	-	-	-

Table 4
Maximum Composite Likelihood Estimate of the Pattern of Nucleotide Substitution in ITS1-5.8S sequences of seven *C. wightii* populations.

	A	**T**	**C**	**G**
A	-	*3.05*	*8.47*	**0**
T	*3.79*	-	**36.22**	*10.04*
C	*3.79*	**13.05**	-	*10.04*
G	**0**	*3.05*	*8.47*	-

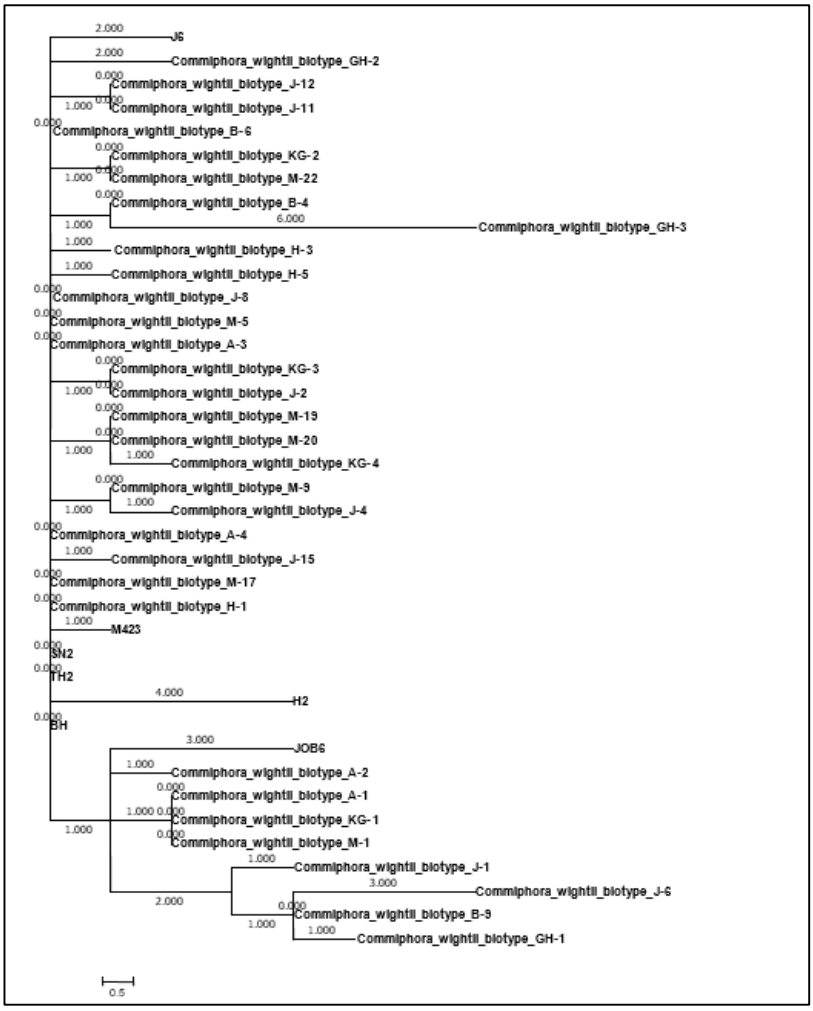

Figure 1: Phylogenetic analysis of ITS1-5.8S region of 39 populations of *C. wightii* by Maximum Parsimony method

Each of the entries shows the probable reason of replacement (r) from row to column. For simplifying, addition of all the r values has been brought equivalent to 100.

Rate of various transitional-replacements are displayed in bold while those of transversions replacements are displayed in italics. The nucleotide-frequencies are 14.96% (A), 12.04% (T/U), 33.41% (C) and 39.60% (G). The transitional substitution was recorded maximum (36.22) between C and T whereas transversional substitution in G-T and G-C was 10.04.

The history of evolution was also collected by the tool Maximum-Likelihood method which is displayed on Tamura-Nei model. The tree which had most log likelihood (-367.8978) has been displayed here. Initial tree(s) for the heuristic generated self by applying the formula of Neighbor-Join and BioNJ algorithms and estimated using the Maximum Composite Likelihood (MCL) approach and from this, the topology which had highest log likelihood value, were considered. The analysis involved 8 nucleotide sequences. ITS1 regions from seven *C. wightii* sample populations belonging to diverse eco-climatic regions were sequenced and analyzed. The amplified ITS1-5.8S regions of seven samples of *C. wightii* varied from 292 - 314 nucleotides (Table 5).

