PCR, sequencing and PCR–RFLP of the 5S-rRNA-NTS region as a tool for the DNA fingerprinting of medicinal and aromatic plants

PCR, sequencing and PCR–RFLP of the 5S-rRNA-NTS region as a tool for the DNA fi ngerprinting of medicinal and aromatic plants†‡

Giorgio Gnavi,[a] Cinzia M. Bertea[a] and Massimo E. Maffei*

ABSTRACT: Molecular genetic methods have several advantages over classical morphological and chemical analyses. For instance, the genetic method requires genotype instead of phenotype, therefore DNA-based techniques have been widely used for rapid identifi cation of herbal medicine and aromatic plants. Using PCR approaches, nanogram quantities of DNA are required to amplify and yield suffi cient amounts of template DNA for molecular genetic analysis. Recently, the molecular discrimination of some higher plant species has been evaluated using sequences of a 5S-rRNA gene spacer region. The variation in the non-transcribed sequence (NTS) region has been used in a number of plant species for studying intraspecifi c variation, mapping 5S rDNA arrays, genome evolution and phylogenetic reconstruction. In this minireview we summarize the potential use of the 5S-rRNA-NTS region as a tool for the DNA fi ngerprinting of medicinal and aromatic plants.

Keywords: DNA fi ngerprinting; 5S-rRNA; non-transcribed sequence region; PCR–RFLP; medicinal and aromatic plants

Introduction

General approaches to herbal identifi cation depend on morphological, anatomical and chemical analyses; such analyses aimed at detecting and quantifying plant samples are aff ected by environmental and/or developmental factors or by methods of sample storage.[1] This means that, according to environmental conditions, the same genotype may express diff erent chemical patterns or, conversely, that diff erent genotypes may respond to the same environmental pressure with the same phenotypic expression.[2] Chemotypes (or chemical phenotypes) are generally considered to be the phenotypical expression of a genotype, although diff erent chemotypes may derive from the same genotype. Very often the identifi cation of plant samples in a mixture is diffi cult to achieve and this problem is particularly exacerbated when plant mixtures are powdered. When toxic/hallucinogenic plants are only present in powder, plant identifi cation is usually achieved by the determination and quantifi cation of active compounds, which often requires time and gas chromatographic (GC)– or liquid chromatographic (LC)–mass spectra determinations. By contrast, DNA analysis is relatively fast and the presence of molecules with higher stability makes quantifi cation easier, provided that specifi c primers for target genes can be used.[3] In this context, molecular genetic methods have recently been shown to be very eff ective in genotypic discrimination.[2] Molecular genetic methods have several advantages over classical morphological and chemical analyses. For instance, the genetic method requires genotype instead of phenotype, therefore DNA-based experiments have become widely employed techniques for rapid identifi cation of herbal medicines. Using PCR approaches, nanogram quantities of DNA are required to amplify and yield suffi cient amounts of template DNA for molecular genetic analysis.[4]

Recently, the phylogenetic relationships of some higher plant species have been evaluated using sequences of a 5S-rRNA gene spacer region. The 5S-rRNA region is a component of all ribosomes, except in the mitochondria of certain species.[5] In all higher eukaryotes, 5S-rRNA is transcribed from hundreds to thousands of genes. Genes encoding 5S-rRNA are located separately from the 18S-26S rRNA gene clusters and organized into tandem repeats, with alternative arrays of sequences coding 5S-rRNA and non-transcribed spacers (NTSs) in one or more sites in the genome.[6] The two gene clusters can be localized either in a linked state, and therefore on the same chromosome,[7] or independently on diff erent chromosomes in the genome.[8] In higher eukaryotes, the 5S-rRNA genes are organized in tandem repeats of a basic unit, 200–900 bp long, with 1000–50 000 copies.[6] The gene is 120 bp long and is associated with spacers of various sizes to form the repeating unit of the tandem arrays. The 120 bp 5S-rRNA gene sequence is highly conserved across species, while the NTS region exhibits variation in base composition and length (in the range 100–800 bp) from species to species,[9] since it is apparently not under the same rigorous selection pressure as in the coding region.[10–12] The high level of conservation of the 5S-rRNA gene is associated with the precise function of 5S-rRNA as a component of the large ribosomal subunit in all eukaryotic organisms. Some regions of the gene are more conserved than others, which is explained by the regulation of 5S-rRNA transcription.[9] Sequence conservation of the coding regions and high divergence in the spacer regions provide a good model for studying the organization and evolution of multigenes in diff erent plant species.[7,13] Based on these assumptions, variation in the NTS region has been used in a number of plant

species for studying intraspecifi c variation, mapping 5S-rDNA arrays, genome evolution and phylogenetic reconstruction.[9,10,14]

