



Advancements in the identification and categorization of medicinal plants via the application of DNA barcodes.

Saptarshi Biswas¹, Dr. Naresh Pratap²

Department of Biochemistry

^{1,2}Shri Venkateshwara University, Gajraula (Uttar Pradesh)

Abstract

Since 2008, medicinal plants have benefited from DNA barcoding, a relatively new molecular identification and classification tool. The use of this method has significantly improved confidence in the quality and efficacy of medicines. In this work, we discuss how the use of DNA barcoding and similar technologies has advanced our understanding of the identification and authenticity of medicinal plants during the past decade. DNA barcodes have expanded their genetic data contribution at every scale from “single locus-based DNA” barcodes to integrated markers to genome-scale levels. DNA barcoding has also been integrated with other methods, “such as high- resolution melting (HRM)”. Next-generation sequencing (NGS) has made it possible for metabarcoding technology to successfully identify species in complex mixtures. DNA barcoding is becoming increasingly important in the study of medicinal plants as a commonly utilized and successful approach.

Keywords: *Medicinal plants, DNA barcoding, metabarcoding.*

Introduction

Herbal remedies and medicinal plants have been vital to human health for thousands of years. The demand for these plants and their byproducts persists today. In time, and especially in today's world customer markets, accurate plant item recognition becomes more and more important in the use and selling of medicinal plants including dietary supplements. DNA fingerprint systems have progressed over the last ten years, starting with isolated genes and moving on to mix genes, genomes in some way and, presently, metabarcoding, to monitor the use, selling, and authenticity of therapeutic plants. The study of herbal products and scientific research and ethnobotanical



recording of such remedies for the benefit of individuals has been greatly aided by DNA barcode technology.

It summarizes important lessons learned and recommendations for the advancement of DNA coding in the field of pharmacopeias and botanical medicines. First, a quick overview of DNA coding techniques and a look at therapeutic plants were given. In the second part, we go over the most widely used gene barcoding techniques as well as new approaches to medicinal plant identification. There were conclusions made, along with predictions for botanical DNA scanning research in the future. This paper is the first comprehensive examination of the applications of DNA barcoding in and herbal therapy.

1.1 Background of DNA barcoding and medicinal plant

1.1.1 DNA barcoding

DNA barcoding, also known as barcoding, was suggested as a method in 2003 (Hebert et al., 2003) to enable precise recognition of species. DNA barcoding is a method that uses a uniform, short DNA sequence as a marker for automatic, precise species identification. DNA barcoding technique has now seen widespread use across a variety of industries and been heralded as a renaissance for taxonomy.

The use of DNA barcoding technology for species identification was first proposed at the 2003 “International Barcode of Life Conference held at Cold Spring Harbor (USA)”. At the third barcode conference in 2009, it was recommended that the matK, rbcL, trnH-psbA, and ITS2 sequences would serve as the fundamental DNA barcode locus for plants. At the eighth quarterly national DNA coding event in 2019, which took place in Trondheim, the Norway, new emerging sequencing techniques for the coding of DNA were extensively explored. DNA barcoding technology is rapidly being adopted by an ever-expanding number of research disciplines (Kress, 2017).

1.1.2 Medicinal plants

The term "medicinal plants" is used to describe all plants that have the potential to benefit human health in some way, whether as medications or nutritional supplements. The market for traditional medicines, teas, and herbal supplements is a sizable and growing sector of the biodiversity industry.



There are 420 identified angiosperm plant families thanks to APG 4. There are 219 plant families known to contain medicinal plants, or 52.1%. Sixty-four percent of the 219 families that contain therapeutic plants have had DNA barcodes applied, and these families contain 832 genera. No DNA barcode sequence data has been recorded or published for the other 77 families of medicinal plants. Table 1 demonstrates that throughout the main branching lines of orchids, the proportion of taxa with recordings of plants for medicinal purposes is remarkably stable (ranging from 43.4% in These plants to 55.6% in Baseline These plants). Nevertheless, there are notable differences in the percentage of families having plants for medicine on which genetic barcodes were initially used that vary from 33.3% in These plants to 78.1% in Early Aquatic plants.

