Saudi Journal of Biological Sciences (2014) **21**, 204–231

King Saud University

Saudi Journal of Biological Sciences

www.ksu.edu.sa
www.sciencedirect.com

REVIEW

The changing epitome of species identification – DNA barcoding

M. Ajmal Ali [a,*], Gábor Gyulai [b], Norbert Hidvégi [b], Balázs Kerti [b], Fahad M.A. Al Hemaid [a], Arun K. Pandey [c], Joongku Lee [d]

[a] *Department of Botany and Microbiology, College of Science, King Saud University, Riyadh 11451, Saudi Arabia*
[b] *Institute of Genetics and Biotechnology, St. István University, Gödöllo H-2103, Hungary*
[c] *Department of Botany, University of Delhi, Delhi 110007, India*
[d] *International Biological Material Research Center, Korea Research Institute of Bioscience and Biotechnology, 111 Gwahangno, Yuseong-gu, Daejeon 305 806, South Korea*

Received 17 February 2014; revised 22 March 2014; accepted 23 March 2014
Available online 31 March 2014

KEYWORDS

DNA barcoding;
Molecular markers;
Species identification;
Plant taxonomy;
Biodiversity;
Conservation genetics

Abstract The discipline taxonomy (the science of naming and classifying organisms, the original bioinformatics and a basis for all biology) is fundamentally important in ensuring the quality of life of future human generation on the earth; yet over the past few decades, the teaching and research funding in taxonomy have declined because of its classical way of practice which lead the discipline many a times to a subject of opinion, and this ultimately gave birth to several problems and challenges, and therefore the taxonomist became an endangered race in the era of genomics. Now taxonomy suddenly became fashionable again due to revolutionary approaches in taxonomy called DNA barcoding (a novel technology to provide rapid, accurate, and automated species identifications using short orthologous DNA sequences). In DNA barcoding, complete data set can be obtained from a single specimen irrespective to morphological or life stage characters. The core idea of DNA barcoding is based on the fact that the highly conserved stretches of DNA, either coding or non coding regions, vary at very minor degree during the evolution within the species. Sequences suggested to be useful in DNA barcoding include cytoplasmic mitochondrial DNA (*e.g. cox*1) and chloroplast DNA (*e.g. rbcL*, *trnL-F*, *matK*, *ndhF*, and *atpB rbcL*), and nuclear DNA (ITS, and house keeping genes *e.g. gapdh*). The plant DNA barcoding is now transitioning the epitome of species identification; and thus, ultimately helping in the molecularization of taxonomy, a need

* Corresponding author. Tel.: +966 11 46 75734.
E-mail addresses: alimohammad@ksu.edu.sa (M. Ajmal Ali), Gyulai.Gabor@mkk.szie.hu (G. Gyulai), Hidvegi.Norbert@mkk.szie.hu (N. Hidvégi), Kerti.Balasz@mkk.szie.hu (B. Kerti), fhemaid@ksu.edu.sa (F.M.A. Al Hemaid), arunpandey79@gmail.com (A.K. Pandey), joongku@kribb.re.kr (J. Lee).

Peer review under responsibility of King Saud University.

ELSEVIER Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.sjbs.2014.03.003

of the hour. The 'DNA barcodes' show promise in providing a practical, standardized, species-level identification tool that can be used for biodiversity assessment, life history and ecological studies, forensic analysis, and many more.

Contents

1. Introduction

There are approximately 1.7 million species identified by using morphological (*i.e.* Linnean) characters including 808 gymnosperm, and 90,000 monocots and about 200,000 dicots of angiosperm. This number may be a gross under-estimate of the true biological diversity of Earth (Blaxter, 2003; Wilson, 2003). Recently, overwhelming landmark publications (Table 1) on DNA barcoding (Hebert et al., 2003) (syn.: profiling, genotyping) based on highly conserved sequence informations provide new tools for systematics (Hebert and Barrett, 2005) and phylogeny (Wyman et al., 2004; Leebens-Mack et al., 2005; Jansen et al., 2006; Hansen et al., 2006). DNA barcodes consist of short sequences of DNA between 400 and 800 base pairs that can be routinely amplified by PCR (polymerase chain reaction) and sequenced of the species studied.

Morphologically distinguishable taxa may not require barcoding; however, subspecies (*ssp.*), cultivars (*cv.*), eco- and morphotypes, mutants, species complex and clones can be diagnosed with molecular barcoding. Barcode of a specimen can be compared with sequences derived from other taxa, and in the case of dissimilarities species identity can be determined by molecular phylogenetic analyses based on MOTU, molecular operational taxonomic units (Floyd et al., 2002).

DNA barcoding was particularly useful for marine organisms (Shander and Willassen, 2005), including fishes (Mason, 2003; Ward et al., 2005); soil meiofauna (Blaxter et al., 2004) and freshwater meiobenthos (Markmann and Tautz, 2005); and extinct birds (Lambert et al., 2005). In the rainforests, rapid DNA-based entomological inventories were so effective (Monaghan et al., 2005; Smith et al., 2005) that tropical ecologists were the most active advocates of DNA barcoding (Janzen, 2004). More pragmatically, DNA barcodes have proved to be useful in biosecurity, *e.g.* for surveillance of disease vectors (Besansky et al., 2003) and invasive insects (Armstrong and Ball, 2005), as well as for law enforcement and primatology (Lorenz et al., 2005).

Barcoding has created some controversy in the taxonomy community (Mallet and Willmott, 2003; Lipscomb et al., 2003; Seberg et al., 2003; DeSalle et al., 2005; Lee, 2004; Ebach and Holdrege, 2005; Will et al., 2005). Traditional taxonomists use multiple morphological traits to delineate species. Today, such traits are increasingly being supplemented with DNA-based information. In contrast, the DNA barcoding identification system is based on what is in essence a single complex character (a portion of one gene, comprising ~650 bp from the first half of the mitochondrial cytochrome c oxidase subunit I gene sometimes called COXI or COI), and barcoding results are therefore seen as being unreliable and prone to errors in identification (Dasmahapatra and Mallet, 2006). Although the mitochondrial cytochrome oxidase subunit I (CO1) is a widely used barcode in a range of animal groups (Hebert et al., 2003), this locus is unsuitable for use in plants due to its low mutation rate (Kress et al., 2005; Cowen et al., 2006; Fazekas et al., 2008). In addition, complex evolutionary processes, such as hybridization and polyploidy, are common in plants, making species boundaries difficult to define (Rieseberg et al., 2006; Fazekas et al., 2009).

The number and identity of DNA sequences that should be used for barcoding is a matter of debate (Pennisi, 2007; Ledford, 2008). The main DNA barcoding bodies and resources are (1) Consortium for the Barcode of Life (CBOL) http://www.