© Triveni Enterprises
Vikas Nagar, Lucknow, INDIA
editor@jeb.co.in
Full paper available on: www.jeb.co.in

J. Environ. Biol.
33, 991-997 (2012)
ISSN: 0254 - 8704
CODEN: JEBIDP

Utilization of intron-flanking EST-specific markers in the genetic characterization of *Artemisia annua* genotypes from the trans-Himalayan region of Ladakh, India

Author Details

Jitendra Kumar	Defence Institute of High Altitude Research, Defence R & D Organization, Leh-194 101, India
Prasad Bajaj	Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Solan-173 234, India
Harvinder Singh	Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Solan-173 234, India
Gyan P. Mishra	Defence Institute of High Altitude Research, Defence R & D Organization, Leh - 194 101, India
Ravi B. Srivastava	Defence Institute of High Altitude Research, Defence R & D Organization, Leh - 194 101, India
Pradeep K. Naik *(Corresponding author)*	Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology, Solan-173 234, India e-mail : pknaik73@rediffmail.com; pradeep.naik@juit.ac.in

Online Copy

Publication Data

Paper received:
16 February 2010

Revised received:
04 August 2011

Accepted:
21 September 2011

Abstract

Genetic variation was assessed utilizing intron-flanking EST-specific markers among genotypes of *Artemesia annua* collected from two sampling sites *viz*. Nubra (9,600 ft) and Leh (11,500 ft) valleys of the trans-Himalayan region, Ladakh, India. The available ESTs (3,60,906) sequences of *A. annua* were aligned with the genomic sequences of *Arabidopsis* to developed 'intron-flanking' EST-PCR based primers. These primers anneal with the conserved region of exon (flanking to the intron) and amplified the introns. Out of the 39 primers selected and tested on 20 genotypes of *A. annua*, we successfully exploited 81 codominant intron length polymorphism (ILP) markers, with an average of 2.08 markers per primer and 92.04% polymorphism detection. Clustering of genotypes revealed distribution of genotypes into 2 distinct clusters with respect to their site of collection. Significantly, this study demonstrates that *Arabidopsis* genome sequence can be useful in developing gene-specific PCR-based markers for other non-model plant species like *A. annua* in the absence of genome sequences.

Key words

Artemisia annua, Intron length polymorphism, Molecular marker, Genome characterization

Introduction

Artemisia annua L. is an annual herb native to Asia, now distributed throughout many countries such as Europe, North America, Central and South America. The plant synthesizes and accumulates substantial quantities of many derivatives of a cadiene skeleton including artemisinin (Klayman, 1993). Artemisinin is currently the most effective anti-malaria agent against multidrug resistant strains of *Plasmodium* species (Phillipson and Wright, 1991). *A. annua* is one of the most important medicinal plants from cold arid regions of India especially in Ladakh region. It is well adapted at high altitudes of 9600-11500 ft above mean sea level and thrives well to the temperature, nutrient and environmental stresses it has

been exposed. Over the year, the *Artemisia* populations in the Ladakh region have developed considerable variability and thus it is essential to study the genetic variations. The analysis of genetic diversity is a prerequisite for effective utilization and protection of plant genetic resources (Weising *et al*., 1995). It also helps in developing DNA based molecular markers for identification of genotypes with better traits.

Many types of molecular markers have been developed since 1980, such as restriction fragment length polymorphism (RFLP; Botstein *et al.*, 1980), random amplified polymorphic DNA (RAPD; Williams *et al.*, 1990), amplified fragment length polymorphism (AFLP; Vos *et al.*, 1995), simple sequence repeat polymorphism (SSR; Becker and Heun, 1995), intron length polymorphism (ILP; Choi *et al.*, 2004), etc. Molecular markers obtained from the non-functional region of the genome are mostly preferred for phylogenetic analysis and inferring taxon relatedness. Particularly the introns (originally thought of as junk DNA) that are widespread and abundant in eukaryotic genomes (Hawkins, 1988; Deutsch and Long, 1999) are nowadays commonly used as sources of DNA polymorphism. Since introns have no functional significance (although may influence the level of gene expression), they are more variable then coding sequences (Wang *et al.*, 1998). Introns between genotypes show polymorphism due to difference in sequence, number of repetition of simple sequences or difference in length. But it is easy to detect intron length polymorphism (ILP) between the populations or genotypes using conventional PCR. To amplify introns by PCR, primers are generally designed based on the flanking exons—called as exon-primed intron-crossing PCR (EPIC-PCR) (Palumbi, 1995). The advantage of this technique is that exon sequences are relatively more conserved across the plant species and the primers designed based on the genome sequences of model plant like *Arabidopsis* could also be used to amplify introns from the non-model plants. Availability of cDNA/EST sequences of *A. annua* provides us with an opportunity to systematically search for DNA polymorphism among the genotypes from Indian origin.