PCR, Sequencing and RFLP

Polymerase chain reaction (PCR) is a rapid and sensitive procedure for the *in vitro* amplifi cation of specifi c segments of DNA. PCR is the standard procedure for detecting the genes of microorganisms, plants and animals. In PCR, repeated synthesis from the same primer sites selectively and exponentially amplifi es the sequence between the primer sites and produces a high concentration of identical amplifi cation products.[15] The ability to perform primer-directed amplifi cation of specifi c sequences of DNA has had an eff ect on research similar to that of the discovery of restriction enzymes and Southern hybridization. Since the introduction of PCR in 1986, an ever-increasing number of scientifi c applications have been reported, including the direct cloning, mutagenesis, sequencing and exact engineering of specifi c genes or gene sequences directly from genomic DNA or complementary DNA (cDNA).[16] Detection of the reacting products is performed by agarose gel electrophoresis, ethidium bromide staining, ultraviolet irradiation and comparison of the product size with a DNA size marker. New methods of PCR product detection, such as the application of capillary gel electrophoresis (e.g. Agilent 2100 Bioanalizer), has greatly improved the accuracy and semi-quantitative determination of PCR products and has opened up new possibilities in applying this technique during routine diagnosis.[17] The success of PCR in generating and detecting specifi c DNA fragments has accelerated the use of molecular biology in many biological systems. Furthermore, the technology to detect PCR products in real-time, i.e. during the reaction, led to the set-up of quantitative real-time PCR (qRT–PCR), which has become a common tool for detecting and quantifying the expression profi les of selected genes.[18] Recent progress in DNA sequencing technology has allowed the generation of large quantities of nucleotide sequence data in a short period of time.[19] Moreover, technological advances in plant gene isolation and identifi cation, such as map-based cloning, insertional mutagenesis and large-scale cDNA sequencing, have accelerated the rate of gene isolation and signifi cantly expanded opportunities for the genetic engineering of crop plants.[20]

The development and use of molecular markers for the detection and exploitation of DNA polymorphism is one of the most signifi cant developments in the fi eld of molecular genetics.[21] DNA markers can identify the organism and its taxonomic association, even from fragmentary remains and even where morphology cannot distinguish strains.[15] Polymorphic markers can defi ne a multilocus genotype characteristic for an individual or a clone; selected markers can be diagnostic for a population or a species. DNA polymorphisms can occur anywhere in the genome, including coding and non-coding, single-copy or repetitive DNA. Sequence polymorphisms in PCR amplifi cation products can be detected using restriction digestion.[22] The bands on gels, which typically serve as molecular markers, may arise from cutting DNA at specifi c sites with restriction enzymes, in order to detect restriction fragment length polymorphisms (RFLPs). Alternatively, bands may be derived by *in vitro* synthesis of a stretch of the target DNA between specifi c sites to which short, single-stranded primers attach and serve as starting (and end) points for PCR.[15]

Using the 5S-rRNA Gene for the DNA Fingerprinting of Medicinal and Aromatic Plants

Relevant results obtained with RFLP analysis of nuclear DNA have been obtained from several medicinal and aromatic plants using the 5S-rRNA-NTS gene. Successful application of NTS comparison has been obtained at both the interspecifi c and the intraspecifi c level. Below are some examples from gymnosperms and angiosperms. In contrast to angiosperms, the structure and organization of the 5S-rDNA locus has only been characterized in a few gymnosperm species.

In Douglas fi r (*Pseudotsuga* spp.), plants known to produce an essential oil characterized by monoterpenes such as bornyl acetate, camphene and α-pinene with proven antifungal activity,[23] sequencing and Southern hybridization of the 5S-rRNA gene revealed repeat units of 888 and 871 bp in length, the latter with a 17 bp deletion in the NTS. A 35 bp region of the NTS immediately upstream of the 52 end of the coding region showed high similarity to other conifers but not to other published plant 5S-rDNA sequences.[24] In the silver fi r (*Abies alba* Mill.), one of the most important conifers in many eastern European mountain forests, which is characterized by the presence of an essential oil containing terpenoids such as limonene, β-phellandrene, α-pinene, β-pinene and camphene,[25] PCR amplifi cation of the gene and NTS region, sequence analysis and Southern hybridization, using a homologous probe, detected DNA sequences of approximately 550 and 700 bp. Sequence analysis of the spacers revealed that the diff erence in length between the sequences occurred in the middle spacer region, as a result of the amplifi cation of a 75 bp sequence of the short unit class, which is organized into four 54–68 bp tandem repeats in the long spacer unit.[26] In *Picea glauca*, a plant producing an essential oil containing δ-3-carene, sabinene, β-pinene, borneol, linalool, β-phellandrene, β-caryophyllene and camphor,[27] and in *Pseudotsuga menziesii*, the 5S-rDNA repeats have a conserved 120 bp transcribed region and an NTS that varies not only in size (from 101 bp in *P. glauca* to 880 bp in *P. menziesii*) but also in the number of diff erent size classes, whereas in Asian pines, the length of the NTS varies (382–401 bp in *Pinus bungeana* and 538– 608 bp in four *diploxylon* pines).[28] In the conifer *Pinus radiata*, a species with an essential oil particularly rich in α- and β-pinene,[29] the 5S-rRNA gene (5S-DNA) has been cloned and characterized at the nucleotide, genomic and chromosomal levels. Sequencing revealed a repeat unit of 524 bp which is present in approximately 3000 copies per diploid genome.[30] Eleven *Pinus* species in the subgenus *Strobus* have only a short 5S-rDNA size class. New World species of the subgenus *Pinus* have both short and long size classes, whereas Old World species only have the longer size class.[31] A molecular analysis of 5S-rDNA of white spruce, *P. glauca*, revealed the presence of two classes of repeating units, one of 221 bp, corresponding to the PCR amplifi cation product, and another of approximately 600 bp.[5]

