DNA Bar-Code for Identification of Microbial Communities: A Mini-Review

Dhiraj Kumar Chaudhary[1]* and Ram Hari Dahal[2]

[1]*Department of Soil Science, Prithu Technical College, Institute of Agriculture and Animal Science, Tribhuvan University, Dang, Nepal*
[2]*Department of Microbiology, Tri-Chandra Multiple Campus, Tribhuvan University, Nepal*

*Corresponding Author: Dhiraj Kumar Chaudhary, Department of Soil Science, Prithu Technical College, Institute of Agriculture and Animal Science, Tribhuvan University, Dang, Nepal.

Abstract

Development of new molecular techniques have enabled scientist for robust identification of species. DNA barcode system provides authentic species identification in large number without delay and error. Several gene markers have been used for the identification of diverse microbial communities in DNA barcoding system. For the identification of bacteria, 16S rRNA gene is an important marker. There is no specific marker for virus. But, some studies have implemented outer core protein and K-mer based barcode system for identification of blue tongue virus (BTV) and human enterovirus (HEVs), respectively. For animal, Mitochondrial Cytochrome c oxidase Subunit I (COI) and mtDNA genes are used. For algae, rbcL and RuBisCo are mostly used whereas for protozoa 18S rRNA gene is used. Identification of fungi is usually done by COI, ITS, LSU, SSU, RPB1, and RBP2 genes. Barcode database are growing very rapidly about 100,000 species per year. This review article summarizes the current molecular techniques for identification of species and efficacy of DNA barcoding for various organisms. DNA has potential to carry digital information too. So, we are hopeful for advancement in digital barcode hologram with error free and appropriate barcode system for identification of various microbial communities in the near future.

Keywords: *DNA Barcode; Microbial Community; Marker Gene; Molecular Technique; Taxonomy*

Introduction

The technological development has made a volcanic eruption of biological knowledge from the covert environmental niches. Computer science has enabled scientist to solve various biological problems and health related problems, such as genetically inherited diseases and drug discovery [1]. In the ecosystem and diversity domain, DNA barcoding is a recent significant effort in this direction [2].

There are stunning quantities of microorganisms in natural habitat. This seems thrilling for researcher to cope exact taxonomical position of microorganism. These microorganisms are still not recognized and only small fractions are cultured in the formulated media remaining huge portion as unculturable [3]. DNA genome-based and DNA sequencing based technologies can be essential component in identifying these microbes. Recently, application of DNA sequence data, the 'DNA barcoding' system is providing a proficient place for species-level classification. Advancement of sequencing and computational technologies have standardized DNA sequence and increased their capabilities in identification of microbes with the help of short sequences [4]. Due to the genetic variation in the short unique sequence, it is useful for differentiating the individual species. Using these sequences, many efforts have been made in identifying different strains. These sequences can also be used for the development of barcode for microbial communities. These fragments of DNA sequences which are implemented for identification of unknown species are referred as DNA barcode and the system involved in recognition of alien strain is called as DNA barcoding [5].

The process of conventional species identification has many shortcomings. First, conventional strain identification methods may lead false characterization due to phenotypic flexibility and genetic variability [6]. Whereas, DNA barcode system proved to be robust in spe-

cies identification, being itself automated gives accurate results. Next, traditional method has problem in identification of cryptic species complex and sometime morphological keys used in identification may vary in particular life cycle arising difficulties in species identification. In contrast, DNA barcode system is rapid tools and can classify large quantities of microorganisms at the same time, without any error [7].

The emerging interest in microbial systematic has built extremely large database in NCBI and other gene bank [8]. DNA barcoding system along with marker genes have crucial role in microbial systematic and taxonomy. In this study, we have reviewed briefly on DNA barcoding system and presented some marker genes which have been applied for species identification.

