INTERNATIONAL ASSOCIATION OF BIOLOGICALS AND COMPUTATIONAL DIGEST
International & Peer-Reviewed Journal
E-ISSN: 2583-3995

A REVIEW METHOD FOR IDENTIFICATION OF RARE AND ENDANGERED PLANTS THROUGH DNA BARCODING

[1]Nikisha Rohit and [2]Hitesh Solanki

Department of Botany, Bioinformatics and Climate Change Impact Management,University School of Sciences, Gujarat University,
Ahmedabad – Gujarat

Email: nikishpurohit@yahoo.com

ABSTRACT

DNA barcoding is a new concept. It has been developed for providing fast, precise and automatable species identification which uses standardized DNA sequences as tags. DNA barcoding can provide the taxonomists; conservationists. The early goal of the DNA barcoding process is to build online libraries of barcode sequences for all known species that can serve as a standard to which DNA barcodes of any identified or unidentified specimens can be matched. This can improve several inherent problems related with traditional taxonomic identification, based onmorphological characters, such as incorrect identification of species due to phenotypic plasticity and genotypic variability of the characters, such as incorrect identification of species due to phenotypic plasticity and genotypic variability of the characters, overlooking cryptic taxa, difficulty in finding reliable characters due to long maturity periods (CBOL Plant Working Group, 2009). It is particularly of much use in areas where species identification with morphological characters is not practicable due to widespread damage or delayed expression. It should be enduring in mind that DNA barcoding is not an alternative to taxonomy, and it cannot replace taxonomy as such, but is a useful tool that creates information on unknown taxa. In this paper, methods of the process of selecting and redefining barcodes for plants evaluation of the factors which manipulate the discriminatory power of the advance with some early applications of DNAbarcoding are discussed and then added the authors' for their views and recommendations.

Keywords: *DNA Barcode,rbcL, matK, Agarose Gel Electrophoresis, Sanger sequencing, Rare, Endangered Plants*

INTRODUCTION

DNA barcoding is a tool for species identification. The need for accurate species identification forboth conserve and utilize plants plant DNA barcoding is most important. DNA barcoding of plantsis already being working in a wide variety of applications. In many regions of the worlds, however,this may be slowed down by a lack of taxonomic expertise (Chase *et al.*, 2009). As well as identifying whole plants, it is also sometimes valuable to be able to identify species from plant material such as seeds, pollen, roots or in mixture of plants sampled from the air, soil and water (De vere *et al.*, 2012). Genetic sequences acquired in the context of DNA barcoding have also been used to create phylogenetic tress which is use in ecological phylogenetic community (Krees *et al.*,2009; Krees *et al.*, 2010). Following the evaluation of several candidate markers, The Plant Working (PWG) of the Consortium for the Barcoding of Life (CBOL) suggested that regions of two plastids genes rbcL and matK be adopted as the standard plant DNA barcodes, with the recognition that supplementary, markers may be required. The use of DNA barcoding as an identification tool is also dependent on the creation of high-quality reference databases of sequence(Hollingsworth *et al.*, 2011). It is crucial that every DNA sequence should be associated with the plant specimen from which it came, along with when, where and by whom it was collected and identified. The lab procedure over which a sample is managed should also be recorded, with the primers used, trace files and character statistics for its DNA sequence all available to end users ofthe information (De vere *et al.*,

2012). All data should be widely available in Gene Bank which provides a source for DNA sequences. In summation, it is urged to deposit data on to the Barcodeof Life Data system (BOLD). BOLD provides a way of managing tasks and allows trace files, scans of herbarium specimens and photographs to be stored alongside.

PERSPECTIVE FOR PLANT DNA BARCODING AND TOMORROW'S VIEWPOINT:

Since the time of their entry into the botanical community over a decade ago DNA barcodes havebeen given to an assortment of investigations in both basic and applied research in plants.

Appearing to the future, plant DNA barcoding will advance in two keyways to attend to the botanical community.

