Caryologia
International Journal of Cytology, Cytosystematics and Cytogenetics

Firenze University Press
www.fupress.com/caryologia

OPEN ACCESS

Citation: A. Giovino, F. Martinelli, A. Perrone (2020) The technique of Plant DNA Barcoding: potential application in floriculture. *Caryologia* 73(2): 27-37. doi: 10.13128/caryologia-730

Received: December 12, 2019

Accepted: April 1, 2020

Published: July 31, 2020

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

The technique of Plant DNA Barcoding: potential application in floriculture

Antonio Giovino[1,*], Federico Martinelli[2,*], Anna Perrone[3]

[1] *Council for Agricultural Research and Economics (CREA), Research Centre for Plant Protection and Certification (CREA-DC), Bagheria, Italy*
[2] *Department of Biology, University of Florence, Sesto Fiorentino, Florence, 50019, Italy*
[3] *Department of Biological, Chemical and Pharmaceutical Sciences and Technologies University of Palermo, Viale delle Scienze, Palermo, 90128, Italy*
*Correspondonding authors. E-mail: antonio.giovino@crea.gov.it; federico.martinelli@unifi.it

Abstract. The objective of this work was to assess the ability of the DNA barcoding approach to identify different taxonomic groups from two flowering plant collections: 1) the most relevant commercial taxa (nursery production) and 2) Mediterranean plants with ornamental attitude (new emerging species). "Core markers", *rbc*L and *mat*K, were adoptedthe identification step of 100 taxa belonging to 20 families. A third marker, the intergenic spacer *trn*H-*psb*A, was also tested, on 74 taxa, when the core markers were not able to discriminate well the analysed germplasm.DNA barcode fragments were recovered for all the total taxa investigated (100%). The rbcL showed the best performances: the greatest amplification success, the best sequencing performance both in terms of the number of sequences obtained and in terms of quality of the sequences obtained. Despite having recorded greater amplification difficulties, according to numerous other studies, matK has shown a good success in sequencing and quality of the obtained sequences (de Vere *et al.* 2012), unlike what is indicated in some protocols that suggests for this region the need for further primers to be adopted for the sequencing phase (Hollingsworth et. al 2011). Results showed that sixty-one taxa overall (61%) were totally resolved at specific or subspecific level, by at least one of the three markers. The *mat*K and *rbc*L locus respectively resolved 44% and 35% of the taxa. The core markers in *multilocus* approach led to the discrimination of a total of 49% taxa. The *trn*H-*psb*A was able to discriminate 52% of taxa analysed and resulting determinant in the discrimination of 14 taxa. Four families, including the major number of taxa (*Arecaeae, Fabaceae, Euphorbiaceae, Asteraceae*), were evaluated in terms of genetic distance (K2P% value). This work highlighted the potential of the barcoding approach for a rapid identification of plant species in order to solve taxonomic disputes and support commercial traceability of floreal products.

Keywords: DNA barcoding, DNA fingerprinting, floriculture,genetic identification.

1 INTRODUCTION

Genetic certification of plant material is, today more than ever, a fundamental requirement to increase the competitiveness of plant nurseries, even