To our knowledge, there is no published information on the use of ILPs for the characterization of genetic diversity in *A. annua*. Therefore attempt has been made to determine the genetic relationships among several genotypes of *A. annua* from Ladakh region (the trans-Himalayan region, India) utilizing candidate ILP markers. This information will be helpful in enriching the available genetic linkage map of *A. annua* and could help further in the development of high yielding genotypes.

Materials and Methods

Intron-flanking EST-primers : The genome sequences (with intron and UTRs) of *Arabidopsis* were downloaded and used as a reference database to find out the homolog sequences corresponding to the ESTs sequences of *A. annua* using local BLASTN search. The entire available ESTs sequences (3,60,906) of *A. annua* were downloaded as FASTA format from the NCBI database (dbEST). We used a high e-value (10^{-20}) for the BLASTN to remove paralogues. Intron-flanking EST-primers were designed based on the "exon/exon" junction site information, inferred from the pair-wise alignments between the ESTs of *A. annua* and their homolog exonic sequences from *Arabidopsis.* As per this pair-wise alignment strategy the predicted "exon/exon" junctions from *A. annua* ESTs are essentially correspond to "exon/intron" junctions (splice sites) in *A. annua* genomic sequences. The forward and reverse primers were designed using Primer3 (Rozen and Skaletsky, 2000) with default settings based on the EST sequences that flanks at least two exon/exon junctions. These primers were expected to span the predicted *A. annua* intron regions in a PCR. The primers used for amplification are listed in Table 1.

DNA extraction and PCR amplification: Twenty genotypes of *A. annua* were used for the analysis, collected from two valleys (ten each): Leh (11,500 ft) and Nubra (9,600 ft), separated from each other through natural mountain barrier (Fig. 1). The interval between plants within a valley was 100-200 m and the pair wise distance between valleys was 50–250 km. Young leaves from the plant were collected, rinsed with distilled water and stored at -80 ^{o}C until DNA extraction. Total genomic DNA was extracted from the frozen leaves by the (CTAB) Cetyl trimethyl ammonium bromide method (Saghai-Maroof *et al.*, 1984) with slight modifications. Samples of 100 mg were ground to powder using liquid nitrogen with a pre-chilled mortar and pestle. The powder was transferred to a 30 ml sterile Falcon tube with 12.5 ml of CTAB buffer. The extraction buffer consisted of 2% (w/v) CTAB, 1.6 M NaCl, 3% polyvinylpyrrolidone (PVP) and 0.5% b mercaptoethanol. After incubating the homogenate for 1.15 hr at 65 ºC an equal volume of chloroform: Isoamyl alcohol mixture (24:1) and 1/10 volume of warm (55 ^{o}C) 10% CTAB was added and centrifuged at 11,000 rpm for 20 min. DNA was precipitated with 1/10 volume of 5 M NaCl and an equal volume of isopropanol followed by centrifugation at 10,000 rpm for 10 min.

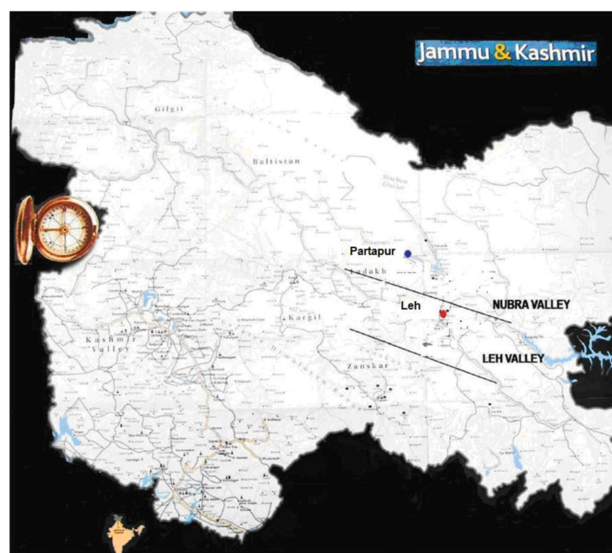

Fig. 1: Collection sites of 20 *Artemisia annua* genotypes from two valleys (Leh and Nubra) and the two collection sites (Leh and Partapur) located in Ladakh, India

Online Copy

Table- 1: List of primers used for Intron length polymorphism (ILP) study, GC content, melting temperature (Tm) and polymorphism