In the genus *Larix*, known to produce essential oils,[32] divergent size classes of 5S-rDNA were identifi ed in *L. decidua* and *L. kaempferi*, using either selective amplifi cation of gene and spacer, sequence analysis or homologous probe hybridization. Two highly divergent unit size classes of approximately 650 and 870 bp were detected in both species.[11]
Many more data are available from angiosperms. Australia is unique in having a single genus of tree, *Eucalyptus* (family Myrtaceae), dominating its forests and woodlands.[14] The trees are cultivated the world over for their oil, gum, pulp, timber, medicine and aesthetic value. The essential oil found in their foliage is the most important one and fi nds extensive use in the food, perfumery and pharmaceutical industries. In addition, the oil possesses a wide spectrum of biological activities, including antimicrobial, fungicidal, insecticidal/insect repellent, herbicidal, acaricidal and nematicidal.[33,34] Sequences of the 5S-rDNA repeat have been determined from two *Angophora* species and nineteen *Eucalyptus* species; the tandemly repeated 5S-rRNA genes were highly conserved, while the non-coding intergenic spacers were variable. A 50 bp repeating element, which has undergone duplication and modifi cation in certain taxa, was identifi able within the spacer and accounted for much of the variability. Based on the modifi cations of the 50 bp element, it is apparent that the spacer from bloodwood (informal subgenera *Blakella* and *Corymbia*) species of *Eucalyptus* were more similar to that of *Angophora* than to non-bloodwood species of *Eucalyptus*.[14] As a basis for further comparative studies, nuclear 5S-rRNA gene repeats from two plants of the family Solanaceae, tobacco (*Nicotiana rustica*) and tomato (*Lycopersicon esculentum*), were isolated and sequenced. *N. rustica* leaves bear glandular trichomes secreting nicotine,[35] while marked diff erences between the composition and yield of the essential oils from diff erent *Lycopersicon* species have been described.[36] The more abundant 5S-rRNA gene repeat in tobacco is 430 bp long, while a second, less common, variant is 521 bp long. In contrast, the 5S-rRNA gene repeat from tomato is only 355 bp long. The results indicated that often observed, but non-conserved, repeating sequence elements probably arise spontaneously by unequal crossover of no functional signifi cance.[37] Another important medicinal and essential oil plant belonging to the family Solanaceae is *Capsicum*. In some species of this genus, the compounds that are primarily responsible for the pungency are capsaicin (8-methyl-*n*-vanillyl-6-nonenamide) and a group of similar substances called capsaicinoids, which includes dihydrocapsaicin and nordihydrocapsaicin.[38] In sequence analysis, the repeating units of the 5S-rRNA genes in the *Capsicum* species were variable in size (278–300 bp). In sequence comparison with other members of the Solanaceae, the coding region was highly conserved but the spacer regions varied in size and sequence. While *C. chinense*, *C. frutescens* and *C. annuum* formed one lineage, *C. baccatum* was revealed to be a species intermediate between the former three species and *C. pubescens*.[6]
The phylogenetic relationship of *Acorus gramineus* and three types of *A. calamus* was analysed by comparing the 700 bp sequences of 5S-rRNA gene spacer regions. Although there was no intraspecifi c variation in the essential oil profi le of *A. gramineus*, which contained a phenylpropanoid (β-asarone) as a predominant constituent, *A. calamus* was classifi ed into two chemotypes: chemotype A, in which β-asarone is a major essential oil constituent, and chemotype B, which contained mainly sesquiterpenoids.[12] The high content of essential fatty acids in fl ax (also known by its synonym, linseed) is responsible for numerous health benefi ts, including cholesterol-lowering and anti-carcinogenic eff ects. The oil has been used for several traditional medicinal purposes.[39] A relatively large proportion of the *Linum* genome (approximately 3%) comprises 5S-rRNA genes. Intraspecifi c sequence variations among fi ve distinct groups of 5S-rRNA genes indicate that 5S-ribosomal genes most closely resemble other angiosperm 5S genes, while groups 3–5 are highly divergent. Sequence variation was higher in the spacer region than in the transcribed region for all pairwise comparisons.[40] The genus *Brassica* is known for its medicinal properties. Sinapine (*O*-sinapoyl choline), a choline ester of sinapic acid, is one of the major phenolic choline esters in oil-extracted rapeseed meal,[41] and a component of Semen Sinapis Albae (white mustard seed), a traditional Chinese medicine.[42] The 5S-rRNA gene from *Brassica campestris* has been cloned, sequenced and characterized; the 5S-rDNA repeat unit is 495 bp in length and consists of a highly conserved 119 bp coding region and a variable non-coding spacer region, which separates it from the coding region of the next repeat unit. Sequences responsible for initiation and termination of transcription of the 5S-rRNA were found to be present within the repeat unit.[43] In *Brassica nigra*, at the 5S-rRNA level the homology to other representatives of the same family was 97–100%. The degree of homology over the gene and spacer level between *B. nigra* and *S. alba*, which are very closely related, is 85–86%.[44] The essential oil from the leaves of *Eruca sativa* (Brassicaceae) is characterized by a high content of sulphur- and nitrogen-containing compounds, such as 4-methylthiobutylisothiocyanate and 5-methylthiopentanonitrile.[45] Two families of the 5S-rDNA, the 0.5 kb-size family and the 1 kb-size family, coexist in the *Eruca sativa* genome. The 0.5 kb-size family consists of the 5S-rRNA genes (S4) that have coding regions similar to those of other reported plant 5S-rDNA sequences, whereas the 1 kb-size family consists of 5S-rRNA gene variants (S1) that exist as 1 kb *Bam*HI tandem repeats.[46] The *Allium* family has over 500 members and, in addition to their nutritional eff ects, several species show antifungal and antibacterial activities. In garlic the antibiotic activity of 1 mg allicin, which is a (+)-*S*-methyl-l-cysteine sulphoxide, has been equated to that of 15 IU penicillin.[47] Using an arbitrary primer sequence within the 5S-rRNA gene, an effi cient probe sequence with a 320 bp NTS fl anking partial 5S coding sequences was obtained in *Allium fi stulosum*.[8] In another study, *A. cepa* and *A. schoenoprasum* were each found to possess 5S-rRNA units of two diff erent sizes. The nucleotide sequence of the long 5S-rRNA unit resulted from partial duplication of a NTS and the insertion of a unique sequence. Although the NTS of the 5S-rRNA of *A. cepa* and *A. schoenoprasum* had quite diff erent nucleotide sequences, the long 5S-rRNA units of *A. cepa* and *A. schoenoprasum* share a common 75 bp sequence.[48]
Sequences of 5S-rRNA gene spacer were used to identify *Epimedium brevicornu*, *E. sagittatum*, *E. wushanense*, *E. pubescens* and *E. koreanum*. These species are listed as source plants of Chinese medicine 'Ying Yang Huo' in the Chinese Pharmacopoeia.
For example, the essential oil producing *E. brevicornum maxim* is a centuries-old medicinal herb, which exerts benefi cial eff ects, viz. preventing bone loss in late postmenopausal women without resulting in a detectable hyperplasia eff ect on the endometrium.[49,50] The neighbour-joining method was used in a sequence analysis of *Epimedium* species. A position-specifi c nucleotide was found in the 5S-rRNA gene spacer for *E. pubescens*, *E. wushanense* and *E. brevicornu*. A 19 bp