Similarity searches of NCBI database sequences showed the similarity of 90 -100% between the different samples of *C. wightii*. The consensus sequence of ITS1-5.8S regions of seven populations was aligned individually with 32 sequences (earlier submitted) from NCBI GenBank database

with accession numbers from EU419958.1 to EU419989.1 (Table 5). These sequences present in the GenBank belonged to various regions coded as: A- Ajmer, B-Bobas, GH- Galta Hills, H- Hirnoda, J-Jobner, KG-Kishan Garh and M-Mangaliyawas.

The results included 39 nucleotide-sequences where the position of codon positions was 1st position, 2nd position, 3rd position along with the Noncoders. The gaps were deleted. 122 positions were reported in the dataset. The phylogenetic analysis of 39 sequences of ITS1-5.8S region of different biotypes of *C. wightii* conducted by maximum likelihood method showed that M423 was grouped with closely related cluster of M5, J8, A3, A4 and M17 whereas J6 and H2 were close only to the biotype H5. Similarly, JOB6 was grouped with biotype GH2.

The history of evolution was also obtained by the tool Maximum Parsimony method. One of the tree out of the total 8 trees which was the most parsimonious trees (length = 40), has been shown (Figure 1). The calculations of consistency index, retention index and composite index are 0.775000 (0.590909), 0.709677 (0.709677) and 0.550000 (0.419355) respectively for all sites and parsimony-informative sites (in parentheses).

The MP tree was generated with the help of the Subtree-Pruning-Regrafting (SPR) algorithm where the initial trees were also generated by summing up the random sequences with 10 replicas. Branch lengths in this tree were also estimated with the help of average-pathway-method. They are shown next to the branches. The results included 39 nucleotide-sequences where the position of codon positions was 1st nucleotide position, 2nd nucleotide position, 3rd nucleotide position along with the Noncoders. The gaps were deleted. 122 positions were reported in the dataset.

The probable reason of replacement (r) from row to column was displayed for every entry. To make things simple, the total of the r values is set to 100. Italics is used to indicate rates of transversional replacements and bold is used to indicate rates with other transitional substitutions. 15.66% (A), 12.08% (T/U), 33.19% (C) and 39.07% (G) are the nucleotide frequencies. A total of 39 nucleotide sequences were analyzed.

First, second, third and noncoding codon locations were covered. The missing positions and gaps were eliminated. The transversional substitution in the ITS1-5.8S region of 39 populations was observed and was found maximum (8.91) between G and T and between G and C (Table 6). The rate of transitional substitution between C and T was maximum i.e. 39.47 whereas the minimum rate of transitional substitution (0.15) was observed between nucleotides A and G.

The Tajima's neutrality test involved 39 nucleotide-sequences where the position of codon positions was 1st nucleotide position, 2nd nucleotide position, 3rd nucleotide position along with the Noncoders. The gaps were deleted. 122 positions were reported in the dataset. The missing positions and gaps were eliminated. The test of ITS1-5.8S region of 39 sequences was performed and a total of 25 segregating sites were observed (Table 7). The ratio of number of segregating-sites and the total number of sites were found to be 0.204918. The nucleotide diversity between all analysed sequences was found to be 0.023915.

Table 5
Per cent similarity of ITS1-5.8s Sequences of 7 populations of *C. wightii* aligned with sequences from NCBI Gene Bank database (Accession no. EU419958.1 to EU419989.1)

Biotype Code	Identity Range (%)	Minimum Identity (%) With Gene Bank database Sequences	Maximum Identity (%) With Gene Bank database Sequences
JOB6	90-95	GH	H,M,J,A,KG
J6	93-97	GH	M,H,B,J
BH	95-100	GH	H,M
H2	89-93	GH,J	H
TH2	95-100	GH	H,M
SN2	94-99	J	H,M,KG,J
M423	9398	GH	H,M

Table 6
Maximum Composite Likelihood Estimate of the Pattern of Nucleotide Substitution in the ITS1-5.8S regions of 39 populations of *C. wightii*

