Identification of a Herbal Powder by Deoxyribonucleic Acid Barcoding and Structural Analyses

Bhavisha P. Sheth, Vrinda S. Thaker

Department of Biosciences, Centre for Advanced Studies in Plant Biotechnology and Genetic Engineering, Saurashtra University, Rajkot, Gujarat, India

ABSTRACT

Background: Authentic identification of plants is essential for exploiting their medicinal properties as well as to stop the adulteration and malpractices with the trade of the same. Objective: To identify a herbal powder obtained from a herbalist in the local vicinity of Rajkot, Gujarat, using deoxyribonucleic acid (DNA) barcoding and molecular tools. Materials and Methods: The DNA was extracted from a herbal powder and selected Cassia species, followed by the polymerase chain reaction (PCR) and sequencing of the rbcL barcode locus. Thereafter the sequences were subjected to National Center for Biotechnology Information (NCBI) basic local alignment search tool (BLAST) analysis, followed by the protein three-dimension structure determination of the rbcL protein from the herbal powder and Cassia species namely Cassia fistula, Cassia tora and Cassia javanica (sequences obtained in the present study), Cassia Roxburghii, and Cassia abbreviata (sequences retrieved from Genbank). Further, the multiple and pairwise structural alignment were carried out in order to identify the herbal powder. Results: The nucleotide sequences obtained from the selected species of Cassia were submitted to Genbank (Accession No. JX141397, JX141405, JX141420). The NCBI BLAST analysis of the rbcL protein from the herbal powder showed an equal sequence similarity (with reference to different parameters like E value, maximum identity, total score, query coverage) to C. javanica and C. roxburghii. In order to solve the ambiguities of the BLAST result, a protein structural approach was implemented. The protein homology models obtained in the present study were submitted to the protein model database (PM0079748-PM0079753). The pairwise structural alignment of the herbal powder (as template) and C. javanica and C. roxburghii (as targets individually) revealed a close similarity of the herbal powder with C. javanica. Conclusion: A strategy as used here, incorporating the integrated use of DNA barcoding and protein structural analyses could be adopted, as a novel rapid and economic procedure, especially in cases when protein coding loci are considered.

INTRODUCTION

Medicinal plants have been used by mankind since ages to cure common as well as serious health ailments. Indian flora comprises of an affluent floristic diversity of medicinal plants. Medicinal values of these plants are exploited in major modes of therapy such as homeopathy, aroma therapy, traditional ayurvedic medicine, Unani, herbal therapy, etc. The process of herbal drug development often experiences (a) deliberate or indeliberate adulteration, and (b) unintentional substitution due to lack of authentic information.[1] Hence, the identification of the plant products are of prime importance in preventing the adulteration and malpractices with the trade of the same.[2] Unambiguous identification and authentication of the plants used for the production is therefore an elementary and critical step at the beginning of an extensive quality assurance process. Many different methods have been used since long for the identification of medicinal plants such as traditional pharmacognostic analysis,[3,4] chromatographic and spectroscopic,[5‑7] physical techniques,[8,9] and molecular markers.[10] The latter are more suitable for identification as deoxyribonucleic acid (DNA), in contrast to ribonucleic acid, is a stable macromolecule that is found in all tissues, including the dead ones.

Among these, molecular marker techniques provide valuable data on diversity through their ability to detect variation at the DNA level, and have been routinely employed as they are easy, rapid and reliable to perform. Many DNA based methods have been used previously for detecting and differentiating plant materials from their contaminants viz. Random Amplified Polymorphic DNA,[11,12] polymerase chain reaction (PCR)-Restriction Fragment Length Polymorphism,[13,14] high-resolution melting analysis,[15,16] DNA barcoding,[17‑21] etc. Among

Key words: Deoxyribonucleic acid barcoding, herbal powder, identification, protein three-dimension structure, rbcL

SUMMARY

• Authentic identification of plants is essential for exploiting their medicinal properties as well as to stop the adulteration and malpractices with the trade of the same. A herbal powder was obtained from a herbalist in the local vicinity of Rajkot, Gujarat. An integrated approach using DNA barcoding and structural analyses was carried out to identify the herbal powder. The herbal powder was identified as Cassia javanica L.