deletion was found for *E. koreanum* in the 5S-rRNA gene spacer. *E. koreanum* was most divergent from the other four endemic Chinese species of *Epimedium*.[51]

Beimu (bulbs of *Fritillaria*) is another important traditional Chinese herbal medicine commonly used as an antitussive and expectorant. There are about 25 species and varieties of *Fritillaria* that carry the name 'Beimu' in commercial markets. The 5S-rRNA spacer region of the extracted DNAs was amplifi ed by PCR with a pair of primers located within the conserved coding region. The isolated DNA clones (similar to 600 bp) covering the 5S-rRNA spacer domain were sequenced. By aligning the isolated nucleotide sequences of the four *Fritillaria* species, sequence diversity was found in the spacer region. Furthermore, a unique *Eco*RI site was used for the rapid identifi cation of diff erent species of *Fritillaria*.[1]

A similar approach was employed for the identifi cation of *Astragalus* species. In this genus, some species produce essential oils characterized by high percentages of monoterpene and sesquiterpene hydrocarbons and alcohols.[52] Huangqi (Radix Astragali, root of *Astragalus*) is a traditional Chinese medicine commonly used as a tonic and diuretic. *Astragalus* has a wide range of immunopotentiating eff ects and has proved effi cacious as an adjunct cancer therapy.[53] About 300 species and varieties of *Astragalus* are determined in China, making the identifi cation of the origin of a particular *Astragalus* species on the consumer market diffi cult. Thus, the correct identifi cation of Huangqi is very important for the modernization of traditional Chinese medicine. The amplifi ed 5S-rRNA spacer regions (~300 bp) of diff erent species of *Astragalus* were sequenced and compared. Diversity in DNA sequence and restriction enzyme mapping among various species was found in their 5S-rRNA spacer domains, leading to the genetic identifi cation of Huangqi[54]

Wild *Saussurea lappa* in the family Asteraceae is a highly endangered plant producing an essential oil characterized by the presence of *trans*-α-ionone and *trans*-α-damascone enantiomers.[55] 5S-rRNA intergenic spacers are eff ective in discriminating *S. lappa* from its substitutes and adulterants. Sequencing results showed that the similarities of ITS-1, ITS-2 and 5S-rRNA intergenic spacers among *S. lappa* and related species were 56.3– 97.8%, 58.597.0% and 26.4–77.9%, respectively. The intraspecifi c variation was much lower. There are also several unique changes in the *S. lappa* sequences that may be used as diff erentiation markers.[56]