1.2 Genome-based DNA barcode markers

All retinal genes are included in the chloroplast genome, which offers a greater amount of genetic data for the identification of species than every single-locus marker. On October 27, 2019, the NCBI released the genomes of 3452 plant chloroplasts. The chlorophyll rnas as a means of plant differentiation and identification is a significant DNA barcode development direction. The complete plastid genome should be employed for DNA barcoding. Genetic fingerprinting can distinguish individuals, species, and communities and detect sample homogeneity and herbal supplement adulterants. Yang et al. (2014) decoded and analyzed the plant chloroplast chromosome of *Lonicera japonica* and discovered *Datura stramonium* species. This approach helps categorize and encode this medicine plant, *L. japonica* cp, by revealing its distinct genetic features (He et al., 2017). Zhou et al. analyzed the molecular frameworks and ancestry of *Papaver rhoeas* and *Papaver orientale* using their full the chloroplasts human genomes. They discovered that a chloroplast genetic might have been applied to get rid of evolutionary theory sets as well as interactions of *Papaveraceae*, a prominent family of herbal remedies (Zhou et al., 2018).

Methodology

The experimental approach was conducted in accordance with established scientific protocols. There are several processes involved in detecting DNA barcoding technology. The first step is to gather samples, which might include processed goods, fresh samples, dried samples, and mixed samples. Second, a variety of techniques, including the CTAB method, SDS process, PVP strategy,



and phenol-chloroform procedure, are used for gathering DNA. Additionally, the amplification step of PCR is performed, with each cycle including the processes of denaturation, heating, and extending. Fourth, DNA is sequenced using the Sanger dideoxy method as an initial stage, moving on to single-molecule approaches for third-generation sequenced and the fourth generation of sequencing by synthesis. Sixth, computational methods such as BLAST are used to execute a similarity-based library searching. Finally, species identification is achieved through the aforementioned steps.

The utilization of computational methods and algorithms to extract meaningful insights from biological data, commonly referred to as bioinformatic analysis, has become. The examination of sequences is a pivotal component within the framework of the technique of DNA a bar code. Genetic barcodes are subject to a wide range of analytical techniques, such as phylogenetic research, DNA sequence investigation, and the genetically distant approach. This research used many biochemical software tools, including ABGD, Taxon DNA, and MEGA 7.0.

1. Results

In the last ten years, a large number of researchers have identified different types of medicinal plants by using whole plant genomes and one or more DNA code identifier sites. The most effective markers and technologies for analyzing medicinal plants have recently come to an agreement⁷. Cost, efficiency, and convenience are crucial when considering DNA barcoding. Single-locus DNA barcode markers are cost-effective, however integrating numerous loci improves identification. When dealing with intraspecific taxa like ecotypes, sequences that cover large information sites, such as chloroplast genomes, are needed. These extensive sequences are super barcodes, improving resolution. The identification of most therapeutic herbs, their substitutes, and adulterants rarely requires such methods. Due to bulk manufacturing and affordability of sequencing technology, medicinal plant DNA barcodes are increasingly targeted at specific taxa. Using the chloroplast genome to identify variation "hotspots" is more efficient than using a universal barcode. This development should boost DNA barcoding approaches.