in the ornamental sector. This represents a unique and effective tool for unambiguous determination of nature of plant species. These improvements will enhancethe floriculture sector through the successful obtainment of the following to objectives: 1) genetic identification (especially for native plants) and particularly to identify the link between genetic resources of ornamental interest and the relative territory of origin, thus promoting the products in harmony with the territory and with sustainability criteria, 2) traceability (native plants and imported plants) through the characterization of autochthonous products, plant material of supply chain and non-native incoming material. Therefore, the genetic certification of plant material is an extremely important aspect for the resolution of related problems:1) taxonomic controversies (synonymy/homonymy) and species of difficult identification, 2) newly introduced programs, genetic improvement, 3) early identification of species with very long phenological cycles, 4)correspondence checks of vegetable species entering the markets, 5) protection of biodiversity, of native or endangered species. Recently, DNA barcoding has emergedas a new molecular tool for taxonomists (Hebert, Ratnasingham & deWaard, 2003). A DNA barcodeis a universally accepted short DNA sequence normally employed for the identification of species (Savolainen *et al.*, 2005), promoted for a variety of biological applications (Hollingsworth, Graham & Little, 2011), including the identification of cryptic species, species discovery (Bickford *et al.*, 2007) and taxonomic revisions (Simeone et al 2013). The genotype is nothing but the set of all the genes that make up the DNA of an organism. Thus DNA-based taxonomy has proved to be a valuable support to the classical taxonomy allowing to face the growing need for accurate and accessible taxonomic information (Tautz *et al.*, 2003). In particular, the advent of molecular markers has marked a remarkable turning point in the world of plant genetics allowing the construction of association genetic maps and the identification of genes responsible for agronomic characters (Giovino *et al.* 2015a). In taxonomic studies, markers are important for botanical classifications and the analysis of phylogenetic relationships (Varshney *et al.*, 2005). Among the molecular techniques, a new approach to the study of biodiversity has become widespread, with all the problems related to this study: the DNA barcoding, literally "DNA barcode". The name of this approach refers to the identification method by which a scanner distinguishes various commercial products using linear bar codes or "UPC" (Universal Product Code).This molecular investigation approach was first proposed to the scientific community in 2003 by the population geneticist Paul Hebert of the University of Guelph (Canada) (Hebert et al 2003). In this work it was used for the identification of species, a gene sequence located in the region of the mitochondrial gene COI, coding for the subunit I of the cytochrome-c oxidase (also known as Warburg's respiratory fragment), therefore the variability of a molecular marker for the identification of biological identities is exploited.Over the years, COI has been successfully used in various animal taxa, including birds (Hebert *et al.*, 2004b), arthropods (Barrett and Hebert, 2005), fish (Ward *et al.*, 2005) and Lepidoptera (Hebert *et al.*, 2004a). In vegetables, COI has not proved to be an excellent marker for phylogenetic studies, due to the low evolutionary rate of the mitochondrial genome. In order to overcome this problem, other markers for DNA barcoding of plants have been identified in recent years. These are DNA sequences present in some sections of the chloroplast genome, such as the trnH-psbA intergenic region, the matK gene or the rbcL gene, which have characteristics similar tocoxI useful for species identification. There are several requirements for a marker to be considered appropriate for DNA barcoding. First of all it is advisable that the marker has a wide taxonomic coverage (also called universality), which would allow the applicability of the gene chosen as barcode markern to the largest possible number of taxa and have a high success rate of PCR and sequencing. A high resolution capacity of the gene is also important, i.e the ability of a given barcode to differentiate species. This is typically based on the amount of interspecific differences between DNA sequences (Polymorphism). Another fundamental assumption is that the molecular marker chosen as a barcode should show a higher interspecific variability than intraspecific variability. Inter- and intra-specific variability are separated by a certain distance (discontinuity between intra and interspecific variability) called "barcoding gap" (Meyer and Paulay, 2005).The ideal marker therefore consists of a highly variable region, which provides for species discrimination, flanked by highly conserved regions for which adequate primers can be designed (Saunders and Kucera, 2010). Therefore, for the plants, the Barcoding protocols refer to the indications of the Plant working Group, which suggests the use of a multi-locus approach (Hollingsworth *et al.*, 2011; Domina *et al.* 2017). The general objective of the research was to use the technique of DNA barcoding to help nursery production thanks to the easy identification of new products, ornamental plants and ornamental-food valueto respond to new and growing market needs.

2 MATERIALS AND METHODS

2.1 Plant collection

Native Sicilian plant species of high ornamental value or dual aptitude for new introduction were selected, collected and morphologically analyzed. Selection was also extended to autochthonous or exotic species already produced at Faro srl (Catania, Italy) in order to gain more insights into: 1) taxonomic controversies (synonymy / homonymy); 2) early identification of species with very long phenological cycles; 3) correct identification of species with significant commercial impact. Samples collection for DNA analysis includes 100 plant species. (Tab. 1). We have included 52 species commercialized by Faro srl in addition to 36 native species that were present in the collection at the CREA-DC (Bagheria, Italy). For all the selected species, a bank of freeze-dried plant material and the respective DNA bank was set up at the CREA-DC of Bagheria for long-term conservation stock. Before proceeding with the application of the molecular characterization protocols, it was necessary to carry out a preliminary characterization at a morphological level,

The plant material under study is represented by 3 replicates for each species (or three distinct plants for each species), specifically from each of them tissue samples were taken, represented by young leaves, which constitute the plant material from which to proceed with DNA extraction . Every single sample was cataloged with an identification code (ID) in order to set up a real germplasm collection, as well as for the establishment of a bank of germplasm DNA (freeze-dried).