DNA barcoding system

Genetic diversity is fluctuating from one species to another species. In molecular barcoding system, the approach for species identification is made usually by retrieving small segment of gene from whole genome of microorganisms. This small segment of gene is considered as barcode sequence which is very specific for a particular gene [5]. This short, consistent sequence can help in differentiating individual species. In barcoding system, genomic DNA is extracted from unknown isolates. Using specific primer, target genes can be retrieved. The short fragment of DNA is amplified by using PCR. The amplicon obtained is then sequenced for bioinformatic analyses. The database using appropriate computer algorithm lead for identification of unknown strains (Figure 1).
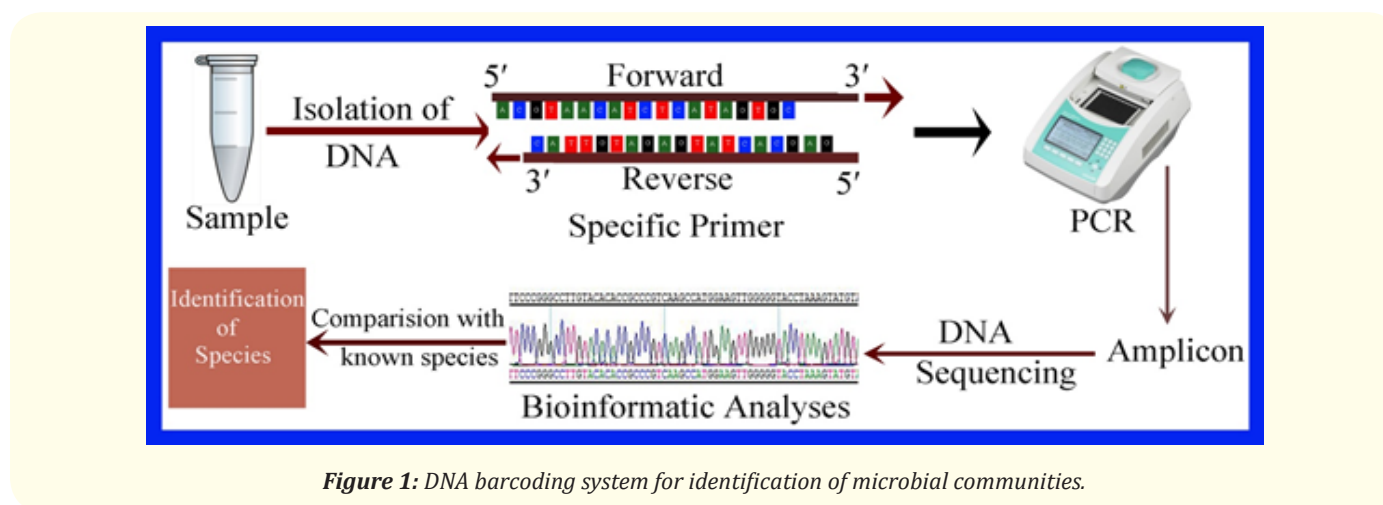

Figure 1: DNA barcoding system for identification of microbial communities.

Beside microbial communities, animal and plant species have been also identified with barcoding system.

There are many molecular markers for animal and plant species. Mitochondrial Cytochrome c oxidase Subunit I (COI) (650 bp) and mtDNA gene have been studied as marker genes for animal [9]. Low rate of nucleotide substitution of mitochondrial genome in plants have created limitation for implementing COI as universal barcode marker. For plants, molecular markers used for species identification are rbcL, 23S rDNA, rpoB, trnH-psbA, rpoC1, atpF-atpH, psbK-psbI, and matK [10]. There is no single promising marker for identification in plants. Many researchers have been involved in attempting to present single barcode gene for identification of plant species. A study conducted by Lahaye., *et al.* (2008) estimated that the matK marker gene works singly as barcode marker for botanical species identification [11]. In case of bacteria, 16S rRNA gene sequence (1500 bp nucleotide in length) is proven marker for species identification [12]. The barcode marker sequence from each unidentified species is compared with a library of reference barcode sequences. The final goal of the DNA barcoding system is to build up a robust and efficient mechanism for the species identification in a rapid manner which should be simple and scalable [4].

DNA bar-code for Microbial Population

Evolution creates biodiversity in microbial communities. Gene sequence from DNA sample noted that the diversity of microbes is almost 100 times higher than what was projected by traditional microbiology. There is vast diversity in the communities of virus, bacteria, algae, fungi, and protozoa [13]. The diversity of microorganisms can be identified with well-established marker genes through DNA barcoding system (Table 1).

S. No.	Microbial Population	DNA Bar-Code genes	References
1.	Virus	K-mer based barcode, Outer core Protein, Oligonucleotides	[15,16,31]
2.	Bacteria	16 S rRNA gene, COI, CPN60.	[17-19,32]
3.	Algae	COI, rbcL, LSU, ITS, RuBisCO	[20-22,33,34]
4.	Fungi	COI, ITS, LSU, SSU, RPB1, RPB2	[23-25]
5.	Protozoa	18s rRNA, 28s rRNA, ITS, COI	[26-29]

Table 1: DNA Bar-Code for Microbial Population.

DNA barcode system for viral communities is still under infancy. Lack of appropriate marker gene seems to be troublesome during viral specimen identification. Some virologists are involved in the establishment of reliable DNA barcode for virus. A study described K-mer based barcode for human enterovirus (HEVs) [14]. For BTV (Blue tongue virus), VP7 outer core protein is used as marker [15]. Similarly, for avian influenza virus (AIV), a fluorescent DNA barcode- based immunoassay has been developed [16].