	A	T	C	G
A	-	*2.76*	*7.57*	**0.37**
T	*3.57*	-	**39.47**	*8.91*
C	*3.57*	**14.37**	-	*8.91*
G	**0.15**	*2.76*	*7.57*	-

Table 7
Results from Tajima's Neutrality Test of ITS1-5.8S region of 39 populations of *C. wightii*

M	***S***	***Ps***	***Θ***	***Π***	***D***
39	25	0.204918	0.048468	0.023915	-1.722494

Abbreviations: m = number of sequences, n = total number of sites, S = Number of segregating sites, ps = S/n, Θ = ps/a1, π = nucleotide diversity and D is the Tajima test statistic

Table 8
Percentage Identity of ITS2-5.8s Sequences of 8 populations of *C.wightii* aligned with sequences from NCBI Gene Bank (from Accession no. EU419958.1 to EU419989.1)

Biotype Code	Identity Range (%)	Minimum Identity (%) With Gene Bank Biotype Sequences	Maximum Identity (%) With Gene Bank Biotype Sequences
A3	91-97	J	H, A, B, M, KG
JOB6	92-98	J	H, A, B, M, KG
J6	91-96	J	H
BH	94-99	J	H, M, KG, A, B, J
TH2	94-99	J	H, M, KG, A
NM1	91-97	J	H, M, KG, A, B
SN2	93-99	J	H, M, KG, A, B, J
M423	92-99	J	H, M, KG, A, B, J

Table 9
Maximum Composite Likelihood Estimate of the Pattern of Nucleotide Substitution in the ITS2-5.8S regions of 40 populations of *C. wightii*

	A	T	C	G
A	-	*2.4*	*6.21*	**18.17**
T	*2.18*	-	**30.48**	*5.17*
C	*2.18*	**11.79**	-	*5.17*
G	**7.65**	*2.4*	*6.21*	-

Table 10
Results from Tajima's Neutrality Test of ITS2-5.8S region of 40 populations of *c. wightii*

m	***S***	***ps***	***Θ***	***π***	***D***
40	31	0.130802	0.030751	0.012923	-2.000780

Abbreviations: m = number of sequences, n = total number of sites, S = Number of segregating sites, ps = S/n, Θ = ps/a1, π = nucleotide diversity and D is the Tajima test statistic.

ITS2-5.8S regions from eight *C. wightii* sample populations belonging to diverse eco-climatic regions were sequenced and analysed. The amplified ITS1- 5.8S regions of eight samples of *C. wightii* varied between 278 - 326 nucleotides. Similarity searches of NCBI database sequences showed 91-99% similarity between different samples of *C. wightii* (Table 8). The consensus sequence of ITS2-5.8S regions of eight populations was aligned individually with 32 sequences from NCBI GenBank database with accession number from EU419958.1 to EU419989.1. These sequences present in the GenBank belonged to various regions coded as: A-Ajmer, B-Bobas, GH- Galta Hills, H- Hirnoda, J-Jobner, KG-KishanGarh and M- Mangaliyawas.

Table 9 displays the maximum likelihood estimation of Nucleotide-Substitution pattern. The probable reason for replacement (r) from row to column is displayed for every entry. To make things simple, the total of the r values is set to 100. The rate of several trans-versional replacements is given in italics and the rates of various transitional substitutions are displayed in bold. The frequencies of the nucleotides are 15.04% (T/U), 32.42% (G), 38.89% (C) and 13.65% (A). There were forty nucleotide-sequences in the analysis. The first sequence, second sequence, third sequence along with the noncoding codon locations were covered. Every position with missing data and gaps was removed. There were 237 positions which were observed in the final dataset. The analysis involved 40 nucleotide sequences.

The results included 39 nucleotide-sequences where the position of codon positions was 1st nucleotide position, 2nd nucleotide position, 3rd nucleotide position along with the Noncoders. The gaps were deleted. 122 positions were reported in the dataset. The transversional substitution in the ITS2-5.8S region of 40 populations was observed and was found maximum 5.17 in between G and T and same in between G and C. The rate of transitional substitution

between C and T is maximum i.e. 30.48 whereas minimum rate of transitional substitution 7.65 was observed between nucleotides A and G (Table 9).