For all the selected species, a bank of freeze-dried plant material and the respective DNA bank was made at CREA in Bagheria, as an important stock for the conservation of the plant material in question.

2.2 Molecular analysis

For the molecular identification of the plants, young leaves, previously subjected to lyophilization, were used as starting material for DNA extraction.DNA was extracted from three biological replicates (lyophilized) for each taxonomic entity using CTAB-related method (Doyle & Doyle, 1987). Amplification and sequencing protocols of three regions of DNA usingrbcL, matK and trnH-psbA were performed,as defined by the Consortium for the Barcode of Life (CBOL). Firstly, these plastid portions, named “core markers”, were used for genetic characterization. For those species in which the core markers were unsuccessful, a third marker was tested based on the trnH-psbA intergenic region. This portion is in fact known to support a greater degree of discrimination between related species. A pipeline of the genetic characterization analysis is shown in Figure 1. Sequences of the rbcL, matK and trnH-psbA primers used in the PCR amplification were the following:

- rbcL-F: ATGTCACCACAAACAGAGACTAAAGC
- rbcL-R: GTAAAATCAAGTCCACCRCG
- matK-3F KIM: CGTACAGTACTTTTGTGTTTACGAG
- 4) matK-1R KIM: ACCCAGTCCATCTGGAAATCTTGGTTC
- 5) trnHf_05: CGCGCATGGTGGATTCACAATCC
- 6) psbA3_f: GTTATGCATGAACGTAATGCTC

In relation to the PCR conditions, the protocol suggested by the CBOL Plant Working Group (Hollingsworth *et al.*, 2009) was followed, and the amplifications were conducted with a Gene®Amp PCR System 9700 thermocycler (Applied Biosystems).The amplicons were run on 2% agarose gel, whose purpose is to ensure the successful amplification of the segments of DNA involved, using the barcode primers used. The gels were analysed using the image acquisition “Gel Doc” of BIO-RAD, which allows to use a special “Quantity One” software, to identify amplified DNA bands.

2.3 Data analysis

The PCR products were purified and sequenced following the DYEnamic™ ET termination kit sequencing kit (Amersham Biosciences) using an automatic sequencer AB3730XL DNA Analyzer (Applied Biosystems). The fragments were sequenced both forward and in reverse, using the same primers adopted for PCR.Through Sequencer software 4.10 (Gene Codes Corporation, USA) the electropherograms were carefully checked and eventually cleaned manually,and assembled in contigs. The obtained sequences were blasted and aligned using MUSCLE software, implemented within Mega 6 program (Tamura *et al.* 2013) used for phylogenetic analysis.

Several parameters have been evaluated to be able to efficiently determine the real discriminating power of the Barcoding markers used. Two categories of parameters were taken into account: 1) thoserelated to technical performances and those useful for assessing the discriminated power. The number of PCR positive samples for each marker was calculated, both for the total number of biological replicates and number of taxa analyzed. Dealing with sequencing success, the number of samples positive for the sequencing procedure was calculated, which concerned only the PCR-positive samples for each marker, both in relation to the total number of biological replicates and to the number of taxa. Quality of the

Table 1. Species selected for molecular investigations.