Sl.No.	Primer sequence	Length of primer	% GC	Tm	Total no. of loci	NPL
1L	3' TCGTTTGGGCATTGGGATCATGC 5'	24	54	58	2	2
1R	5' TTGAGACGTGGTCTAACTGCCA 3'	22	50	54		
2L	3' ATTTCAGTTTTGGCCTCGTG 5'	20	45	49	2	1
2R	5' CGTATCACCGCGAGATTCTT 3'	20	50	51		
3L	3' CTAAGGATGCCTAACCGCCGCA 5'	22	59.1	58	2	2
3R	5' AGCTTGAGCCTACCTTCGCACAC 3'	23	56.5	58		
4L	3' AGGCTCAAGCTAAGATTCTGCTCG 5'	24	50	57	-	-
4R	5' GAGGATCGCTAATCTCGTCTGG 3'	22	54.6	56		
5L	3' TGACACCTGCCCGGTGCTGGAA 5'	22	63.6	60	2	2
5R	5' TGATCGGTGCGGCGGTTAGGCA 3'	22	63.6	60		
6L	3' TGGGGAAAATCAGCCTGTTA 5'	20	40.9	49	3	3
6R	5' ACGAAAGTCGGCCTTAGTGA 3'	20	50	51		
7L	3' GCCTTTGCACTCGAGGGCCAA 5'	21	61.9	58	2	2
7R	5' GCCATCAGTGAGATACCACTCTGG 3'	24	54.2	58		
8L	3' AACCGTGGCTCCAAAGCTCTCAGC 5'	24	58.3	60	2	2
8R	5' CGCACTTGGTTTAGCCAACGTGCA 3'	24	50	58		
9L	3' CTCGTATTCTGCACCCATGA 5'	20	50	51	3	3
9R	5' TTGCACCCTCCACTACCTTT 3'	20	50	51		
10L	3' TACCAGACGGGATTGGGGTTCCA 5'	23	56.5	58	3	3
10R	5' GCAACATGGGAGCCGAACCCTCCA 3'	24	62.5	62		
11L	3' TCCAACCTCAAGTGGAACAC 5'	20	50	51	2	2
11R	5' GCATTAGATGGTCTCCTTGACGGA 3'	24	50	57		
12L	3' CTCAACAGTGATCAAGACCTCGTG 5'	24	50	57	2	2
12R	5' TGATAGACCGAGCTGTTTCCGGTA 3'	24	50	57		
13L	3' TCTGCAGGATTAGGGATGGCTGTG 5'	24	54.2	58	2	2
13R	5' CCATCACATATAGCTGCAACCGTA 3'	24	45.8	55		
14L	3' GGAGTCAACCCAGCTTCCGCTA 5'	22	59	58	2	2
14R	5' TGGAGACCGCTAGTGCTTCCAGAC 3'	24	58.3	60		
15L	3' AGAGAATGGTCTAATGGCTTGC 5'	22	45.5	52	2	2
15R	5' TGATAGACCGAGCTGTTTCCGGTA 3'	24	50	57		
16L	3' TCTGCAGGATTAGGGATGGCTGTG 5'	24	54.2	58	2	1
16R	5' CCATCACATATAGCTGCAACCGTA 3'	24	45.8	55		
17L	3' GCCATGCAACTCTGAACGGCTG 5'	22	59	58	2	2
17R	5' TGATAGACCGAGCTGTTTCCGGTA 3'	24	50	57		
18L	3' GTTTTGGCACGGGTCATAGC 5'	20	55	53	2	2
18R	5' ACAATGTCGATCCAGAACCAC 3'	21	47.6	52		
19L	3' AGAGTCCTGTGTCTCAACCATCGA 5'	24	50	57	2	1
19R	5' TTAGTTCCGGGGTCCACTAACGCA 3'	24	54.2	58		
20L	3' ACATTGTCTTGGCCGGATAG 5'	20	50	51	3	2
20R	5' TATTACACCCCTTGCCCAAA 3'	20	54.2	49		
21L	3' GAACTCCGGTCATTGTACCACCAG 5'	24	58.3	58	2	2
21R	5' AGGATGGCCCGACATGCGTTTACC 3'	24	50	60		
22L	3' TGCTTGATGCTTCTCACACG 5'	20	55	51	3	3
22R	5' GGCTCGATCTTTGCAAGCTAGG 3'	22	56.5	56		
23L	3' CTCTAAAAGCACCTCCGTGCAGG 5'	23	52.2	58	3	3

