The Application of DNA Molecular Approaches for the Identification of Herbal Medicinal Products

Alice Gutteridge[a], Malcolm Burns[a,b]
(a) Molecular and Cell Biology, LGC, Queens Road, Teddington, Middlesex, TW11 0LY, UK
(b) To whom correspondence should be addressed. E-mail:Malcolm.Burns@lgcgroup.com
Telephone: (+44) (0)208 943 7000, Fax: (+44) (0)208 943 2767

Summary

Traditional herbal medicines, inclusive of those originating from China, India and Korea, incorporate a wide range of plant species with beneficial properties that have conventionally been relied upon to treat ailments and disease. In recent years these medicines and treatments have gained popularity in the West as non-prescribed alternatives to conventional pharmaceuticals, as consumers are reportedly more comfortable taking products that are deemed natural, and assume that side-effects are less severe. Production of herbal medicines is often not strictly regulated, leading to a multitude of quality control problems that are difficult to identify once the plants have been processed into powders, tinctures and other preparations. Adulteration with morphologically indistinguishable species is a particular problem that can arise from mistaken identification upon harvesting, ambiguous taxonomy due to confusion between Latin nomenclature and local terminology, fraudulent substitution using less potent/valuable species, or masking of endangered species. As well as misleading consumers, adulteration of medicines can put consumers at risk of unknown chemical interactions with other drugs, and allergic reactions.

Traditional approaches for identification of herbal products include chemical analysis and microscopy but the former may lack accuracy when species or population determination is required and the latter requires skills and experience that are increasingly scarce. Molecular DNA approaches for speciation can be capitalised upon to facilitate detection and identification of herbal species. This report outlines the major developments in molecular methods for species identification and their applicability for use with unknown or disputed herbal medicines. The increased availability of accurate species-specific DNA sequence information facilitated by the establishment of open-access databases has laid the foundations for DNA molecular approaches to impart specific, sensitive and reliable identification of herbal medicinal plants and their products.

Herbal Medicinal Products and Traceability

A herb can be defined as any plant "*with leaves, seeds, or flowers used for flavouring, food, medicine, or perfume*"[1]. Directive 2004/24 of the European Parliament and Council[2] provides a definition of a herbal medicinal product as *"exclusively containing as active ingredients one or more herbal substances or one or more herbal preparations"*. Medicinal practices such as Indian Ayurveda and Traditional Chinese Medicine (TCM) utilise naturally-occurring herbal and animal ingredients in combination with meditation, dietary regulation and hygienic living to create a holistic therapeutic experience for the patient. These treatments have been relied upon for thousands of years, becoming an intrinsic part of Asian cultures.

Repute of traditional medicines has been passed down and accumulated over generations and often has little or no grounding in conventional scientific evidence.

St. John's Wort (*Hypericum perforatum*) medicines are a popular alternative treatment for depression in Europe and are used for their antimicrobial and anti-inflammatory properties in TCM[3]. It is now known that St. John's Wort has contra-indications associated with many synthetic drugs including conventional antidepressants and oral contraceptives[4]. There is no clear consensus on the degree of efficacy of St. John's Wort, and the outcomes of clinical trials, and meta-analyses of these clinical trials, have often been conflicting. A multi-centre double-blind study conducted in the US showed no significant difference in antidepressant activity of St. John's Wort when compared to a placebo. However, beneficial effects have been noted in smaller German clinical trials that tended to exclude patients with major depression. St. John's Wort medicines as a treatment for depression are available on prescription in Germany[5].

There is a common misconception that natural products are harmless[6], despite their potency and unknown mechanisms of action. Ginger (*Zingiber officinale*) is often used in Indian and Chinese medicine as a remedy for colds, nausea, arthritis, migraines and hypertension. The declaration of any *Zingiber* species incorporated into a product is a necessity, as consumption can adversely affect diabetic patients or those taking anticoagulants[7]. Less potent species from the same genus *Z montanum* and *Z zerumbet* are substituted into medicine, despite their different biologically active constituents and pharmacological effects[8].

Herbal medicinal products such as these are an increasingly popular choice for Western people, either taken alongside prescribed treatment or relied on completely for ailments ranging from acne to cancer. A 2008 survey commissioned by the MHRA reported that 35% of adults had used herbal medicines purchased over-the-counter, on the internet, or from a herbal practitioner or clinic in the previous 2 years[6]. In 2008, the global market value was estimated at US$ 83 billion[9].

