Genetic Analysis of *Artemisia annua* with the Help of Intron Length Polymorphism Marker

By

EKTA (061508)

NEETI POKHREL (061527)

Submitted in partial fulfillment of the Degree of Bachelor of Technology

DEPARTMENT OF
BIOTECHNOLOGY & BIOINFORMATICS
JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY
WAKNAGHAT, SOLAN, HP, INDIA

Dr. Pradeep Kumar Naik
Senior Lecturer
Biotechnology & Bioinformatics

Jaypee University of
Information Technology
Waknaghat-173 215
Solan, Himachal Pradesh
Phone No.: 91-1792-239227
Fax No.: 91-1792-245362

CERTIFICATE

This is to certify that the thesis entitled, "Genetic diversity of different cultivars of Artemisia annua using ILP markers" submitted by Ekta and Neeti Pokhrel in partial fulfillment for the award of Bachelor of technology in bioimformatics of Jaypee University of Information Technology has been carried out under my supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of this or any other degree or diploma.

Place : Waknaghat
Date : May,2010

(Dr. Pradeep Kumar Naik)
Project Guide

ACKNOWLEDGEMENTS

The success of our research would never have happened without the support and valuable advice from many people who contributed to our valuable experience.

We would like to express our sincere gratitude and deep appreciation to our advisor, Senior Lecturer, Dr. P.K Naik for his guidance, invaluable advice, supervision, kindness and encouragement throughout this study.

Our sincere appreciation is also extended to Dr. Pradeep Naik, our co-advisor for his helpful criticism, suggestions and for his constructive comments.

A special acknowledgment is extended to the Faculty of Biotechnology, Jaypee University of Information Technology (JUIT), Waknaghat for giving us the opportunity to study the Bachelor of Teachnology in Biotechnology, Department of Biotechnology and JUIT for providing us with the best research facilities. Our great appreciation is also extended towards our Head of Department, Dr. R. S. Chauhan.

A special thanks to our Vice-Chancellor, Dr. Y. Medury and Registrar, Brig. (Retired) Balbir Singh, for providing us with the right environment to work.

We would also like to express out our thanks to all our friends for their warm
friendship and support and assistant until we finished our study.

Finally, we would like to express our deepest gratitude to our dearest parents, our family for their encouragement, understanding, care, love and continuous support throughout our life.

Place : Waknaghat, (HP)
Dated : June,2007

Ekta
061508

Neeti Pokhrel
061527

INDEX

LIST OF FIGURES

List of tables

ABSTRACT

The present study was undertaken with the aim of molecular identification and biochemical characterization of different genotypes of *Artemisia annua* using intron length polymorphic DNA (ILP) and Artemisinin estimation with the help of GAS chromatography. The phylogenetic relationships of 20 *artemisia annua* genotypes which are collected from two valleys *viz.* Partappur (9,600 ft) and Leh (11,500 ft) of trans-Himalayan region were analyzed using 40 PCR markers. This is the first report of molecular genetic diversity studies in *Artemisia annua* from this region of the world. ILP analysis yielded 74 fragments, of which 69 were polymorphic, with an average of 7.4 polymorphic fragments per primer. ILP markers were found to be efficient with regards to polymorphism detection, as they detected 93.74 % of polymorphism. The results of PCA analysis were comparable to the cluster analysis. These analyses, allowed us to identify the groups corresponding to the *Artemisia annua* collection sites.

CHAPTER 1

INTRODUCTION

Artemisia annua L. belongs to the family Asteraceae. It comprises hardy herbs and shrubs known for their volatile oils. The genus *Artemisia* classification is as follows:

- Kingdom → Plantae
- Subkingdom → Tracheobionta
- Superdivision → Spermatophyta
- Division → Magnoliophyta
- Class → Magnoliopsida
- Subclass → Asteridae
- Order → Asterales
- Family → Asteraceae
- Genus → *Artemisia* L.
- Species → *Artemisia annua* L.

It is a diploid plant with chromosome number, 2n=18. Artemisia annua is an annual herb native to Asia and China. Yields can be maximised at higher *altitudes* . Artemisia thrives well in dry cold region of Ladhak having marginal rocky and sandy soils. It is well adapted to survive at high altitudes of 2,500-6,000 m above mean sea level (MSL) and the temperature, nutrient and environmental stress that they are subjected to under the cold arid conditions. . Traditionally, the plant is used for crafting of aromatic wreaths, flavouring of spirits, and its essential oil is distilled for perfumery and industrial use (Ferreira et al., 1996). The essential oil also possesses strong insecticidal potential (Charles et al., 1991). Besides the volatile terpenoids, the plant synthesizes and accumulates substantial quantities of many derivatives of a cadiene skeleton including artemisinin (an endoperoxide seco sesquiterpene lactone) which is produced by the artemisinin pathway and mevalonate pathway (figure 1 and 2). Artemisinin is currently the most effective agent against multidrug resistant strains of *Plasmodium* species, the malarial parasites (Phillipson and Wright, 1991; Klayman, 1993). Due to this discovery, *Artemisia annua* is now rated as one of the top ten industrial crops of the modern world.

The leaves of Artemisia annua (fig 3 and 4) has many medicinal properties such as Antibacterial, Antiperiodic, Antiseptic, Carminative, Digestive, Febrifuge. An aromatic anti-bacterial plant.It is often used in the Tropics as an affordable and effective anti-malarial. The plant contains artemisinin, this substance has proved to be a dramatically effective anti-malarial against multi-drug resistant Plasmodium species (Phillipson and Wright, 1991; Klayman, 1993). The seeds are used in the treatment of flatulence, indigestion and night sweats.

Over the year, the Artemisia populations in the Ladhak region have developed considerable variability. No reports are available so far regarding the genetic characterization of this plant from Ladhak region and detailed investigation is required. Moreover, the analysis of genetic diversity and relatedness between or within different genotypes is a prerequisite towards effective utilization and protection of plant genetic resources (Weising et al., 1995). It will then be possible to achieve rational conservation, and the identification of diagnostic or agronomic traits linked to molecular markers. In principle, genetic uniqueness is brought about by two factors, inheritance and mutations and since all genetic differences between individuals are laid down in the primary sequence of their genomic DNA, the most straight forward method is identifying an individual sequence for genomes under comparison (Krawczak and Schmidtke, 1994).

Unlike the morphological and biochemical markers which may be affected by environmental factors and growth practices (Xiao et al., 1996; Ovesna et al., 2002; Higgins 1984), DNA markers portray genome sequence composition, thus, enabling to detect differences in the genetic information carried by the different individuals. A wide variety of DNA-based markers have been developed in the past few years. Restriction fragment length polymorphism (RFLP) was the first molecular marker (Bostein et al., 1980), generated for genome analysis and mapping. However, the development of the polymerase chain reaction (PCR) technology has introduced a considerable number of useful molecular markers, e.g., random amplified polymorphic DNA (RAPDs) (Williams et al., 1990) and inter simple sequence repeats (ISSR) (Zeitkiewicz et al., 1994) which are independent of environmental factors and unaffected by developmental stages of the plant. Currently, there are several methods of DNA isolation available as variants of the few principal protocols (Dellaporta et al., 1984; Saghai-Maroof et al, 1984; Doyle and Doyle, 1987; Webb and Knapp, 1990). Plant protocols need to be tailored to each plant species (and sometimes even each tissue) due to the presence of secondary metabolites that vary in nature and/or composition. These phytochemicals may not only hinder the *su moto* application of the other methods, but they may also interfere with subsequent amplification and/or restriction-digestion of the isolated DNA. Sometimes, the contaminants are not detectable but are sufficiently high to interfere in the analysis of polymorphisms. Particularly, medicinal and aromatic plants like *Artemisia annua* are rich in the myriad of natural products and they require additional efforts to find an appropriate procedure for reliable and consistent results.

The objectives of this study also to compare the effectiveness of both the PCR-based molecular approaches to determine the genetic relationships among several genotypes of *Artemisia annua*

from two valleys i.e. Leh and Partapur of Ladhak region in India. This project also aims at identifying the high and low artemisinin content genotypes.The objective of the present work is to

1) Standardization of the protocol for ILP analysis in Artemisia annua
2) Study of genetic polymorphism of Artemisia annua population from different altitudes.
3) Quantification of secondary metabolites.

NOVELTY

This is the first reported protocol for isolation of DNA from the metabolite-rich species like *Artemisia annua* as well as genetic characterizations of Artemisia from the trans-Himalayan region with ILPs molecular marker.

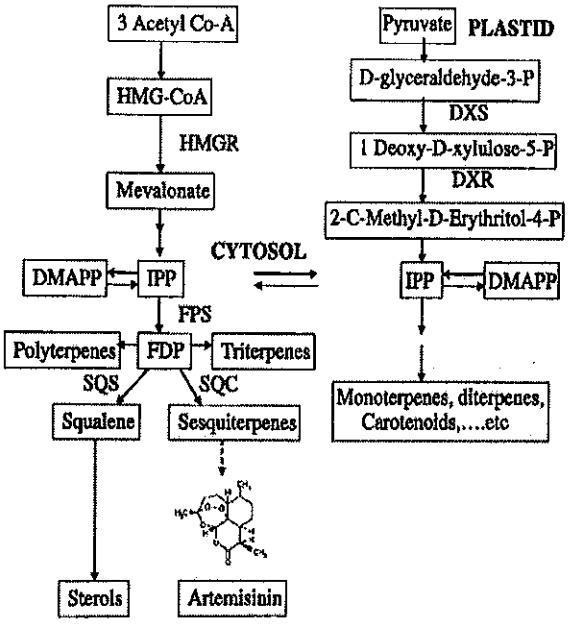


Fig. 1 Mevalonate pathway

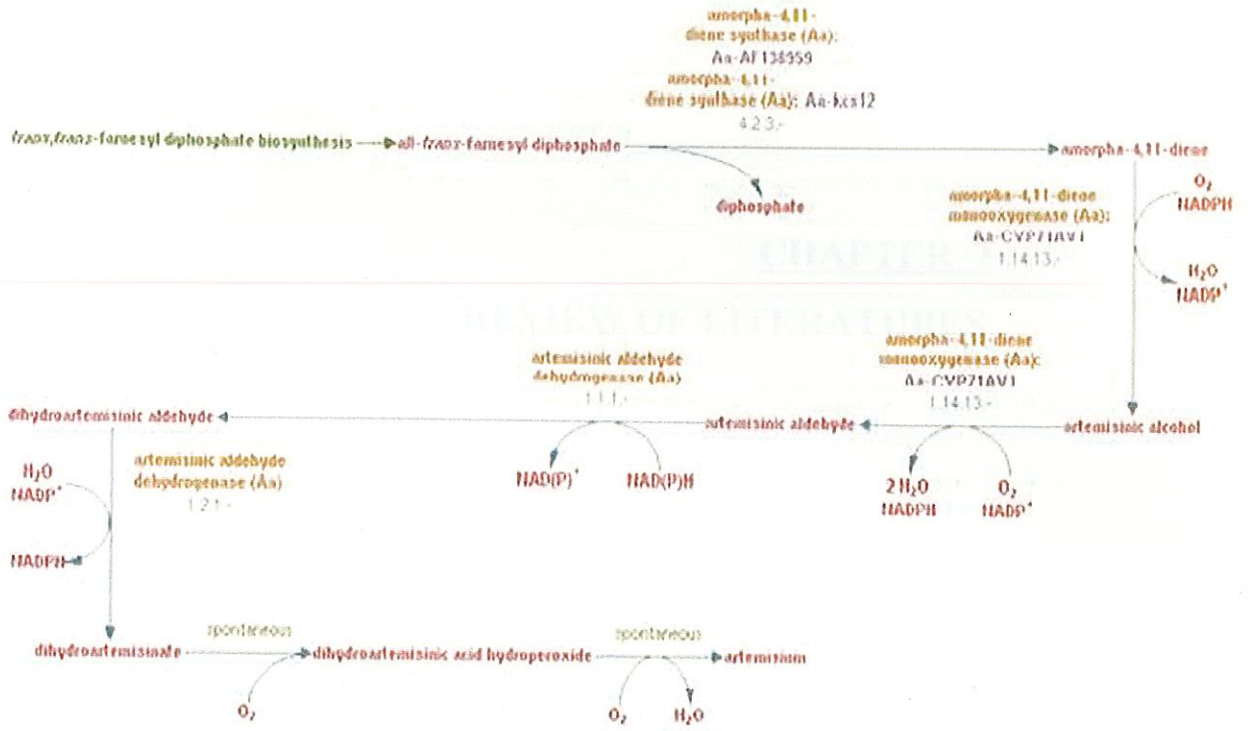


Fig.2 Artemisinin pathway

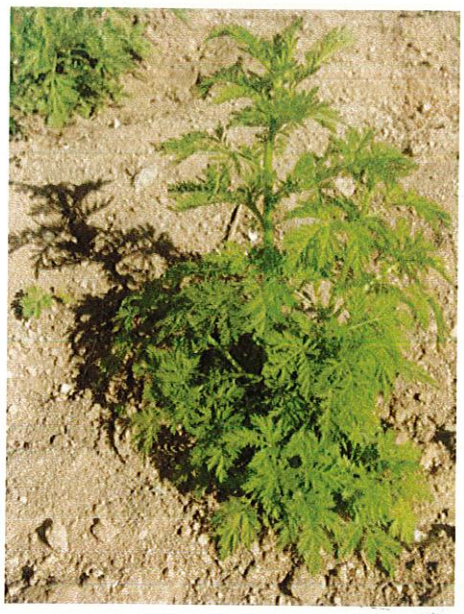


Fig. 3 Artemisia annua leaves is anti nature.