Online Copy

23R	5' TACGCCTAAGAGCACCTTCAGGA 3'	23	54.2	56		
24L	3' GTCGGGTTCTTTGAACGTCAGCGA 5'	24	56.5	58	3	3
24R	5' TGCGGCCATTGTATTCCGGCCAA 3'	23	52.2	58		
25L	3' ACGAAGAACCGTAAAGACCCAC 5'	22	54.2	54	3	2
25R	5' AAGATCACGGTTGACTCTGCCA 3'	22	56.5	54		
26L	3' GCGTATTCTGATGGTTGGTCTCGA 5'	24	50	57	3	3
26R	5' TTGGATTGCGCCTCGTCACCAA 3'	22	54.5	56		
27L	3' GCTAACAAGGCCTAAAACGGACTC 5'	24	50	57	2	2
27R	5' CTTGGAGCATCATCACCAGCAA 3'	22	50	54		
28L	3' GATGCTTACGTTGGTGACGAAGCA 5'	24	50	57	2	2
28R	5' CAACATCCCACACAGTGAAGCTG 3'	23	52.2	56		
29L	3' TTTGTGGTCGACAGCAATGACC 5'	22	50	54	2	2
29R	5' TGCAGTGTTGCTTGTTTTTGCCA 3'	24	41.7	53		
30L	3'GGGTCAACGTTCTTTGGGCACTG 5'	23	56.5	58	2	1
30R	5'TGTGGGTGGAAATCTGCCCAAACC 3'	24	54.2	58		
31L	3'TGACGACCTCTCGTGTATTGTT 5'	22	45.5	52	2	2
31R	5' CCATTTATTTCCCCGTGTCA 3'	20	45	49		
32L	3' GGTTCTTTGCACTGACACGA 5'	20	50	51	2	2
32R	5' TTGTACCCGTCTTGACACCA 3'	20	50	51		
33L	3' AGCTTTTGCGCAACTCAAGCAG 5'	22	50	54	2	2
33R	5' ATCCGAGGGTGTGTGCACCTG 3'	21	61.9	58		
34L	3' AGCTTTTGCGCAACTCAAGCAG 5'	22	50	54	2	2
34R	5' GAACATCAGGAAAGATTGTGCTGC 3'		24	45.8	55	
35L	3' TTTGTTCAGGGGTGTGATGC 5'	20	50	51	3	2
35R	5' CTCAACGTTCTTTGGGCACTG 3'	21	52.4	54		
36L	3' TAGCTTTTGCGCAACTCAAG 5'	20	45	49	2	2
36R	5'TTGATATCATTGAGCGAGATTG 3'	22	36.4	49		
37L	3' TTGAAAACTGGTTGCAATGG 5'	20	40	47	2	2
37R	5'TGTGCCATCATATTCAGTGGA 3'	21	42.9	50		
38L	3' AGCTTTTGCGCAACTCAAGCAG 5'	22	50	54	2	2
38R	5' ATCCGAGGGTGTGTGCACCTG 3'	21	61.9	58		
39L	3' AGCTTTTGCGCAACTCAAGCAG 5'	22	50	54	2	2
39R	5'GAACATCAGGAAAGATTGTGCTGC 3'	24	45.8	55		
40L	3' ACCATGGATGATCTTGGTTCCCA 5'	23	47.8	55	2	2
40R	5'TCTGTAAGACATGGTAGCTCACGA 3'	24	45.8	55		
Total					**88**	**81**

NPL = Number of polymorphic loci

Online Copy

RNA was removed by RNase treatment. DNA was quantified by comparing with uncut l DNA on the agarose gel.

Amplification reactions were performed by 'touchdown PCR' in volumes of 25 µl reaction mixture containing 50 ng template DNA, 0.5 mM primer, 200 mM of each dNTPs, 15 mM $MgCl_2$, 0.1% Triton X-100, and 1unit of Taq polymerase. To the reaction mixture, 1.5 ml of 10X PCR buffer was added and incubated at 94 °C for 4 min to denature the template DNA followed by 10 cycles of denaturation at 94 °C for 45 sec, primer annealing at 59 °C minus 0.3 °C per cycle for 30 sec and primer extension at 72 °C for 1 min. In the next 20 cycles, the annealing temperature was reduced to 56 °C for 30 sec while the denaturation and primer extension time remained the same as in first 10 cycles. The last cycle consisted of only primer extension at 72 °C for 5 min. The amplification products were separated in 3% metaphor-agarose gel electrophoresis (100 V, 2 hr) and band sizes were determined against '1 kb plus DNA ladder' (Invitrogen Life Technologies, Carlsbad, CA, USA).