1. Building a more comprehensive global plant DNA barcode library for universal use.

2. Developing new markers and adopting new sequencing technologies.

matK gene

matK is one of the most quickly evolving genes. It has a length 43 of about 1550 bp and encodes the enzyme maturase which is involved in the splicing of type-2 introns from RNA transcripts (Neuhaus *et al.*, 1987). Since matK is entrenched in the group 2 introns of the lysine gene trnK which can be easily PCR amplified with a primer set designed from the conserved regions of the genes trnK, rps16 and psbA. Considering the high evolutionary rate of matK, it has been tested byseveral workers for suitability as a plants barcode and has been proposed either alone or in combination with other loci. For example, after testing matK along with 11 other cpDNA loci in 98 land plant taxa, suggested a combination of rpoC1 + rpoB + matK as the most promising combination for barcoding of land plants (Ford *et al.*, 2009). After testing matK, rbcL, rpoB, rpoC1, and trnH-psbA as barcodes in Cyperaceae, also advocated the use of matK alone as a universal barcode for land plants (Starr *et al.*, 2009).
matK has been used as a marker to build plant phylogenies because of its fast evolution and the universal presence in plants (Hilu *et al.*, 1997; Kelchner *et al.*, 2000). Nevertheless, failure of PCRamplification of matK in some taxonomic groups was also reported. So that avoid this problem, new sets of primers were developed which works well in most of the major taxonomic groups (Cuenoud *et al.*, 2002). This primer set amplifies a DNA fragment of~930 bp between positions 429 and 1313 of the matK sequence (Schmitz-Linneweber *et al.*, 2001; Cuenoud *et al.*, 2002). Phylogenetically, the rate of evolution of matK was establish appropriate for resolving intergenericas well as interspecies relationships in many angiosperms (Soltis *et al.*, 1998; Johnson *et al.*, 1995).The CBOL Plant Working Group tested matK in nearly 550 plant species and found that nearly 90% of the angiosperm's samples were simply amplified and sequenced using a single primer pair,though the success was limited in gymnosperms (83%) and much worse in cryptogams (10%).
Because of this high universality and species discrimination, the CBOL Plant Working Group recommended matK in combination with rbcL as the standard two locus barcode for plants.

rbcL gene
rbcL is the best characterized gene sequence. Consequently, most of the investigating group testedits suitability in barcoding. It encodes the large subunit of rubilose 1,5- bisphosphate carboxylase/oxygenase (RUBISCO). RUBISCO is a critical photosynthetic enzyme. rbcL was thefirst gene that was sequenced from the plants (Zurawski *et al.*, 1981). Additionally, in order to attain enough species discrimination the entire~1430 bp needs to be sequenced, which acts as a limiting factor for its use as a barcoding sequence because an ideal DNA barcoding region shouldbe short enough to amplify from degraded DNA and it analyzed via single pass sequencing one solution for this was to amplify short sequences with sufficient variability (Chase *et al.*, 2007). Primers for PCR amplification and sequencing for such short sequence within rbcL gene have beenestablished accordingly for most of the taxa because of the ease in PCR amplification across a wide range of plant groups (Krees *et al.*, 2007; Fay *et al.*, 1997). Most of the phylogenetic studiesrecommend that rbcL is best suited to reconstruct

the relationships down to the generic levels butis not useful for specific levels (Soltis *et al.*, 1998).

Agarose Gel Electrophoresis

Agarose gel electrophoresis has proven to be a well-organized and effective way of separating nucleic acid. Agarose's high gel strength allows for the handling of low percentage gels for the separation of large DNA fragments. The size of pores generated by the bundles of agarose is determined by molecular sieving in the gel matrix (Devor, 2010). DNA fragments smaller than 100 bp are more effectively separated which use polyacrylamide gel electrophoresis. Overall, thehigher the concentration of agarose, the smaller the pore size. Agarose can be adjusted to create low melting agarose through hydroxyethylation. When the isolation of separated DNA fragmentsis anticipated, low melting agarose is mostly used. Hydrooxyethylation eases the packing densityof the agarose bundles which can be effectively reduced their pore size (Serwer, 1983). This meansthat a DNA fragment of the same size will take longer to move through a low melting agarose gel as opposed to a standard agarose gel because the bundles associate with one another through non-covalent interactions, it is probable to re-melt an agarose gel after it has set (Dea *et al.*, 1972). Nevertheless, in certain situations, such as when hazardous waste disposal is difficult or when young students are performing an experiment, a less toxic dye may be preferred.

