DNA Molecular Markers Based Authentication of Herbal Drugs - A Review
Biswa K1*, Biswas R2
1Department of Biotechnology, DAV College, Sector-10, Chandigarh, India.
2Department of Zoology, Government Home Science College, Sector-10, Chandigarh, India.
Manuscript No: IJPRS/V3/I1/00131, Received On: 14/03/2014, Accepted On: 20/03/2014

ABSTRACT
Herbal drugs mainly comprising of the crude plant parts, extracts, oils, gums, mixtures of extracts, is an integrated part of the traditional system of medicine especially in the developing countries. About 75% of the world population is dependent on this system of medicine and there has a global market of US $62 billion and is expected to grow to US $5 trillion by the year 2050. However, there is no proper method of standardization to authenticate these drugs which are adulterated. The adulteration lies mainly with the starting material. This adulteration may be intentional or non-intentional. The authentication of these medicinal plants are carried out up to a certain extent by chemical fingerprinting methods which however may not give the correct identification due to variation in the chemical composition arising from age and genotype of the plants and due to geographical variation. Hence the most desirable way to authenticate these plants is by the genome based methods. Developing DNA molecular markers and bar-coding these plants by sequencing a standard region of the DNA are best way to identify the adulterants as well as authenticate the desired species of plant. Through this review an attempt has been made to present in a nutshell the various genome based methods used by scientists throughout the world mainly in the last decade to successfully develop and establish various DNA markers for authentication of different medicinal plants and herbal drugs.

KEYWORDS
World Health Organization, Genotype, Restriction Fragment length polymorphism

INTRODUCTION
Herbal drugs are integrated part of both modern medicine and traditional system of medicine since pre historic days and known to be the oldest health care products. According to World Health Organization (WHO) total global herbal drug market is estimated as US $62 billion and is expected to grow to US $5 trillion by the year 2050. In India Ayurveda alone contributes Rs 3500 crores (US $813 million) annually to the internal market.

The other Indian medicinal plants-based industry is growing at the rate of 7–15% annually and is expected to grow by 20% in the near future. According to findings of the Associated Chambers of Commerce and Industry of India (Assocham) “Indian herbal market is registering an extremely significant growth and is likely to reach Rs.15,000 crore, in the domestic market by 2015 and increase in demand from present US$ 1.5 billion to US$ 3 billion in the international market”.

According to WHO “Herbal Medicines include herbs, herbal materials, herbal preparations and herbal finished products”. Herbs include crude plant material, such as leaves, flowers, fruit, seeds, stems, wood, bark, roots, rhizomes or

*Address for Correspondence:
Dr. Kakoli Biswas
Department of Biotechnology,
DAV College, Sector-10, Chandigarh, India.
E-Mail Id: kakolibiswas14@gmail.com
other plant parts, which may be entire, fragmented or powdered. Herbal materials include herbs, fresh juices, gums, fixed oils, essential oils, resins and dry powders of herbs. In some countries, these materials may be processed by various local procedures, such as steaming, roasting or stir-baking with honey, alcoholic beverages or other materials.

Herbal preparations are the basis for finished herbal products and may include comminuted or powdered herbal materials, or extracts, tinctures and fatty oils of herbal materials. They are produced by extraction, fractionation, purification, concentration, or other physical or biological processes. They also include preparations made by steeping or heating herbal materials in alcoholic beverages and/or honey, or in other materials. Finished herbal products consist of herbal preparations made from one or more herbs. If more than one herb is used, the term “mixture herbal product” can also be used. Finished herbal products and mixture herbal products may contain excipients in addition to the active ingredients. However, finished products or mixture herbal products to which chemically defined active substances have been added, including synthetic compounds and/or isolated constituents from herbal materials, are not considered to be herbal\(^1\).

According to WHO, 75-80% of the world population is dependent on herbal medicines especially in the developing countries in the primary health care mainly due to their better cultural acceptability, better compatibility with the human body and lesser side effects\(^2\). India officially recognizes over 3000 plants for their medicinal value. It is generally estimated that over 6000 plants in India are in use in traditional, folk and herbal medicine, representing about 75% of the medicinal needs of the Third World countries\(^3\). There are about 120 distinct chemical substances present in plants which are presently used as medicines in the whole world, while several other drugs are simple modifications of these substances\(^4\).