Famiglia	Specie
Acanthaceae	*Acanthus mollis* L.
Arecacea	*Acoelorraphe wrightii* H. Wendl. ex Becc.
	Arengaengleri Becc.
	Caryota urens L.
	Chamaerops humilis var. *humilis/ Chamaerops humilis L.*
	Chamaerops humilis var. *argentea* André
	Chamaeropshumilis L. "Vulcano"
	Chamaeropshumilis L. "Etna star"
	Howeaforsteriana (F. Muell.) Becc.
	Livistonachinensis (Jacq.) R.Br. ex Mart.
	Phoenix canariensis Chabaud
	Phoenix dactylifera L.
	Phoenix reclinata Jacq.
	Phoenix roebelenii O'Brien
	Sabal minor (Jacq.) Pers.
	Sabal palmetto (Walter) Lodd. ex Schult. & Schult.f.
	Trachycarpus fortune (Hook.) H. Wendl.
	Washingtonia robusta H. Wendl.
	Washingtonia filifera (Linden ex André) H. Wendl. ex de Bary
	Butia capitata (Mart.) Beccari
	Bismarckia nobilis Hildebr. & H. Wendl.
	Brahea armata S. Watson
	Brahea edulis H.Wendl. ex S.Watson
	Trithrinax campestris (Burmeist.) Drude&Griseb.
	Arecastrum romanzoffianum (Cham.) Becc. *Syagrus romanzoffiana* (Cham.)
Xanthorrhoeaceae	*Aloe arborescens* Mill.
	Aloe vera (L.) Burm.f.
	Aloe plicatilis (L.) Mill..
	Aloe × *spinosissima* Jahand.
Fabaceae	*Spartium junceum* L.
	Ceratonia siliqua L
	Genista madoniensis Raimondo
	Genista demarcoi Brullo, Scelsi & Siracusa
	*Genista tyrrhena*Vals.
	Genista cupanii Guss.
	Genista aetnensis (Biv.) DC.
	Genista aristata C.Presl
Cistaceae	*Cistus albidus* L
	Cistus salvifolius L.
	Cistus x pulverulentus Pourr.
	Cistus × *skanbergii* Lojac.
Cycadaceae	*Cycascircinalis* L.
	*Cycas revoluta*Thunb.
Myrtaceae	*Myrtus luma* Molina
	Metrosideros excelsa Sol. ex Gaertn.
	Myrtus communis L.
Lamiaceae	*Rosmarinus officinalis* L.
	Salvia leucantha Cav.
	Lavandula angustifolia Mill.
	Lavandula stoechas L.
	Sideritis italica (Mill.) Greuter&Burdet
	Salvia officinalis L.
Ericaceae	*Arbutus unedo* L.
	Erica siculaGuss.
	Erica peduncularis C.Presl
	Erica multiflora L.
Asteraceae	*Helichrysum italicum* (Roth) G. Don
	Helichrysum hyblaeum Brullo
	Helichrysum nebrodense Heldr.
	Helichrysum scandens Guss.
	Anthemis cupaniana Tod. ex Nyman
	Centaurea sphaerocephala L.
	Jacobaea gibbosa (Guss.) B.Nord. &Greuter
	Pallenis maritime (L.) Greuter
	Ptilostemon greuteri Raimondo & Domina
	Senecio candidus (Presl.) DC. /*Jacobaea candida* (C.Presl) B.Nord. & Greuter
	Jacobaea ambigua(Biv.) Pelser&Veldkamp
	Anthemis maritima L.
	Hieracium cophanense Lojac.
Iridaceae	*Iris pseudopumila* Tineo
	Iris germanica L.
Strelitziaceae	*Strelitzia augusta* Thunb
	Strelitzia Nicolai Regel&K.Koch
	Strelitzia reginae Banks
Tamaricaceae	*Tamarix gallica* L.
Convolvulaceae	*Calystegia soldanella* (L.) R. Br.
Amaranthaceae	*Diotis maritima* (L.) Desf. ex Cass./*Achillea maritima* (L.) Ehrend. &YPGuo
Liliaceae	*Tulipa radii* Reboul
Brassicaceae	*Brassica insularis* Moris
	Brassica villosa subsp. *tinei* (Lojac.) Raimondo & Mazzola
	Brassica rupestris subsp. *hispida* Raimondo & Mazzola
Rosaceae	*Rosa sicula* Tratt.
	Rosa sempervirens L.
	Rosa canina L.
	Rosa corymbifera Borkh.
Caryophyllaceae	*Dianthus busambrae* Soldano & F. Conti
	Dianthus rupicola subsp. *aeolicus* (Lojac.) Brullo&Miniss.
	Dianthus rupicola Biv. subsp. *rupicola*
	Dianthus rupicola subsp. *lopadusanum* Brullo & Miniss.
	Dianthus siculus C. Presl

