Prospects of barcoding the Italian wild dendroflora: lights and shadows

Piredda R[1,2], Attimonelli M[2], Bellarosa R[1], Schirone B[1], Simeone MC[1,2]

Motivation

DNA barcodes may be particularly relevant to impact the fundamental crisis facing biodiversity, as a standardized, species-level identification tool for biodiversity assessment, life history and ecological studies. Trees play important roles in the conservation of numerous land ecosystems, in the wood trade, and in the definition of bio-geographic processes; nevertheless, peculiar biological, evolutionary and taxonomical features are likely to constitute an intriguing challenge to barcoders. The objectives of this work are: to test universal application of barcoding in a set of tree species among the most representative of the forest ecosystems in a core area of the Mediterranean region, and to investigate the main potential drawbacks of barcoding certain taxonomic groups which might prove exceptionally changeling due to hybridization, phenotypic similarity, recent speciation, low plastid mutation rate. This study involves both a floristic approach [1] to estimate the advantage to use four plastid regions (trnH-psbA, rbcL, rpoC1, matK) in 52 tree taxa of Italian woodland. In the second part, we perform a taxon-based approach [1] to deepen barcoding applicability to Quercus, an emblematic study case for its complex evolutionary issues, with a taxonomy still difficult to be assessed .

Methods

Plant material - The sampling design included taxa spanning from the alpine forest, to the mountain, the temperate, the riparian, and the Mediterranean evergreen forest/maquis. Collectively, 50 native Italian and 2 naturalised North American woody species (24 genera, 11 plant families one individual per species) were investigated. Three families resulted thoroughly sampled due to their abundancy: Oleaceae, Aceraceae and Fagaceae. Moreover, we performed a taxon-based investigation focused on Quercus involving 30 oak individuals (three individuals per nine species and three rare endemic taxa) sampled all over Italy. DNA was extracted, amplified and sequenced according to standard CBOL protocols. The obtained sequences have been analysed as follows: in the floristic dataset, haplotypes were defind with BLASTClust (ftp://ftp.ncbi.nlm.nih.gov/blast/); BLASTClust is an identification method that not require multiple sequence alignment. It is used to cluster either nucleotide or protein sequences and accepts a number of parameters that can be modified to control the stringency of clustering.

[1] Dipartimento Ambiente e Foreste (D.A.F.), University of Tuscia, 01100 Viterbo, Italy [2] Dipartimento di Biochimica e Biologia Molecolare “E. Quagliariello”, University of Bari, 70125 Bari, Italy

We used parameters allowing to cluster together only identical sequences; in the taxon-based dataset, sequences were aligned with MUSCLE and edited with Seaview. Haplotype lists and networks were generated with DnaSP. Inter- and intraspecific divergences were calculated with K2P nucleotide evolution model using MEGA. The same package was used to produce NJ dendrograms reporting bootstrap values too.

Results

The Mediterranean basin constitutes one of the world's major biodiversity hotspots. In the floristic dataset matK did not amplify in 44% of the investigated taxa and was then discarded. The 3 markers trnh-psbA, rbcL and rpoC1 were tested either single or in different combinations. Results indicate that trnH-psbA resolves a much higher percentage of samples (61%) than 2 coding regions rbcL (13%) or rpoC1 (27%). rpoC1 performed very differently in taxa, indeed in Fraxinus rpoC1 exhibited good discrimination power (3/3), whereas failed at genus-level in Castanea. Overall there were 14 species in our dataset that could not be resolved. Of these, 10 belong to genus Quercus and if we exclude Quercus spp. from the total dataset, the percentage of taxa showing single haplotypes reaches 17% with rbcL, 43% with rpoC1 and 82% with trnH-psbA. Multiregion combinations did not identify further haplotypes and no percentages of species discrimination higher than trnH-psbA alone were provided. We assume that the investigated species set conforms well with the taxonomic diversity of Italian dendroflora. In Quercus taxon-based investigation discrimination species matched only the acknowledged Quercus subgeneric classifications (i.e. Subgenus Quercus, Subgen. Cerris and Subgen. Sclerophyllodrys). Inside subgenus certain alleles in one species appeared to be more closely related to alleles from different species than to other conspecific alleles [2]. Such polyphyletic pattern is present in all markers, and it could be due to limited DNA variation at the chosen marker, polyploidy, hybridization, reticulate evolution and shared ancestral polymorphisms, not excluding imperfect species definitions and taxonomies. We consider satisfactory for future applications assigning barcodes to related groups of species. However, as suggested by CBOL-Plant Working Group, the rapidly evolving internal transcribed spacer region of the ribosomal DNA (nrITS) may represent a useful supplementary barcode in difficult genera.

Contact e-mail

m.attimonelli@biologia.uniba.it

Supplementary information

References

[1] Hollingsworth ML, et al. (2009). Mol. Ecol. Res. 9, 439-457 [2] Funk D.J. et al. (2003. Annu. Rev. Ecol. Evol. Syst. 34, 397-423