Loading dyes used in gel electrophoresis serve three major purposes:

1. Add density to the sample, allowing it to sink into the gel
2. The dyes offer color and simplify the loading process. Finally, the dyes move at standard rates through the gel, permitting for the assessment of the distance that DNA fragments have migrated.

Polymerase Chain Reaction (PCR)

Polymerase Chain Reaction (PCR) is a technique to estimate a short sequence of DNA or RNA even in samples having only minute quantities of DNA or RNA. PCR is utilized to reproduce selected sections of DNA or RNA. Before, amplification of DNA intricate cloning the segments of interest into vectors for expression in bacteria and took weeks. But with PCR done in test tubes,it occupies just a few hours today. PCR is extremely effectual in that untold numbers of copies canbe cleared of the DNA.

Additionally, PCR uses the same molecules which use for copying DNA by nature:

1. Two "primers" and short single stranded DNA sequences that are combined to resemble tothe beginning and ending of the DNA stretch to be copied.
2. An enzyme called polymerase that travels along the segment of DNA, reading its code andaccumulating a copy and a pile of DNA building block that the polymerase needs to make that copy.

Three chief steps are involved in PCR. These three steps are repeated for 30 or 40 cycles. The cycles are prepared on an automated cycler, a device which quickly has heats and cools the test tubes containing the reaction mixture.

1. Denaturation: At 94˚C, the double standard DNA melts and opens into two pieces of singlestranded DNA.
2. Annealing: At medium temperatures, around 54˚C, the primers pair up (anneal) with the single-stranded "template" (The template is the sequence of DNA to be copied). On the small length of double stranded DNA (the joined primer and template), the polymerase attaches and starts copying the template.
3. Extension: At 72˚C, the polymerase works best and DNA building blocks complementaryto the template are coupled to the primer, making a double stranded DNA molecule.

With one cycle, a single segment of double stranded DNA template is amplified into two separatepieces of double stranded DNA. These two pieces are than accessible for amplification in the nextcycle. As the cycles are repeated, more and more copies are generated and the number of copies of the template is increased exponentially. The reaction is easy to execute. It requires no more thana test tube, a few simple reagents and a source of heat".

Sanger sequencing for DNA

Sanger sequencing is also known as the "Chain termination method", this method is for determining the nucleotide sequence of DNA. The method was developed by two-time Noble Laureate Frederick Sanger and his colleagues in 1977, hence the name the Sanger sequence. There are three main steps to Sanger sequencing

1. Generating n DNA fragments of varying length, each terminated with a labeled nucleotide,where n is the number of nucleotide bases in the target DNA sequence. This is done by combiningDNA primer, nucleotides (dATP, dCTP, dGTP, dTTP), DNA polymerase, the DNA sequence of interest and labeled dideoxynucleotides (ddATP, ddCTP, ddGTP, ddTTP). No nucleotide can be added to the DNA chain once a dideoxynucleotides has been combined, so each fragment will end with a labeled nucleotide. A much smaller amount of dideoxynucleotides is used than the amountof regular nucleotides.
2. Separate the n DNA sequences are length using capillary gel electrophoresis. The shorter fragments move faster than the longer fragments. The result is that the DNA pieces are fed into third step from shortest to longest sequence.
3. A laser excites the label on the nucleotide at the end of each sequence. Each base is taggedwith a different label, so the light emitted by each excited nucleotide can be tied to the correct base. The laser generates a chromatogram showing the fluorescent peak of each nucleotide. The chromatogram has the nucleotides in the correct order because of the electrophoresis.

If the target sequence is 1,000 nucleotides long and there is only one copy of the template, it is going to take longer to generate the 1,000 tagged fragments. Nevertheless, if there are several copies of the template, in theory it will take less time to generate all 1,000 of the tagged fragments(sigma Aldrich Sanger sequencing).

CONCLUSION

Here DNA barcoding comes to liberate to test recovery of standard DNA barcodes for rare and endangered plants In future, the large scale implementation of plant DNA barcoding require algorithms that place search results in context of standard plant names and character-based keys for distinguishing closely-related species. It can be a better alternative and useful in identifying rare and endangered plant leaves even when flowers or fruit are not available, identifying products in commerce for example, herbal medicinal supplements, identification of rice varieties, timber wood identification.