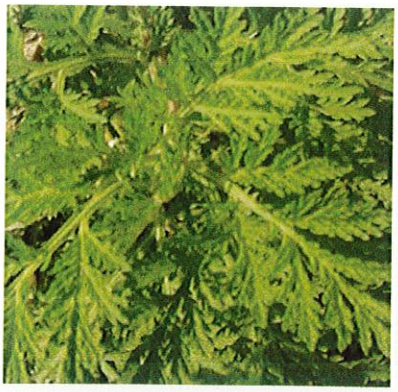


Fig. 4 *Artemisia annua* leaves malarial in

CHAPTER 2

REVIEW OF LITERATURES

Molecular markers are powerful tools for genetic research and breeding. Many types of molecular marker 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), single-nucleotide polymorphism (SNP; Kruglyak, 1997) and intron length polymorphism (ILP; Choi *et al.*, 2004).

2.1 ILPs

Introns are noncoding sequences interspersed in genes. In comparison with exons, introns are more variable because in general selective pressure in intronic regions is much less than exonic regions. For example, the average number of SNPs per 1000 bp in introns is over three times as high as that in exons among eight varieties in rice (*Oryza sativa* L.; Feltus *et al.*, 2006). Recently, Fredslund (Fredslund *et al.*, 2006) developed a web program GeMprospector that allows automatically designing cross-species candidate ILP markers in legumes or grasses. Length polymorphism is the most intuitive variation in introns. So, ILPs have been exploited as molecular markers, which have many desirable properties, including specific, codominant, neutral, convenient and reliable. However, ILP has not been widely utilized because it is newly developed as a kind of molecular marker and the number of ILP markers having been exploited in plants is still very limited. To date, studies of exploiting ILP markers have been restricted to a few species (Choi *et al.*, 2004; Wang *et al.*, 2005; Wei *et al.*, 2005). ILP is detected by PCR with primers designed on exons flanking the target intron. Obviously, the key point of developing ILP markers is to identify suitable introns. A general method for identifying introns is to compare cDNA/EST sequences with the genome sequence. Therefore, introns can be easily identified in many model organisms. However, this method is not applicable to most other organisms because they have only cDNA/EST sequences available. Fortunately, studies have indicated that the exon-intron structures are largely conserved among homologous genes from different species (Batzoglou *et al.*, 2000). Therefore, the joints between adjacent exons after splicing (or termed 'intron positions' for short) in a cDNA/EST of an organism can be deduced according to the homologous genes from related model organisms. This provides a way of developing ILP markers in any organisms. Nowadays, many organisms have got a large number of cDNA/EST sequences available in public databases. Therefore, large-scale exploitation of ILP markers in various organisms becomes possible.

The main point of their work is to develop conserved ILP primers that are usable in different species of legumes or grasses. Such cross-species ILP markers would be useful for the research of comparative genomics. However, the requirement of primer conservation would probably make many ILPs not exploitable as markers. There could be two reasons: (i) the flanking exon sequences of some introns are not conserved enough for designing cross-species primers and (ii) many ESTs do not have homologs available in multiple species at present so that the conserved regions in these ESTs for designing primers cannot be identified.

Introns are non-coding sequences in a gene that are transcribed but spliced out of the precursor mRNA. Introns are widespread and abundant in eukaryotic genomes. For example, introns constitute ~11 and 24% of the fruit fly and human genomes, respectively. Generally speaking, introns have little functional significance, although some introns may influence the level of gene expression.Therefore, introns are more variable than coding sequences

Intron polymorphisms can also be exploited as genetic markers. They have been successfully utilized in population genetics surveys and gene mapping. There could be various polymorphisms in introns, but intron length polymorphism (ILP) is the most easily recognizable type. It can be conveniently detected by the PCR. To amplify introns by PCR, primers can be designed in flanking exons. This approach is called exon-primed intron-crossing PCR (EPIC-PCR). The advantage of EPIC-PCR is that exon sequences are relatively more conservative and therefore the primers designed in exons may have more extensive applications than those designed in non-coding sequences. Using this approach, Bierne et al. developed several ILP markers in penaeid shrimps. To date, however, studies of exploiting intron polymorphism markers have been restricted to a few genes. No efforts on genome-wide exploitation of intron polymorphism markers have been reported

ILPs are caused by InDels, but many of them cannot be simply considered equivalent to the generally defined InDels because an ILP may contain several (instead of only one) InDels. In addition, even if we take all ILPs as InDels, they are at least a special subset of InDels exhibiting polymorphisms in the non-coding regions of genes. This may make ILPs possess special characteristics and usefulness. In the work described here, we performed a genome-wide search for ILPs and a large-scale exploitation of candidate ILP markers via electronic EPIC-PCR based on the released genomic and cDNA sequence data in rice. A set of ILP markers selected from the

candidates was developed and investigated their characteristics by experiment. Moreover, a web-accessible database was established for rice ILP marker.

We got the idea from the following research papers :-

Marc Ohresser, Philippe Borsa, Claude Delsert published a paper in which a novel intron-length polymorphism at the actin gene locus mac-1 is reported and used as a genetic marker for population studies in mussels of the genus Mytilus.

Hui Wei (Department of Horticulture, Iowa State University, Ames, IA 50011), Yan Fu (Department of Genetics and Developmental and Cellular Biology, Iowa State University, Ames, IA 50011), Rajeev Arora (Department of Horticulture, Iowa State University, Ames, IA 50011). Intron length polymorphisms were used to investigate relationships among eight Rhododendron L. species.

Xusheng Wang,[y] Xiangqian Zhao, Jun Zhu, and Weiren Wu.(Institute of Bioinformatics, Huajiachi Campus, Zhejiang University, Hangzhou, 310029, P. R. China) 23 February 2006. A genome-wide search of ILPs between two subspecies (indica and japonica) in rice using the draft genomic sequences of cultivars 93-11 (indica) and Nipponbare (japonica) and 32 127 full-length cDNA sequences of Nipponbare obtained from public databases. By testing 51 pairs of the ILP primers on five Gramineae plants and three dicot plants, we found another desirable characteristic of rice ILP markers that they have high transferability to other plants.

2.2 ARTEMESIA ANNUA

N.S. Sangwan, R.S. Sangwan and Sushil Kumar. A protocol is described to obtain high molecular weight, restrictable, and amplifiable genomic DNA from the antimalarial plant *Artemisia annua.* The method is a CTAB (cetyl-trimethylammonium) procedure that includes a rapid micro-column chromatography through DE-52 ion-exchange resin. The template should be pure enough for PCR and RFLP based applications. Restriction of the purified DNA with *Dra* I and *Hin*d III resulted in discrete bands that may have arisen from repeat sequences equally spaced between the restriction site(s). This approach might be useful for direct identification of repeat sequences as putative species-specific probes.

A review by Olliaro PL, involved developing artemisinin based drug combinations for the treatment of drug resistant falciparum malaria.The emergence and spread of drug resistant malaria represents a considerable challenge to controlling malaria. To date, malaria control has relied heavily on a comparatively small number of chemically related drugs, belonging to either the quinoline or the antifolate groups. Only recently have the artemisinin derivatives been used but mostly in south east Asia. Controlling resistance is key to ensuring that the investment put into developing new antimalarial drugs is not wasted. Current efforts focus on research into new compounds with novel mechanisms of action, and on measures to prevent or delay resistance when drugs are introduced. Drug discovery and development are long, risky and costly ventures. Antimalarial drug development has traditionally been slow but now various private and public institutions are at work to discover and develop new compounds. Today, the antimalarial development pipeline is looking reasonably healthy. Most development relies on the quinoline, antifolate and artemisinin compounds. There is a pressing need to have effective, easy to use, affordable drugs that will last a long time. Drug combinations that have independent modes of action are seen as a way of enhancing efficacy while ensuring mutual protection against resistance. Most research work has focused on the use of artesunate combined with currently used standard drugs, namely, mefloquine, amodiaquine, sulfadoxine/pyrimethamine, and chloroquine. There is clear evidence that combinations improve efficacy without increasing toxicity. However, the absolute cure rates that are achieved by combinations vary widely and depend on the level of resistance of the standard drug. From these studies, further work is underway to produce fixed dose combinations that will be packaged in blister packs. This review will summarise current antimalarial drug developments and outline recent clinical research that aims to bring artemisinin based combinations to those that need them most.

Delabays N, Simonnet X, Gaudin M (Federal Agricultural Research Station For Plant Production, Nyon, CH-1960, Switzerland).Artemisinin, the endoperoxide sesquiterpene lactone produced by the Chinese medicinal herb Artemisia annua, is very difficult to synthesise. Moreover, its production by mean of cell, tissue or organ cultures is very low. Presently, only its extraction from cultivated plants is viable. A large variation in artemisinin content has been observed in the leaves of plants from different origins. The genetic basis of this variation has been assessed and evidence for a quantitative inheritance of the artemisinin concentration presented. Additive genetic components were predominant, resulting in a high narrow-sense heritability estimate.

Biochem Mol Biol Int. (1999 June) Artemisia annua as a self-reliant treatment for malaria in developing countries by de Ridder S, van der Kooy F, Verpoorte R. The antimalarial chloroquine, which was in the past a mainstay of malaria control, is now ineffective in most malaria areas and resistance to other antimalarials is also increasing rapidly. The discovery and development of artemisinins from Artemisia annua have provided a new class of highly effective antimalarials. ACTs are now generally considered as the best current treatment for uncomplicated Plasmodium falciparum malaria. This review gives a short history of the malaria disease, the people forming a high risk group and the botanical aspects of A. annua. Furthermore the review provides an insight in the use of ART and its derivatives for the treatment of malaria. Its mechanism of action and kinetics will be described as well as the possibilities for a self-reliant treatment will be revealed. This self-reliant treatment includes the local production practices of A. annua followed by the possibilities for using traditional prepared teas from A. annua as an effective treatment for malaria. Finally, HMM will be described and the advantages and disadvantages discussed.