A standard authenticated process is available for quality control of the synthetic drugs, but no such standard process exists for raw materials or finished products in the traditional drugs. Herbal medicinal products may vary in composition and properties, unlike conventional pharmaceutical products, which are usually prepared from synthetic, chemically pure materials by means of reproducible manufacturing techniques and procedures. As defined by American Herbal Products Association, “Standardization refers to the body of information and controls necessary to produce material of reasonable consistency. This is achieved through minimizing the inherent variation of natural product composition through quality assurance practices applied to agricultural and manufacturing processes\(^5\). Correct identification and quality assurance of the starting material is, therefore, an essential prerequisite to ensure reproducible quality of herbal medicine, which contributes to its safety and efficacy\(^6,7,8\) and the root cause lies mainly with the starting material where the adulteration takes place. These drugs are frequently found to be adulterated with other drugs. Adulteration in herbal drugs may be intentional for monetary gains or unintentional due to lack of awareness and knowledge about the authentic plants, confusion in vernacular names between indigenous systems of medicine and local dialect, similarity in morphological and aromatic features of the different plant sources, non-availability of the authentic plants, careless collection and other unknown reasons\(^9\). Also there is a potential for adulteration of plants with extracts from plants that have lower drug content. Thereby this requires scientific methods for their authentication and standardization.

Methods for Identification, Limitations and Alternative Techniques

Most of the regulatory agencies and pharmacopoeias suggest various methods like macroscopic and microscopic evaluation and chemo-profiling using chromatographic (TLC and HPTLC) and spectroscopic (MS and LC-MS) techniques for quality control and standardization for the purpose\(^10,11\). Macroscopic and microscopic evaluation has its...
DNA Molecular Markers Based Authentication of Herbal Drugs - A Review

own limitations because of the similar morphological features and cell types. Use of molecular markers which are generally referred to biochemical constituents, including primary and secondary metabolites (phenotype) and other macromolecules such as nucleic acids (Genotype) can be relied upon. However, variations in the chemistry of the secondary metabolites due to the age of the plants and their varied geographical distribution limits its use as marker compounds as the secondary metabolites tend to vary with these two factors and therefore, restricts it only to their qualitative studies. DNA as molecular markers have several advantages over typical phenotypic markers and are reliable for informative polymorphisms as the genetic composition is unique for each species and is not affected by age, physiological conditions as well as environmental factors. DNA can be extracted from fresh or dried organic tissue of the botanical material hence the physical form of the sample for assessment does not restrict detection. Based on the specificity of the genotype of a system, a particular DNA profile being unique can be ascribed to a particular organism. Hence, various DNA markers based methods can be used for species characterization and adulteration detection in medicinal plants. In the past authentication by DNA profiling and their patents have been reviewed by several authors.

1. Hybridization Based
- Restriction fragment length polymorphism (RFLP)
- Variable number tandem repeat (VNTR)
- Microarray or chip based

2. Polymerase Chain Reaction (PCR) Based
- Random amplification of polymorphic DNA (RAPD)
- Amplified fragment length polymorphism (AFLP)
- Short tandem repeat (STR) or Simple Sequence Repeats (SSR) or Micro satellite polymorphism
- Polymerase Chain reaction-Restriction fragment Length Polymorphism (PCR-RFLP)
  - (ITS, 5s rRNA, rbcl, matK coxI and coxII genes can be used)
  - Direct amplification of length polymorphism (DALP)
  - Sequence characterized amplified regions (SCAR)
  - Inter Simple Sequence Repeats (ISSR)
  - Amplification Refraction Mutation System (ARMS)
  - Sequencing and Bar Coding

Hybridization Based

Restriction Fragment length polymorphism (RFLP) and Variable Number Tandem Repeats (VNTRs) finds its application majorly in forensic science and are not used for plants identification.