40 nucleotide sequences were involved in the Tajima's neutrality test. The results included 39 nucleotide-sequences where the position of codon positions was 1st position, 2nd position, 3rd position along with the Noncoders. The positions consisting of missing data and gaps were removed. The test of ITS1-5.8S of 40 sequences was performed and total of 31 segregating sites were observed. The ratio of total segregating sites and total sites was found to be 0.130802 (Table 10). The nucleotide diversity between all the analysed sequences was found to be 0.012923.

Discussion

The internal transcribed spacer is inserted in 18S-5.8S-26S region which separated the rDNA locus elements. The internal transcribed spacer region consisted of three parts: The internal transcribed spacer-1 and internal transcribed spacer-2 and conserved highly i.e. 5.8S rDNA exon located in between[53]. The total length of this region varying between 500 bp to 750 bp in angiosperms[6,7] while it may vary to be longer in some other seed plants i.e. 1,500-3,500 bp[32-34]. Both internal spacers are incorporated into the mature-ribosome, but during its maturation, it undergoes a specific cleavage during in ribosomal RNAs[10,11,21,22]. It is now certain that the internal transcribed spacer-2 is capable of producing the large subunit (LSU) rRNA as part of the ribosome biogenesis process Hadjiolova.

The accurate larger order configuration of both spacers is essential for directing endo-nucleolytic enzymes to appropriate cleavage sites[35]. The length of the ITS2 sequence varies significantly among various organisms although the study done by Hadjiolova et al[21] observed structural homologous-domains within *Saccharomyces cerevisiae* and mammals. To overcome the coding-regions, spacers evolve more quickly, it is widely employed as phylogenetic-reconstruction marker at various levels. The initial implementation by Porter and Collins[10] has gained widespread popularity for phylogeny reconstruction. The ITS, which is a component of the rDNA transcriptional unit, is found in almost all organisms.

The region in question offers several advantages: (1) It exhibits biparental inheritance, as opposed to the chloroplast which was inherited maternally along with its marker of mitochondria. (2) Amplification of PCR is straightforward, with numerous primers accessible universally for a wide range of organisms. (3) It has a multicopy-structure. (4) Its average size facilitates easy sequencing. (5) Published studies have demonstrated sufficient variation at the species, making it suitable for evolutionary research Alvarez.

According to Alvarez and Wendel [2] and Baldwin et al[6], the observed variation is a result of frequent nucleotide polymorphisms or it may be common insertion along with deletions in its sequence. The significant rate of its divergence is a crucial factor for examining population differentiation or phylogeography[6,49,54]. The large number of copies enables very consistent amplification and sequencing outcomes, as well as the opportunity to investigate coordinated and intertwined evolution. The utilisation of Internal Transcribed Spacer (ITS) in phylogenetic investigations is on the rise, with the range of publically available ITS sequences having tripled since 2003. The plant families that have been extensively researched are Asteraceae, Fabaceae, Orchideaceae, Poaceae, Brassicaceae and Maggini Apiaceae.

The variability observed in ITS (Internal Transcribed Spacer) including NRTU (Nuclear Ribosomal Transcriptional Unit) families is typically influenced by factors such as the range, copies of gene, mutation rates, its concerted-evolution, the placement of NRTU clusters on chromosomes and sexual:asexual reproduction [13, 14]. Polymorphism can result from the loss of sexual recombination or by coordinated evolution failing quickly considerable to homogenise reoccurrence in the face of large mutation rates[3,10,39]. Also, in agamospermous plants, concerted evolution has been retarded[37]. As *C. wightii* is apomictic, its sexual reproduction shortage could be a main reason for its polymorphism.

In its simulation model, apomixis spread in its population is never always associated with a decline in its variation in genetic and majority of it is retained. All mentioned factors elevate the polymorphism in *C. wightii* populations, unfortunately the converse is always reported.

A variety of different forms of ITS1 have been explained, within and among individuals e.g. Drosophila[44] and Aedes species[52], the ortho-pteran Melanopline species[30], the mammal *Homo sapiens*[19]. The equivalence between processes creating variation and homogenizing processes is called as degree of sequence variation[16,44]. Variation could be due to sexual reproduction, dispersing of rDNA loci on different chromosomes (>1 NOR)[50], i Schlotterer nter-breeding with their own sibling species and the pressure of environment, which may be the reason for minor alleles. Known homogenizing processes are those processes which are indulged in concerted evolution[13,14,24,25,31].