Recently, our group has studied some medicinal and aromatic plants in order to better characterize, from a genotypic point of view, those species whose taxonomical identifi cation is not possible through morphological investigations or is critical (e.g. long analytical procedures) through chemical analysis. Extending the work of Sugimoto *et al.*,[12] the 5S-rRNA spacer region of both diploid (β-asarone-free) and triploid (β-asaronerich) *A. calamus* were amplifi ed by PCR, using a pair of primers located at the 32 and 52 ends of the coding sequence of the 5S-rRNA gene. The PCR products were digested using Eco*RI* and the restriction profi le of the spacer domain was shown to be different for the two cytotypes. Along with chemical analysis of alcoholic extracts, sequence analysis coupled to restriction mapping was demonstrated to represent a powerful tool to distinguish the *A. calamus* diploid cytotype from the others[4] (Figure 1, lanes 1 and 2).

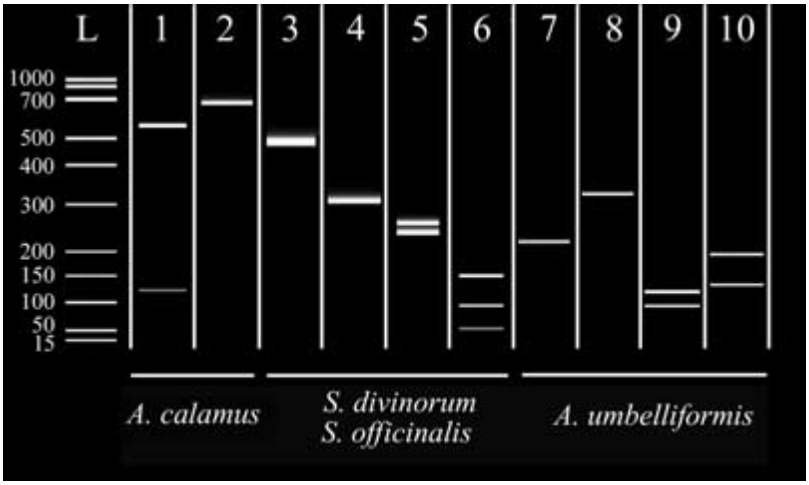

Figure 1. Some examples of DNA fi ngerprinting by the use of PCR– RFLP of the 5S-rRNA-NTS region. *Acorus calamus*: PCR–RFLP analysis using *Eco*RI; digested products from triploid (lane 1) and diploid (lane 2) 5S-rRNA gene spacer regions (modifi ed from ref. [4]). *Salvia divinorum* and *Salvia offi cinalis*: PCR–RFLP analysis using *Taq*I; *S. divinorum* undigested PCR products (lane 3), *S. offi cinalis* undigested PCR products (lane 4), *S. divinorum Taq*I PCR-digested products (lane 5), *S. offi cinalis Taq*I-digested PCR products (lane 6) (adapted from [59]). *Artemisia umbelliformis*: PCR product of the NTS spacer of chemotypes of *A. umbelliformis* containing thujone (Au1, lane 7) and without thujone (Au2, lane 8), Au1 *Rsa*I PCR-digested products (lane 9), Au2 *Taq*I PCR-digested products (lane 10) (adapted from [2]). L, bp ladder. The PCR products were separated using the Agilent 2100 Bioanalizer and the DNA 1000 LabChip® Kit (Agilent Technologies)

Salvia divinorum Epling & Játiva-M. is a perennial herb belonging to the family Lamiaceae and is most recognized for its hallucinogenic properties. The active ingredient of *S. divinorum* is the neoclerodane diterpene salvinorin A, a psychotropic molecule that produces hallucinations.[57] For this reason, *S. divinorum* is a frequently used hallucinogen, with a potency in producing hallucinations similar to that of LSD.[58] Molecular fi ngerprinting using the 5S-rRNA-NTS region allowed the rapid and precise identifi cation of *S. divinorum*. By aligning the isolated nucleotide sequences, great diversities were found in the spacer regions of *S. divinorum* when compared to those of *S. offi cinalis*. Specifi c *S. divinorum* primers were designed on the sequence of the 5S-rRNA gene spacer region. In addition, a PCR–restriction fragment length polymorphism (PCR–RFLP) method was applied using *Nde*I and *Taq*I restriction enzymes. An *Nde*I site, absent in *S. offi - cinalis*, was found in the *S. divinorum* NTS region at 428–433 bp. For *Taq*I, multiple sites (161–164, 170–173 and 217–220 bp) were found in *S. offi cinalis*, whereas a unique site was found in *S. divinorum* (235–238 bp)