Bacterial diversity can be distinguished with 16S rRNA gene which is a universal marker for bacteria [17]. COI gene is another DNA barcode developed for bacteria which is 650 bp in length [18]. Chaperonin-60 (cpn60) (known as GroEL 7 Hsp60), is a molecular chaperone conserved in bacterial strains that could be used as barcode marker for bacterial species identification [19].

Numerous marker genes have been suggested for algal species recognition. COI, rbcL (the rubisco operon), internal transcribed spacer (ITS), tufA, COX1, and large subunit (LSU) 28S of the ribosomal cistron, 23S Universal Plastid Amplicon (UPA) are some of the examples of barcode gene for algal communities [20-22].

There are problematic boundaries for multitaxon evolutionary and biodiversity studies of fungi due to lack of simple DNA barcode marker. In recent days, fungal gene sequences are elevating in NCBI database. This made urge to scientist to perform taxonomical studies

based on genomic criteria. These day mycological researches are also very keen for DNA barcoding of fungal species. ITS, LSU, SSU, RPB1, and COI are some of the examples of marker genes which can be introduced for fungus taxonomical studies [23,24]. In a study made by Stockinger., *et al.* (2010) illustrated a larger 1500 bp gene segment for arbuscular mycorrhiza. The study presents SSUmCf-LSUmBr 1500 bp sequence as proven barcode fragments for identification of arbuscular mycorrhizal fungi [25].

Rapid identification of pathogenic protozoan is utmost necessary in medical field for diagnosis of protozoan diseases. Molecular techniques have made convenient for identification of protozoa species. Further, combining with DNA barcoding system, this field has become more easy and interesting for protozoan classification. Amoeba is extensively used in research for developing appropriate DNA barcode. Due to cosmopolitan in distribution, several attempts have been made to build proper molecular markers for protozoa. Three genes, namely, SSU, ITS, and COI genes have been cloned to develop the molecular marker. In many studies, COI gene is concluded as the best barcode marker for identification of amoebic strains. Likewise, 18S rRNA, 28S rRNA, ITS regions have been studied as marker genes for *Piroplasma* [26-29].

Concluding Remarks

DNA barcoding is an emerging field in taxonomy. This technique provides a rapid solution for 'cryptic species' differentiation. Genetic distance boundary between two strains can be established easily. This criterion should be considered during development of specific DNA barcode. Even though the boundary seems to be taxon related, it was noted that the value of genetic distance between two DNA barcode sequences equivalent to or more than 3% identifies distinct species [4,7]. But, in defining DNA barcodes, genetic distance approach imposes certain limitations. First, DNA which are selected as a barcode marker undergo evolution that varies substantially between and within species and between difficult groups of species, consequently resulting in large overlaps of intra-specific and inter-specific distances. Next, numerous gene fragments of different markers have created challenges in appropriate selection of gene sequence for species identification. Thirdly, marker chosen for DNA-barcode also possesses challenge to separate morphological characters from species to species [5,30].

The barcode databases are increasing rapidly (> 100,000 specimens per annum). International Barcode of Life explains appropriately that short sequence of gene from whole genome can be utilized for taxonomical studies of microbial species. Eventually, scientific world are in deep concerns to develop a digital barcode hologram. The digital barcode hologram, will lead us to apply barcode reader for species identification [4,5]. In conclusion, the development of digital barcode hologram for life is still in its primitive stage. The first phase of such work is to develop the barcode with digital information, i.e., DNA-based digital information. DNA has many potential advantages for unchallengeable, high information storage which can also carry digital information. Finally, we are very hopeful to notice condensed, digital barcode hologram with error-free and appropriate barcode for microbial communities in the near future which will enable us for robust identification of microbial species.

Bibliography

1. Iranbakhsh A and Seyyedrezaei SH. "The impact of information technology in biological sciences". *Procedia Computer Science* 3 (2011): 913-916.

2. Savolainen V., *et al.* "Towards writing the encyclopedia of life: an introduction to DNA barcoding". *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 360.1462 (2005): 1805-1811.

3. Schloss PD and Handelsman J. "Metagenomics for studying unculturable microorganisms: cutting the Gordian Knot". *Genome Biology* 6.8 (2005): 229-232.

4. Lebonah DE., *et al*. "DNA Barcoding on Bacteria: A Review". *Advances in Biology* (2014).

5. Chakraborty C., *et al*. "DNA barcoding to map the microbial communities: advances and future directions". *Applied Microbiology and Biotechnology* 98.8 (2014): 3425-3436.

6. Hawksworth DL and Kalin-Arroyo MT. "Magnitude and distribution of biodiversity". In: Heywood VH (ed) Global biodiversity assessment. Cambridge University Press. Cambridge (1995): 107-191.

7. Hebert PD., *et al*. "Biological identifications through DNA barcodes". *Proceedings of Biological Sciences* 270.1512 (2003): 313-321.