Data collection and analysis : Fragments were scored as a binary unit character (present = 1; absent = 0). Only unambiguously resolved and reproducible bands were scored. A genetic similarity

matrix using simple matching coefficient was generated using NTSYSpc2.1. A dendrogram was computed using the unweighted pair-group method with arithmetic average (UPGMA) clustering method. POPGENE software was used to calculate Nei's unbiased genetic distance among different genotypes. Data for observed number of alleles (Na), effective number of alleles (Ne), Nei's genetic diversity (H), Shannon's information index (I), number of polymorphic loci (NPL) and percentage polymorphic loci (PPL) were also analyzed (Zhao *et al.*, 2006). The ILP data were subjected to a hierarchical analysis of molecular variance (AMOVA) (Excoffier *et al.*, 1992), using two hierarchical levels; among valleys and among genotypes within each valley. GenAlex software was used to calculate a principal coordinates analysis (PCA) that plots the relationship between distance matrix elements based on their first two principal coordinates (Peakall and Smouse, 2001).

Results and Discussion

Candidate ILP markers in *Artemisia annua* : ILP markers were applied to characterize and compare genetic diversity among two populations of *A. annua*. ESTs sequences of plant species are valuable sources of information for PCR-based gene-specific markers studies and gene mapping. However, as a result of unavailability of genomic DNA sequence in many non-model plants, the potential utility of PCR-based EST markers (in terms of their ability to generate high frequency of band polymorphism) has not been fully exploited. It is generally believed that intron regions are more divergent than exons (Choi *et al*., 2004). Therefore, the EST–PCR primers that are designed to anneal to exons, amplify across intron regions and resulted in relatively higher band polymorphism. In the current study, EST-PCR primers were designed based on the predicted exon/intron junction sites in *A. annua*. These splice sites were identified by aligning the ESTs of *A. annua* with the genomic sequence of *Arabidopsis*. Using BLASTN, out of 3,60,906 ESTs sequences from *A. annua* we detected 40,439 ESTs homologs with *Arabidopsis*. We successfully obtained 19,885 (73.86%) e-PCR products as putative ILPs in *A. annua* based on the EST-PCR primers designed. However, we failed to acquire e-PCR products from ~1/3 putative ILPs in *Artemisia*, probably due to the several constraint conditions set during primer designing and e-PCR. It is noted that the 19,885 ILP loci detected by e-PCR in *Artemisia* were located only on 13,254 ESTs. This result also reflects the nonrandom distribution of ILPs in *Artemisia* genome.

ILP analysis: Forty intron-flanking EST-primers were used for PCR amplification and subsequent polymorphism detection. Out of the 40 primers tested, 39 primers produced PCR amplification with the success rate of 98% (Table 1). To increase PCR specificity in the amplification of ILP loci, we adopted a Touchdown-PCR (Td-PCR) program. Most of the primers (23) amplified well by the Td-PCR with the default initial annealing temperature (59°C). Some other primers (8) required a lower (57°C) or higher (60°C) initial annealing temperature. The remaining 8 primers produced good amplification at a constant annealing temperature of 54 ^{o}C. A representative gel demonstrating the successful amplification of introns from the genomic DNA isolated from different genotypes is shown in Fig. 2. The genotypes collected from two different valleys were clearly distinguished and separated—revealed existence of higher genetic variations. All the amplified fragments varied in sizes from 200–500 bp. Out of 88 amplified bands, 81 (92.04%) were found polymorphic.

Phylogenetic analysis across twenty genotypes: Dendrogram in the present study shows very clear pattern of clustering for among population samples as well as within population in both the cases (Fig. 3). Similar results were obtained in azukibean (Fernandez *et al.,* 2002) and apricot (Kumar *et al.,* 2009). The dendrogram was generated by UPGMA cluster analysis considering 88 ILP bands amplified by 39 primers. Genotypes from Leh valley (L1-L10) were distinctly separated into a single cluster (Jaccard's similarity coefficient ranging from 0.5 to 0.78). Similarly, the genotypes from Nubra valley were distributed within a single cluster with Jaccard's similarity coefficient ranging from 0.47 to 0.78. The genotypes from Leh valley were more diverse in comparison to Nubra valley. The results of PCA analysis were comparable to the cluster analysis. All the genotypes from the Leh and Nubra valley were distributed within two distinct clusters (Fig. 4). It is clearly shown that the genotypes from the Leh valley (1 to 10) were widely separated from each other (the genetic variation among genotypes is high) in comparison to genotypes from the Nubra valley (11 to 20, genetic variation among genotypes is low).

Genetic diversity analysis: Relatively high genetic variation was detected among the *A. annua* genotypes. Genetic diversity analysis in terms of Na, Ne, H, I, Ht, Hs, NPL and PPL between both the valleys, revealed higher values for Leh valley indicating more variability among the genotypes in comparison to Nubra valley (Table 2).