Abdin MZ, Israr M, Rehman RU, Jain SK. Artemisinin, a sesquiterpene lactone containing an endoperoxide bridge, has been isolated from the aerial parts of Artemisia annua L. plants. It is effective against both drug-resistant and cerebral malaria-causing strains of Plasmodium falciparum. The relatively low yield (0.01-0.8 %) of artemisinin in A. annua is a serious limitation to the commercialization of the drug. Therefore, the enhanced production of artemisinin either in cell/tissue culture or in the whole plant of A. annua is highly desirable. It can be achieved by a better understanding of the biochemical pathway leading to the synthesis of artemisinin and its regulation by both exogenous and endogenous factors. Furthermore, genetic engineering tools can be employed to overexpress gene(s) coding for enzyme(s) associated with the rate limiting step(s) of artemisinin biosynthesis or to inhibit the enzyme(s) of other pathway

competing for its precursors. These aspects which may be employed to enhance the yield of artemisinin both in vitro and in vivo are discussed in this review.

Liu C, Zhao Y, Wang Y. Artemisinin isolated from the aerial parts of Artemisia annua L. is a promising and potent antimalarial drug which has a remarkable activity against chloroquine-resistant and chloroquine-sensitive strains of Plasmodium falciparum, and is useful in treatment of cerebral malaria. Because the low content (0.01-1 %) of artemisinin in A. annua is a limitation to the commercial production of the drug, many research groups have been focusing their researches on enhancing the production of artemisinin in tissue culture or in the whole plant of A. annua. This review mainly focuses on the progresses made in the production of artemisinin from A. annua by biotechnological strategies including in vitro tissue culture, metabolic regulation of artemisinin biosynthesis, genetic engineering, and bioreactor technology.

Since 1960s and 1970s. Chinese scientists have put considerable effort and resources into the search for new antimalarial compounds extracted from Chinese traditional herbs. Archaeological findings indicate that qinghao (Artemisia annua L.) has been used as a traditional remedy in China for over two thousand years. Its antimalarial principle was finally isolated in 1971 and named artemisinin or qinghaosu (meaning the principle of qinghao in Chinese). Its rapid action, low toxicity and powerful effect against falciparum malaria made it a favored subject for research. In 1976, the unique structure of the molecule, characterized by an endoperoxide and an alternative O-C-O-C segment, was identified. The Subsequent studies of the structure/activity relationship led to the discovery of dihydroartemisinin, artemether and artesunate. Now qinghaosu and these three derivatives are being used around the world as effective new antimalarial drugs in the fight against falciparum malaria, including multi-drug-resistant Plasmodium falciparum. At the present time new qinghaosu analogues or derivatives are being developed and studies of their structure/activity relationships, their antimalarial mechanisms, their interaction with ferrous ions and the DNA damage associated with these processes are being actively pursued. Research into qinghaosu and its derivatives has already produced and will no doubt continue to produce results of the utmost importance in the fight against malaria and other diseases.

Quantitation of artemisinin and its biosynthetic precursors in *Artemisia annua* L. by high performance liquid chromatography–electrospray quadrupole time-of-flight tandem mass spectrometry. Artemisinins are derived from extracts of sweet wormwood (*Artemisia annua*) and are well established for the treatment of malaria, including highly drug-resistant strains. Their

efficacy also extends to phylogenetically unrelated parasitic infections such as schistosomiasis. In this review, recent advances in defining the role of artemisinins in medicine, with particular focus on their controversial mechanisms of action. This safe and cheap drug class that saves lives at risk from malaria can also have important potential in oncology. Artemisinin and its derivatives are a potent new class of antimalarials, originated from *Artemisia annua, L.* The clinical efficacy of these drugs is characterized by an almost immediate onset and rapid reduction of parasitaemia. Their efficacy is high in such areas as well where multidrug-resistance is rampant, but in these areas, their combination with other (effective) antimalarials (e.g., mefloquine) is highly recommended. In this short review, the chemical structures, pharmacological properties, and clinical uses of artemisinin drugs are discussed.

CHAPTER 3

MATERIALS AND METHODS

The present investigation was undertaken with a view to study the molecular diversity of twenty different genotypes obtained from two different regions of India. The details of materials used and methods adopted in the present investigation are described hereunder.

3.1 Sources of sequence data

The dicot model plant Arabidopsis and monocot model plant rice were taken as subject species; whereas the available EST sequence data from *Artemisia annua* were taken as query species. The genome, cDNA and CDS (coding sequence) data of rice (Oryza sativa L. ssp. japonica cv. Nipponbare) and Arabidopsis (Arabidopsis thaliana ecotype Columbia) were downloaded from http:// www.tigr.org/ and http://www.arabidopsis.org/, respectively. The EST sequences of *A. annua* (~85,282) were downloaded from http://www.plantgdb.org/. The redundant ESTs sequences were removed using BLASTCLUST with sequence similarity 60% and 100% sequence coverage. After removing the redundant ESTs the total numbers of ESTs left were 68,974.

3.2 Development of putative intron polymorphic (PIP) markers

An in-house pipeline Perl script was used to predict PIP markers from Artemisia ESTs. The procedure consisted of three steps. The first step was to identify intron positions and lengths in each subject species by aligning its CDSs with its genome sequence using program SIM4. A threshold of 100% identity was used in the alignment. The second step (Fig. 5) was to identify possible intron positions in each query species by aligning query EST sequences with subject CDSs using BLASTN. A query EST was thought to be homologous to a subject CDS only if there were at least 100 bp overlapping and 80% similarity between them. The third step (Fig. 5) was to design primers for those query ESTs containing possible positions of introns. For each of those query ESTs, a pair of primers was designed using program Primer3 from the query EST with 100 bp on each side of the target intron. The designed primers were tested by electronic PCR (e-PCR) on the EST sequences of corresponding species. A putative intron was taken as a PIP marker if the e-PCR yielded the unique product as expected. To increase the quality and usability of the in silico exploited ILP markers, we required exact matches between primers and templates and set a high margin value on the product size for the e-PCR. We took a putative ILP locus as a candidate ILP marker when it was successfully and uniquely detected by the e-PCR.

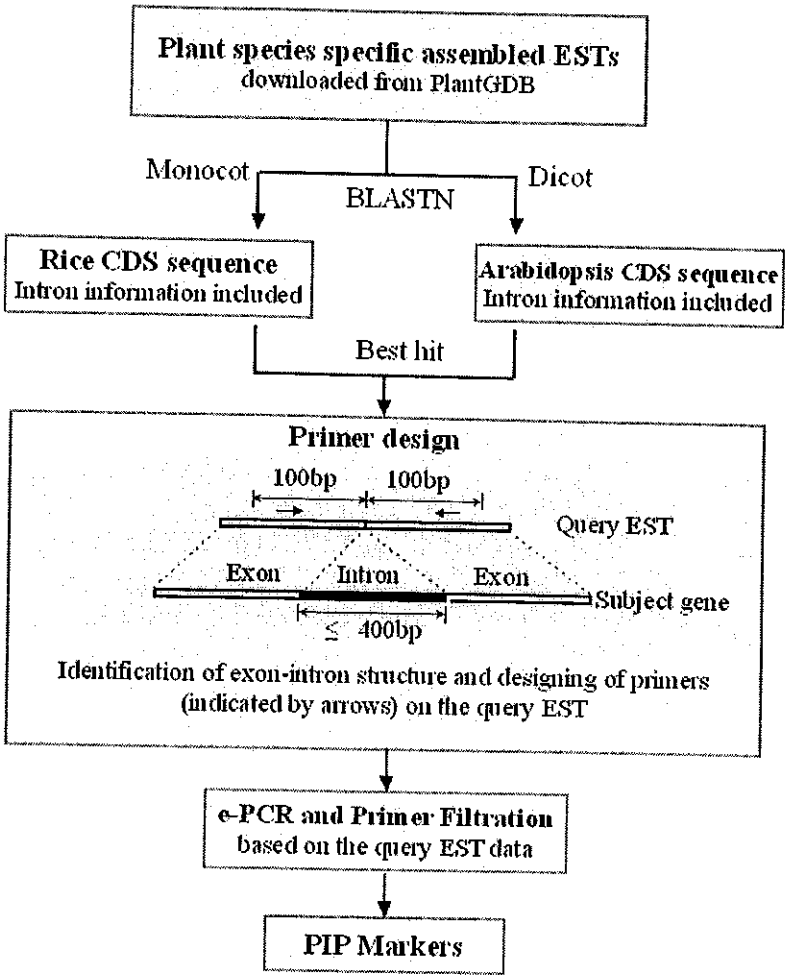


Figure 5. Protocol for developing ILP markers from ESTs sequences of *Artemisia annua*.

S.NO	GENE NAME	PRIMER SEQUENCE	G+C	Tm
i1L	aasqs	3' GTCGTTTGGGCATTGGGATCATGC 5'	54%	58
i1R		5' TTGAGACGTGGTCTAACTGCCA 3'	50%	54
i2L		3' ATTTCAGTTTTGGCCTCGTG 5'	45%	49
i2R		5' CGTATCACCGCGAGATTCTT 3'	50%	51
i3L	ads	3' CTAAGGATGCCTAACCGCCGCA 5'	59.10%	58
i3R		5' AGCTTGAGCCTACCTTCGCACAC 3'	56.52%	58
i4L	ads	3' AGGCTCAAGCTAAGATTCTGCTCG 5'	50%	57
i4R		5' GAGGATCGCTAATCTCGTCTGG 3'	54.55%	56
i5L	ads	3' TGACACCTGCCCGGTGCTGGAA 5'	63.63%	60
i5R		5' TGATCGGTGCGGCGGTTAGGCA 3'	63.63%	60
i6L	ads	3' TGGGGAAAATCAGCCTGTTA 5'	40.90%	49
i6R		5' ACGAAAGTCGGCCTTAGTGA 3'	50%	51
i7L	ads	3' GCCTTTGCACTCGAGGGCCAA 5'	61.90%	58
i7R		5' GCCATCAGTGAGATACCACTCTGG 3'	54.17%	58
i8L	cyp71av1	3' AACCGTGGCTCCAAAGCTCTCAGC 5'	58.33%	60
i8R		5' CGCACTTGGTTTAGCCAACGTGCA 3'	50%	58
i9L	cyp71av1	3' CTCGTATTCTGCACCCATGA 5'	50%	51
i9R		5' TTGCACCCTCCACTACCTTT 3'	50%	51
i10L	dxs	3' TACCAGACGGGATTGGGGTTCCA 5'	56.52%	58
i10R		5' GCAACATGGGAGCCGAACCCTCCA 3'	62.50%	62
i11L	dxs	3' TCCAACCTCAAGTGGAACAC 5'	50%	51
i11R		5' GCATTAGATGGTCTCCTTGACGGA 3'	50%	57
i12L	dxs	3' CTCAACAGTGATCAAGACCTCGTG 5'	50%	57
i12R		5' TGATAGACCGAGCTGTTTCCGGTA 3'	50%	57
i13L	dxs	3' TCTGCAGGATTAGGGATGGCTGTG 5'	54.17%	58
i13R		5' CCATCACATATAGCTGCAACCGTA 3'	45.83%	55
i14L	dxs	3' GGAGTCAACCCAGCTTCCGCTA 5'	59%	58
i14R		5' TGGAGACCGCTAGTGCTTCCAGAC 3'	58.33%	60
i15L	dxs	3' AGAGAATGGTCTAATGGCTTGC 5'	45.45%	52
i15R		5' TGATAGACCGAGCTGTTTCCGGTA 3'	50%	57
i16L	dxs	3' TCTGCAGGATTAGGGATGGCTGTG 5'	54.17%	58
i16R		5' CCATCACATATAGCTGCAACCGTA 3'	45.80%	55
i17L	dxs	3' GCCATGCAACTCTGAACGGCTG 5'	59%	58
i17R		5' TGATAGACCGAGCTGTTTCCGGTA 3'	50%	57
i18L	dxs	3' GTTTTGGCACGGGTCATAGC 5'	55%	53
i18R		5' ACAATGTCGATCCAGAACCAC 3'	47.60%	52
i19L	dxraa	3' AGAGTCCTGTGTCTCAACCATCGA 5'	50%	57
i19R		5' TTAGTTCCGGGGTCCACTAACGCA 3'	54.16%	58
i20L	dxraa	3' ACATTGTCTTGGCCGGATAG 5'	50%	51
i20R		5' TATTACACCCCTTGCCCAAA 3'	54.16%	49
i21L	dxraa	3' GAACTCCGGTCATTGTACCACCAG 5'	58.33%	58
i21R		5' AGGATGGCCCGACATGCGTTTACC 3'	50%	60
i22L	dxraa	3' TGCTTGATGCTTCTCACACG 5'	55%	51
i22R		5' GGCTCGATCTTTGCAAGCTAGG 3'	56.52%	56