This colossal demand and an unstable financial environment have provided the impetus behind adulteration of herbal medicinal products with other species of plant, heavy metals, and pharmaceutical substances. The distribution of herbal products within the UK is regulated under the Traditional Herbal Medicines Registration Scheme (THR)[10] but adulteration of herbal medicinal products, whether intended or not, may occur in the country of origin and pose a risk to consumers if undetected.

The booming TCM industry has resulted in a scarcity of wild species that are commonly incorporated into medicinal products. American ginseng (*Panax quinquefolius*) and Asian ginseng (*Panax ginseng*) are at risk from extinction by unregulated trade and were recently included in the Convention on International Trade in Endangered Species (CITES), appendix II[11]. Wild ginseng is believed to be more potent than cultivars, provoking unsustainable and illegal harvesting in Far East Asia. *P ginseng* is of a higher value than *P quinquefolius* on the Korean market but roots are indistinguishable when sliced or processed, leading to fraudulent substitution[12].

Whilst there are numerous purported benefits of herbal medicines, this has also led to an increase in the propensity for adulteration of such products, and this has also resulted in a number of serious cases that have detrimentally affected the lives of consumers. In 1990, the toxic herb *Aristolochia manshuriensis* was found to be present in a slimming aid supplied by

a Belgian clinic instead of another species of the *Aristolochia* genus, used in TCM under the common name "Mu Tong". Over 100 women developed kidney failure and many were subsequently diagnosed with urothelial carcinomas[13].

For these reasons, it is of utmost importance that only the intended herb species with the desired therapeutic effects are taken from harvest to medicine and identification and traceability of the correct species is fundamental in helping address any authenticity or adulteration issues.

Authentication of Herbal Medicinal Products

Traditional approaches to identification of herbal plants include morphological examination of phenotypic characteristics. However, this can sometimes prove to be subjective and experts in the field often do not always agree upon the same species identity. Microscopy[14] is also a powerful technique for the identification of herbal medicines but the necessary skills and experience are increasingly scarce. Current chemical detection methods identify the contents of herbal medicines by using the absorbance spectra of chemically characteristic metabolites to produce a distinctive chemical fingerprint[15]. High-Performance Liquid Chromatography (HPLC) is the most commonly used method for analysis of herbal medicine components, and research into improving component separation is on-going. Developing approaches that help combine different techniques (e.g. HPLC in tandem with mass spectroscopy[16]) has led to a more thorough analysis of herbal medicines.

With these methods, resolution at species level is dependent upon chemical fingerprints; however these can often be heterogeneous within one species, for example when plants are of different ages, or have been grown in diverse soil types or at different altitudes[17]. These methods are highly reliable for quality assessment and detection of adulteration with heavy metals and pharmacological compounds. However, analytical situations requiring resolution of trace amounts of a single herbal species in a mixed sample might be unsuitably complex for currently-available metabolomic methods.

The solution to help resolve both the ambiguity of species identification using morphological examination, and the requirement to be able to detect trace amounts of a herbal species, may lie in the study of the molecular composition of a herbal plant. DNA is an ideal analyte as it is present in all parts of the plant and has the potential to be detected even after mechanical processing and heat treatment.

Molecular DNA Approaches for Speciation of Herbal Medicinal Products

A significant historical milestone in the hunt for molecular polymorphism was reached in the 1970's through the advent of restriction fragment length polymorphism (RFLP). This technique exploits the presence of potentially informative nucleotide substitutions or insertions/deletions within restriction sites of enzymes isolated from bacteria and archaea. Application of these restriction endonucleases to amplified genomic DNA produces fragments that vary in length if a polymorphism between samples is present. The extent of discriminative ability is dependent on the genomic similarity and this varies according to the samples studied. The fragments can then be visualised by gel electrophoresis, or for large genomes, Southern blot[18]. Although RFLP assays are highly reproducible, their

implementation can be time-consuming and labour intensive. With the advent of modern molecular instrumentation, such as capillary electrophoresis, some of these disadvantages have been lessened.