Famiglia	Specie
	Dianthus rupicola subsp. *hermaensis* (Coss.) O. Bolòs& Vigo
Euphorbiaceae	*Euphorbia ceratocarpa* Ten.
	Euphorbia characias L.
	Euphorbia dendroides L.
	Euphorbia meuselii Geltman
	Euphorbia myrsinites L.
	Euphorbia helioscopia L.
	Euphorbia bivonae Steud.
	Euphorbia pithyusa subsp. *cupanii* (Guss. ex Bertol.) Radcl.-Sm.
	Euphorbia amygdaloides L.

sequence was given by the quality of the peaks present on the electropherograms to indicate the precision and reliability of the sequences obtained. Sequences with quality over 70% were considered suitable. The reported value indicates the average of biological replicates. Fragment length was determined and referred to the average length of the fragments obtained for each marker, in relation to the total of biological replicates, following the analysis and cleaning of the electropherograms. The value of thepower of discrimination parameter was given by the number of taxa that have been univocally discriminated on the level of species (or subspecies).The discriminating power was assessed both for single locus and in multi-locus approach. The discrimination power of each locus was evaluated by phylogenetic analysis with Mega6, conducted by comparing all the sequences generated in this study and using a subset of referring sequences related to each taxa found by BOLD Database / GenBank. The level of genetic divergence was determined and indicated the degree of variability between a group of sequences, obtained from the distance matrices calculated according to the parameter K2P% (Kimura, 1980). It was calculated within some families considered most representative by number of species. Number of variable sites was determined. It indicated the number of bases subject to variations within the gel phylogenetic group considered on the total length of the fragments obtained for each locus. Like the previous one, it was calculated within some families considered most representative of the entire collection of analyzedplant species.

3 RESULTS AND DISCUSSION

Results of discrimination outputs for each of the three markers are reported in Tab. 2. Using a total of

Table 2. Technical performancesof markers used in DNA barcoding techniques referred to the total of biological replicates (a) and tested taxa (b).

(a)

	rbcL	matK	trnH-psbA
Number of tested samples*	300	300	222
Successful amplification	(93%) 279/300	(70%) 210/300	(80%) 177/222
Successful sequencing (contigs)	(95%) 265/279	(93%) 195/210	(91%) 161/177
High quality sequence (contigs)	90%	80%	85%
Fragment length (average in bp)	569	766	518

(b)

	rbcL	matK	trnH-psbA
Number of tested samples*	100	100	74
Successful amplification	(97%) 97/100	(81%) 81/100	(89%) 66/74
Number of taxa successfully sequenced	(99%) 96/97	(96%) 78/81	(94%) 62/66

300 samples (including biological replicates), rbcLobtained 93% PCR success, 95% sequencing success, with 90% sequence quality and an average fragment length of 569 bp. MatK showed a success of PCR and sequencing, respectively of 70% and 93% and a quality of sequences of 80% with an average length of fragments of 766 bp. The use of trnH-psbA marker showed PCR and sequencing success respectively of 80% and 91% and a sequence quality of 85% with an average fragment length of 518 bp. Consideringa total number of 100 taxa tested, rbcL showed higher values than the other two markers, with PCR success of 97% and a success of sequencing of 99%, for matK the recorded values were of 81% for successful apmification and 96% for sequencing success, while trnH-psbA marker showed respectively PCR and sequencing successof 89% and 94%. In relation to the above results, rbcL showed the best performances: the greatest amplification success, the best sequencing yield both in terms of the number of sequences obtained and in terms of the quality of the sequences obtained. The matK, despite having experienced greater amplification difficulties agreeing with numerous other studies (de Vere *et al.* 2012), it showed a good success of sequencing and good quality of obatined sequences. This does not agree withprevious works that suggest the need to use matK with additional primers for sequencing purposes (Hollingsworth et. to 2011). Taxa identification was firstly carried out using "core markers" (rbcL and matK). The use of the third marker, the IGS trnH-psbA was reserved for those situations in which both core mark-