Microarray and Chip Hybridization

Other than these methods Microarray hybridization can be used for authentication of plant based drugs. Microarray also known as the DNA chip or biochip is a collection of thousands of microscopic spots of DNA molecules arranged in an array. Each DNA spot contains picomoles ($10^{12}$) of known probe
labeled with specific fluorescent dyes which, are allowed to hybridize with several other DNA or RNA molecules at stringent experimental condition. The hybridization pattern is then interpreted through various algorithms. Microarray has been widely used to authenticate traditional Chinese herbal drugs\textsuperscript{16,17}. DNA microarray was also used to identify \textit{Dendrobium} species from Chinese medicinal formulation\textsuperscript{18}. In a pioneer work Chen \textit{et al.}, applied molecular biotechnology to the identification of Chinese Patent Medicine (CPM), particularly well refined oral liquids and injections\textsuperscript{19}. They used a PCR based method can also be developed to an easy to use and cost-effective visual chip by taking advantage of G-quadruplex based Hybridization Chain Reaction. This method was carried out on a CPM product of ginseng \textit{Panax ginseng}. Some dealers substitute the ginseng root with that of codonopsis (\textit{Codonopsis pilosula} (Franch.) Nannf.) Which, is less curative in effect and much cheaper than ginseng. This study demonstrates that DNA identification of specific Medicinal materials is an efficient and cost-effective way to audit highly processed CPM and will assist in monitoring their quality and legality.

Polyme...
and authenticating *Swerta chirata* from its adulterated counterpart *Andrographis paniculata*. US Patent 6803215 was issued for the use of SCAR to differentiate plant and animal medicinal materials, including *Panax* and medicinal snakes. Several sets of SCAR primers were then designed to distinguish *Panax ginseng* and *Panax quinquefolius* and their adulterants.

**PCR-Restriction Fragment Length Polymorphism (PCR-RFLP)**

PCR-RFLP uses restriction endonucleases to digest PCR products of regions with sequence polymorphisms. Endonuclease which recognizes and cleaves at the polymorphic sites, digest a longer PCR fragment into smaller fragments and change the banding pattern. It is preferable to have conserved regions for constructing the universal primers and two such are ribosomal DNA (rDNA) and large subunit of ribulose-1,5, biphosphate carboxylase L (rbcL) genes. PCR-RFLP (ITS-RFLP) has been used for authentication of *Panax* species, *Fritillaria pallidiflora*, *Atractylodes* species, and others from their adulterants. ITS-RFLP has been used by us to distinguish between two *Boerhavia* species and *Trianthema portulacastrum* which has similar morphological features and the roots of these plants are used for the active metabolite and also to identify six out of ten plants in Dasamula-an Indian ayurvedic drug. RFLP of the internal transcribed spacer (ITS) gene is a common method to detect polymorphism to establish species specific patterns. The internal transcribed spacer is a sequence of RNA in a primary transcript that lies between precursor ribosomal subunits. These sequences are coded by nuclear ribosomal DNA. Eukaryotic organisms have two internal transcribed spacers; ITS-1 located between the 18S gene and the 5.8S gene, while ITS-2 is located between the 5.8S and the 28S gene. It has typically been most useful for molecular systematics at the species level, and even within species (e.g., to identify geographic races) because of its higher degree of variation than other genic regions of rDNA (for small- and large-subunit rRNA).

Variation among individual rDNA repeats can sometimes be observed within the ITS. In the ITS region restriction digestion shows specific patterns for a species and a variation in the pattern within a species, which can be detected as polymorphism. Thin-layer chromatography (TLC) and (PCR-RFLP) methods were used to differentiate the three medicinal plants, *Thunbergia laurifolia*, *Crotalaria spectabilis*, and *Curcuma aff. amada*, share the common name ‘Rang Chuet’ a thai medicine for the treatment of poisoning in which *matK* gene was used for PCR-RFLP. The ITS sequences of several medicinal species, including *Panax ginseng* (*Renshen*), *Panax quinquefolius* (*Xiyangshen*), and *Codonopsis* (*Dangshen*) species, have been patented and their distinct RFLP pattern was developed to distinguish them. US Patent 5876977 was issued for a method of fingerprinting ginseng by taking advantage of the unique PCR-RFLP patterns of different ginsengs. Besides the ITS region, 5S rRNA intergenic spacer (5SrRNA) has also been employed to authenticate medicinal material. US Patent 6569625 was issued for a method of differentiating four medicinal *Fritillaria* species using this DNA region. *Fritillaria cirrhosa* could be distinguished from *F. thunbergio* by PCR-RFLP of the 5s rRNA gene when cleaved by EcoRI but *F. anhuiensis* and *F.puquiensis* were not cleaved. 5s rRNA-NTS (Non Transcribed Spacer) gene RFLP analysis has given successful results from several medicinal and aromatic plants. Successful comparison of NTS region has been obtained at both inter-specific and intra-specific levels. Certain essential oil producing plants such as *Picea glauca* and *Pseudotsuga menziesii*, the NTS region shows variation not only in size but also in the number of different size classes. PCR-RFLP of the 5s-rRNA- NTS gene with *EcoRI* of both diploid (β-asarone-free) and triploid (βasarone rich) *Acorus calamus* showed 44 different size for both the cyto-types. Further phyto-chemical analysis and cytotype studies were performed to confirm the results. In the genus *Fritillaria* there are around 25 species and varieties which carry the name *Beimu* in commercial markets. Rapid
identification of these species and varieties was carried using PCR amplification of the ITS region by designing primers to the conserved coding region followed by digestion with EcoRI. A similar technique was used for the identification of thujone free chemotypes of Artemisia umbelliformis. By PCR-RFLP method using Rsal and TaqI restriction enzymes, the two chemotypes were clearly distinguished. In the Salvia species DNA fingerprinting of the 5S-rRNA-NTS was used for the identification of S. divinorum. PCR-RFLP method was performed using NdeI and TaqI. An NdeI site absent in S. officinalis was found in S. divinorum NTS region. For TaqI, multiple restriction sites were found in S. officinalis, whereas a unique site was found in S. divinorum.