REFERENCES

1) Chase M.W., Fay M.F. (2009) Barcoding of plants and fungi. Science 325:682-683.
2) Cuénoud P., Savolainen V., Chatrou L. W., Powell M., Grayer R. J. and Chase M. W. (2002) Molecular phylogenetics of caryophyllales based on nuclear 18S rDNA and plastid rbcL, atpB, and matK DNA sequences. American journal of Botany 89; 132-144.
3) Dea I.C.M., McKinnon A.A., Rees D.A. (1972) Tertiary and qualitative structure and aqueous polysaccharide systems which model cell wall adhesion: reversible changes in conformation and association if agarose, carrageenan and galactomannans. Journal of Molecular Biology 68:153-172.
4) Devor E.J. IDT tutorial: gel electrophoresis (2010) Available from: http://cdn.idtdna.com/Support/Technical/TechnicalBulletinPDF/Gel Electrophoresis.pdf.
5) De Vere N., Rich T.C.G., Ford C.R et al (2012) DNA barcoding the native flowering plants and conifers of Wales. PLoS One 7:e37945
6) Ebach M. C., Holrege C., (2005) DNA barcoding is no substitute for taxonomy. Nature, 434; 697.
7) Fay M. F., Swensen S. M., Chase M. W., (1997) Taxonomic affinities of Medusagyneoppositifolia (Medusagynaceae). Kew Bull. 52; 111-120.
8) Hollingsworth P.M., Graham S.W., Little D.P (2011) Choosing and using a plant DNA barcode. PLoS One 6:e19254.
9) Hilu K. W. Liang H., (1997) The matK gene: sequence variation and application in plant
10) systematics. American Journal of Botany 84; 830-839.
11) Johnson L. S. Soltis D. E. (1995) Phylogenetic inference in Saxifragaceaesensu strict and

Gilia (Polemoniaceae) using matK sequences. Annals of Missouri Botanical Garden 82; 149-175.
12) Kress W.J., Wurdack K.J., Zimmer E.A., Weigt L.A., Janzen D.H. (2005) Use of DNA barcodes to identify flowering plants. Proceedings of the National Academy of Sciences USA 102:8369-8374.
13) Kress W.J., Erickson D.L., Swenson N.G et al., (2010) Advances in the use of DNA barcodes to build a community phylogeny for tropical trees in a Puerto Rican forest dynamics plot.PLoS One 5:e15409.
14) Kress W.J., Erickson D.L., Jones F.A et al., (2009) Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. Proceedings of the National Academy ofSciences USA 106; 18621-18626.
15) Lahaye R. et al., (2008) DNA barcoding the floras of biodiversity hot-spots. Proceedings of National Academy of Sciences USA 105; 2923-2928.
16) Neuhaus H., Link G. (1987) The chloroplast tRNA(UUU) gene from mustard (Sinapsis alba) contains a class II intron potentially coding for a maturase-related polypeptide. Current Genetics 11; 251-257.
17) Ratnasingham S., Hebert P.D.N (2007) BOLD: the barcode of life data system (www.barcodinglife.org). Molecular Ecology Notes 7; 355-364.
18) Starr J. R., Naczi R. F. C., Chouinard B. N. (2009) Plant DNA barcodes and species resolution in sedges (Carex, Cyperaceae). Molecular Ecological Resources 9(Suppl. 1); 151-163.
19) Schmitz-Linneweber C., Maier R. M., Alcaraz J. P., Cottet A., Herrmann R. G., Mache R., (2001) The plastid chromosome of spinach (Spinaciaoleracea): complete nucleotide sequence andgene organization. Plant Molecular Biology. 45; 307-315.
20) Soltis D. E., Soltis P. S. (1998) Choosing an approach and an appropriate gene for phylogenetic analysis. In Molecular Systematics of Plants II: DNA Sequencing Kluwer, Dordrecht21-24.
21) Serwer P. (1983) Agarose gels: properties and use for electrophoresis. Electrophoresis. 4; 375-382.
22) Zurawski G., Perrot B., Bottomley W., Whitefield P. R. (1981) The structure of the gene for the large subunit of ribulose-1, 5- bisphosphate carboxylase from spinach chloroplast DNA. Nucleic Acids Research. 9; 3251-3270.