In present investigation, all the 12 samples of *C. wightii* were subjected to primers ITS1+ITS2 and ITS3+ITS4 in combinations to amplify ITS1-5.8S and ITS2 -5.8S respectively. One amplified segment of 400bp was generated in A3, JOB6, TH2, SN5, 423M and PB1 by a combination of ITS1+ITS2 primers whereas two segments of 300bp and 400bp were amplified in BH, H2 and NM1. Primers set ITS3+ITS4 amplified one segment of each (400bp) A3, JOB6, J6, KG2, KVT2, NM1, TH2, SN5, 423M and PB1 of ITS2-5.8S regions of nrDNA. The same primers set amplified one segment in each BH (300bp) and H2 of size 200-300bp. The consensus sequence of ITS1 region

generated by assembly of forward and reverse sequences of sample JOB6 comprised of 171 nucleotides without ambiguous sites. The consensus sequence of length 171 bases possessed 70.76% GC content.

The consensus sequence generated by forward and reverse sequences of sample J6 comprised of 154 unambiguous sites with 71.97% GC content. By comparing nucleotide with the corresponding trace peak of ITS1 region of sample BH, the consensus sequence of size 189 nucleotides was generated with 73.01% GC content. ITS sequences have proved to be of tremendous utility to angiosperm phylogeny[23]. ITS, thus, clearly holds a promise for inferring phylogenetic relationships among a very large number of taxa. In angiosperms, in general, ITS is best suited for reconstructing relationships among closely related genera and intrageneric groups[6,7]. Given their level of polymorphism at low taxonomic levels, ITS regions are commonly used to compile evidence of reticulate evolution.

Sequence additivity[26] and PCR-RFLP patterns[8,38] of the ITS have been used to find evidence for the origin of hybrids and to determine their progenitor species. In present study, sequence alignments of sense and antisense strands of ITS2-5.8S were performed subsequently and adjusted manually. Insertion along with the deletions was scored as single characters when we expected confidence in positional homology. Sequence alignment of ITS2-5.8S of sample JOB6 was performed and insertion and deletion were scored to the position homology by observing trace peaks. The consensus sequence of length 283 unambiguous sites was generated with 73.14% GC content. The homologous positioning of nucleotides of sample J6 was also scored on SeqAssem 07 and a nucleotide sequence of 293 sites without any ambiguity was obtained. The GC content of J6 consensus was 69.62 percent.

Conclusion

Based on the observations through analysis of ITS regions by the latest software, it can be concluded that the plant of *C. wightii* might have evolved during their reproductive isolation, where continuous habitat loss and fragmentation of population may be the probable reasons. Especially, population of Jodhpur (J6), which acted outliner in the analysis and showed significant diversity in comparison to the rest of the samples, probably due to variation in climatic conditions in the area as compared to the rest of populations. Also, in our study, we observed detrimental effect on genetic variation when apomictic behaviour and reproductive isolation were coupled.

Natural population of cross breeding plant species carried higher genetic diversity levels in comparison to self-breeders leading to genetic polymorphism justified in this study which was detected by using ITS1 and ITS2 makers. It also seems that because of high differentiation in population, there is a disruption in population continuum where over exploitation, anthropogenic activities and unsustainable utilization will be the reasons. As per our study, the molecular characterisation of *C. wightii* was performed by using Internal Transcribed Spacer (ITS1-5.8S-ITS2) regions of nrDNA in the vast and diverse phytogeographical regions of Rajasthan.

Acknowledgement

We gratefully acknowledge Desert Medicine Research Centre, ICMR. Jodhpur (Rajasthan) (presently National Institute for Implementation Research in Non-Communicable Diseases, Jodhpur) for providing facilities of DNA sequencing and bioinformatic analysis.