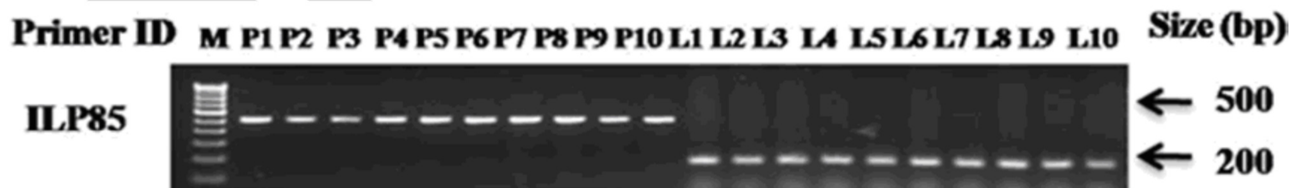

Fig. 2: ILP amplification products obtained from the 20 genotypes of *Artemisia annua* studied. L1 to L10 are the genotypes collected form Leh valley and P1 to P10 are the genotypes collected from Nubra (Partapur) valley. M = 100 bp λ DNA ladder. Representative gel showing polymorphism in genotypes between valleys with ILP 85 marker

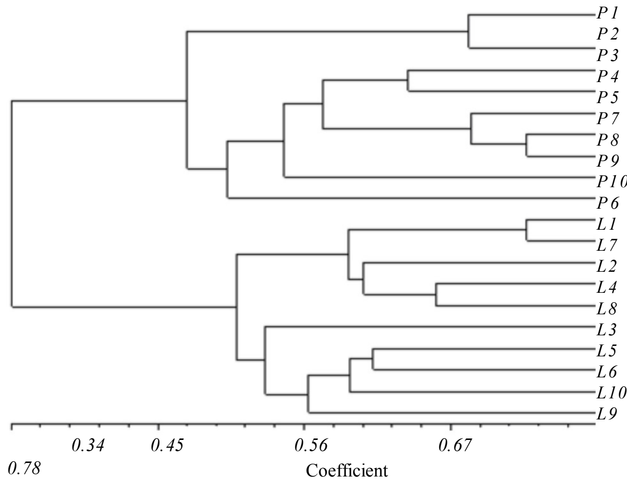

Fig. 3: Dendrogram illustrating genetic relationships among 20 genotypes of *Artemisia annua* collected from 2 different valleys: Leh and Nubra of Ladakh region. The dendrogram was generated by UPGMA cluster analysis calculated from 88 ILP bands produced by 39 primers

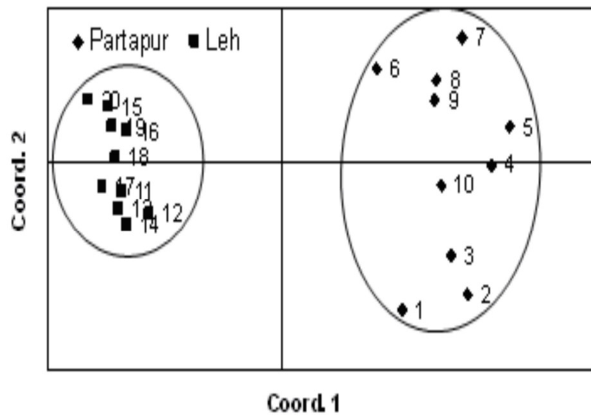

Fig. 4: Two-dimensional plot of principal component analysis of twenty local *Artemisia annua* using ILP analysis. The number plotted represents individual genotypes. The number 1 to 10 are from the Leh valley and from 11 to 20 are from Nubra valley

Table- 2: Summary of genetic variation statistics for all loci of ILPs among the *Artemisia annua* genotypes with respect to their distributions among two valleys

S.No.	Sample size	Na	Ne	H	I	Ht	NPL	PPL	Gst
Leh	10	1.7500±0.4359	1.5141±0.3583	0.2958±0.1899	0.4343±0.2688	0.2958±0.0361	57	75	1
Nubra	10	1.6184±0.4890	1.5171±0.4249	0.2792±0.2242	0.3975±0.3171	0.2792±0.0502	47	61.84	1

Na = Observed number of alleles; Ne = Effective number of alleles; H = Nei's gene diversity; I = Shannon's information index; Ht = Total genetic diversity; Gst = Genetic diversity between population; NPL = Number of polymorphic Loci; PPL = Percentage of polymorphic Loci

Table- 3: Overall genetic variability across all the 20 genotypes of *Artemisia annua* based on ILPs analysis

Na	Ne	H	I	Ht	Hs	Gst	NPL	PPL	Nm
1.934	1.747	0.405	0.581	0.405	0.288	0.290	71	93.42	0.612
(0.250)	(0.296)	(0.138)	(0.183)	(0.019)	(0.033)				

Hs = Genetic diversity in population; Nm= Estimate of gene flow from Gst; Nm = 0.25 (1-Gst)/Gst