A similar approach was employed for the identification of Astragalus species. Astragalus genus which is mainly found in China has around 300 species and varieties and thus making the identification of the plant difficult. After amplification, sequencing and comparison of the 5S-rRNA of around 300 bp restriction mapping showed diversity in the spacer region leading to the genetic identification of Huangqi.

**Simple Sequence Repeats (SSR) Analysis/Microsatellites**

SSR analysis is also referred to as simple sequence length polymorphism (SSLP). SSR is also known as microsatellites or short tandem repeats (STRs) which are short tandem repeats of 2–8 nucleotides widely and abundantly dispersed in most nuclear eukaryotic genomes. The different numbers of repeating units (alleles) in polymorphic loci lead to variation in band sizes. Amplification of microsatellites helps to generate a unique pattern and establish a specific profile for an individual. This can be extensively used to identify and authenticate the plants showing inter-specific variation. SSR analysis has been applied in authentication of ginseng. In SSR analysis, Panax quinquefolius showed different allele patterns compared with those of Panax ginseng (Renshen). Moreover, cultivated and wild Panax quinquefolius (Xiyangshen) can be distinguished from each other. New microsatellite markers were identified in Panax ginseng and Acanthopanax senticosus.

**Inter Simple Sequence Repeats Markers (ISSR)**

Another molecular marker technique called inters simple sequence repeats (ISSRs) was developed which are semi arbitrary markers amplified by PCR in the presence of one primer complementary to a target microsatellite. Such amplification does not require genome sequence information and leads to multilocus and highly polymorphic patterns. Each band corresponds to a DNA sequence delimited by two inverted microsatellites. The primers used can be either unanchored or more usually anchored at 3’ or 5’ end with 1 to 4 degenerate bases extended into the flanking sequences. Like RAPDs, ISSRs markers are quick and easy to handle, however they have the reproducibility of SSR markers because of the longer length of their primers. ISSR markers were undertaken with the objective of establishing specific molecular markers to authenticate and distinguish between three genuine species of rhubarb (Rheum officinale Baill., Rheum palmatum L., and Rheum tanguticum Maxim. ex Balf) whose roots and rhizomes are prescribed as medicine for various ailments in China.
fragments are carried out\textsuperscript{70,71}. It has been used for authentication and studying genetic diversity of Chinese medicines\textsuperscript{33,72,75}.

**Direct Amplification of Length Polymorphism (DALP)**

DALP uses a selective forward primer containing a 5' core sequence (e.g. M13 universal sequencing primer) plus additional bases at the 3' end and a common reverse primer (e.g. M13 reverse primer) to generate multi banded patterns in denaturing polyacrylamide gel\textsuperscript{74}. DALP has been used to detect polymorphisms between species\textsuperscript{74,75} and between strains\textsuperscript{74} and to authenticate Panax ginseng and Panax quinquefolius\textsuperscript{72}.

**Amplification Refraction Mutation System (ARMS) and Multiplex Amplification Refraction Mutation System (MARMS)**

Amplification Refraction Mutation System (ARMS) is a variation of PCR that exploits the fact that the primers only bind to their target sequence when their 3' ends are complementary. Oligonucleotides having mismatched 3' end ("mutated") residues will not bind to the "normal" target sequences and no amplification will take place. Multiplex Amplification Refraction Mutation System (MARMS) is the multiplex PCR using a common primer and multiple mutation specific primers as used in ARMS. Species identification from Ginseng drugs was carried out using MARMS\textsuperscript{76}. ARMS were used for molecular authentication of Alisma orientale\textsuperscript{77}.