Table 3. Discriminating power of Barcoding markers.

	risoluzione a livello di specie	tot.%	
singolo locus	rbcL*	**35%**	(34/96)
	matK*	**44%**	(34/78)
	tmH-psbA*	**52%**	(32/62)
multi-locus	rbcL + matK**	**49%**	(38/77)
	rbcL + trnH-psbA**	**53%**	(32/60)
	matK + tmH-psbA**	**54%**	(25/46)

ers presented difficulties, due to lack of amplification, failure of sequencing reactions or insufficient discriminating power. The overall identification results at species level for each tested taxawere reported in Tab. S1. Out of a total of 100 taxa tested, 61% of taxa were successfully identified at the species level with at least one of the three locus, while 37% remained at the genus level. Only the remaining 2% of the taxa remained undetermined due to the failure of all three markers employed.Considering the individual markers, rbcL allowed a unique identification at the species level of 34 taxa (35%), matK of 34 taxa (44%) and trnH-psbA of 32 taxa (52%) (Tab. 3). MatK showed greater percentage values of resolving power in terms of discrimination of taxa than rbcL, confirming the trends indicated by other studies (Chen *et al.* 2010). When rbcL and matKwere not able to discriminate species (belonging to 14 taxa), trnH-psbA was decisive in the identification of them, allowing to increase the total number of discriminated taxa from 47 to 61 taxa. The core markers, used in multi-locus, rbcL + matK, allowed the unambiguous identification at the species level of 38 taxa. Further combinations of the two markers rbcL + trnH-psbA and matK + trnH-psbA allowed the discrimination of 32 taxa and 25 taxa respectively.The use of the multi-locus approach based on core markers appeared to be the most efficient, with a good compromise between the high technical performance of the rbcL and the best resolving power supported by the matK.The following families showed the highest success rate of species discrimination: *Asteraceae* (9 uniquely discriminated taxa out of 13, *Caryophillaceae* with 4 taxa of 6, *Fabacecae* with 8 taxa out of 8, *Euphorbiaceae* with 9 taxa out of 9, *Brassicaceae* with 3 taxa out of 3, *Ericaceae* with 4 taxa out of 4. Minor successes in terms of unambiguous resolution at the species level, have been found for *Arecaceae*, (7 taxa discriminated at the species level on a total of 24), and for the *Cistaceae* (none). Levels of genetic divergence for larger families were reported in Tab. 4. The rbcL marker showed the lowest values of genetic divergence for *Arecaceae*, with 0.7%, and the highest values for *Asteraceae*, with 2.1%, while matK showed the lowest values for the *Arecaceae* with 1.5% and the highest for *Fabaceae* with 6.4%. TrnH-psbA showed the highest values for *Euphorbiaceae* with 9.1% and the lowest for *Arecaceae* with 2.9%. TrnH-psbA has confirmed high variability values and its ability to discriminate within very similar taxonomic groups (Chase *et al.* 2007).In the *Arecaceae* family, which in our study included 24 species from 15 different genera, the trnH-psbA marker recorded the highest genetic divergence value with a percentage of 2.9%. The lowest values occurred with rbcL with a percentage of 0.7%, while matK showed intermediate values compared with the first two with a value of 1.5% (Tab. 4). The rbcL was able to identify two species of *Arecaceae* (*Acoelorraphe wrightii* and *Caryota urens L.).* When rbcL failed, matK was decisive for identification of 4 taxa (*Arenga engleri Becc., Phoenix roebelenii O'Brien, Sabal minor (Jacq.) Pers., Bismarckia nobilis Hildebrandt & H.Wendl., 1881).* Other authors indicated rbcL and matK as highly decisive phylogenetic analysis of this family (Asmussen *et al.* 2006).Only in the case of *Washingtonia robusta H. Wendl.*, the discrimination was possible through the use of both core markers. Relating to *Fabaceae* (8 species investigated from 3 different genera), the lowest values of genetic divergence were recorded with rbcL with 1.5% and the highest with trnH-psbA with values of 7.4%. Using matK a genetic divergence of 6.4% was obtained, discriminating 4 species out of 8. The matK was determinant for 1 taxa (*Genista aristata C.Presl*), while the trnH-psbA was determinant for 2 taxa (*Genista tyrrhena Vals., Genista demarcoi Brullo, Scelsi & Siracusa).* Considering that Genista was the most represented genus (with 6 species), rbcL showed a better result than matK within this group, discriminating 5 species (*Spartium junceum L., Ceratonia siliqua L., Genista madoniensis Raimondo, Genista cupanii Guss.*, *Genista aetnensis Raf. ex Biv.*). This result appears to be in contrast with the potential expressed by matK within the Fabaceae in other studies (Gao et al 2011; Gao and Chen 2009). Here, the Genista group showed excellent levels of discrimination with this marker. Relating to *Asteraceae* (13 investigated species belonging to 8 different genera),matK showed values of genetic divergence of 4.4% and rbcL 2.1%. As for the trnH-psbA, given the excessive variability shown by the analyzed sequences, a subdivision into genera. The lowest genetic divergence values were recorded for *Anthemis* with 1% and higher for *Jacobaea* with 3.3%. (Tab. 4).