Table- 4: Summary of nested analysis of molecular variance (AMOVA) based on ILPs, among the populations of *Artemisia annua.* Levels of significance are based on 1000 iteration steps

Source of variance	d.f	SSD	Variance component	Percentage	P-value
Among valley	1	89.25	7.711	39	< 0.01
Among genotypes/valley	18	218.5	12.139	61	< 0.01

The respective values for overall genetic variability for Na, Ne, H, I, Ht, Hs and Gene flow (Nm) across all the 20 genotypes are given in Table 3. AMOVA helps in partitioning of the overall ILP variations among valley and among genotypes within the valley. Molecular variance among valley (39 %) and among genotypes within valley (61%) (Table 4) revealed higher variations within the populations. All the components of molecular variations were significant (P < 0.01). The occurrence of more genetic variance of *A. annua* within population as reported in our study is a pattern often described for mountain plants (Cotrim *et al.,* 2003), for endemic plants from Tibet (Chen *et al.,* 2005) and from central Asian desert plants (Ge *et al.,* 2003) and several other studies of out crossing endemic species (Jaquemyn et al., 2004; Juan *et al.,* 2004). Similar reports have been made in ISSR studies of populations of *R. crenulata* from Hengduan Mountain Region, China, where the within population diversity was 55.14% while among population diversity was 44.8 (Lei *et al.*, 2006). This is helpful in making strategy for germplasm collection and evaluation. The rate of gene flow estimated using Gst value was found to be 0.612. In general, dispersal resulting in colonization and gene flow into existing populations is very important for both the persistence and genetic success of a species (Hamrick and Godt, 1996). In population genetics, a value of a gene flow (Nm) < 1.0 (less than one migrant per generation into a population) or, equivalently, a value of gene differentiation (Gst) > 0.25 is generally regarded as the threshold quantity beyond which significant population differentiation occurs (Slatin, 1987).

The study indicates that *A. annua* populations in the north-western Himalayan region are genetically highly diverse which should be exploited for future conservation and breeding of *Artemisia* from

this region. Genetic variation among elite genotypes of *A. annua* based on ILP analysis could be useful in selecting parents to be crossed for generating appropriate populations intended for both genome mapping and breeding purposes. Considering the high genetic differentiation among the wild populations of *A. annua*, conservation of only a few populations may not adequately protect the genetic variations within the species in the northwestern Himalayan region. This species, or at least a large part of its genetic diversity, may be lost in the near future, owing to its importance and consequent exploitations as a medicinal plant, if appropriate conservation measures will not be adopted. Since no single, or even a few plants, will represent the whole genetic variability in *A. annua*, there appears to be a need to maintain a sufficiently large population in natural habitats to conserve genetic diversity in *A. annua* to avoid genetic erosion. Based on polymorphic feature, genetic diversity, genetic similarity, and gene flow among the populations of *Artemisia* based on ILP study, we recommend that any future conservation plans for this species should be specifically designed to include representative populations with the highest genetic variation for both *in situ* conservation and germplasm collection expeditions.

Acknowledgments

The authors are thankful to Defense Research and Development Organization HQ for its support and funding the project. We are greatly indebted to the anonymous reviewers of this manuscript for extremely helpful suggestions.

References

Becker, J. and M. Heun: Barley microsatellites: Allele variation and mapping. *Plant Mol. Biol.*, **27**, 835–845 (1995).

Botstein, D., R.L. White, M. Skolnick and R.W. Davis: Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American J. Human Genet.*, **32**, 314–331 (1980).

Chen, S., T. Xia, S. Chen and Y. Zhou: RAPD profiling in detecting genetic variation in endemic *Coelonema* (Brassicaceae) of Qinghai-Tibet Plateau of China. *Biochem. Genet*., **43**, 189-201 (2005).

Choi, H.K., D. Kim, T. Uhm, E. Limpens, H. Lim, J.H. Mun, P. Kalo, R.V. Penmetsa, A. Seres and O. Kulikova: A sequence-based genetic map of *Medicago truncatula* and comparison of marker colinearity with *M. sativa*. *Genetics*, **166**, 1463–1502 (2004).

Cotrim, H.C., M.W. Chase and M.S. Pais: *Silene rothmaleri* (Caryophyllaceae), a rare, fragmented but genetically diverse species. *Biodiversity Conserv.*, **12**, 1083-1098 (2003).

Deutsch, M. and M. Long: Intron–exon structures of eukaryotic model organisms. *Nucleic Acids Res.*, **27**, 3219–3228 (1999).

Excoffier, L., P.E. Smouse and J.M. Quattro: Analyses of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479-491 (1992).

Fernandez, M.E., A.M. Figueiras and C. Benito: The use of ISSR and RAPD marker for detecting DNA polymorphism, genotype identification and genetic diversity among barley cultivars with known origin. *Theor. Appl. Genet.*, **104**, 845-851(2002).