i23L	dxraa	3' CTCTAAAAGCACCTCCGTGCAGG 5'	52.17%	58
i23R		5' TACGCCTAAGAGCACCTTCAGGA 3'	54.16%	56
i24L	dxraa	3' GTCGGGTTCTTTGAACGTCAGCGA 5'	56.52%	58
i24R		5' TGCGGCCATTGTATTCCGGCCAA 3'	52.17%	58
i25L	dxraa	3' ACGAAGAACCGTAAAGACCCAC 5'	54.17%	54
i25R		5' AAGATCACGGTTGACTCTGCCA 3'	56.52%	54
i26L	idi1	3' GCGTATTCTGATGGTTGGTCTCGA 5'	50%	57
i26R		5' TTGGATTGCGCCTCGTCACCAA 3'	54.54%	56
i27L	idi1	3' GCTAACAAGGCCTAAAACGGACTC 5'	50%	57
i27R		5' CTTGGAGCATCATCACCAGCAA 3'	50%	54
i28L	idi1	3' GATGCTTACGTTGGTGACGAAGCA 5'	50%	57
i28R		5' CAACATCCCACACAGTGAAGCTG 3'	52.17%	56
i29L	idi1	3' TTTGTGGTCGACAGCAATGACC 5'	50%	54
i29R		5' ATGCAGTGTTGCTTGTTTTTGCCA 3'	41.67%	53
i30L	pod1	3' GGGTCAACGTTCTTTGGGCACTG 5'	56.52%	58
i30R		5' TGTGGGTGGAAATCTGCCCAAACC 3'	54.17%	58
i31L	pod1	3' TGACGACCTCTCGTGTATTGTT 5'	45.45%	52
i31R		5' CCATTTATTTCCCCGTGTCA 3'	45%	49
i32L	pod1	3' GGTTCTTTGCACTGACACGA 5'	50%	51
i32R		5' TTGTACCCGTCTTGACACCA 3'	50%	51
i33L	pod1	3' AGCTTTTGCGCAACTCAAGCAG 5'	50%	54
i33R		5' ATCCGAGGGTGTGTGCACCTG 3'	61.90%	58
i34L	pod1	3' AGCTTTTGCGCAACTCAAGCAG 5'	50%	54
i34R		5' GAACATCAGGAAAGATTGTGCTGC 3'	45.83%	55
i35L	pod1	3' TTTGTTCAGGGGTGTGATGC 5'	50%	51
i35R		5' CTCAACGTTCTTTGGGCACTG 3'	52.38%	54
i36L	pod1	3' TAGCTTTTGCGCAACTCAAG 5'	45%	49
i36R		5' TTGATATCATTGAGCGAGATTG 3'	36.36%	49
i37L	pod1	3' TTGAAAACTGGTTGCAATGG 5'	40%	47
i37R		5' TGTGCCATCATATTCAGTGGA 3'	42.86%	50
i38L	pod1	3' AGCTTTTGCGCAACTCAAGCAG 5'	50%	54
i38R		5' ATCCGAGGGTGTGTGCACCTG 3'	61.90%	58
i39L	pod1	3' AGCTTTTGCGCAACTCAAGCAG 5'	50%	54
i39R		5' GAACATCAGGAAAGATTGTGCTGC 3'	45.83%	55
i40L	eps	3' ACCATGGATGATCTTGGTTCCCA 5'	47.82%	55
i40R		5' TCTGTAAGACATGGTAGCTCACGA 3'	45.80%	55

Table 1. List of 40 ILPs primers developed and custome synthesized for experimental validation.

3.3 Plant materials

Twenty genotypes of Artemisia annua were collected from two valleys (Leh and Nubra) with altitude ranging from 9,600 m (Nubra) to 11,500 m (Leh) from the cold arid desert of Trans-Himalayas (Ladhak) (figure 6) . Ten genotypes from each valley were obtained from the nursery of Defence Institute of High Altitute Research (DIHAR). The leaves were stored in laboratory at –20°C until further analysis. The interval between samples was 100-200m, the pair wise distance between valley divisions was 50– 250 Km.

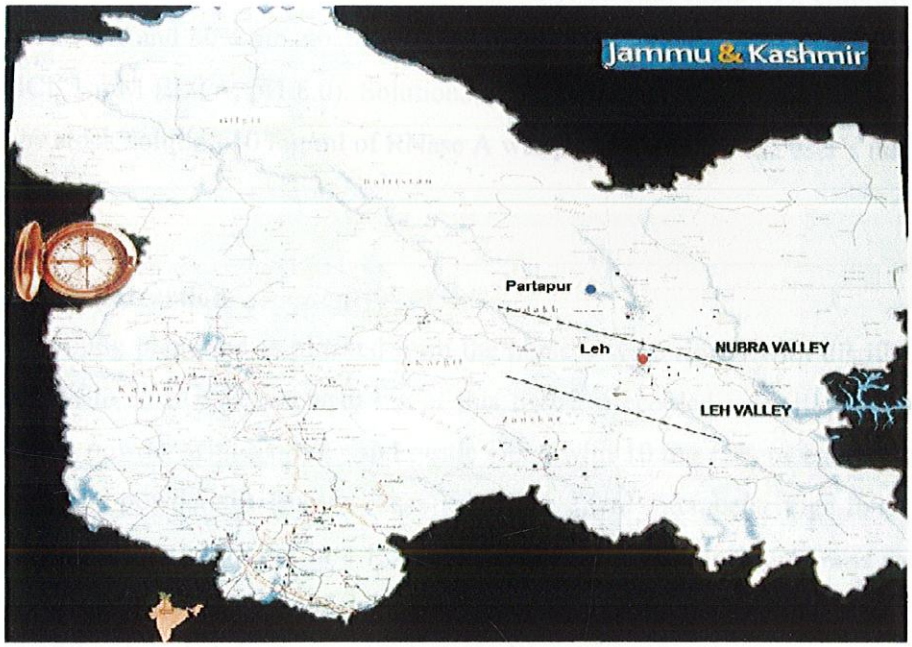


Fig. 6. Collection sites of 20 *Artemisia annua* genotypes from two valleys (Leh and Nubra) and the two collection sites (Leh and Partapur) located in Ladakh (Jammu and Kashmir, India).

3.4 DNA extraction and PCR amplification

Several experiments based on the available protocols (Doyle and Doyle 1990; Stewart and Via 1993; Stange et al 1998) were performed using fresh plant materials for (i) incubation time of buffer and tissue mixture at 65^0C, (ii) buffer to tissue ratio and (iii) extraction with phenol:chloroform:isoamyl alcohol vs. Tris saturated phenol followed by chloroform:isoamyl alcohol extraction in extraction and purification phases. All the experiments were repeated 3-4 times to check reproducibility.

Extraction Buffer

The extraction buffer (pH 8.0) contained 2% CTAB, 100 mM Tris-HCl, 20 mM EDTA, 1.4 M NaCl, 3% PVP and 0.2% β-mercaptoethanol.

Reagent and Solution:

Tris saturated phenol, phenol:chloroform:isoamyl alcohol (25:24:1), chloroform:isoamyl alcohol (24:1), 70% and 80% ethanol, 4M NaCl, 3M sodium acetate (pH 5.2) and TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). Solutions and buffers were autoclaved at 121^0C at 15 psi pressure. The stock solution 10 mg/ml of RNase A was prepared as per the user's manual (Sigma USA).

3.4.1 DNA Extraction

Fresh young plant leaves plucked form the nursery were rinsed with distilled water and blotted gently with soft tissue paper, 0.1 g of this tissue, precooled using liquid nitrogen, was ground to a fine powder with a mortar and pestle along with 10 mg (2% of extraction buffer) of PVP (Sigma). The powdered tissue was scraped into a 2.0ml microcentrifuge tubes containing preheated (65^0C) extraction buffer in a 1:5 ratio (0.5 ml). β-Mercaptoethanol was then added to the final concentration of 0.2 M and mixed well. The mixture was incubated in water bath at 65^0C for 90 min and cooled for 5 min. An equal volume of chloroform:isoamyl alcohol mixture (24:1) was added to the extract and mixed by gentle inversion for 5 to 10 min to form an uniform emulsion. The mixture was centrifuged at 8000 rpm for 8 min at room temperature. Chloroform: isoamyl alcohol extraction was repeated again. The aqueous phase was pipetted out gently, avoiding the interface. To the above solution, 5 M NaCl (to final concentration 2M) and 0.6 volume of isopropanol of the total solution was added and incubated at room temperature for 1 h. To the above solution, two volumes of 80% ethanol was added and incubated again for 10 min at room temperature for DNA precipitation. After incubation, the mixture was centrifuged at 10,000 rpm for 15 min. The white/translucent pellet was washed with 70% ethanol, dried and resuspended in 200 µL of TE buffer.

3.4.2 Purification phase

The sample was then treated with 20 µL of 10 mg/ml of RNase and incubated at 37^{0}C for 60 min. After incubation with RNase, equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) was added and mixed gently by inverting the microcentrifuge tube followed by centrifuged at 10,000 rpm for 5 min at room temperature. The supernatant was pipetted out into a fresh tube. The sample was then extracted with equal volumes of chloroform:isoamyl alcohol (24:1) twice. The DNA was precipitated with 0.6 volumes of isopropanol, 2.0 M NaCl (final concentration) and incubated for 10 min. To the above, 20 µL of sodium acetate and 1 volume of 80% ethanol were added, incubated for 30 min and centrifuged at 5,000 rpm for 3 min to pellet the DNA. The pellet was then washed with 70% ethanol twice; air-dried and finally suspended in 40-50 µL of TE buffer.

3.4.3 Evaluation of resultant DNA

The yield of the extracted DNA and purity was checked by running the samples on 0.8% agarose gels along with standard (non restriction enzyme digested), lamda DNA marker (Biogene, USA).

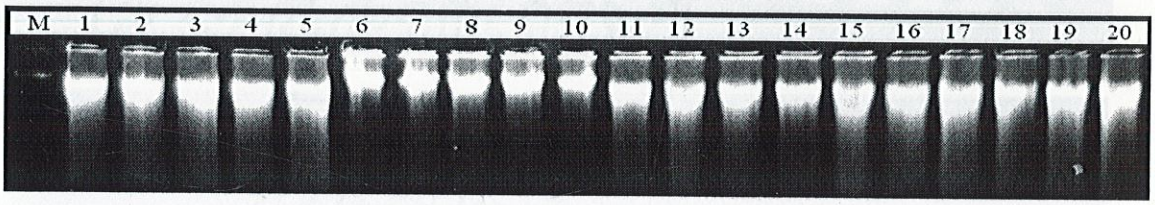


Fig. 7. Agrose gel electrophoresis showing purified high molecular weight *Artemisia annua* genomic DNA of different Species, more than 200 ng genomic DNA from each genotype was electrophoresed on 0.8% agarose gel at 65 V for 2 hr and stain with ethidium bromide. Lane M, 100 ng mol. wt. uncut λ DNA. Lane 1 to 10; are the genomic DNA extracted from Leh valley and Lane 11 to 20; are from Nubra valley (Partapur).