The discovery of RFLP technology led to a flurry of innovation that aimed to utilise the core biological mechanism of restriction digestion. In 1995, Vos and colleagues[19] combined RFLP with selective PCR amplification to create Amplified Fragment Length Polymorphisms (AFLP). After digestion, adaptors are ligated to the ends of the newly-cleaved fragments and primers are designed to anneal to these and extend into the restriction site so that these sequences are selectively amplified. Ghosh *et al* successfully designed an AFLP assay to discriminate between *Zingiber officinale* and two commonly substituted adulterant *Zingiber* species. Seven primer pairs produced 837 fragments, 99.7% of which were polymorphic. This level of polymorphism between species provides ample information to identify unknown *Zingiber* samples based on their unique AFLP fingerprints. Although most studies use ethidium bromide gels for detection, lab-on-a-chip gel electrophoresis techniques provide clearer and more informative data[8].

Random Amplified Polymorphic DNA (RAPD) methods were developed by Williams *et al* to overcome the economic and temporal constraints of RFLP. A single pair of short arbitrary primers is used to amplify genomic DNA, producing a set of fragments for each sample[20]. The concept is similar to that of RFLP, but primers are less stringent than restriction enzymes so RAPDs lacks reproducibility. Shinde *et al* successfully identified the three main components of an Ayurvedic medicine by applying RAPD. However, 120 primer sets were screened before one was selected for discriminating between three known species[21]. This degree of rigour may prove to be too arduous when a larger number of species needs to be identified.

Inter-Simple Sequence Repeats (ISSR)[22] refer to stretches of DNA that reside between repeat motifs present in the DNA sequences. PCR primers can be designed to anneal to the repeat motifs and amplify the intervening genomic regions. The level of polymorphism and the number of amplified intervening genomic regions can be high, producing a large number of PCR products, the analysis of which is often referred to as DNA fingerprinting. Italian researchers were able to group geographically distinct asparagus (*A acutifolius*) plants into a phylogenetic distance tree, demonstrating that ISSR-PCR is appropriate for closely-related accessions[23]. In contrast to ISSR, Simple Sequence Repeat (SSR) primers are designed to amplify the repetitive DNA sequences, which can vary in size between closely related plant varieties based on the number of repeat motifs. This method therefore obligates the availability of sequence information for the plant species in question; a limitation that is lessening with the introduction of accessible and up-to-date DNA sequence databases.

Polymorphic fragments highlighted by RAPD and ISSR can be sequenced to design primers for SCAR (sequence characterized amplified region)-PCR, a technique that has become popular for the detection of adulterants in herbal medicines[24]. Amplification of a single target results in a clear presence/absence result that is more reproducible than predecessor RAPD and ISSR methods. A unique pair of primers has to be designed for each species requiring identification; for this reason SCAR-PCR is mostly utilised for the detection of a single species.

It is not recommended that low-quality DNA is analysed using RFLP, AFLP, RAPD or ISSR-PCR, as genetic artefacts caused by degradation may create or remove polymorphisms within primer target sequences or restriction sites.

Universal Primers

Molecular DNA approaches for speciation often capitalise upon genes that are universally present in eukaryotic genomes, such as the 18S ribosomal RNA subunit gene. Within these sequences are regions of DNA that are highly conserved between different species, to which universal PCR primers can be designed in order to amplify flanking hyper-variable sequences[25].

The main advantage of this universal priming approach is the wide applicability to a range of organisms to facilitate species identification, without the need for prior in-depth and detailed sequence information. Potential disadvantages to this approach include that expertise in aligning multiple DNA sequences from different species is often required so that a consensus sequence can be identified in order to generate the universal primers. This also often necessitates a good knowledge of a range of DNA databases and access to appropriate primer design software that allows advanced primer design (e.g. generation of degenerate primers). In addition, the application of universal primers to a complex matrix consisting of ingredients derived from multiple species can run the risk of generating multiple PCR products that cannot be easily resolved and identified using traditional DNA sequencing approaches.

Species-Specific Primers

The use of universal primers often facilitates rapid and detailed species information from a DNA sample. However, by their very nature, universal primers do not exclusively target species-specific DNA sequences, often precluding accurate species identification if more than one species is present in a sample. A second general approach for species identification using molecular DNA analysis is to use species-targeted primers. Such species-specific primers are often designed by accessing taxonomic databases that house DNA sequence information specific to particular species (e.g. sequence information from the mitochondrial Cytochrome c Oxidase subunit I gene from members of the animal kingdom).