References

1. Ainouche M.L. and Bayer R.J., On the origins of the tetraploid Bromus species (section Bromus, Poaceae). Insights from the internal transcribed spacer sequences of nuclear ribosomal DNA, *Genome*, **40(5)**, 730-743 **(1997)**

2. Alvarez I. and Wendel J.F., Ribosomal ITS sequences and plant phylogenetic inference, *Mol. Phylogenet Evol*., **29(3)**, 417-434 **(2003)**

3. Appels R. and Honeycutt R.L., rDNA Evolution over a billion years, In DNA Systematics, ed. Dutta S.K., CRC Press, Boca Raton, FL, 81-135 **(1986)**

4. Avise J.C. and Hamrick J.L., eds., Conservation Genetics: Case histories from nature, Chapman & Hall, New York, 512 **(1996)**

5. Baldridge G.D., Dalton M.W. and Fallon A.M., Is higher-order structure conserved in eukaryotic ribosomal DNA intergenic spacers, *J. Mol. Evol*., **35**, 514-523 **(1992)**

6. Baldwin B.G., Molecular phylogenetics of Calcydenia (Compositae) based on ITS sequences of nuclear ribosomal DNA: Chromosomal and morphological evolution re-examined, *Am. J. Bot.*, **80**, 222-238 **(1993)**

7. Baldwin B.G., Sanderson M.J., Porter J.M., Wojciechowski M.F., Campbell C.S. and Donoghue M.J., The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny, *Ann. Miss. Bot. Gard*., **82(2)**, 247-277 **(1995)**

8. Barji Isfaniza and Awang Adeni Dayang Salwani, Diluted Sulphuric Acid Hydrolysis of Destarched Sago Fibre assisted with Selected Pre-treatments for Glucose and Xylose Production, *Res. J. Chem. Environ.*, **28(1)**, 13-20 **(2024)**

9. Borsch T., Hilu K.W., Quandt D., Wilde V., Neinhuis C. and Barthlott W., Noncoding plastid trnT-trnF sequences reveal a well resolved phylogeny of basal angiosperms, *J. Evol. Biol*., **16(4)**, 558-576 **(2003)**

10. Campbell C.S., Wojciechowski M.F., Baldwin B.G., Alice L.A. and Donoghue M.J., Persistent nuclear ribosomal DNA sequence polymorphism in the *Amelanchier* agamic complex (Rosaceae), *Mol. Biol. Evol*., **14**, 81-90 **(1997)**

11. Clarke G.M. and Young A.G., Genetics, Demography and Viability of Fragmented Population, eds., Young A. and Clarke G.M., Cambridge University Press, Cambridge, UK, 1-6 **(2004)**

12. Clegg M.T., Gaut B.S., Learn G.H. and Morton B.R., Rates and patterns of chloroplast DNA evolution, Proc. Nat. Acad. Sc., USA, **91**, 6795-6801 **(1994)**

13. Dover G., Molecular drive: a cohesive mode of species evolution, *Nature*, **299**, 111-117 **(1982)**

14. Dover G.A., Linares A.R., Bowen T. and Hancock J.M., Detection and quantification of concerted evolution and molecular drive, *Methods Enzymol*., **224**, 525-541 **(1993)**

15. Fatope M.O., Al-Burtomani S.K., Ochei J.O., Abdulnour A.O., Al-Kindy S.M. and Takeda Y., Muscanone: a 3-O-(1", 8", 14"-trimethylhexadecanyl)naringenin from Commiphora wightii, *Phytochem*., **62(8)**, 1251–1255 **(2003)**

16. Fritz G.N., Conn J., Cockburn A. and Seawright J., Sequence analysis of the ribosomal DNA Internal Transcribed Spacer 2 from populations of *Anopheles nuneztovari* (Diptera: Culicidae), *Mol. Bio. Evol*., **11(3)**, 406-416 **(1994)**

17. Godt M.J.W. and Hamrick J.L., Allozyme diversity in the endangered pitcher plant Sarracenia rubra ssp. Alabamensis (Sarraceniaceae) and its close relative S. rubra ssp. rubra, *Am. J. Bot*., **85(6)**, 802 **(1998)**

18. Goffinet B. and Buck W.R., Systematics of the Bryophyta (Mosses): from molecules to a revised classification, *Monographs Systema. Bot. Misso. Botan. Garden*, **98**, 205-239 **(2004)**

19. Gonzalez I.L., Chambers C., Gorski J.L., Stambolian D., Schmickel R.D. and Sylvester J.E., Sequence and structure correlation of human ribosomal transcribed spacers, *J. Mol. Biol*., **212(1)**, 27-35 **(1990)**

20. Hadjiolova K.V., Georgiev O.I., Nosikov V.V. and Hadjiolov A.A., Localization and structure of endonuclease cleavage sites involved in the processing of the rat 32S precursor to ribosomal RNA, *Biochem J*., **220(1)**, 105-16 **(1984)**