Correspondence:
Prof. Vrinda S. Thaker,
Department of Biosciences,
Saurashtra University, Rajkot - 360 005,
Gujarat, India.
E-mail: thakervs@gmail.com
DOI : 10.4103/0973-1296.172963

Access this article online
Website: www.phcog.com
Quick Response Code:

Cite this article as: Sheth BP, Thaker VS. Identification of a herbal powder by deoxyribonucleic acid barcoding and structural analyses. Phcog Mag 2015;11:570-4.

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as the author is credited and the new creations are licensed under the identical terms.

For reprints contact: reprints@medknow.com
these, DNA barcoding is a comparatively recent technique, which uses DNA sequences from a small fragment of the genome to identify organisms.[22] Furthermore, bioinformatics approaches have been of immense help for the in-depth understanding of the plant systems in various ways.[23,24] The rbcL gene codes for the large subunit of Rubisco enzyme in plants, is exclusively chloroplastic in origin and is a very useful molecule in phylogenetic analysis[25] as well as designated as a core barcode in plants.[26] It has been used in identification of medicinal materials earlier.[18,27,28] In this study, a herbal powder was obtained from a herbalist in the local vicinity of Rajkot, Gujarat. Thereafter, the DNA was extracted, followed by the PCR and sequencing of the rbcL barcode. The basic local alignment search tool (BLAST) results, showed an equal sequence similarity to Cassia roxburghii and Cassia javanica. Hence, the comparative protein structural analysis was carried out to solve the ambiguity observed in the BLAST results.

**MATERIALS AND METHODS**

In this study, an easy, inexpensive, and rapid approach was used for the identification of a herbal powder, using the molecular and bioinformatics tools.

**Deoxyribonucleic acid isolation from the medicinal extract and Cassia species**

The DNA was isolated from the unknown medicinal extract using the cetyltrimethylammonium bromide method.[29] The same was done for the Cassia sp. available in the local vicinity viz. Cassia tora, Cassia fistula and C. javanica.

**Polymerase chain reaction amplification, nucleotide sequencing of the rbcL gene and basic local alignment search tool analysis**

The rbcL gene was amplified from the herbal powder as well as the selected species of Cassia (mentioned earlier) with the primer pair: rbcL_F (5’‑ATG TCA CAA ACA GAG ACT AAA GC‑3’) and rbcL_R (5’‑GTA _R (5’‑GTA AAA TCA AGT CCA CCR CG‑3’).[30] The PCR amplification was conducted using 25 µL of reaction mixture containing 10 mM Tris·HCl, pH 8.3, 50 mM KCl, 25 mM MgCl2, 10 mM of each dNTPs, 10 pmol each of forward and reverse primers, 2 µL of template DNA and 1 units of Taq DNA polymerase (Genet, India) (without proof-reading activity) on a Veriti (96 Well Fast Thermal Cycler), Applied Biosystems, USA. 5 µl of each PCR product was electrophoresed through 2% agarose gel to determine the size of fragment. Positive products were then purified using the Big Dye Terminator V3.1 Cycle Sequencing Kit (Applied Biosystems, NY, USA) as per manufacturer’s protocol. After purification, the reaction products were analyzed on an ABI PRISM Genetic Analyzer 3130 (Applied Biosystems, NY, USA). Sequence editing was performed using Bioedit.[32] The nucleotide sequences from the samples of: C. tora, C. fistula, C. javanica were submitted to the National Center for Biotechnology Information (NCBI) Genbank database using the standalone submission tool Sequin. The obtained rbcL sequence of the herbal powder was compared with sequences in the NCBI-Genbank using the BLAST. In order to increase the taxonomic coverage for the further bioinformatics/analyses, the rbcL sequences of Cassia abbreviata and C. roxburghii were retrieved from the GenBank and used in addition to those plants for which the wetlab studies were carried out. All the nucleotide sequences were translated computationally using Bioedit[32] in order to obtain the protein sequences, which were further used for the three-dimension (3D) structural analyses.