(Figure 1, lanes 3–6).[59] Thus, even in this case, molecular analysis of the 5S-rRNA-NTS led to the precise and unequivocal identifi cation of a species. *S. divinorum* is often sold, in legal or illegal markets, as a powder that can be easily adulterated by adding dried leaves of other species, thus making it hard to establish the purity of samples.[60] We recently developed a new mathematical model for the quantitative analysis of *S. divinorum* in a biological mixture by quantifying DNA by means of SYBR Green I fl uorescence dye quantitative real-time PCR (qRT–PCR).[3] This model is based on relative quantifi cation of DNA extracted from a mixture vs. a reference DNA extracted from a known amount of the pure species. The results of this work showed an almost perfect correspondence between qRT–PCRcalculated weight and the weight estimated by an analytical weighted method, proving the eff ectiveness of this method for the quantitative analysis of a given species in a plant mixture.[3] Thujone is a natural terpenoid also associated with common wormwood (*Artemisia absinthium* L.) and Roman wormwood (*Artemisia pontica* L.), absinthe's most widely used ingredients.[61,62] There is currently a heated debate on the toxicity of absinthe and thujones (see ref. 61 and references therein), but European Union legislation has imposed a limit of 35 ppm on the total amount of these compounds in alcoholic beverages.[63] To overcome this issue, thujone-free chemotypes of *A. umbelliformis* have been selected by horticultural techniques.[64] Two chemotypes of *A. umbelliformis* (with and without thujone) used to prepare a local liqueur, genepi, were studied and specifi c *A. umbelliformis* primers were designed on the sequence of the 5S-rRNA gene spacer region. When a PCR–RFLP method was applied, using *Rsa*I and *Taq*I restriction enzymes, the two chemotypes were clearly distinguished[2] (Figure 1, lanes 7–10). As recently presented at the 40th ISEO in Italy, our research is ongoing and we will soon report on new genotypic characterizations of medicinal, aromatic and food plants using the 5S-rRNA gene NTS region.

Advantages and Limits of the 5S-rRNA-NTS Region as a Tool for DNA Fingerprinting

The information gathered in this minireview shows that the analysis, sequencing and PCR–RFLP of the 5S-rRNA gene NTS region can be used in diff erent species producing bioactive compounds or essential oils for a rapid and unequivocal molecular fi ngerprinting of these plant species and their chemo/genotypes. The method has been applied to several taxa belonging to systematically distant families and orders, and in all cases studied and here reported it showed a large potential for practical applications. The direct advantages are mainly applied in solving taxonomic uncertainties; however, its application can be extended to establishing the quality/origin of herbs and food/feed plants. In some cases, as demonstrated for *S. divinorum*, this technique may help to speed up forensic investigations and its application may allow the detection and identifi cation of specifi c herbs in plant mixtures. Because the 5S-rRNA genes are highly conserved, their sequence analysis is used for inferring phylogenetic relationships among deep branches of eukaryotes. Furthermore, the spacer region is more informative for the study of phylogenetic relationships at the interspecifi c and intergeneric levels, due to the faster rate of divergence in comparison to the highly conserved coding region.[26] All these considerations play in favour of the sustainability of this method for DNA fi ngerprinting of medicinal and aromatic plants.

By comparing taxonomic methods for plant identifi cation, chemical methods such as GC–MS or LC–MS, and biomolecular methods such as PCR or RFLP–PCR, we can try to understand the advantages and limits of the biomolecular method here reviewed. On the one hand, analytical methods help the taxonomic distinction of species when plant identifi cation is diffi cult to achieve, but conclusions are often deceived by the phenotypic plasticity of plants responding to environmental biotic and abiotic conditions, which aff ects gene expression and product (molecule) formation. On the other hand, DNA analyses and techniques are not so frequently used in laboratories where phytochemical analyses are performed and require the availability of (specifi c) primer sequences and the sequencing of regions of interest (such as the NTS described in this minireview). Although DNA analysis is currently considered to be cutting-edge technology, it has certain limitations that make its use mainly limited to academia. In order to establish a marker for the identifi cation of a particular species, DNA analysis of closely related species and/or varieties and common botanical contaminants and adulterants is necessary, and this is a costly process. Isolation of good-quality DNA suitable for analysis from semi-processed or processed botanicals is also a challenge. DNA fi ngerprinting might ensure the identifi cation of the correct genotype but does not reveal the contents of the active principle or chemical constituents. Hence, DNA analysis and phytochemical analyses (e.g. GC–MS, HPLC, LC–MS, etc.) should be used hand-in-hand rather than in isolation. Furthermore, it cannot be excluded that, whenever the two chemotypes for which a DNA fi ngerprinting has been assessed would mix, their DNA markers might lose their value due to meiotic rearrangements. Therefore, verifi cation of the system with samples used in the chemical profi ling is fundamental, as well as all statements addressing the relative stability of the defi ned DNA markers. Presently, some controversy exists over the value of DNA barcoding,[65] largely because of the perception that this new identifi cation method would diminish rather than enhance traditional morphology-based taxonomy. However, more and more gene sequences are now suitable for DNA barcoding of fl owering plants. Recently, Kress and co-workers[66] pointed out the necessity of employing more than one locus to attain species-level discrimination across all fl owering plant species, and stressed the need to look for algorithms for combining barcoding sequences from two or more DNA regions to yield species-level unique identifi ers.

Concluding Remarks

The use of chromatographic techniques and marker compounds to standardize botanical preparations has limitations because of their variable sources and chemical complexity. As the science of plant genetics has progressed, researchers have tried to explore these molecular marker techniques for their applications in commercially important plants, such as food crops and horticultural plants, and for the authentication of medicinal plants. DNA sequencing can be used as a defi nitive means for identifying species and, as briefl y reviewed in this paper, one of the most useful sequencing-based markers as diagnostic tools for authentication purpose is the spacer region of 5S-rRNA. The relative rapidity of the use of specifi c markers obtained from target species in PCR and/or RLFP–PCR analyses makes the spacer region of 5S-rRNA an important tool that deserves further research and application to better characterize important medicinal and essential oil-producing plants.