Advantages of this approach include the availability of taxonomic DNA reference sequences from validated systems, and that no in-depth knowledge of advanced primer design, DNA databases or sequence alignment tools is always necessary. Equally well, the approach is relatively simple to apply and there is a reasonable chance of producing just one PCR product that is ideal for standard DNA sequencing. Potential disadvantages of this approach include that the primers chosen must usually be different depending on the target species the analyst wishes to identify, and additionally, depending upon how the primers are designed, specificity is not assured and needs to be theoretically and experimentally validated.

Bioinformatics and DNA Sequence Information for Herbal/Medicinal Plants

A number of molecular biology approaches have been described above, to help facilitate accurate species identification. However, all these approaches are dependent upon the

availability of DNA sequence information from particular species, and the confidence with which this DNA sequence information is known.

The largest and most well-known source of DNA sequences is GenBank, maintained by the National Center for Biotechnology Information (NCBI)[26]. Sequences are directly submitted from scientific institutions and can be used directly in associated applications such as the Basic Local Alignment Search Tool (BLAST) for sequence comparison and primer design. However, the overwhelming number of sequence entries may sometimes make it difficult to select appropriate sequence information. Specialist species-specific databases such as Oryzabase[27], Maize DB[28] and ArkDB[29] provide detailed information on the species of interest, and are therefore useful for those with focused research aims.

The concept of "barcoding" samples, the process whereby a sample or product is given a unique identification tag, has been well received and implemented in the global retail industry, and its application to DNA molecular approaches for taxonomic classification can prove equally informative. The Consortium of the Barcode of Life (CBOL) endeavours to combine the efforts of researchers around the world to standardise a system of DNA barcoding that can be applied to all species[30], and is the largest biodiversity genomics initiative currently being undertaken.

The Barcode of Life Database[31] is based on such a DNA barcoding approach, and uses variations in short standardised gene regions to identify new species. The taxonomic marker used for members of the animal kingdom is predominantly the Cytochrome Oxidase subunit I (COI) of the mitochondrial genome, whilst for plants the taxonomic markers tend to be *rbcL* or *matK*. The Barcode of Life Database is a project aimed at creating a public collection of reference sequences from validated specimens of all species of life, and consists of an integrated web platform for use of DNA barcode data. The approach of DNA barcoding is increasingly being used as a way to help identify species involved in legal and regulatory disputes.

There has been notable success with the heralding of mitochondrial cytochrome oxidase I (COI) as a near-universal barcoding region for animals. Cytochrome oxidase I has been used for identification of numerous species within the animal kingdom, including fish, mammalian and avian species. However the rate of evolution in plant mtDNA is much slower, limiting the utility of COI for speciation of herbal medicinal products[32]. Various other regions in chloroplast, nuclear and ribosomal genomes have been suggested as a barcode for plants, with successes and failures that are often specific to the taxon in question.

In 2009, the CBOL Plant working group published a paper[33] proposing plastid genes *rbcL* and *matK* as the most appropriate out of the seven shortlisted. Universality, sequence quality and discriminatory power were assessed, and though *rbcL* and *matK* scored highly, no "ideal" barcode or barcode combination that fulfilled all three criteria could be ascertained. The efficiency of amplifying and sequencing *matK* varies between taxa, but it is generally acknowledged that rigorous optimisation of PCR conditions is required. Once obtained, sequence quality is particularly high for this region. The discriminatory power of *matK* and *rbcL* has been found to be lower than other candidate barcodes; for this reason it is advised that additional regions are considered when resolution below family level is required. Stoeckle and colleagues utilised *rbcL* and *matK* to identify the population of plants present in commercial tea, and were able to sequence at least one region for 90% of samples. This study showed the ability of DNA barcoding to highlight adulterant plant species, as many species

other than tea (*Camellia sinensis*) were detected. The pivotal determinant of scientific advances in this field is the availability of prior sequence information; 62% of the sequences detected in the tea samples were left undetermined[34].