Ge, X.Y., Y. Yu, N.X. Zhao, H.S. Chen and W.Q. Qi: Genetic variation in the endangered inner Mangolia endemic shrub *Tetraena mongolica* (Zygophyllaceae). *Biol. Conserv*., **111**, 427-434 (2003).

Hamrick, J.L. and M.J.W. Godt: Conservation genetics of endemic plant species. In: Conservation genetics-case histories from nature (Eds.: J.L. Avise and J.L. Hamrick). Chapman & Hall, London, 281–304 (1996).

Hawkins, J.D.: A survey on intron and exon lengths. *Nucleic Acids Res.*, **16**, 9893–9908 (1988).

Klayman, D.L.: *Artemisia annua* I: From weed to respectable antimalarial plant. In: Human medicinal agents from plants (Eds.: A.D. Kinghorn and M.F. Balandrin). Am. Chem. Soc. Symp. Ser. Washington D.C., 242–255 (1993).

Jacqemyn, H., O. Honnay and P. Galbusera: Genetic structure of the forest herb *Primula elatior* in a changing landscape. *Mol. Ecol*., **13**, 211-219 (2004).

Juan, A., M.B. Crespo and R.S. Cowan: Patterns of variability and gene flow in *Medicago citrine*, an endangered endemic of islands in the western Mediterranean, as revealed by amplified fragment length polymorphism (AFLP*). Mol. Ecol.*, **13**, 2679-2690 (2004).

Kumar, M., G.P. Mishra, R. Singh, J. Kumar, P.K. Naik and S.B. Singh: Correspondence of ISSR and RAPD markers for comparative analysis of genetic diversity among different apricot genotypes from cold arid deserts of trans-Himalayas. *Physiol. Mol. Biol. Plants*, **15**, 25-236 (2009).

Lei, H.G., T. Tashi, S. Suhua and Z. Yang: Determination of genetic variation in *Rhodiola crenulata* from the Hengduan mountains region, China using inter simple sequence repeats. *Yidong Genet. Mol. Biol*., **29**, 339-344 (2006).

Palumbi, S.R.: Nucleic acids II: The polymerase chain reaction, In: Molecular Systematics (Eds.: D. Hillis and C. Moritz). 2nd Edn. Sinauer Associates Inc., Sunderland, MA, pp. 205–247 (1995).

Peakall, R. and P.E. Smouse: GenAlEx V5: Genetic Analysis in Excel. Population genetic software for teaching and research. Australian National University,Canberra, Australia (2001).

Phillipson, D.J. and C.W. Wright: Antiprotozoal agents from plant sources. *Plant Medica*, **57**, 553–559 (1991).

Rozen, S. and H. Skaletsky: Primer on the WWW for general users and for biologist programmers. *Methods Mol. Biol.*, **132**, 365–386 (2000).

Saghai-Maroof, M.A., K.M. Soliman, R.A. Jorgensen and R.W. Allard: Ribosomal spacer length in barley: Mendelian inheritance, chromosomal location and population dynamics. *Proc. Natl. Acad. Sci. USA*, **81**, 8104-8118 (1984).

Slatin, M.: Gene flow and geographic structure of natural populations. *Science*, **236**, 787–792 (1987).

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper and M. Zabeau: AFLP: A new technique for DNA fingerprinting. *Nucleic Acid Res.*, **23**, 4407–4414 (1995).

Wang, D.G., J.B. Fan, C.J. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins, E. Winchester, J. Spencer, L. Kruglyak, L. Stein, L. Hsie, T. Topaloglou, E. Hubbell, E. Robinson, M. Mittmann, M.S. Morris, N. Shen, D. Kilburn, J. Rioux, C. Nusbaum, S. Rozen, T.J. Hudson, R. Lipshutz, M. Chee, E.S. Lander: Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science*, **280**, 1077–1082 (1998).

Weising, K., G. Atkinson and C. Gardner: Genomic fingerprinting by microsatellite-primed PCR: A critical evaluation. *PCR Methods and Appl.*, **4**, 249-255 (1995).

Williams, J.G.K., A.R. Kubelik, K.J. Livak, J.A. Rafalski and S.V. Tingey: DNA polymorphism amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res.*, **18**, 6531–6535 (1990).

Zhao, W.G., J.Q. Zhang, Y.H. Wangi, T.T. Chen, Y. Yin, Y.P. Huang, Y. Pan and Y. Yang: Analysis of genetic diversity in wild populations of mulberry from western part of Northeast China determined by ISSR markers. *J. Genet. Mol. Biol.*, **7**, 196–203 (2006).

Online Copy