**Homology modeling of the rbcL proteins**

The protein 3D structures of rbcL proteins of different Cassia species along with that from the herbal powder were elucidated using the homology modeling approach. The 3D models were checked for their validity using the Ramachandran plot analysis. The models were also stereochemically checked with the help of Verify 3D server.[33] Furthermore, the protein 3D models were submitted to the protein model database (PMDB)[34] with the PMDBIDs: PM0079748-PM0079753.[Table 1]

**Structural analyses using bioinformatics’ tools**

The multiple structural alignment was carried out using PROMALS3D server.[35] Further in order to clear ambiguity of the previously obtained BLAST results, the pair-wise structural alignment analyses was also carried using herbal rbcL protein as template and that of C. javanica and C. roxburghii as targets.

**RESULTS AND DISCUSSION**

In the present investigation, a herbal powder, procured from a local herbalist, was tested with the plant DNA barcoding approach. rbcL is a core barcode locus[36] and otherwise has been used in the phylogenetic analysis of plants since decades.[36‑41] The DNA was successfully isolated from the herbal powder evidenced in the form of sharp bands on the 1% agarose gel [Figure 1]. The rbcL barcode locus was amplified (~550 bp) [Figure 2] and sequenced on the Applied Biosystems 3130 Genetic Analyzer. The sequences obtained in this study were submitted to NCBI (Accession No. JX141397, JX141405, JX141420). The sequence was then compared with other sequences in the database using the NCBI-BLAST tool [Figure 3]. The query sequence showed an identity of 89% and query coverage of

![Figure 1: Agarose gel electrophoresis of the genomic deoxyribonucleic acid from the herbal powder and other Cassia species (Lane 1: Herbal powder; Lane 2: Cassia fistula; Lane 3: Cassia javanica; Lane 4: Cassia tora)](image-url)
96% (with the same maximum score, total score and e-value) to two species of the *Cassia* genus: *C. javanica* and *Cassia roxburghii*. In addition, the nucleotide sequences of *Cassia roxburghii* (with accession no. JQ301861) and *C. abbreviata* (with accession no. JF265329.1) were retrieved from Genbank and used for the further bioinformatics' analyses.

The nucleotide sequences were computationally translated to amino acid sequences and the homology modeling of the protein sequences was carried out for the herbal powder as well as the selected *Cassia* species [Figure 4]. The hypothetical protein 3D structures were validated using bioinformatics tools. To assess stereochemical quality and structural integrity of the model, RAMPAGE [42] and verify 3D [33] servers were used. The homology models obtained so were checked for the stereochemical viability using RAMPAGE server [42] (http://mordred.bioc.cam.ac.uk/rapper/rampage.php) [Table 1]. The Rampage server assesses the accuracy of the generated hypothetical models based on the Ramachandran plot analyses of the same. The Ramachandran diagram plots phi versus psi dihedral angles for each residue in the protein. The predicted models were found to be acceptable since all had more than 90% of the residues in the favored region. Verify 3D analyzes the compatibility of an atomic model (3D) with its own amino acid sequence (1D) and hence tests the accuracy of the model. The models were also checked for the 3D-1D profiles using Verify 3D server [25] [Table 2]. Residues with a score over 0.2 are considered reliable. In the present study, the profiles for all the predicted models were >0.2 indicating the correctness of the predicted models. This indicates the correctness of the models predicted in this study. Hence, the predicted models were used for the further bioinformatics analysis. The multiple structural alignment of the predicted homology models was carried out using PROMALS3D server [35] using the default parameters in order to find the structural distance between the *rbcL* proteins of the herbal powder with that of the *Cassia* species. The promals 3D server analyses the sequence-structure relationships in form of multiple structural alignment and generates a phylogenetic tree. In this study, the phylogenetic tree [Figure 5] shows, the structural similarities between the *rbcL* protein of the medicinal extract with that of *C. javanica*, *C. roxburghii* and *C. abbreviata*. Furthermore, in order to clear ambiguity of the NCBI-BLAST results, the pairwise structural alignment analyses was carried out using the RCSB Protein Databank Protein Comparison Tool [43]. In this study, the medicinal extract *rbcL* protein was used as a template and the *rbcL* proteins of *C. javanica* and *C. roxburghii* were used as targets. A lower

| Plant name       | Three‑dimension‑one‑dimension score |
|------------------|-------------------------------------|
| *Cassia abbreviata* | 0.61                                 |
| *Cassia fistula*  | 0.46                                 |
| *Cassia javanica* | 0.35                                 |
| *Cassia roxburghii* | 0.62                                |
| *Cassia tora*    | 0.35                                 |
| Herbal powder    | 0.38                                 |