Research groups such as those led by Shaw[35], Scarcelli[36] and Dong[37] have combed the chloroplast genome for highly variable sequences that could serve as new barcoding regions. Shaw advocates further investigation into non-coding sequences such as *rpl32-trnL* and these findings have been well-received; a PubMed search for "*rpl32-trnL*" returns 31 studies published after the original article in 2007[38]. Introns and intergenic spacers can be amplified relatively easily, as primers are designed to anneal to well-conserved exons adjacent to these polymorphic regions[35]. In contrast to this, Dong *et al.* cite a gene of unknown function, *ycf1*, as the most variable locus but also highlight the potential of a cohort of non-coding sequences[37]. In spite of these efforts, not all studies that have utilised these promising regions have reached expectations, particularly for species with low rates of nucleotide substitution[39].

Plastid genomes are uniparentally inherited, so reliance on these sequences may lead to incomplete analysis when alloploid (an organism containing the genetic information of two or more species) or hybrid species are encountered[40]. The internal transcribed spacer regions (ITS1 and 2) in ribosomal DNA have garnered a great deal of interest as alternatives or supplements to plastid barcodes. Chen *et al* investigated the applicability of seven barcoding regions for 6600 plant samples and concluded that the ITS2 region currently holds the most promise as a universal plant barcode. ITS2 sequences exhibited higher interspecies discriminatory power than other high-performing plastid regions and enabled identification of 92.7% and 99.8% of samples at the species and genus level, respectively[41]. Despite these promising data, Chen admits that the universal barcode "Holy Grail" has yet to be found. As repetitive sequences, ITS1 and ITS2 are more likely to be affected by genetic anomalies commonly seen in plants, such as gene multiplication, pseudogenes and introgression[42].

As part of the CBOL initiative, all sequence information is open-access and accessible via NCBI GenBank, with around 191 entire chloroplast genome entries[26,37]. Sequence coverage on this public database is dependent on the strength of academic interest in the taxa, such that commercially useful crops have an abundance of data whereas lesser known crops may only have had one region of interest sequenced by a single research group. Table 1 lists some commonly-used herbal species in Western, Chinese and Ayurvedic medicine and illustrates the availability of DNA sequence information, as published on NCBI GenBank. Using accurate sequence information, molecular approaches have the potential to provide a method for absolute determination of species identity, based on statistical analysis of the genetic distance between two accessions. Although difficulties are still experienced with segregating closely-related species, DNA barcoding regions contain a wealth of information; to fully capitalise upon this may contribute towards true definition of the genetic boundaries that exist between species. The race for cheaper, faster and better quality sequencing methods will eventually negate the problem of incomplete sequence coverage, and the benefits of a comprehensive barcode sequence library will be unparalleled for researchers, those working in the herbal medicine industry, regulatory authorities, and consequently consumers.

Table 1 – Sequence Information for Some Popular Medicinal Herb Species

The following table details the availability of sequence information on NCBI GenBank, with particular emphasis on barcoding regions that are established in the literature: the ribosomal ITS regions and chloroplast regions *psbA-trnH*, *rbcL*, *matK* and *trnL*. (Searches correct as of August 2012).

Common name(s)	Binomial name	ITS	psbA-trnH	rbcL	matK	trnL	Entire chloroplast	Total results
Aloe Vera	Aloe vera	X	x	x	x	x		127
Boswellia	Boswellia serrata							2
Cat's Claw	Uncaria tomentosa	X		x		x		12
Chamomile	Matricaria recutita	X	x					27
Chinese Angelica, Dong	Angelica sinensis	X		x	x			28
Chinese Liquorice	Glycyrrhiza uralensis	X	x	x	x	x		104
Clove	Syzygium aromaticum	X			x			6
Devil's Claw	Harpagophytum	X						1
Echinacea	Echinacea purpurea	X	x	x				52
Ephedra	Ephedra sinica	X	x	x	x	x		65
Eucalyptus	Eucalyptus globulus	X	x	x	x	x	x	385
Evening Primrose	Oenothera biennis	X	x	x	x	x	x	48
Fenugreek	Trigonella foenum-graecum	X	x		x	x		47
Feverfew	Tanacetum parthenium	X	x	x	x	x		27
Garlic	Allium sativum	X		x		x		276
Ginger	Zingiber officinale	X	x	x	x	x		201
Gingko	Ginkgo biloba	X	x	x	x	x	x	633
Ginseng	Panax ginseng	X	x	x	x	x	x	15,926
Goldenseal	Hydrastis Canadensis	X		x	x	x		29
Gotu Kola	Centella asiatica	X	x	x	x			83
Guggul	Commiphora wightii	X						41
Huáng qí	Astragalus propinquus	X						3
Milk Thistle	Silybum marianum	x	x	x	x	x		26