**Sequencing and Bar Coding**

Sequence based techniques can give information at the species level of known as well as unknown plants. Much information can be obtained from DNA sequences\textsuperscript{78}. The technique of sequencing based identification is also known as Bar Coding where a small part of the DNA, normally a small gene is sequenced and matched with the existing DNA sequences from the DNA data bases. These data bases are intended to consist of sequences of the standardized region of entire organisms existing on the Earth. Hebert \textit{et. al.}, proposed that sequence from a small standardized region of the genome could serve as a species recognition tag\textsuperscript{79,80}. Thus, an unidentified organism or tissue could be ascribed to a species when such a sequence from it is compared with those available in a database\textsuperscript{79}. Another advantage of using sequencing for species identification is that the identities of adulterants can be identified by performing sequence searches on public sequence databases such as GenBank. However, prior sequence knowledge is required for designing primers for amplification of the region of interest\textsuperscript{78}. The more variable is a particular region, the more closely related individuals can be differentiated by this region. Regions commonly used include nuclear rRNA genes, mitochondrial genes and chloroplast genes\textsuperscript{81}. These are mainly ITS rRNA, 5s rRNA gene, rbcl, matK, rpoB, rpoC\textsubscript{1}, trnH-psbA, atpA, atpB, atpF, chlB, rps16, rp14, rpL16 etc. By database searching, Mihalov \textit{et. al.}, successfully identified soybean substituted for ginseng (\textit{Panax} species) in commercial samples\textsuperscript{82}. A sequence of the spacer region between atpF and atpA of the plastid genome allows discrimination of three varieties of the Chinese medicinal plant \textit{Angelica}\textsuperscript{83}. It was later established that sequencing of the ITS region can also reveal information within the species and thus help in authentication of the medicinal plant \textit{Angelica sinensis} which is adulterated by other morphologically species of \textit{Angelica}\textsuperscript{84}. Based on various previous studies and recommendations (5, 20-33), Singh \textit{et. al.}, used four loci from chloroplast genome namely rbCl, rpoC\textsubscript{1}, rpoB and matK and one locus from nuclear genome (nuclear ribosomal ITS) to establish the inter-specific and intra-specific variations among \textit{Dendrobium} species an Orchid of high medicinal and ornamental value\textsuperscript{85}. Studies on the bar-coding of various medicinal plants based on sequencing of both nuclear genes and chloroplast genes have been reviewed in detail by Sucher and Carles\textsuperscript{86}, Singh \textit{et. al.}\textsuperscript{85}.

No specific conclusion has been reached on which gene combination could decipher the perfect identity of any plant species, it would be
more appropriate to standardize a specific gene or gene combination for each experimental plant of study. Yip et al., proposed a model for selection of appropriate molecular identification method for species and strain level has been

CONCLUSION

Adulteration is a major problem in the herbal drugs market and therefore authentication and standardization is prerequisite for them to minimize the unfair trade as well for modernization, industrialization and internationalization of all traditional herbal medicines. Much of this is carried out by chemical fingerprinting which is not always reliable due to variation in the metabolite content. DNA markers are more reliable and can authenticate plants up to the varietal level. However some of these techniques like microarray and sequencing are expensive but other DNA marker based methods can be used for reliable authentication. From the number of choices of methods for genome based marker it would be appropriate to choose and start with more feasible techniques like RAPD, Microsatellite and/or PCR-RFLP and then move on to the more complex techniques as these techniques can reveal the identity even at the species level. Another aspect has to be taken care that the active metabolite may not be present in desirable quantity in the authenticated plants. This can be therefore combined with the chemical analysis and also the authenticated plants can be cultivated in the geographical regions where it is known to produce high amount of the required active metabolite.

REFERENCES


48. Suwanchaikasem, P., Phadungcharoen, T., & Sukrong, S. (2013). Authentication of the Thai medicinal plants sharing the same common name ‘Rang Chuet’: Thunbergia laurifolia, Crotalaria spectabilis, and


single primers in the polymerase chain reaction to distinguish strains of Cryptococcus neoformans. *Journal of Clinical Microbiology*, 31(9), 2274-2280.