3.5 ILP analysis

We conducted PCR in a 15 ml reaction mixture containing 50 ng template DNA, 0.5 mM of each primer, 200 mM of each dNTP, 1.5 mM MgCl2, 0.1% Triton X-100 and 1 U Taq polymerase and 1.5 ml of 10· PCR buffer. We first tested all primer pairs with a touchdown PCR (Td-PCR) 36 program: 5 min at 94 °C; 10 cycles of 30 s at 94 °C, 30 s at 59 °C minus 0.3°C/cycle, 1min at 72 °C; 20 cycles of 30 s at 94 °C, 30 s at 56 °C, 1 min at 72 °C; and 5 min at 72 °C for final extension. For primer pairs that did not generate good amplification results, we

adjusted the initial annealing temperature (59 °C) to 60 or 57 °C. The purpose of using Td-PCR was to increase the specificity of amplification, but some primer pairs only required a routine PCR program: 5 min at 94 °C; 35 cycles of 30 s at 94 °C, 30 s at 54 °C, and 1 min at 72 °C; and 5 min at 72 °C for final extension. For most primers, we used 6% non-denaturing PAGE (250 V, 2 h) for separating PCR products and silver staining for visualizing DNA bands following Xu et al.37 with modification. After staining, the banding patterns were scanned; data collected from reproducible and successful amplification were stored and band sizes were estimated by comparison to molecular mass standards included in each gel.

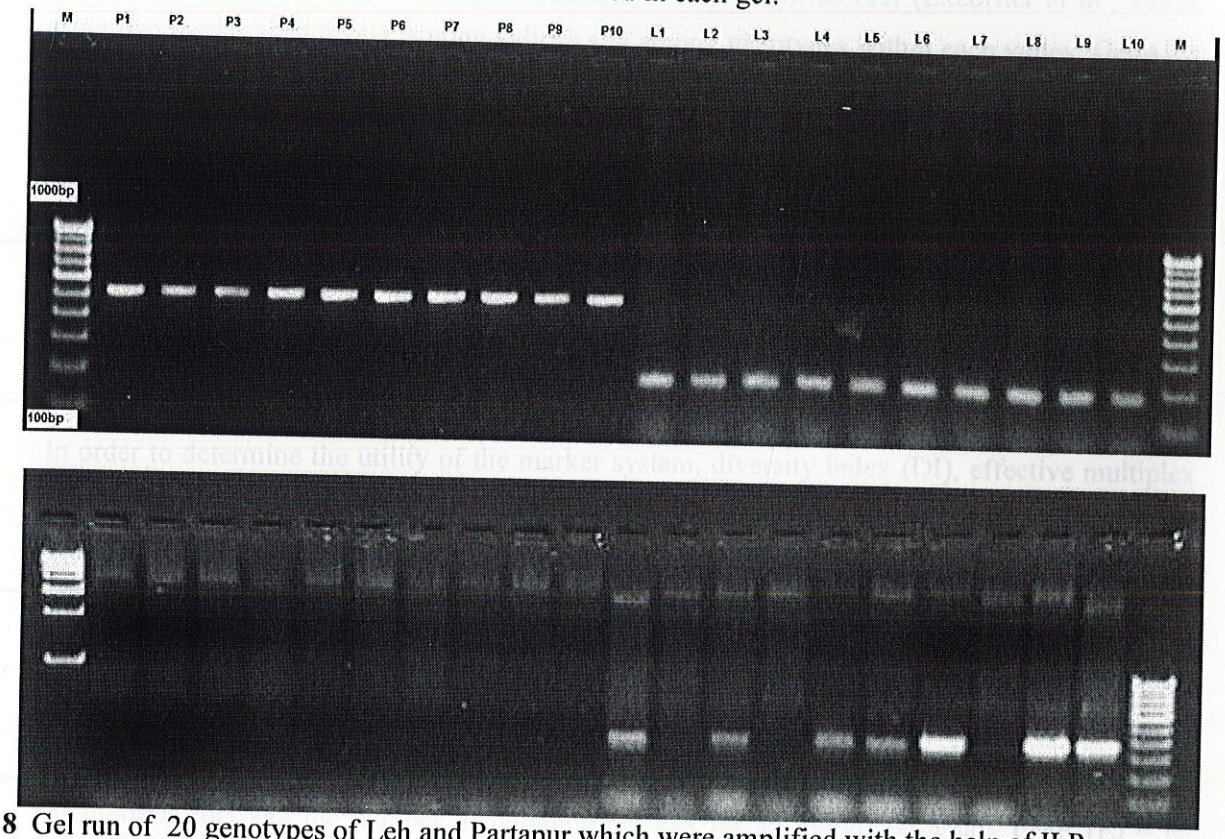


Fig . 8 Gel run of 20 genotypes of Leh and Partapur which were amplified with the help of ILPs.

3.6 Data collection and analysis

Fragments were scored as a binary unit character (present = 1; absent = 0). Only unambiguously resolved and reproducible bands were scored. The reproducibility of banding patterns was confirmed by running many duplicate experiments with samples randomly loaded. The mean reproducibility, calculated as identical bands in two repeats, was between 97 and 100%. 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 averages (UPGMA) clustering method (Rohlf 1998). POPGENE software was used to calculate Nei's unbiased genetic distance among different genotypes with all markers. 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). Within species diversity (Hs) and total genetic diversity (Ht) (Nei, 1978) were calculated within the species and within two major groups (as per their collection site) using POPGENE software. The RAPD and ISSR 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). The mean Polymorphism Information Content (PIC) value for ILP markers was calculated using the following formula:

$$PIC_i = 1 - \sum_{j=1}^{n} p_{ij}^2,$$

where p_{ij} is the frequency of the jth pattern for the ith marker.38

In order to determine the utility of the marker system, diversity index (DI), effective multiplex ratio (EMR) and marker index (MI) were calculated according to Powell *et al.* (1996). DI for genetic markers was calculated from the sum of the squares of allele frequencies: $DI_n = 1\text{-}\Sigma\, pi^2$ (where 'pi' is the allele frequency of the ith allele). The arithmetic mean heterozygosity, Di_{av}, was calculated for each marker class: $Di_{av} = \Sigma\, Di_n/_n$, (where 'n' is the number of markers (loci) analyzed). The DI for polymorphic markers is: $(Di_{av})p = \Sigma\, Di_n/n_p$ (where 'n_p' is the number of polymorphic loci and n is the total number of loci). EMR (E) is the product of the fraction of polymorphic loci and the number of polymorphic loci for an individual assay. EMR (E) = n_p (n_p/n).

MI is defined as the product of the average diversity index for polymorphic bands in any assay and the EMR for that assay, MI = DI_{avp} * E.

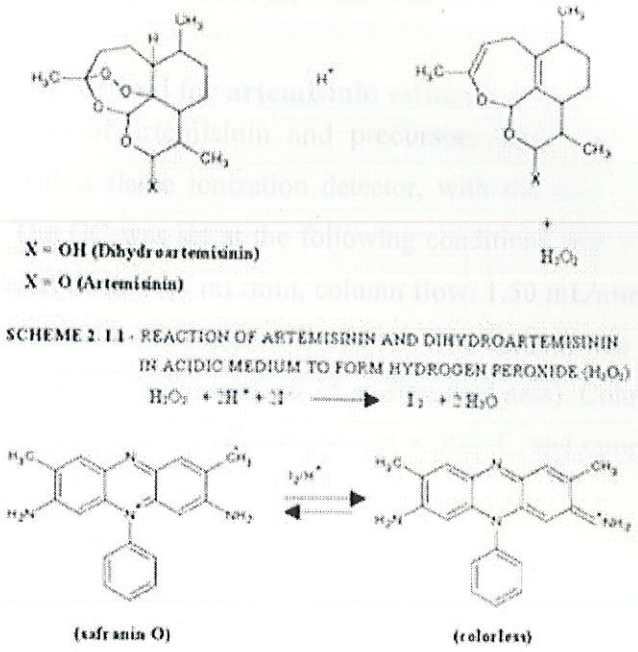
3.7 Extraction of artemisinin

The samples of Artemisia annua used for artemisinin extract were Partapur leaf, Leh leaf, Leh flower. Artemisinin was extracted following the procedure of Broomhead et al., (1990).

Weight of each sample was noted and 100 mg of the powdered form of sample was taken. Dried leaves and inflorescences were ground to a powder in a pestle and mortar. Artemisinin was extracted following the procedure of Broomhead et al., (1990). About 5 ml of ethanol was added and heated at 30°C for 2 hours. The extract was filtered through Whattman filter paper No.1. Again 5ml of ethanol was added to the residue and heated at 30°C for 2 hours. Second, third and fourth extractions of the same samples were done with 10 ml ethanol for 10 min. under condition mentioned above. All the extracts were pooled and ethanol evaporated to dryness in a water bath shaker at 60 ^{0}C. The resultant residue was dissolved in 10 ml acetonitrile (HPLC grade) and filtered with 0.22 µm durapore membrane filter (Millipore) for GC analysis.

3.8 Spectrophotometric determination of Artemisinin (ART)

A simple and rapid method for determination of ART is described. The method is based on the reaction of H_2O_2 generated by cleavage of endoperoxide linkage of ART and its reaction with potassium iodide (KI) to liberate iodine (I). Liberated I bleaches red colored safranine O to colorless species and is measured at 521nm. Beer's law is obeyed in the range of 16-112 µg/ml for both ART & DHA. The molar absorptivity, Sandell's sensitivity, detection limit and quantitation limit for ART were found 0.3401×10^{4} l/mol/em, 1.43×10^{-2} µg/cm^2, 0.0625 µg/ml and 0.2075 µg/ml. The optimum reaction conditions and other parameters were evaluated. The statistical evaluation of the method was examined by determining intraday and inters day precision.



SCHEME 2.11 - REACTION OF ARTEMISININ AND DIHYDROARTEMISININ IN ACIDIC MEDIUM TO FORM HYDROGEN PEROXIDE (H_2O_2)

Fig. 9 Reaction between hydrogen peroxide-potassium iodide systemwith safranin o

Experimental procedure:

Apparatus used was spectrophotometer. Reagents involved were HCL(5M), 2% Potassium Iodide(KI), Sodium Acetate(CH3COONa)(2M) and Safranin(0.01%). To prepare 0.01% safranin, 0.01gm of safranin was dissolved in 50% ethanol and the volume as raised to 100ml using water. Standard ART solution: stock - 1000µg/ml (0.1g ART was dissolved in ethanol and the volume was raised to 100ml), working solution - stock solution was diluted appropriately to get the working solution.

Method of determination of Artemisinin:

A 10ml calibrated flask was taken and different aliquots of ART in a series (16.0-112.0µg/ml) were transferred to it. Then 1ml of KI and 1ml of HCl to each flask was added and shaken gently until a yellow color appeared due to liberation of iodine. The reaction mixture was shaken for 5 minutes. Contents were diluted with distilled water and mixed well.The absorbance of each solution was measured at 521nm against corresponding reagent blank (made by replacing ART by distilled water). Absorbance corresponding to the bleached color, which in turn corresponds to analyte ART concentration, was obtained by subtracting Absorbance of test solution from Absorbance of blank. The method involves reaction of H_2O_2, generated by cleavage of endoperoxide linkage of corresponding ART solution in acidic medium, with KI to liberate iodine. This liberation of iodine bleaches the red colored safranine O which were

measured at 521nm. The decrease in absorbance is proportional to Artemisinin concentration.

3.9 GC-FID method for artemisinin estimation

Analysis of artemisinin and precursors was performed in a gas chromatographer equipped with a flame ionization detector, with the data collection through the GC-Solution software. The GC was set at the following conditions with helium as the carrier gas: pressure: 15.6 psi, total flow: 37.4 mL/min, column flow: 1.50 mL/min, linear velocity: 50.5 cm/s, purge flow: 3.0 mL/min, and a split ratio of 21:9. The column was a Rtx-5 crossbond 100% dimethyl polysiloxane, (15m×0.25mm ID, 0.25_m film thickness). Column temperature was set at 195 ◦C, injector at 240 ◦C, and FID temperature set at 300 ◦C, and sampling rate of 40 ms.