Common name(s)	Binomial name	ITS	psbA-trnH	rbcL	matK	trnL	Entire chloroplast	Total results
Pot Marigold	Calendula officinalis	x			x	x		56
Saint John's Wort	Hypericum perforatum	x	x	x	x	x		82
Saw Palmetto	Serenoa repens			x	x	x		34
Valerian	Valeriana officinalis	x	x	x	x	x		59

Molecular methods for detection and identification of species in herbal medicinal products have clear potential to be fit for purpose, as samples can be identified regardless of appearance and chemical constituents (which may change according to the tissue type, growth conditions or age of the sample), and the sample can be a taxonomic mixture. The sample is rarely required in abundance; this factor is particularly valued when expensive or rare herbal medicines are needed.

This review has discussed some of the advantages and limitations of some of the techniques available, and often these are dependent on the experimental setting: does the analyst wish to ascertain the presence or absence of a single, or set of, known species? Alternatively do they require an in-depth audit of herbal and animal constituents in the medicine? The latter scenario can only achieve confident results with a highly comprehensive database of barcoding sequences, and the CBOL Plant Working Group is working toward this goal. The knowledge base on which molecular identification of plants depends upon is steadily expanding, and barcoding success stories are published in open-access journals such as PLOS One and are making headlines in mainstream news outlets[43].

Conclusion

This review has highlighted the clear potential that the application of molecular DNA approaches may afford in the area of detection and identification of herbal medicinal plants and their products. Phenotypic characterisation of herbal plants can be subjective, microscopy skills are scarce and traditional chemical characterisation of herbal products is dependent upon the generation of reliable chemical profiles that can be confounded by tissue type, growth conditions and location. As with many areas of speciation and food authenticity, molecular DNA approaches have the potential to be universally applied as the DNA molecule is ubiquitous, DNA is expressed at all developmental stages of a cell's life cycle, and the DNA molecule is robust enough to remain relatively intact even when subjected to high temperatures and highly processed chemical reactions. DNA approaches can be completely specific and demonstrate good sensitivity, ideal for when the target herbal species are at very low concentrations. The only current limitation is the availability and accuracy of DNA sequence information for particular species. Through such initiatives as CBOL, the amount of species-specific DNA sequence information is ever increasing, and DNA molecular approaches should be capitalised upon in order to realise some of the potential that they can offer for the detection and identification of herbal medicinal plants and associated products.

This review has provided evidence for the applicability of using molecular DNA techniques for the purpose of ascertaining the genetic content of "herbal medicinal products". Successes in this area will help support EU legislation and potentially provide regulatory bodies and analytical laboratories with the means to expose adulteration of herbal products that could otherwise be detrimental to public health, the economy and the environment.

Acknowledgements

The work described in this review was funded by the UK National Measurement Office as part of the Government Chemist Programme 2011-2014.

References

1 Erin McKean, editor, (2005), The New Oxford American Dictionary (2nd ed.), Oxford University Press, ISBN 978-0-19-517077-1

2 Directive 2004/24 of the European Parliament and Council, amending, as regards traditional herbal medicinal products, Directive 2001/83/EC on the Community code relating to herbal products for human use

3 C Howard, E Socratous, S Williams S, E Graham, Fowler, *et al*, PlantID – DNA-based identification of multiple medicinal plants in complex mixtures. Chinese Medicine, 2012 **7(18)**

4 F Borrelli, AA Izzo, Herb-drug interactions with St John's Wort (Hypericum perforatum): an update on clinical observations. AAPS, 2009, **11(4),** 710-727

5 KM Klemow, A Bartlow, J Crawford, *et al*, 2011. Ch11: Medical Attributes of St. John's Wort (Hypericum perforatum). In IFF Benzie, S Wachtel-Galor, 2011, Herbal Medicine: Biomolecular and Clinical Aspects, 2nd edition, Boca Raton (FL): CRC Press. Bookshelf ID: NBK92750, PMID: 22593920