21. Hadjiolova K.V., Normann A., Cavaille J., Soupene E., Mazan S., Hadjiolov A.A. and Bachellerie J.P., Processing of truncated mouse or human rRNA transcribed from ribosomal minigenes transfected into mouse cells, *Mol. Cell Biol*., **14(6)**, 4044-4056 **(1994)**

22. Hedrick P.W., Conservation genetics: where are we now?, *Trends Ecol. Evol*., **16**, 629-636 **(2001)**

23. Hershkovitz M.A., Zimmer E.A. and Hahn W.J., Ribosomal DNA sequences and angiosperm systematics, In Hollingsworth P.M., Bateman R.M. and Gornall R.J., eds., Molecular systematics and plant evolution, Taylor & Francis London, UK, 268-326 **(1999)**

24. Hillis D.M. and Davis S.K., Ribosomal DNA: Intraspecific polymorphism, concerted evolution and phylogeny reconstruction, *Systematic Zoo*, **37(1)**, 63-66 **(1988)**

25. Hillis D.M., Moritz C., Porter C.A. and Baker R.J., Evidence for biased gene conversion in concerted evolution of ribosomal DNA, *Sc*., **251(4991)**, 308-310 **(1991)**

26. Hoot S.B. and Taylor W.C., The utility of nuclear ITS, a LEAFY homolog intron and chloroplast atpB-rbcL spacer region data in phylogenetic analyses and species delimitation in Isoetes, *Am. Fern J*., **91**, 166-177 **(2001)**

27. Houseley J., Kotivic K., Hage A.E. and Tollervey D., Trf4 targets ncRNAs from telomeric and rDNA spacer regions and functions in rDNA copy number control, *EMBO J*., **26**, 4996-5006 **(2007)**

28. Houseley M., Martin-Bravo S., Dobes C., Gong W., Jordon-Thaden I., Kiefer C., Kiefer M., Paule J., Schmickl R. and Koch M.A., Non-coding nuclear DNA markers in phylogenetic reconstruction, *Plant Systems. Evol*., **282**, 257-280 **(2009)**

29. Kalia R.K., Singh R., Rai M., Mishra G.P., Singh S.R. and Dhawan A.K., Biotechnological interventions in sea buckthorn (*Hippophae* L.): Current status and future prospects, *Curr. Status Future Prospect*, **25**, 559-575 **(2011)**

30. Kuperus W.R. and Chapco W., Usefulness of Internal Transcribed Spacer regions of ribosomal DNA in Melanopline (Orthoptera: Acrididae) systematics, *Annals Entomol. Soc. Am*., **87(6)**, 751-754 **(1994)**

31. Linares A.R., Bowen T. and Dover G.A., Aspects of non-random turnover involved in the concerted evolution of intergenic spacers within the ribosomal DNA of Drosophila melanogaster, *J Mol Evol*., **39(2)**, 151-159 **(1994)**

32. Liston A., Robinson W.A., Oliphant J.M. and Alvarez-Buylla E.R., Length variation in the nuclear ribosomal DNA internal transcribed region of non-flowering seed plants, *Syst Bot*., **21(2)**, 109-120 **(1996)**

33. Maggini F., Frediani M. and Gelati M.T., Nucleotide sequence of the internal transcribed spacers of ribosomal DNA in Picea abies Karst, *DNA Seq*., **11(1-2)**, 87-89 **(2000)**

34. Maggini F., Marrocco R., Gelati T.M. and De Dominicis R.I., Length and nucleotide sequences of the internal spacers of nuclear ribosomal DNA in gymnosperms and pteridophytes, *Plant Syst. Evol*., **213(3-4)**, 199-205 **(1998)**

35. Mai J.C. and Coleman A.W., The Internal Transcribed Spacer 2 exhibits a common secondary structure in green algae and flowering plants, *J. Mol. Evol*., **44**, 258-271 **(1997)**

36. Martinez-Gomez P., Sanchez-Perez R., Rubio M., Dicenta F., Gradiziel T.M. and Sozzi G.O., Application of recent biotechnologies to *Prunus* tree crop genetic improvement, *Cien. Inv. Agr*., **32**, 67-90 **(2005)**

