C. R. Biologies xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Comptes Rendus Biologies

www.sciencedirect.com

Taxonomy / Taxonomie

Effectiveness of the DNA barcoding approach for closely related conifers discrimination: A case study of the *Pinus mugo* complex

Konrad Celiński *, Hanna Kijak, Aleksandra Wojnicka-Półtorak, Katarzyna Buczkowska-Chmielewska, Joanna Sokołowska, Ewa Chudzińska

Adam Mickiewicz University in Poznań, Faculty of Biology, Institute of Experimental Biology, Department of Genetics, 89, Umultowska, 61-614 Poznań, Poland

ARTICLE INFO

Article history:
Received 24 April 2017
Accepted after revision 9 June 2017
Available online xxx

Keywords:
The *Pinus mugo* complex
Chloroplast DNA
Barcoding
Taxonomy
Species identification

ABSTRACT

DNA barcoding is a standard and efficient method, frequently used for identification, discrimination and discovery of new species. Although this approach is very useful for classifying the world's biodiversity, little is known about its usefulness in barcoding at lower taxonomic level and its discrimination rate for closely related species, like conifers. In this study, we compared the genetic variation of eight chloroplast DNA barcode regions (*matK, rbcL, trnH-psbA, trnL-trnF, rpl20-rps18, trnV, ycf1, ycf2*) in 17 conifers - three closely related pines from *Pinus mugo* complex and 14 more distant conifers representing two genera and four sections of the Pinaceae family. The discrimination rate for a single and for multiple DNA barcode regions analyzed in this study was estimated using the Tree-Building and PWG-Distance methods. The usefulness of the DNA barcoding approach for analyzing and resolving taxonomic inconsistency among closely related and more phylogenetically distant conifers was evaluated and discussed.

1. Introduction

Unambiguous taxonomic identification of biological objects is fundamental in any biological investigation. Inappropriate or uncertain description of a research object may not only lead to false results and inadequate conclusions, but also introduce serious misunderstandings in biological databases. These errors can be further replicated by the other researchers, causing even more discrepancies.

DNA barcoding is a standard, rapid and efficient method for identifying, discriminating, and discovering new species. This method may rely on a single or a combination of multiple DNA regions in the genome of the object under study [1,2]. Undoubtedly, this approach is convenient and very useful in identifying and classifying the world's biodiversity, in species delimitation, in food authenticity testing, in monitoring the illegal trade of wildlife or forensic investigations [3,4], as demonstrated in a number of previously conducted studies on plants [5,6] and animals [7,8].

Currently, much attention is paid to refining the DNA barcoding technique by searching for the most universal or most variable DNA barcode regions, discovering and characterizing new candidate regions, and determining the factors possibly affecting the final discrimination rate of the method. Successful application of DNA barcoding strongly depends on a selected region, the type of organism under study, phylogenetic distance and richness of the species in clades, but also on the size and completeness of the DNA barcode databases [9]. Moreover, in plant species, the discrimination rate depends on their lineages and equals 70% for angiosperms [10], and only 32% for gymnosperms

* Corresponding author.
E-mail address: celinski@amu.edu.pl (K. Celiński).

http://dx.doi.org/10.1016/j.crvi.2017.06.002
1631-0691/

[2]. DNA barcoding of gymnosperms is particularly challenging due to incomplete reproductive isolation among species, predominantly paternal inheritance of the chloroplast genome, hybridization and introgression, recent radiation and slow rate of molecular variation [11].

The vast majority of plant taxonomic investigations conducted with the DNA barcoding approach (including those mentioned above) were conducted on phylogenetically distant angiosperms, for which the species discrimination rate is relatively high. To date, little is known about the effectiveness of this method for discriminating closely related species (like conifers) and about the usefulness of DNA barcoding for investigations at the lower taxonomic level.

Usually, closely related conifers with similar morphology and lack of evident species determinants are grouped into larger units called complexes. The most well-known and intensively studied for years are the *Pinus halepensis* complex [12], the *Pinus contorta-Pinus banksiana* complex [13], the *Pinus kesiya* complex [14], and the *Pinus mugo* complex [15]. Research within such aggregates is particularly difficult, mainly due to complex taxonomic problems, with determination of origin, rank, presence of many synonymous names, similar ecological niches, as well as the occurrence of hybridization phenomena, combined with the presence of cross-species mixes. Species discrimination, delimitation or identification in such complexes is both scientifically interesting, but also practically very important with a view to the protection of endangered species, encompassed by such complexes.