6 Ipsos MORI report results on herbal medicines, November 2008 – "Public Perceptions of Herbal Medicines: General Public Qualitative & Quantitative Research"
Press release:
http://www.mhra.gov.uk/NewsCentre/Pressreleases/CON036071
Full report:
http://www.mhra.gov.uk/home/groups/comms-po/documents/news/con036073.pdf

7 AM Bode, Z Dong, The Amazing and Mighty Ginger. In IFF Benzie, S Wachtel-Galor, 2011, Herbal Medicine: Biomolecular and Clinical Aspects. 2nd edition, Boca Raton (FL): CRC Press. Bookshelf ID: NBK92775, PMID: 22593941

8 S Ghosh, PB Majumder, S Sen Mandi, Species-specific AFLP markers for identification of Zingiber officinale, Z. montanum and Z. zerumbet (Zingiberaceae). Genetics and Molecular Research, 2011, **10(1)** 218-229

9 MM Robinson, X Zhang, The World Medicines Situation 2011, 3rd Edition - "Traditional Medicines: Global Situation, Issues and Challenges" WHO/EMP/MIE/2011.2.3

http://digicollection.org/hss/documents/s18063en/s18063en.pdf

10 "Using herbal medicines safely" a webpage published by the MHRA, last modified 13/09/2012
http://www.mhra.gov.uk/Safetyinformation/Generalsafetyinformationandadvice/Herbalmedicines/Usingherbalmedicinessafely/index.htm

11 WWF report on the status of *Panax ginseng* and *P. quinquefolius*
http://wwf.panda.org/what_we_do/endangered_species/ginseng/
Addition of *Panax ginseng* to CITES appendix II:
http://www.cites.org/eng/cop/11/prop/54.pdf

12 K Leem, SC Kim, CH Yang, J Seo, Genetic identification of *Panax ginseng* and *Panax quinquefolius* by pyrosequencing methods, Bioscience, Biotechnology and Biochemistry, 2005, **69(9)** 1771-1773

13 Report by the MHRA policy division, 2008 "Public Health Risk with Herbal Medicines: An Overview"
http://www.mhra.gov.uk/home/groups/es-herbal/documents/websiteresources/con023163.pdf

14 BP Kackson, DW Snowdon, 1968, Powdered Vegetable Drugs, An atlas of Microscopy for use in the Identification and Authentication of some Plant materials Employed as Medicinal Agents, J & A Churchill Ltd, London

15 L Giri, HC Andola, VK Purohit, MSM Rawat, RS Rawal, ID Bhatt, Chromotographic and Spectral Fingerprinting Standardisation of Traditional Medicines: an overview as modern tools, Research Journal of Photchemistry, 2010, **4(4)** 234-241

16 YZ Liang, P Xie, K Chan, Quality control of herbal medicines. Journal of chromatography, Analytical technologies in the biomedical and life sciences, 2004, **812** 53–70

17 DP Singh, SK Srivastava, R Govindarajan, AKS Rawat, High performance liquid chromatographic determination of bergenin in different Bergenia species. Acta Chromatographica 2007, **19** 246-252

18 D Botstein, RL White, M Skolnick, RW Davis, Construction of a genetic linkage map in man using restriction fragment length polymorphisms, American Journal of Human Genetics, 1980, **32** 314-331

19 P Vos, R Hogers, M Bleeker, M Reijans, T van de Lee, *et al,* AFLP: a new technique for DNA fingerprinting, Nucleic Acids Research, 1995, **23(21)** 4407-4414

20 JGK Williams, AR Kubelik, KJ Livak, JA Rafalski, SV Tingey, DNA polymorphisms amplified by arbitrary primers are useful as genetic markers, Nucleic Acids Research, 1990 **18(22)** 6531-6535

21 VM Shinde, K Dhalwal, KR Mahadik, KS Joshi, BK Patwardhan, RAPD analysis for determination of components in herbal medicine, Evidence-Based Complementary Alternative Medicine, 2007, **4(1)** 21-23

22 E Zietkiewicz, A Rafalski, D Labuda, Genome fingerprinting by simple sequence repeat (SSR)-anchored polymerase chain reaction amplification, Genomics, 1994 Mar 15, **20(2)** 176-83