CHAPTER 4

RESULTS AND DISCUSSION

Many protocols were attempted for extraction of good quality DNA from *Artemisia annua* but were not possible to obtained good quality DNA. The Artemisia species obtained from high altitude (Ladhak and Partapur) contain high amount of polysaccharides, polyphenols, essential oils & other secondary metabolites that interfere with DNA isolation. These secondary metabolites entangled to nucleic acid during DNA isolation & interfere with subsequent isolation procedure. We developed a simple and efficient method of genomic DNA extraction from leaf of *Artemisia annua* by modifying CTAB protocol (Khanuja et al 1999). By using 3.5 M NaCl in extraction buffer and 80% ethanol with 2.0 M NaCl (final concentration) during precipitation and further purification with Tris saturated phenol during purification phase the quality and quantity (200-400 ng) of DNA was improved significantly without contamination of polysaccharides and secondary metabolites. In the present protocol, the use of 3.5 M NaCl in the extraction buffer reduced 90% of polysaccharides contamination and very little or no jelly like precipitate was found during precipitation of DNA. One of the most significant steps of our protocol was the use of only Tris saturated phenol (pH 8.0), followed by chloroform: isoamyl alcohol extraction. Most of the protocols in the literature used phenol: chloroform: isoamyl alcohol (25:24:1) or chloroform: isoamyl alcohol (24:1) for protein removal (Doyle and Doyle 1990; Dellaporta et al 1983), whereas in our experiments use of either phenol: chloroform: isoamyl alcohol (25:24:1) or chloroform: isoamyl alcohol (14:1) gave yellowish pellet (Figure 4b), which confirmed the presence of protein contamination. By using Tris saturated phenol (pH 8.0) followed by chloroform: isoamyl alcohol (24:1) extraction, protein impurities could be successfully removed, without affecting DNA yield. It was also observed that buffer to tissue ratio and incubation time were also important factors for obtaining higher yields of DNA and in case of Artemisia annua 5:1 buffer to tissue ratio and 90 min incubation at 76^0C gave best results. On agarose gel electrophoresis, DNA gave sharp bands (Figure 7). Use of high concentration of NaCl in the extraction buffer decreased contamination of polysaccharides (Danshwar P and Sher-ullah 2004). Compared to precipitation at -20 ^{0}C (Khanuja et al 1999), we could achieve the same at room temperature without compromising the quality and quantity of DNA suitable for ligation, PCR amplification and other downstream processes necessary for DNA fingerprinting (Figure 9).

4.1 ILP Analysis

Genetic relationship among the 20 genotypes of *Artemisia annua* has been carried out using ILP. Twenty primers generated reproducible, informative and easily scrabble ILP profiles were preselected. These primers produced_multiple band profiles (Figure 9) with a number of amplified DNA fragments varying from 1 to 4, with a mean number 7.4 markers per primer. All

the amplified fragments varied in size from 200 – 2,500 bp. Out of 74 amplified bands, 69 were found polymorphic. The highest number of polymorphic bands (4) was observed with primers 7, 31, and 32 with 100 % polymorphism. The average number of polymorphic bands per primer among 20 genotypes was 7.4 (Table 2). The observed high proportion of polymorphic loci suggests that there is a high degree of genetic variation in the *Artemisia* genotypes. The resolving power of the 20 ILP primers ranged from 2 for primer 3, 8, 13, 14, 17, 21, 22, 27, 29, 34, 39 to a maximum of 4.3 for primer 32. In addition to its high resolving power, ILP primer 32 has the ability to distinguish all 20 *Artemisia annua* genotypes.

Table2 . List of primers used for ILP study, total number of loci, the level of polymorphism and resolving power.

S. No.	Total no. of loci	Number of polymorphic loci	Percentage of polymorphic loci (%)	Total number of fragments amplified	Resolving power
3	2	2	100	20	2
5	2	2	100	21	2.1
6	3	3	100	30	3
7	4	4	100	40	4
8	2	2	100	20	2
9	3	3	100	28	2.8
10	3	3	100	29	2.9
11	2	2	100	23	2.3
13	2	2	100	20	2
14	2	2	100	20	2
17	2	2	100	20	2
18	2	2	100	21	2.1
19	3	2	66.66	40	4
20	3	2	66.66	36	3.6
21	2	2	100	20	2
22	3	3	100	20	2
23	3	3	100	23	2.3
24	4	3	75	38	3.8
25	3	2	66.66	40	4
26	3	3	100	29	2.9
27	2	2	100	20	2
29	2	2	100	20	2
30	2	1	50	33	3.3
31	4	4	100	40	4
32	4	4	100	43	4.3
34	2	2	100	20	2
35	3	3	100	26	2.6
39	2	2	100	20	2
TOTAL	74	69	93.74	760	-

4.2 ILP derived dendrogram analysis

A dendrogram based on UPGMA analysis grouped the 20 genotypes into two main clusters (I and II) with Jaccard's similarity coefficient ranging from 0.55 to 0.70 (Figure 10). Cluster I represents the genotypes from Partappur while, cluster II have all the genotypes from Leh valley valley. The results of PCA analysis were comparable to the cluster analysis.

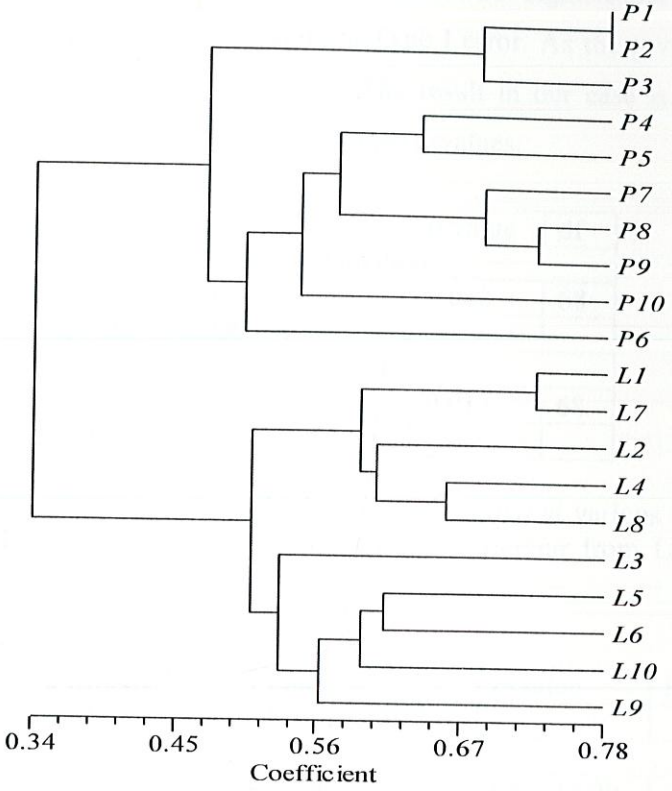


Fig. 10. Dendrogram illustrating genetic relationships among 20 genotypes of *Artemisia annua* collected from 2 different valleys: Leh and Nubra Ladhak of Ladhak region. The dendrogram was generated by UPGMA cluster analysis calculated from ILP bands produced by 40 primers.

4.3 Genetic Diversity Analysis

A relatively high genetic variation was detected among the *Artemisia annua* genotypes. Genetic diversity analysis in terms of Na, Ne, H, I, Ht, Hs, NPL and PPL for both the valleys (i.e. Leh and Partappur valleys) revealed that Leh valley and Partapur valley had equal variability among the genotype (Table 5). The respective values for overall genetic variability for Na, Ne, H, I, Ht, Hs, Gene flow (Nm) across all the 20 genotypes were given in Table 5. The Gst value of 1 indicated that 0 % of the genetic diversity resided within the populations. quantified artemisinin

data from both Leh and Partapur region. Mean values have been given from each stage of development of inflorescence. F1 and F2 are the continuous probability distribution between the stages of development and site of sample collection. The F-value between the interaction of sites and stages of inflorescence development is 0.03 and 0.02 respectively using both the estimation methods and are not statistically significant. The statistical test used for the analysis is 2-way ANOVA. The p-value is the observed significance level of a hypothesis test which equals the probability of making a Type I error. As the p-value is very also is nearly all the cases, null hypothesis is rejected. The result in our case is statistically significant because the F-values in all the cases are greater than p-values.

	Mean ± SEM	t-value	p-value	df
(a) Artemisinin % (g/100g) by GC-FID method				
Leh	0.490 ± 0.038	2.57	0.012	68
Partappur	0.367 ± 0.030			
(b) Artemisinin % (g/100g) by OD at 521nm				
Leh	0.401 ± 0.035	2.46	0.017	68
Partappur	0.300 ± 0.022			

Table 3. Quantification of artemisinin from the leaves of *A. annua* at various stages of vegetative growth of the plant collected from Leh and Parteppur from Ladhak region.

	Emergence	Middle	Late	F-value	P-value
	Mean ± SEM	Mean ± SEM	Mean ± SEM		
(a) Artemisinin % (g/100g) by GC-FID method					
Leh	0.064 ± 0.008	0.078 ± 0.002	0.086 ± 0.003	F1 = 13.20	< 0.001
Partappur	0.057 ± 0.004	0.069 ± 0.003	0.078 ± 0.002	F2 = 5.00	0.030
(b) Artemisinin % (g/100g) by OD at 521nm					
	0.054 ± 0.006	0.063 ± 0.001	0.067 ± 0.002	F1 = 8.71	0.001
	0.047 ± 0.004	0.057 ± 0.002	0.062 ± 0.007	F2 = 4.55	0.038

Table 4. Quantification of artemisinin from the different stages of inflorescence development of *A. annua* collected from Leh and Partappur from Ladhak region.

F1 and F2 are the F-value between the stages of development and site of sample collection. The F-value between the interaction of sites and stages of inflorescence development is 0.03 and 0.02 respectively using both the estimation methods and are not statistically significant. The statistical test used for the analysis is 2-way anova.

S.No.	N	Na	Ne	H	I	Ht	NPL	PPL	Hs	Gst	Nm
1	10	1.7500±0.4359	1.5141±0.3583	0.2958±0.1899	0.4343±0.2688	0.2958±0.0361	57	75	0	1	0
2	10	1.6184±0.4890	1.5171±0.4249	0.2792±0.2242	0.3975±0.3171	0.2792±0.0502	47	61.84	0	1	0

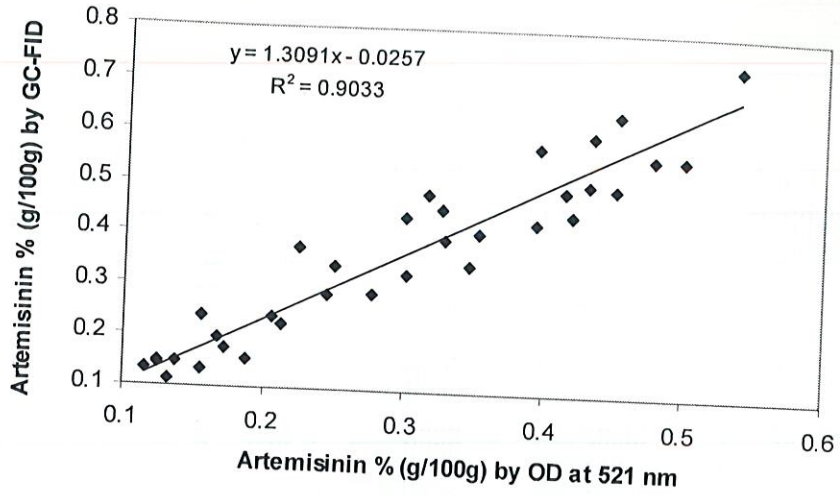
Table 5. Summary of genetic variation statistics for all loci of ILPs among the *Artemisia annua* genotypes with respect to their distributions among two valleys. Na = Observed number of alleles; Ne = Effective number of alleles; H = Nei's gene diversity; I = Shannon's Information index; Ht = Total genetic diversity; Hs = Genetic diversity in population; Gst = Genetic diversity between population; NPL = Number of Polymorphic Loci; PPL = percentage of Polymorphic Loci.