Table 2: Three‑dimension‑one‑dimension profile of the homology models (using verify three‑dimension)
Table 3: Pairwise structural alignment of rbcL protein from the medicinal extract with that of Cassia spp

| Template       | Target               | RMSD |
|----------------|----------------------|------|
| Herbal powder  | Cassia javanica      | 0.81 |
|                | Cassia roxburghii    | 1.42 |

RMSD: Root mean square deviation

root mean square deviation (RMSD) value was observed in the pairwise rbcL protein structural alignment of the herbal powder with C. javanica, than that with C. roxburghii [Table 3]. A lower RMSD value implies, a closer relationship between the molecules under study. Hence, the herbal powder had a closer relationship with the C. javanica, which is also evidenced in the phylogenetic tree obtained from the multiple structural alignment. Hence, the herbal powder was identified as C. javanica.

There has been quite a discussion about the use of the sequence databases and similarity searching in the literature.[44-46] Little and Stevenson[46] found that BLAST can give accurate identification at the genus, but not the species, when DNA barcode sequences are considered. Similarly, there have been identification related problems, in doing similarity searching with BLAST.[45] Hence, a strategy to solve such cases, is imperative, especially when protein coding loci are considered. The sequence similarity results, can be further confirmed by the protein structure data, as in this study; particularly when ambiguous BLAST results are obtained. Thus, the combined use of DNA barcoding and protein 3D structural analyses can help to identify the samples. The present study can therefore be used as an example of this new approach in conjunction with routine approaches of species’ identification.

CONCLUSION

Hence this strategy, incorporating the integrated use of DNA barcoding and protein structural analyses could be adopted, as a novel, rapid and economic procedure for identification of plant species, especially in cases when protein coding loci are considered.

Acknowledgments

The first author, Sheth B. P. thankfully acknowledges University Grants Commission, New Delhi for the financial support in form of RFSMS fellowship. The authors would like to thank Dr. Kiran Chudasama for the necessary help as well as the CAS program, Govt. of Gujarat and Department of Biosciences, Saurashtra University, Rajkot for the instrumentation and logistics.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