The *Pinus mugo* complex comprises 16 species, 91 varieties, and 19 other forms [15]. This large and polymorphic complex of closely related pines natively occurs in the main European mountains, including the Pyrenees, the Alps, and the Carpathians [16]. Most of researchers agree that the *Pinus mugo* complex is comprised of three major components, i.e. *Pinus mugo* Turra (dwarf mountain pine), *Pinus uncinata* Rammond (moutain pine), and *Pinus uliginosa* Neumann (peat bog pine). Taxonomically, these three pines are considered to be either three independent species or subspecies inside *Pinus mugo sensu lato* [15–17].

Pinus mugo is a medium-sized shrub characterized by long and curved branches with symmetrical female cones [18,19], while *P. uncinata* is a tree reaching up to 12–20 m in height, with straight trunk and strongly asymmetrical female cones [20]. *Pinus uliginosa* is the most morphologically variable taxa out of the three pines under study, as it can occur as either a medium-size shrub (still usually much higher than *P. mugo*) or a mono- oligo- or polycormic (multi-trunk) tree with asymmetrical female cones [16,21].

In general, *P. mugo* is distributed in the eastern part of the Alps and in the Carpathian Mountains, while *P. uncinata* is commonly found in the western part of the Alps and in the Pyrenees. In turn, *P. uliginosa* forms several isolated and small populations in lowland peat bogs in Central Europe [16]. Moreover, in some unique stands, i.e. the Zieleniec reserve (the Sudety Mountains, southwestern Poland), several taxa from the *Pinus mugo* complex occur. Within such sympatric populations or contact zone due to the overlapping of phenological phases of the different pine taxa, the natural gene flow among them is observed as resulting in the generation of hybrid individuals [22–26].

The reciprocal genetic relationship among *P. mugo*, *P. uncinata* and *P. uliginosa* was extensively investigated using serological methods [27], allozymes [28,29] or, more recently, RAPD markers [20], molecular cytogenetics and flow cytometry [30] approaches. The obtained data showed a conserved genome organization and an absence of distinct genetic differentiation among them. Moreover, as recently demonstrated, these three pines share the same haplotypes of chloroplast and mitochondrial DNA [31,32]. Additionally, comparative study on needles and cones characteristics implied a lack of any powerful morphological marker for the straightforward identification of *P. mugo*, *P. uncinata* and *P. uliginosa* [33]. On the other hand, recent chemotaxonomic studies revealed some differences in composition of essential oils extracted from needles of *P. uncinata* and *P. uliginosa* [34]. Furthermore, based on mass spectrometry-assisted volatile compounds analysis, species-specific chemotaxonomic markers were proposed for these three pines [35].

Despite many investigations carried out and a number of various techniques used, no species-specific DNA markers have been developed for the three pines. Thus, the origin, species distinctiveness and relationship among *P. mugo*, *P. uncinata* and *P. uliginosa*, as well as their taxonomic status within *Pinus mugo* complex are still enigmatic and require further investigation to reach consensus. For these reasons, we decided to apply a DNA barcoding approach to get a further insight into this scientific problem. The current study covers:

- sequencing eight chloroplast DNA barcode loci for *P. mugo*, *P. uncinata* and *P. uliginosa* individuals to enrich the barcodes database (deposited in GenBank);
- evaluating a genetic variation level for these eight DNA barcodes among the three closely related pines in comparison to an external group of fourteen, more distant conifer species;
- determining a species discrimination rate in *Pinus mugo* complex for a single and multiple DNA barcode regions;
- gaining more insight into taxonomic relationships within the *Pinus mugo* complex;
- assessing if any of the analyzed chloroplast DNA regions could be useful in resolving similar taxonomic ambiguities in the other pine complexes.

Consequently, we were able to verify the hypothesis about genetic distinctiveness of *P. mugo*, *P. uncinata* and *P. uliginosa*, and answer the question about the relevance of the DNA barcoding approach for discriminating closely related conifer species. The answer to the former question is particularly important – not only from a purely taxonomic point of view, but mainly in the context of the protection of those endangered taxa.

2. Material and methods

2.1. Taxon sampling

In this study, we analyzed seventeen different conifer taxa from the Pinaceae family, i.e. 14 species belonging to the *Pinus* genus and three species from the *Picea* genus

Please cite this article in press as: K. Celiński, et al., Effectiveness of the DNA barcoding approach for closely related conifers discrimination: A case study of the *Pinus mugo* complex, C. R. Biologies (2017), http://dx.doi.org/10.1016/j.crvi.2017.06.002