Artemisinin was extracted from the leafs and flowers of the plant which yielded analytical results stating that content of Artemisinin with the stage of the flower, in early stages(budding stage) content was found to be between 0.057-0.064% in 100gm, whereas in late stages, it was 0.078-0.086%. Also that the content found in plants from Leh region (0.065-0.086%) is more than that of Partapur region (0.057-0.079%) at each stage of the flower. Highly significant correlation(>0.9) was observed in artemisinin content (w/w%) obtained by GC-FID and spectrophotometric methods.

This section tells us about the estimation of artemisinin by GC-FID and spectrophotometry results of both have been plotted over a graph which gives highly significant correlation of artemisinin content in plant material collected from Leh and Nubra (Partapur) valleys of Ladakh region.

This result is the estimation of artemisinin from 35 different cultivars of Artemisia annua. The regression is very high which implies a positive correlation between the values obtained from GC-FID and SPECTROPHOTOMETRIC methods. The graph here shows a linear increase in atremesinin content obtained by using these methods.

(a) Partappur



(b) Leh

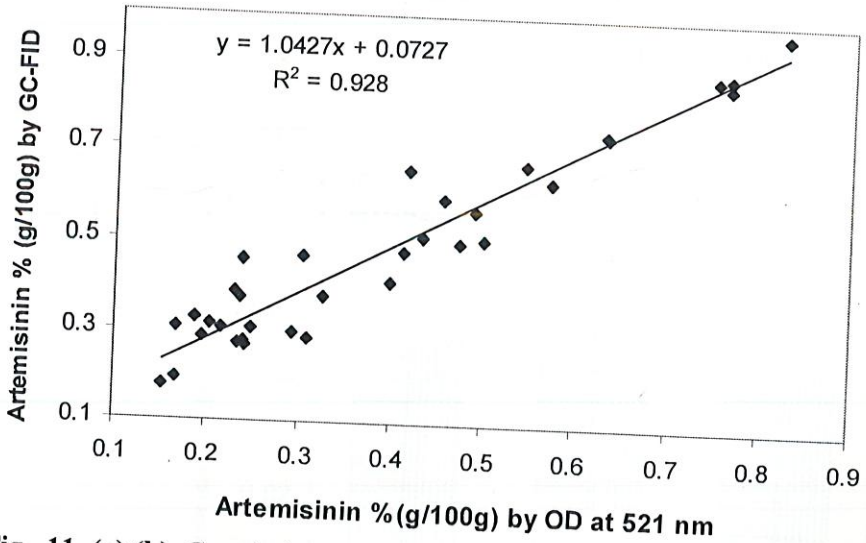
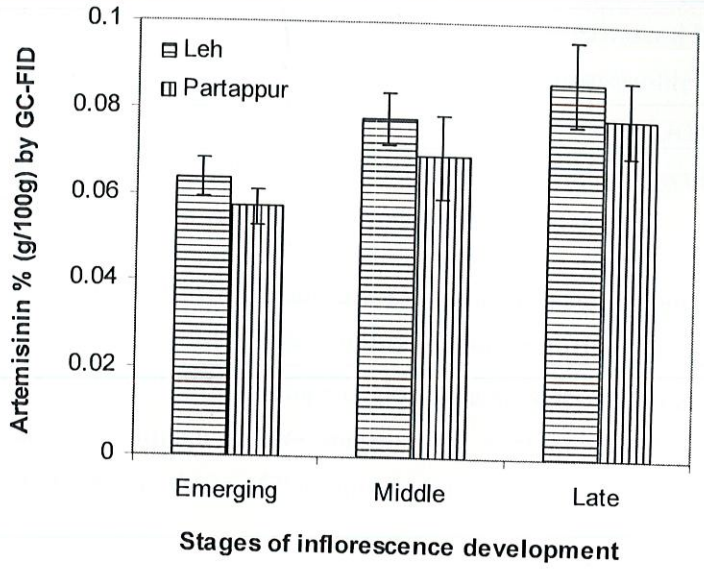


Fig. 11 (a),(b) Correlation of artemisinin content (w/w%) of 35 plant materials analyzed by both GC-FID and spectrophotometric methods. These plant materials consisted of leaves from different parts of 35 cultivars of *A. annua*. Samples were collected from 2 different sites (a) Partappur and (b) Leh from Ladhak region.

The Artemisinin content present during 3 different stages of inflorescences, when analyzed by both GC-FID and spectrophotometric methods, the three different stages are emerging (budding stage), middle (flowering stage) and late (later stage).The values in the bar graph are means of ten replicates and vertical bars represent the standard deviation. Individual bars

for different regions have been plotted. The artemisinin content was found to be more in samples from the Leh region as with those compared to the Partappur samples. It was also observed that the artemesinin content was found to be very high in the later stages (Late) of inflorescence than the emerging and the middle stages (figure 12).

(a)



(b)

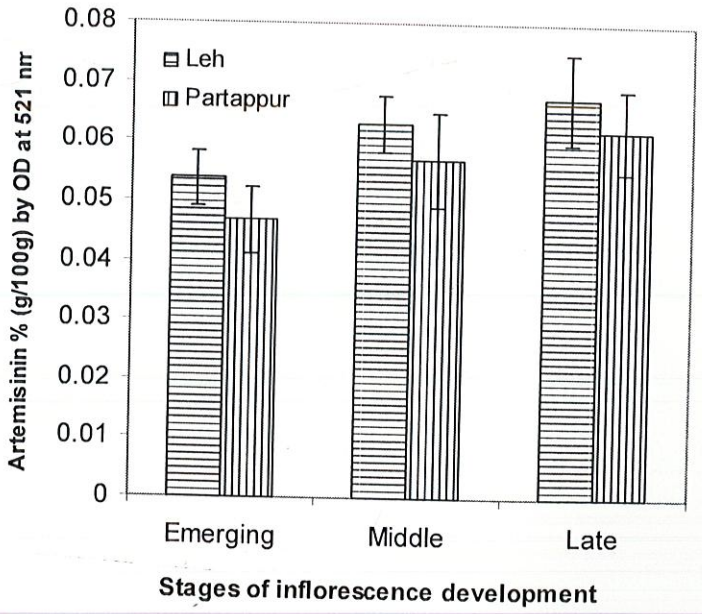


Fig. 12 Artemisinin content (w/w%) of inflorescences analyzed by both (a) GC-FID and (b) spectrophotometric methods at 3 stages of development of *A. annua*. The values are means of 10 replicates. Vertical bar represent ± standard deviation. Samples were obtained from 2 different sites Partappur and Leh from Ladhak region.

CONCLUSION

DNA extraction protocol was standardized, which was a must as quality of DNA was very bad because of the presence of polyphnolic compounds, the extracted DNA was used for amplification of regions of Artemisia annua genes which are responsible for production of secondary metabolite, Artemisinin. These genes were of those enzymes which play a key role in the formation of Artemisinin. The primers designed for amplification were designed after a thorough study of DOXP pathway for Artemisinin production.

Artemisinin was also extracted from the samples collected from partapur and leh valley of ladakh region. These samples were analyzed and quantified, results of which state that Artemisinin content increases with the stage of flower, that is, its low in early stage (budding stage) whereas comparatively higher in late stages, also, content of Artemisinin in samples from Leh region were higher than those of Partapur region.

BIBLIOGRAPHY

Ahmad A, Misra LN (1994) Terpenoids from *Artemisia annua* and constituents of its essential oil. *Phytochemisty.* 37 (1): 183-186.

Asawamahasakda W, Benakis A,. Meshnick SR, (1994) The interaction of artemisinin with red cell membranes. *J Lab Clin Med* 123: 757-62.

Augustijns P, D'Hulst A, Van Daele J, Kinget R, (1996) Transport of artemisinin and sodium artesunate in Caco-2 intestinal epithelial cells. *J Pharm Sci* 85:577-9.

Bagchi GD (2003). Essential oil constituents of *Artemisia annua* during different growth periods at monsoon conditions of subtropical north Indian plains. *J. Essent. Oil Res.* 15(4): 248 - 250.

Bannister LH, Hopkins JM, Fowler RE, Krishna S, Mitchell GH (2000). A brief
ultrastructural guide to asexual blood stages of Plasmodium falciparum. *Parasitol Today* 16: 427-433.

Bhakuni RS et al. (2001). Secondary metabolites of *Artemisia annua* and their biological activity. *Curr. Sci.*, 80 (1): 35-48.

Brossi A, Venugopalan B, Dominguez Gerpe L, Yeh HJ, Flippen-Anderson JL, Buchs P, Luo XD, Milhous W, Peters W (1988). Arteether, a new antimalarial drug: synthesis and antimalarial properties. *J Med Chem* 31:645-50.

Chawira AN, Warhurst DC (1987). The effect of artemisinin combined with standard antimalarials against chloroquine-sensitive and chloroquine-resistant strains of Plasmodium falciparum in vitro. *J Trop Med Hyg.* 90: 1-8.

Cox-Singh J, Davis TM, Lee KS, Shamsul SS, Matusop A, Ratnam S, Rahman HA, Conway DJ, Singh B (2008). Plasmodium knowlesi malaria in humans is widely distributed and potentially life threatening. *Clin Infect Dis* 46: 165-171.

Davis TME, Karunajeewa HA, Ilett KF (2005) Artemisinin-based combination therapies for uncomplicated malaria. *Med. J. Austr.* 182: 181-185.

Deng DA et. al. (1981) Studies on the structure of artemisic acid. *Kexue Tonbao.* 19: 1209-1211.

Eckstein-Ludwig U, Webb RJ, Van Goethem ID, East JM, Lee AG, Kimura M, O'Neill P M, Bray PG, Ward SA, Krishna S, (2003) Artemisinins target the SERCA of Plasmodium falciparum. *,ature* 424: 957-961.

Eckstein-Ludwig U, Webb RJ, van Goethem IDA, East JM, Lee AG, Kimura M, O'Neill PM, Bray PG, Ward SA, Krishna S (2003). *,ature*. 424: 957

Foote SJ, and Cowman AF (1994). The mode of action and the mechanism of resistance to antimalarial drugs. *Acta Trop* 56: 157-71.

Francis SE, Sullivan Jr. DJ, and Goldberg DE, (1997). Hemoglobin metabolism in the malaria parasite Plasmodium falciparum. *Annu Rev Microbiol* 51: 97-123.

Gupta S, Singh M, Madan AK, (1999). Superpendentic index: a novel topological descriptor for prediction of biological activity, *J. Chem. Inf. Comput. Sci.* 39: 272-277.

Hien TT, White NJ (1993). Qinghaosu. *Lancet* 341: 603-8.

Homewood CA, Warhurst DC, Peters W, Baggaley VC (1972). Lysosomes, *p*H and the Antimalarial Action of Chloroquine ,*ature* 235: 50-52.

Hong YL, Yang YZ, Meshnick SR, (1994). The interaction of artemisinin with malarial hemozoin. *Mol. Biochem. Parasitol.* 63: 121-128.

http:://rbm.who.int/partnership/wg/wg_management/docs/ RBMConsensusStatement_act.pdf.

Idury R.M., Cardon L.R., A simple method for automated allele binning in microsatellite markers, Genome Res. 7 (1997) 1104_1109.

Jain N et. al. (2002). Essential oil composition of *Artemisia annua* L. 'Asha' from the plains of northern India *J. Essen. Oil Res.* 14 (4): 305 - 307.

Jung M,(1990). Artemisinic acid: a versatile Chiral synthon and bioprecursor to natural products. *Planta Medica.* 56: 56.

Jung M,(1997). Synthesis and cytotoxicity of novel artemisinin analogs. *Bioorg. Med.Chem.Lett* 7 (8):1091-1094.

Jung M,(1994) Current developments in the chemistry of artemisinin and related compounds. *Curr Med Chem* 1:35-49

Kamchonwongpaisan S, and Meshnick SR,(1996) The mode of action of the antimalarial artemisinin and its derivatives. *Gen Pharmacol* . 27: 587-92.

Kamchonwongpaisan S, Samoff E, Meshnick SR,(1997) Identification of hemoglobin degradation products in *Plasmodium falciparum. Mol. Biol. Parasitol.*, 86: 179-186.