REFERENCES

1. Deyana K, Sakumbar B. Molecular marker based adulteration detection in traded food and agricultural commodities of plant origin with special reference to spices. Curr Trends Biotecnol Pharm 2010;4:454-89.
2. Tchen N, Crockett SL, Khan IA, Scheffler BE. Authentication of medicinal plants using molecular biology techniques to compliment conventional methods. Curr Med Chem 2004;11:1391-401.
3. Tan Y, Gao TT, Ma Y, Ding X, Cheng YH. Pharmacognostic identification of Ferula ayneschakwiovi ZhongYao Cai 2011;34:1694-6.
4. Wu HZ, Ai LO, Wu HM, Lu Y, Yang YF. Pharmacognostic identification of Centepidea minima. Zhong Yao Cai 2009;32:200-2.
5. Kurz C, Carie R, Schiebe R. Characterisation of cell wall polysaccharide profiles of apricots (Punus armeniaca L.), peaches (Punus persica L.), and Pumpkins (Cucumis spp.) for the evaluation of fruit product authenticity. Food Chem 2008;106:421-30.
6. Hilt P, Schieber A, Yldirim C, Arnold M, Klaiber I, Conrad J, et al. Detection of phloridzin in strawberries (Fraganaenanassa Duch.) by HPLC-PDA/MS/MS and NMR spectroscopy. J Agric Food Chem 2003;51:2896-9.
7. Hernandez A, Martin A, Aranda E, Bartolome T, Cordoba Mde G. Detection of smoked paprika “Pimenton de La Vera” adulteration by free zone capillary electrophoresis (FZCE). J Agric Food Chem 2006;54:4141-7.
8. Xia Y, Wang O, Pu Z. Identification of Kochia scopana and its substitutes by scanning electron microscopy and UV spectrum. ZhongYao Cai 2003;26:253-6.
9. Joshi VC, Sinivas PV, Khan IA. Rapid and easy identification of illicium verum Hook. f. and its adulterant Illicium anisatum Limn. by fluorescent microscopy and gas chromatography. JAOAC Int 2005;88:703-7.
10. Agarwal M, Shrivastava N, Padh H. Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Rep 2008;27:617-31.
11. Weder JK. Identification of plant food raw material by RAPD-PCR: Legumes. J Agric Food Chem 2002;50:4456-63.
12. Khan S, Mirza KJ, Al-Qurainy A, Abdin MZ. Authentication of the medicinal plant Senna angustifolia by RAPD profiling. Saudi J Biol Sci 2011;18:287-92.
13. Wang CZ, Li F, Ding JY, Peng X, Yuan CS. Simultaneous identification of Bubus fritillariae zinzhuo using PCRFPLP analysis. Phytopathology 2007;14:628-32.
14. Chu BH, Ding XY, Li XK, He J, Ding G, Liu DY. Molecular authentication of Allisma orientalis (Sam.) jusep food from its adulterants by PCRFPLP(J). Food Sci 2007;1-2.
15. Jakola L, Suokas M, Häggman H. Novel approaches based on DNA barcoding and high-resolution melting of amplicons for authenticity analyses of berry species. Food Chem 2010;123:494-500.
16. Ganopoulos I, Bosmali I, Madesis P, Tsatfaris A. Microsatellite genotyping with HRM (high resolution melting) analysis for identification of the PGI common bean variety pleke megalosperma prespor. Ital Food Res Technol 2012;234-50-48.
17. Tchen N, Parveen I, Pan Z, Khan IA. DNA barcoding of medicinal plant material for identification. Curr Opin Biotecnol 2014;25:103-10.
18. Herrmann F, Wink M. Use of rbcL sequences for DNA barcoding and authentication of plant drugs used in Traditional Chinese Medicine. PeerJ PrePrints 2014;1:e1968.
19. Ma HL, Zhu ZB, Zhang XM, Miao YY, Guo QS. Species identification of the medicinal plant Tulipa edulis (Liliceae) by DNA barcode marker. Biochem Syst Ecol 2014;55:362-8.
20. Chen S, Pang X, Song J, Shi L, Yao H, Han J, et al. A renaissance in herbal medicine identification: From morphology to DNA. Biotecnol Adv 2014;32:1237-42.
21. Sarvat M, Yarmagni MV. DNA barcoding, microarrays and next generation sequencing: Recent tools for genetic diversity estimation and authentication of medicinal plants. Crit Rev Biotecnol 2014;0:1-13.
22. Hebert PD, Cywinska A, Ball SL. Biological identifications through DNA barcodes. Proc R Soc Lond B Biol 2003;270:313-21.
23. Sheth BP, Thaker VS. Plant systems biology: Insights, advances and challenges. Planta 2014;230:33-44.
24. Sharma V, Sarkar IN. Bioinformatics opportunities for identification and study of medicinal plants. Brief Bioinform 2013;14:238-50.
25. Sheth BP, Thaker VS. rbcL: A key to C-value paradox in plants. Plant Arch 2012;12:915-9.
26. Hollingsworth PM, Forrest LL, Spouge JL, Haji-babaei M, Ratnasingham S, van der Bank M, et al. A DNA barcode for land plants. Proc Natl Acad Sci 2009;106:12794-7.
27. Vidal AO, Schnerr H, Rojmyr M, Lysholm F, Knight A. Quantitative identification of plant genera in food products using PCR and pyrosequencing technology. Food Control 2007;18:921-7.
28. Li M, Cao H, But PP, Shaw PC. Identification of herbal medicinal materials using DNA barcodes. J Syst Evol 2011;49:271-83.
29. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 1987;19:11-5.
30. Levin RA, Wagner WL, Hoch PC, Nepokroeff M, Pires JC, Zimmer EA, et al. Family-level relationships of Onagraceae based on chloroplast rbcL and ndhF data. Am J Bot 2003;90:107-15.
31. Kress WR, Erickson DL, Jones FA, Swenson NG, Perez R, Sanjur O, et al. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. Proc Natl Acad Sci U S A 2009;106:18621-6.
32. Hall TA. Bio Edit: A user-friendly biological sequence alignment editor and analysis program.