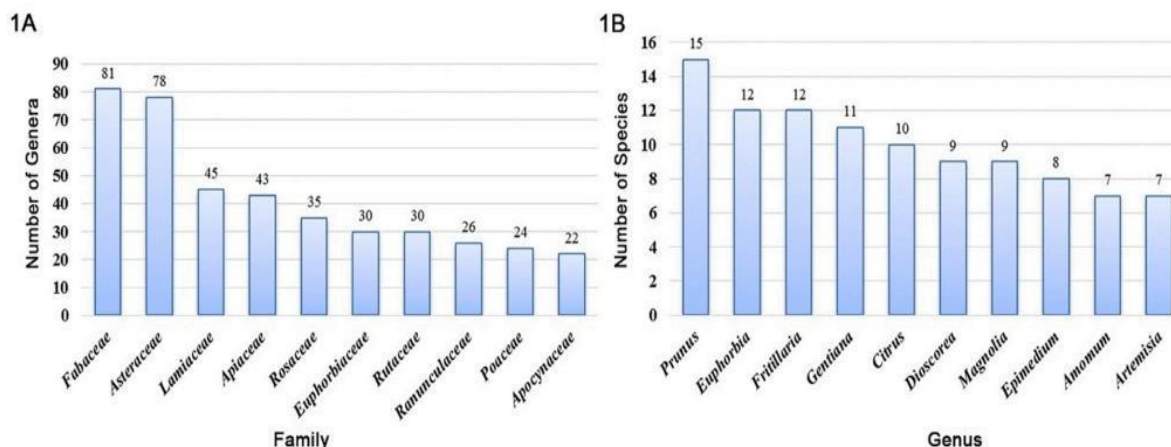


Fig. 1. Top-10 families (1A) and Top-10 genera (1B) recorded in the Pharmacopoeia throughout the world. The Y-axis represents the number of genera in that family or the number of species in that genus recorded in the Pharmacopoeia. We can intuitively see that some important families or genera have made great contributions to the medicinal plants used by human populations.

Table 1

The numbers and proportion of angiosperm families by major evolutionary lineages with reported medicinal plants and medicinal plants with DNA barcode data.

| Types | Total number of families | Number of families with medicinal plants | % families with medicinal plants | Number of families with medicinal plants and DNA barcodes | % families with medicinal plants and DNA barcodes |
|-------------------|--------------------------|--|----------------------------------|---|---|
| Basal angiosperms | 27 | 15 | 55.6% | 9 | 60.0% |
| Basal eudicots | 77 | 38 | 49.3% | 29 | 76.3% |
| Monocots | 17 | 8 | 47.1% | 3 | 37.5% |
| Rosids | 152 | 87 | 57.2% | 57 | 65.5% |
| Asterids | 147 | 71 | 48.3% | 44 | 62.0% |
| Total | 420 | 219 | 52.1% | 142 | 64.8% |

Conclusion

DNA barcoding, a method that has gained widespread acceptance, has significantly contributed to the categorization of therapeutic plants, the identification of additives and replacements, and the regulation of the drug business. The current technological capabilities are proficient in the identification and utilization of a wide range of plant raw materials and partially processed products within the medicinal materials market. However, there remains a deficiency in the availability of efficient, expeditious, and standardized identification techniques for certain refined pharmaceutical



products, such as tablets, pills, oral liquids, and injections. This is particularly evident in the case of the intricate botanical constituents found in Chinese patent medicine, posing a significant challenge for the majority of researchers. Furthermore, the determination of distinct The development of the sequencing of genomes is closely related to the identification of DNA sequences in plants used for medicine. Notwithstanding the developments in last- and following-generation sequencing capabilities, there remains a significant gap in the studies of chloroplast (or plastid) genetics in medicinal plants. The current lack of sequencing for numerous chloroplast genomes in medicinal plants poses a constraint on our capacity to formulate novel DNA barcodes with specificity.

References

1. Hebert, P.D.N., Cywinska, A., Ball, S.L., Dewaard, J.R., 2003. Biological identifications through DNA barcodes. *Proc. R. Soc. B Biol. Sci.* 270, 313–321.
2. Kress, W.J., 2017. Plant DNA barcodes: applications today and in the future. *J. Syst. Evol.* 55, 291–307.
3. Yang, Y., Dang, Y.Y., Li, Q., Lu, J.J., Li, X.W., Wang, Y.T., 2014. Complete chloroplast genome sequence of poisonous and medicinal plant *Datura stramonium*: organizations and implications for genetic engineering. *PLoS One* 9, e110656.
4. He, L., Qian, J., Li, X.W., Sun, Z.Y., Xu, X.L., Chen, S., 2017. Complete chloroplast genome of medicinal plant *Lonicera japonica*: genome rearrangement, intron gain and loss, and implications for phylogenetic studies. *Molecules* 22, 249.
5. Zhou, J., Cui, Y., Chen, X., Li, Y., Xu, Z., Duan, B., Li, Y., Song, J., Yao, H., 2018. Complete chloroplast genomes of *Papaver rhoeas* and *Papaver orientale*: molecular structures, comparative analysis, and phylogenetic analysis. *Molecules* 23, 437.