8. Riesenfeld CS., *et al*. "Metagenomics: genomic analysis of microbial communities". *Annual Review of Genetics* 38 (2004): 525-552.

9. Ratnasingham S and Hebert PDN. "A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System". *PLoS One* 8.7 (2013): e66213.

10. Fazekas AJ., *et al*. "Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well". *PLoS One* 3.7 (2008): e2802.

11. Lahaye R., *et al*. "DNA barcoding the floras of biodiversity hotspots". *Proceedings of the National Academy of Sciences of the United States of America* 105.8 (2008): 2923-2928.

12. Tian Q., *et al*. "DNA barcoding for effiecientspecies-and pathover-level identification of the quarantine plant pathogen Xanthomonas". *PLoS One* 11.11 (2016): e0165995.

13. Tringe SG., *et al*. "Comparative metagenomics of microbial communities". *Science* 308 (2005): 554-557.

14. Wei C., *et al*. "Identification and typing of human enterovirus: a genomic barcode approach". *PLoS One* 6.10 (2011): e26296.

15. Yin HQ., *et al*. "Nanoparticle-based bio-barcode assay for the detection of bluetongue virus". *Journal of Virological Methods* 178.1-2 (2011): 225-228.

16. Cao C., *et al*. "Detection of avian influenza virus by fluorescent DNA barcode-based immunoassay with sensitivity comparable to PCR". *Analyst* 135.2 (2009): 337-342.

17. Tindall BJ., *et al*. "Notes on the characterization of prokaryotic strains for taxonomic purposes". *International Journal of Systematic and Evolutionary Microbiology* 60.1 (2010): 249-266.

18. Michel H., *et al*. "Cytochrome c oxidase: structure and spectroscopy". *Annual Review of Biophysics and Biomolecular Structure* 27 (1998): 329-356.

19. Links MG., *et al*. "The chaperonin-60 universal target is a barcode for bacteria that enables de novo assembly of metagenomic sequence data". *PLoS One* 7.11 (2012): e49755.

20. Zou S., *et al*. "How DNA barcoding can be more effective in microalgae identification: a case of cryptic diversity revelation in Scenedesmus (Chlorophyceae)". *Scientific Reports* 6 (2016): 36822.

21. Sasikumar K and Anuradha C. "DNA barcoding as a tool for algal species identification and diversity studies". *Double Helix Research* 7 (2012): 2250-3668.

22. Tan J., *et al*. "Assessment of four molecular markers as potential DNA barcodes for red algae Kappaphycus Doty and Eucheuma J. Agardh (Solieriaceae, Rhodophyta)". *PLoS One* 7.12 (2012): e52905.

23. Xu J. "Fungal DNA barcoding". *Genome* 59.11 (2016): 913-932.

24. Cho Y., *et al*. "Explosive invasion of plant mitochondria by a group I intron". *Proceedings of the National Academy of Sciences of the United States of America* 95.24 (1998): 14244-14249.

25. Stockinger H., *et al*. "DNA barcoding of arbuscular mycorrhizal fungi". *New Phytologist* 187.2 (2010): 801-812.

26. Gou H., *et al*. "A DNA barcode for Piroplasmea". *Acta Tropica* 124.1 (2012): 92-97.

27. Hoef-Emden K. "Pitfalls of establishing DNA barcoding systems in Protists: the Cryptophyceae as a test case". *PLoS One* 7.8 (2012): e52905.

28. Karnati SK., *et al*. "Technical note: specific PCR amplification of protozoal 18S rDNA sequences from DNA extracted from ruminal samples of cows". *Journal of Animal Science* 81.3 (2003): 812-815.

29. Nassonova E., *et al*. "Barcoding amoebae: comparison of SSU, ITS and COI genes as tools for molecular identification of naked lobose amoebae". *Protist* 161.1 (2010): 102-115.

30. Lorenz JG., *et al*. "The problems and promise of DNA barcodes for species diagnosis of primate biomaterials". *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 360.1462 (2005): 1869-1878.

31. Pnathee DR., *et al*. "Novel molecular marker associated with Tm2[a] gene conferring resistance to tomato mosaic virus in tomato". *Plant Breeding* 132.4 (2013): 413-416.

32. Woo PC., *et al*. "Then and now: use of 16S rDNA gene sequencing for bacterial identification and discovery of novel bacteria in clinical microbiology laboratories". *Clinical Microbiology and Infection* 14.10 (2008): 908-934.

33. Hadi SIIA., *et al*. "DNA Barcoding Green Microalgae Isolated from Neotropical Inland Waters". *PLoS One* 11.2 (2016): e0149284.

34. Meyer MT., *et al*. "Rubisco small-subunit α-helices control pyrenoid formation in Chlamydomonas". *Proceedings of the National Academy of Sciences of the United States of America* 109.47 (2012): 19474-19479.