23 M Sica, G Gamba, S Montieri, L Gaudio, S Aceto, ISSR markers show differentiation among Italian populations of *Asparagus acutifolius*, L. BMC Genetics, 2005, **6(17)**

24 G Heubl, New aspects of DNA-based authentication of Chinese medicinal plants by molecular biological techniques, Planta Medica, 2010, **76** 1963-1974

25 CR Linder, LA Moore, RB Jackson, A Universal Molecular Method for Identifying Undergound Plant Parts to Species, Molecular Ecology, 2000, **9** 1549-1559

26 DA Benson, I Karsch-Mizrachi, DJ Lipman, J Ostell, DL Wheeler, GenBank, Nucleic Acids Res, 2008, **36**(Database issue), D25-30

27 Oryzabase – an online rice database http://www.shigen.nig.ac.jp/rice/oryzabaseV4

28 CJ Lawrence, Q Dong, ML Polacco, TE Seigfried, V Brendel, (2004) MaizeGDB, the community database for maize genetics and genomics, Nucleic Acids Research **32** D393-D397 http://www.maizegdb.org

29 ArkDB – database for commonly farmed species hosted at the Roslin Institute, Edinburgh, http://www.thearkdb.org/arkdb

30 Consortium for the Barcode of Life (CBOL) website – “About” page http://www.barcodeoflife.org/content/about/what-cbol

31 S Ratnasingham, PDN Hebert, BOLD: The Barcode of Life Data System (www.barcodinglife.org), Molecular Ecology Notes, 2007 May 1; **7(3),** 355–364, doi: 10.1111/j.1471-8286.2007.01678.x

32 V Savolainen, RS Cowan, AP Vogler, GK Roderick, R Lane, Towards writing the encyclopaedia of life: an introduction to DNA barcoding, Philosophical Transactions of the Royal Society B: Biological Sciences, 2005, **360(1462)** 1805-1811

33 CBOL Plant Working Group, communicated by DH Janzen, A DNA barcode for land plants, Proceedings of the National Academy of Sciences of the United States of America, 2009 **106(31)** 12794-12797

34 MY Stoeckle, CC Gamble, R Kirpekar, G Young, S Ahmed, DP Little, Commercial teas highlight plant DNA barcode identification successes and obstacles, Scientific Reports, 2011, **1(42)**

35 J Shaw, EB Lickey, EE Schilling, RL Small, Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: The tortoise and the hare III, American Journal of Botany, 2007 **94(3)** 275-288

36 N Scarcelli, A Barnaud, W Eiserhardt, UA Treier, M Seveno M *et al,* A Set of 100 Chloroplast DNA Primer Pairs to Study Population Genetics and Phylogeny in Monocotyledons, PLoS ONE, 2011, **6(5)** e19954

37 W Dong, J Liu, J Yu, L Wang, S Zhou, Highly variable chloroplast markers for evaluating plant phylogeny at low taxonomic levels and for DNA barcoding, PLoS ONE, 2012, **7(4)** e35071

38 PubMed search ran 07/05/2013 for "rpl32-trnL"[All Fields]

39 M Arca, DD Hinsinger, C Cruaud, A Tillier, J Bousquet, N Frascaria-Lacoste, Deciduous trees and the application of universal DNA barcodes: A case study on the circumpolar Fraxinus. PLoS ONE, 2012, **7(3)** e34089

40 MW Chase, N Salamin, M Wilkinson, JM Dunwell, RP Kesanakurthi *et al*, Land plants and DNA barcodes: short-term and long-term goals, Philosophical Transactions of the Royal Society B: Biological Sciences, 2005, **360(1462)** 1889-1895

41 S Chen, H Yao, J Han, C Liu, J Song *et al*, Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS ONE, 2010, **5(1)** e8613

42 A Alvarez, JF Wendel, Ribosomal ITS sequences and plant phylogenetic inference. Molcular Phylogenetics and Evolution, 2003, **29(3),** 417-434

43 Story on the BBC News website: http://www.bbc.co.uk/news/uk-wales-18590298
Covering publication of: N de Vere, TGG Rich, CR Ford, SA Trinder, C Long *et al*, DNA Barcoding the Native Flowering Plants and Conifers of Wales, PLoS ONE, 2012, **7(6)** e37945