References

1. Z. H. Cai, P. Li, T. T. X. Dong, K. W. K. Tsim. *Planta Med.* **1999**, *65*, 360.
2. P. Rubiolo, M. Matteodo, C. Bicchi, G. Appendino, G. Gnavi, C. M. Bertea, M. E. Maff ei. *J. Agric. Food Chem.* **2009**, *57*, 3436.
3. P. Luciano, C. M. Bertea, G. Temporale, M. E. Maff ei. *Forens. Sci. Int. Genet.* **2007**, *1*, 262.
4. C. M. Bertea, C. M. M. Azzolin, S. Bossi, G. Doglia, M. E. Maff ei. *Phytochemistry* **2005**, *66*, 507.
5. G. R. Brown, J. E. Carlson. *Theoret. Appl. Genet.* **1997**, *95*, 1.
6. Y. K. Park, K. C. Park, C. H. Park, N. S. Kim. *Mol. Cell* **2000**, *10*, 18.
7. G. I. Scoles, B. S. Gill, Z.-Y. Xin, B. C. Clarke, C. L. McIntyre, C. Chapman, R. Appels. *Plant Syst. Evol.* **1988**, *160*, 105.
8. S. H. Lee, G. S. Do, B. B. Seo. *Chromatogr. Res.* **1999**, *7*, 89.
9. M. S. Negi, J. Rajagopal, N. Chauhan, R. Cronn, M. Lakshmikumaran. *Genome* **2002**, *45*, 1181.
10. W. J. Baker, T. A. Hedderson, J. Dransfi eld. *Mol. Phylogen. Evol.* **2000**, *14*, 218.
11. J. F. Trontin, C. Grandemange, J. M. Favre. *Genome* **1999**, *42*, 837.
12. N. Sugimoto, F. Kiuchi, M. Mikage, M. Mori, H. Mizukami, Y. Tsuda. *Biol. Pharmacol. Bull.* **1999**, *22*, 481.
13. A. V. Cox, M. D. Bennett, T. A. Dyer. *Theoret. Appl. Genet.* **1992**, *83*, 684.
14. F. Udovicic, G. I. Mcfadden, P. Y. Ladiges. *Mol. Phylogen. Evol.* **1995**, *4*, 247.
15. K. Bachmann. *New Phytol.* **1994**, *126*, 403.
16. L. M. Foster, K. R. Kozak, M. G. Loftus, J. J. Stevens, I. K. Ross. *Mycol. Res.* **1993**, *97*, 769.
17. T. Stadejek. *Med. Weteryn.* **2006**, *62*, 390.
18. S. A. Deepak, K. R. Kottapalli, R. Rakwal, G. Oros, K. S. Rangappa, H. Iwahashi, Y. Masuo, G. K. Agrawal. *Curr. Genom.* **2007**, *8*, 234.
19. H. Kotani, S. Tabata. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1998**, *49*, 151.
20. S. Gibson, C. Somerville. *Trends Biotechnol.* **1993**, *11*, 306.
21. K. Semagn, A. Bjornstad, M. N. Ndjiondjop. *Afr. J. Biotechnol.* **2006**, *5*, 2540.
22. T. Harada, K. Matsukawa, T. Sato, R. Ishikawa, M. Niizeki, K. Saito. *Euphytica* **1993**, *65*, 87.
23. V. Tesevic, S. Milosavljevic, V. Vajs, I. Dordevic, M. Sokovic, V. Lavadinovic, M. Novakovic. *J. Serbian Chem. Soc.* **2009**, *74*, 1035.
24. V. Amarasinghe, J. E. Carlson. *J. Hered.* **1998**, *89*, 495.
25. E. Duquesnoy, V. Castola, J. Casanova. *Flavour Fragr. J.* **2007**, *22*, 293.
26. V. Besendorfer, I. Krajacic-Sokol, S. Jelenic, J. Puizina, J. Mlinarec, T. Sviben, D. Papes. *Theoret. Appl. Genet.* **2005**, *110*, 730.
27. E. Puchalska, B. Czajkowska, M. Kielkiewicz. *Acta Physiol. Plant* **2008**, *30*, 225.
28. Z. L. Liu, D. M. Zhang, X. Q. Wang, X. F. Ma, X. R. Wang. *Am. J. Bot.* **2003**, *90*, 17.
29. P. V. Petrakis, C. Tsitsimpikou, O. Tzakou, M. Couladis, C. Vagias, V. Roussis. *Flavour Fragr. J.* **2001**, *16*, 249.
30. S. W. Gorman, R. D. Teasdale, C. A. Cullis. *Plant Syst. Evol.* **1992**, *183*, 223.
31. G. F. Moran, D. Smith, J. C. Bell, R. Appels. *Plant Syst. Evol.* **1992**, *183*, 209.
32. A. S. Timeryanov. *Khimiya Prirodn. Soed.* **1994**, 640.