Pharmacognosy Magazine, Oct-Dec 2015, Vol 11, Issue 44 (Supplement 4)
for windows 95/98/NT. Nucleic Acids Symp Ser 1999;41:95-8.

33. Eisenberg D, Lüthy R, Bowie JU. VERIFY3D: Assessment of protein models with three-dimensional profiles. Methods Enzymol 1997;277:396-404.

34. Casirgini T, De Meo P, Cozzetto D, Talamo IG, Tramontano A. The PMDB protein model database. Nucleic Acids Res 2006;34:D306-9.

35. Pei J, Kim BH, Grishin NV. PROMALS3D: A tool for multiple protein sequence and structure alignments. Nucleic Acids Res 2008;36:2295-300.

36. Hasabe M, Kofuji R, Ito M, Kato M, Iwatsuki K, Ueda K. Phylogeny of gymnosperms inferred from rbcL gene sequences. Bot Mag Tokyo 1992;105:673-9.

37. Cosner ME, Jansen RK, Lammers TG. Phylogenetic relationships in the Campanulales based on rbcL sequences. Plant Syst Evol 1994;190:79-98.

38. Muasya AM, Simpson DA, Chase MW, Culham A. An assessment of suprageneric phylogeny in Cyperaceae using rbcL DNA sequences. Plant Syst Evol 1998;211:251-71.

39. Murakami N, Nagami S, Watanabe M, Iwatsuki K. Phylogeny of Aspleniaceae inferred from rbcL nucleotide sequences. Am Fern J 1999;89:232-43.

40. Frye AS, Kron KA. rbcL phylogeny and character evolution in Polygonaceae. Syst Bot 2003;28:326-32.

41. Kim JS, Hong JK, Chase MW, Fay MF, Kim JH. Familial relationships of the monocot order Liliaceae based on a molecular phylogenetic analysis using four plastid loci: MatK, rbcL, atpB and atpF-H. Bot J Linn Soc 2013;172:5-21.

42. Lovell SC, Davis IW, Arendall WB 3rd, de Bakker PI, Word JM, Prisant MG, et al. Structure validation by Calpha geometry: Phi, psi and Cbeta deviation. Proteins 2000;50:437-50.

43. Priti A, Bliven S, Rose PW, Bluhm WF, Bizon C, Godzik A, et al. Pre-calculated protein structure alignments at the RCSB PDB website. Bioinformatics 2010;26:2983-5.

44. Atoschul SF, Boguski MS, Gish W, Wootton JC. Issues in searching molecular sequence databases. Nat Genet 1994;6:119-29.

45. Krauthammer M, Rzhetsky A, Morczov P, Friedman C. Using BLAST for identifying gene and protein names in journal articles. Gene 2002;269:246-52.

46. Schuler GD. Sequence alignment and database searching. Methods Biochem Anal 1998;39:145-71.

47. Menlove KJ, Clement M, Crandall KA. Similarity searching using BLAST. In: Bioinformatics for DNA Sequence Analysis. New York: Humana Press; 2009. p. 1-22.

48. Little DP, Stevenson DW. A comparison of algorithms for the identification of specimens using DNA barcodes: Examples from gymnosperms. Cladistics 2007;23:1-21.

49. Halgren RG, Felden MR, Fong CJ, Zacharewski TR. Assessment of clone identity and sequence fidelity for 1189 IMAGE cDNA clones. Nucleic Acids Res 2001;29:582-8.

ABOUT AUTHORS

Miss. Bhavisha P. Sheth, is currently working as a Ph. D student at the UGC-CAS Department of Biosciences, Saurashtra University, Rajkot. She is a recipient of the Senior Research Fellowship from the University Grants Commission, New Delhi, India. Her research interests include plant biotechnology and systems biology. She has authored nine publications in peer reviewed national and international reputed journals.

Dr. Vrinda S. Thaker, is a senior Professor at the Department of Biosciences, Saurashtra University, Rajkot. Her research interests include plant physiology and molecular biology. She has authored more than 100 publications in peer reviewed journals of international repute. She has guided 13 Ph. D and 24 M. Phil students till date.