37. Nei M. and Li W.H., Mathematical model for studying genetic variation in terms of restriction endonucleases, Proc. Natl. Acad. Sci. USA., **76(10)**, 5269-5273 **(1979)**

38. Nwakanma D.C., Pillay M., Okoli B.E. and Tenkouano A., PCR-RFLP of the ribosomal DNA internal transcribed spacers (ITS) provides markers for the A and B genomes in Musa L., *Theor. Appl. Genet*., **108(1)**, 154-159 **(2000)**

39. Palmer J.D. et al, Chloroplast DNA variation and plant phylogeny, *Ann. Mo. Bot. Gard*., **75**, 1180-1206 **(1988)**

40. Porter C.H. and Collins F.H., Species-diagnostic differences in the ribosomal DNA internal transcribed spacer from the sibling

species Anopheles freeborni and Anopheles hermsi (Diptera: Culicidae), *Am J Trop Med Hyg*., **45(2)**, 271-279 **(1991)**

41. Reddy C.S., Meena S.L., Krishna P.H., Charan P.D. and Sharma K.C., Conservation threat assessment of *Commiphora wightii* (Arn.) Bhandari - An Economically Important Species, *Taiwania*, **57(3)**, 288-293 **(2012)**

42. Rzhetsky A. and Nei M., A simple method for estimating and testing minimum evolution trees, *Mol. Biol. Evol*., **9**, 945-967 **(1992)**

43. Saitou N. and Nei M., The neighbor-joining method: A new method for reconstructing phylogenetic trees, *Mol. Biol. Evol.*, **4**, 406-425 **(1987)**

44. Schlotterer C. and Tautz D., Chromosomal homogeneity of Drosophila ribosomal DNA arrays suggests intrachromosomal exchanges drive concerted evolution, *Curr. Biol*., **4(9)**, 777-783 **(1994)**

45. Sneath P.H. and Sokal R.R., Numerical Taxonomy - The principle and practice of numerical classification, 1st Edition, W. H. Freeman and San Francisco **(1973)**

46. Sobti S.N. and Singh S.D., A chromosome survey of Indian medical plants: Part-1, Proc Ind. Acad. Sci., **54**, 138-144 **(1961)**

47. Tajima F., Statistical methods to test for nucleotide mutation hypothesis by DNA polymorphism, *Genetics*, **123**, 585-595 **(1978)**

48. 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, *Mol Biol Evol*., **28**, 2731-2739 **(2011)**

49. Vargas P., Morton C. and Jury S.L., Biogeographic patterns in Mediterranean and Macaronesian species of Saxifraga (Saxifragaceae) inferred from phylogenetic analyses of ITS sequences, *Am J Bot*., **86(5)**, 724-34 **(1999)**

50. Vogler A.P. and DeSalle R., Evolution and phylogenetic information content of the ITS-1 region in the Tiger Beetle (*Cindela dorsalis*), *Mol. Bio. Evol*., **11(3)**, 393-405 **(1994)**

51. Vyas P., Jaiswal S.K., Joshi V., Angel B., Angel A., Manzoor U., Yadav A. and Joshi R., Analysis of genetic diversity among populations of an endangered medicinal plant *Commiphora wightii* (Arn.) Bhandari (Guggul) using Molecular Markers, *Jour Pl Sci Res*., **39(2)**, 141-152 **(2023)**

52. Wesson D.M., Porter C.H. and Collins F.H., Sequence and Secondary structure comparisons of internal transcribed spacer rDNA in mosquitoes (Diptera: Culicidae), *Mol. Phylogen. Evol*., **1(4)**, 253-269 **(1992)**

53. Wheeler W.C. and Honeycutt R.L., Paired sequence difference in ribosomal RNAs: evolutionary and phylogenetic implications, *Mol. Biol. Evol*., **5(1)**, 90-96 **(1988)**

54. Yamaji H., Fukuda T., Yokoyama J., Pak J.H., Zhou C., Yang C., Kondo K., Morota T., Takeda S., Sasaki H. and Maki M., Reticulate evolution and phylogeography in Asarum sect. Asiasarum (Aristolochiaceae) documented in internal transcribed spacer sequences (ITS) of nuclear ribosomal DNA, *Mol. Phylogenet. Evol*., **44(2)**, 863-884 **(2007)**.

(Received 12th September 2024, accepted 20th November 2024)