33. D. R. Batish, H. P. Singh, S. K. Kohli, S. Kaur. *Forest Ecol. Manag.* **2008**, *256*, 2166.
34. D. J. Boland, J. J. Brophy. *ACS Symp. Ser.* **1993**, *525*, 72.
35. M. Meyberg, S. Krohn, B. Brummer, U. Kristen. *Flora* **1991**, *185*, 357.
36. R. M. Smith, J. A. Marshall, M. R. Davey, K. C. Lowe, J. B. Power. *Phytochemistry* **1996**, *43*, 753.
37. K. Venkateswarlu, S. W. Lee, R. N. Nazar. *Gene* **1991**, *105*, 249.
38. E. S. Poyrazoglu, O. Yemis, C. Kadakal, N. Artik. *J. Sci. Food Agric.* **2005**, *85*, 1435.
39. S. Krist, G. Stuebiger, S. Bail, H. Unterweger. *Eur. J. Lipid Sci. Technol.* **2006**, *108*, 48.
40. R. G. Schneeberger, C. A. Cullis. *Plant Syst. Evol.* **1992**, *183*, 265.
41. H. Zhou, Y. X. Huang, T. Hoshi, Y. Kashiwagi, J. Anzai, G. X. Li. *Anal. Bioanal. Chem.* **2005**, *382*, 1196.
42. L. F. Liu, T. Liu, G. X. Li, Q. Wang, T. Ng. *Anal. Bioanal. Chem.* **2003**, *376*, 854.
43. S. Bhatia, K. Singh, V. Jagannathan, M. Lakshmikumaran. *Plant Sci.* **1993**, *92*, 47.
44. I. Capesius. *J. Plant Physiol.* **1993**, *142*, 112.
45. M. Miyazawa, T. Maehara, K. Kurose. *Flavour Fragr. J.* **2002**, *17*, 187.
46. K. Singh, S. Bhatia, M. Lakshmikumaran. *Genome* **1994**, *37*, 121.
47. N. Benkeblia. *Food Sci. Technol.* **2004**, *37*, 263.
48. F. Shibata, M. Hizume. *Theoret. Appl. Genet.* **2002**, *105*, 167.
49. G. Zhang, L. Qin, Y. Y. Shi. *J. Bone Miner. Res.* **2007**, *22*, 1072.
50. R. H. Hui, D. Y. Hou, T. C. Li, C. X. Guan, Y. Q. Zhu, X. Y. Liu. *Chin. J. Anal. Chem.* **2004**, *32*, 695.
51. Y. Sun, K. P. Fung, P. C. Leung, D. W. Shi, P. C. Shaw. *Planta Med.* **2004**, *70*, 287.
52. H. Akhlaghi, A. Rustaiyan, K. Larijani, A. Shafaghat, N. Masnabadi, S. Masoudi. *J. Essent. Oil Res.* **2007**, *19*, 269.
53. S. Sinclair. *Altem. Med. Rev.* **2009**, *3*, 338.
54. X. Q. Ma, J. A. Duan, D. Y. Zhu, T. T. X. Dong, K. W. K. Tsim. *Phytochemistry* **2000**, *54*, 363.
55. P. Werkhoff , W. Bretschneider, M. Guntert, R. Hopp, H. Surburg. *Z. Lebens. Unt. Forsch.* **1991**, *192*, 111.
56. F. Chen, H. Y. E. Chan, K. L. Wong, J. Wang, M. T. Yu, P. P. H. But, P. C. Shaw. *Planta Med.* **2008**, *74*, 889.
57. K. Babu, E. W. Boyer, C. Hernon, E. Brush. *Clin. Ped. Emerg. Med.* **2005**, *6*, 81.
58. D. J. Siebert. *J. Ethnopharmacol.* **1994**, *43*, 53.
59. C. M. Bertea, P. Luciano, S. Bossi, F. Leoni, C. Baiocchi, C. Medana, C. M. Azzolin, G. Temporale, M. A. Lombardozzi, M. E. Maff ei. *Phytochemistry* **2006**, *67*, 371.
60. R. Bucheler, C. H. Gleiter, R. Schwoerer, I. Gaertner. *Pharmacopsychiatry* **2005**, *38*, 1.
61. D. W. Lachenmeier, S. G. Walch, S. A. Padosch, L. U. Kroner. *Crit. Rev. Food Sci. Nutr.* **2006**, *46*, 365.
62. D. W. Lachenmeier, J. Emmert, T. Kuballa, G. Sartor. *Forens. Sci. Int.* **2006**, *158*, 1.
63. Council Directive (EEC) No 88/388. *J. Eur. Commun.* **1998**, *L184*, 61.
64. C. Rey, I. Slacanin. *Rev. Suisse Viticult. Arboricult. Horticult.* **1997**, *39*, 1.
65. R. Holmes. *New Sci.* **2004**, *2453*, 32.
66. W. J. Kress, K. J. Wurdack, E. A. Zimmer, L. A. Weigt, D. H. Janzen. *Proc. Natl Acad. Sci. USA* **2005**, *102*, 8369.