Kimura M, Yamaguchi Y, Takada S, Tanabe K, (1993) Cloning of a Ca(2+)- ATPase gene of Plasmodium falciparum and comparison with vertebrate Ca(2+)-ATPases. *J Cell Sci* 104 (Pt 4): 1129-1136.

Klayman DL (1985) Qinghaosu (artemisinin): an antimalarial drug from China. *Science* 228: 1049-1055.

Kondrachine AV, Trigg PI (1997). Global overview of malaria. *Indian J Med Res* 106: 39-52.

Kumar S, Srivastava S, (2005) Establishment of artemisinin combination therapy as first line treatment for combating malaria: *Artemisia annua* cultivation in India needed for providing sustainable supply chain of artemisinin .*Curr. Sci.*, , 89: 1097-1102.

Lai H, SinghNP(1995).Selective cancer cell cytotoxicity from exposure to dihydroartemisinin and holotransferrin. *Cancer Letters,* 91: 41-46.

Le WJ et al. (1982). Studies on the efficacy of artemether in experimental schjitosomiasis. *Acta Pharmaceu.Sin.* 17: 187-193.

Le WJ et. al. (1980) Artemisin derivatives against *Schistosoma japonicum* in animals. *Chin.*

Lee IS, Hufford CD (1990). Metabolism of antimalarial sesquiterpene lactones. *Pharmacol Ther* 48: 345-55.

Lee MR, (2002) Plants against malaria Part I: Cinchona or the Peruvian bark. *J. R. Coll. Physicians Edinburgh*, 32: 300-305.

Li GQ, Guo XB, Fu LC, Jian HX, Wang XH (1994). Clinical trials of artemisinin and its derivatives in the treatment of malaria in China. *Trans R Soc Trop Med Hyg* 88 Suppl 1:S5-6.

Liu JM, (1979). Structure and reaction of arteannuin. *Acta Chim. Sin.* 37:129

Liu SS, Cao CZ, Li ZL, (1998). Approach to estimation and prediction for normal boiling point (NBP) of alkanes based on a novel molecular distance edge (MDE) vector λ, *J. Chem. Inf. Comput. Sci.* 38: 387-394.

Luo XD, Shen CC (1987). The chemistry, pharmacology, and clinical applications of qinghaosu (artemisinin) and its derivatives. *Med Res Rev* 7: 29-52.

Maeno Y, Toyoshima T, Fujioka H, Ito Y, Meshnick SR, Benakis A, Milhous WK, Aikawa M, (1993) Morphologic effects of artemisinin in Plasmodium falciparum. *Am J Trop Med Hyg* 49:485-91.

Meshnick SR, (2002). Artemisinin: mechanisms of action, resistance and toxicity. *Int. J. Parasitol.* 32: 1655-1660.

Meshnick SR, Thomas A, Ranz A, Xu CM, Pan HZ, (1991) Artemisinin (qinghaosu): the role of intracellular hemin in its mechanism of antimalarial action. Mol Biochem Parasitol. 49(2):181-189.

Meshnick SR, Tsang TW, Lin FB, Pan HZ, Chang CN, Kuypers F, Chiu D, Lubin B, (1989). Activated oxygen mediates the antimalarial activity of qinghaosu. *Prog Clin Biol Res.* 313:95-104.

Meshnick SR, Yang YZ, Lima V, Kuypers F, Kamchonwongpaisan S, Yuthavong Y, (1993) Iron-dependent free radical generation from the antimalarial agent artemisinin (qinghaosu). *Antimicrob Agents Chemother* 37: 1108-14.

Meshnick, SR, Thomas A, Ranz A, Xu CM, Pan HZ (1991) Artemisinin (qinghaosu): the role of intracellular hemin in its mechanism of antimalarial action. *Mol. Biochem. Parasitol.* 49 (2): 181-189.

Milhous WK, Kyle DE, Sherman IW (1998). ed. Malaria Parasite Biology, Pathogenesis, and Protection. Washington, D.C. *ASM Press*, pp 303-316.

42

Misra LN et. al. (1993). Bisnor-cadinanes from *Artemisia annua* and definitive C-13 NMR assingements of beta- arteether. *Phytochemisrty.* 33 (6): 1461- 1464.

Moore DV, Lanier JE (1961). Observations on two Plasmodium falciparum infections with an abnormal response to chloroquine *Am. J. Trop. Med. Hyg.* 10: 5-9.

Olliaro PL, Trigg PI (1995). Status of antimalarial drugs under development. *Bull World Health Organ* 73: 565-71.

Padmanaban G, Rangarajan, PN (2001). Emerging targets for antimalarial drugs *Expert Opin. Ther. Targets* 5: 423-441.

Price RN, Nosten F, Luxemburger C, ter Kuile FO, Paiphun L, Chongsuphajaisiddhi T, White N J (1996). Effects of artemisinin derivatives on malaria transmissibility. *Lancet* 347: 1654-8.

Rasooli L et. al. (2003) Microbial sensitivity to and chemical properties of the essential oil of *Artemisia annua* L. *J. Essen. Oil Res.* 15 (1): 59 - 62.

Ridley RG (2002). Medical need, scientific opportunity and the drive for antimalarial drugs. *,ature* 415: 686-93.

Ronn AM, Msangeni HA, Mhina J, Wernsdorfer W H, and Bygbjerg I C (1996). High level of resistance of Plasmodium falciparum to sulfadoxine-pyrimethamine in children in Tanzania. *Trans R Soc Trop Med Hyg.* 90: 179-81.

Sachs J, Malaney P (2002). The economic and social burden of malaria. *,ature.* 415: 680-685.

Sharma VP, (1996) Reemergence of malaria in india *Indian J. Med. Res.* 103: 26-45.

Sharma VP, (1998) Fighting malaria in India *Curr. Sci.* 75: 1127-1140.

Sharma VP, (1999) Current scenario of malaria in India *Parasitologia.* 41: 349-353.

Sharma, VP, (2003) Malaria and poverty in India. *Curr. Sci.* 84: 513-515.

Sherman IW (1998) (Ed.) In Malaria. *Am Soc Microbiol Press, Washington, DC.* pp. 1-556.

Singh NP, Lai H, (2001) Selective toxicity of dihydroartemisinin and holotransferrin toward human breast cancer cells. *Life Sci.* 70 (1): 49-56.

Sowunmi A, Oduola A M, Ogundahunsi O A, and Salako LA (1998). Comparative efficacy of chloroquine plus chlorpheniramine and pyrimethamine/sulfadoxine in acute uncomplicated falciparum malaria in Nigerian children. *Trans R Soc Trop Med Hyg* 92: 77-81.

Surolia N, RamachandraRao S R, Surolia A (2002). Paradigm shifts in malaria parasite biochemistry and anti-malarial chemotherapy. *BioEssays*, 24: 192-196.

Sy LK et. al. (1998) A novel endoperoxide and related sesquiterpines from *Artemisia annua* which are possibly derived from allylic hydroperoxides. *Tetrahedron.* 54(17): 4345-4356.

Sy, LK et. al. (2001) Structure elucidation of artannuin-O, a novel cadinane diol from *Artemisia annua*, and the synthesis of arteannuin- K, arteannuin- L, arteannuin- M, arteannuin- O. *Tetrahedron*, 57 (40): 8481-8493.

Tewari A, Bhakuni RS (2003). Terpenoid and lipid constituents from *Artemisia annua*. *Indian J. Chem*. Sect B, 42 (7):1782-1785.

Tu YY et. al. (1981) Chemical constituents in *Artemisia annua* L. and derivatives of artemisinin. *Zhong Yao Tong Bao*. 6 (2):31.

Tu YY et. al. (1981) Studies on the constituents of *Artemisia annua* L. *Acta Pharmacuetica Sinica*, 16 (5): 366-370.

Van Agtmael MA, Eggelte TA, and van Boxtel CJ (1999). Artemisinin drugs in the treatment of malaria: from medicinal herb to registered medication. *Trends Pharmacol Sci.* 20: 199-205.

Varadi A, Tusnady GE, Sarkadi B, (2003) Membrane topology of the human ABC
transporter proteins. In: ABC Proteins From Bacteria to Man. I. B. Holland, S. P. C. Cole, K. Kuchler & C. F. Higgins (eds). London/San Diego: *Academic Press*, pp. 37-46.

Vathsala, P. G. *et al.*, Widespread occurrence of the *plasmodium falciparum* chloroquine resistance transporter (pfcrt) gene haplotype svmnt in *p. falciparum* malaria in India *Am. J. Trop. Med. Hyg*., 2004, 70, 256-259.

Wallaart E (2000) "Investigations on biosynthesis of the novel antimalarial drug artemisinin in the plant *Artemisinin annua*" Ph. D. thesis Rjkniversiteit Groningen, 16.

Ward SA, Bray PG, Mungthin M, Hawley SR (1995). Current views on the mechanisms of resistance to quinoline-containing drugs in Plasmodium falciparum. *Ann Trop Med Parasitol* 89: 121-4.

Wernsdorfer WH, and Payne D (1991). The dynamics of drug resistance in Plasmodium falciparum. *Pharmacol Ther* 50: 95-121.

White N (1999a). Antimalarial drug resistance and combination chemotherapy. *Philos Trans R Soc Lond B Biol Sci.* 354: 739-49.

White NJ (2008). Plasmodium knowlesi: the fifth human malaria parasite. *Clin Infect Dis* 46: 172-173.

WHO (2008a). Global malaria control and elimination: report of a technical review. In.: World Health Organisation, pp. 1-47.

WHO (2008b). Treatment policies. In: World Health Organisation. pp. 1-52.

WHO, (2008c). World Malaria Report 2008. In.: World Health Organisation, pp. 1-190.

WHO, First Meeting of Regional Technical Advisory Group on Malaria, Manesar, Haryana, 15-17 December 2004, World Health Organization, Regional Office for South-East Asia, New Delhi, SEA-MAL-239, pp. 1-38.

WHO, Report, Regional Office for South-East Asia, New Delhi, SEA-Mal- 237, October 2004, pp. 1-25.

WHO,Report,2001,

WHO. Roll Back Malaria. *A Global Partnership.* Geneva: World Health Organization. (1998).
Wirth DF (2002). Biological revelations. ,*ature.* 419: 495-496.

Wu LJ et al. (1995) The experimental studies of early treatment of artesunnate against schistosomiasis. *Chin .J. Schistoson. Control,* 7 (3): 129-133.

Wu ZH, Wang YY (1984). Structure and synthesis of arteannuin and related compounds XI. Idenification of epoxy arteannuninic acid. *Acta Chim. Sinica.* 42: 596.

Zheng GQ et. al. (1994). Cytotoxic terpenoids and from artemisia. *Planta Medica.* 60: 54-57. Zhu

DY et. al. (1984) Structure of artemisilactone. *Acta Chim.Sinica.* 42 (9): 937-939.

Zhu DY. et. al. (1982). Study on antibacterial constituents of Qinghaosu (*Artemisa annua* L.) *Zhongcaoyao.* 13 (2): 54.

María e. Maggi,1 arnaldo mangeaud,2 maría c. Carpinella,1 Carlos g. Ferrayoli,1 graciela r. Valladares,2 and sara m. Palacios1, genome-wide investigation of intron length polymorphisms and their potential as molecular markers in rice (oryza sativa l.)

Xusheng wang, xiangqian zhao, jun zhu, and weiren wu institute of bioinformatics, huajiachi campus, zhejiang university, hangzhou, 310029, p. R. China

Development of intron polymorphism markers weiren wu college of crop science

Fujian agriculture & forestry university.

Pip: a database of potential intron polymorphism markers

Long yang, gulei jin, xiangqian zhao, yan zheng, zhaohua xu and weiren wu*

Department of agronomy, college of agriculture and biotechnology, zhejiang university, hangzhou, 310029, china. Received on march 18, 2007; revised on may 12, 2007; accepted on may 25, 2007

Quantification of artemisinin: analysis of presence of artemisinin by spectrophotometric and gas-chromatographic techniques (gc- fid).