Relating to *Asteraceae,* rbcL has allowed us to identify at the species level 4 taxa (*Centaurea sphaerocephala L., Helichrysum nebrodense Heldr., Ptilostemon greuteri Raimondo & Domina, Pallenis maritima (L.) Greu-*

Figure 1. Flowchart summarizing steps for the genetic identification of samples using DNA Barcoding and selected markers.

Figure 2. Workflow used for the molecular characterization of the plant species usable by companies using international CBOL standards.

Figure 3. Proposed type of genetic labels for traceability of plant species at commercial level.

ter), while the matK has discriminated 5 taxa resulting in particular in the discrimination of 2 species (*Helichrysum italicum (Roth) G. Don, Hieracium cophanense Lojac.*). The trnH-psbA was determinant in the resolution of a further 3 taxa (*Jacobaea gibbosa (Guss.) Peruzzi, Jacobaea ambigua (Biv.) Pelser & Veldk., Senecio candidus (C. Presl) DC. Jacobaea gibbosa (Guss.) Peruzzi* showed a wide variability compared to the other species of the genus *Jacobaea*, departing from these in all three markers used. This highlighted the presence of different clusters within the species. For *Asteraceae*, the discrimination was rather high in agreement with other studies that indicated high levels of discrimination success (Gao et al 2010).Within family *Euphorbiaceae* (9 species investigated of a single genus) the lowest values of genetic divergence occurred with rbcL with 1.2%, the highest with trnH-psbA with 9.1% and intermediate values with matK(4%). (Tab. 4). Figures S1-S2 showed the phylogenetic relationships using the three markers for *Euphorbiaceae*. The rbcL identified 6 taxa (*Euphorbia bivonae Steud., Euphorbia ceratocarpa Ten., Euphorbia dendroides L., Euphorbia helioscopia L., Euphorbia myrsinites L., Euphorbia pithyusa subsp. Cupanii Guss.*). MatK correctly identified 3 taxa and was decisive for 1 taxa (*Euphorbia amygdaloides L.*), while the trnH-psbA was determinant for 2 taxa (*Euphorbia characias L., Euphorbia meuselii Raimondo & Mazzola*). For the genera *Brassica, Erica, Cistus, Chamaerops, Dianthus, Euphorbia and Genista*, the work of molecular identification was performed with the use of referring species found specifically for this study. This was due to the absence in the international databases of species similar to those selected in this study (Aubriot *et al.* 2013; Domina *et al.* 2017; Giovino *et al.* 2015b). Therefore, these species and their respective sequences are new will be added into international databases.

For taxa discriminated on a species level with the DNA Barcoding methodology (green colour in Figs 4 and 5), our data open the possibility of a real "identity certification" card for these plant species in order to trace their commercial products at marketing stage, in order to guarantee their unique identification and traceability, to protect both biodiversity and economic aspects of nursery productions as well as end-users.The certification and traceability system may follow a very precise path (Fig. 2; Fig. 3).This traceability can begin with the use of DNA Barcoding protocols for the identification of the species. Consequently, the realization of a label where, in addition to the generic species, it will be possible to include molecular results, translated into a barcode swhich, by scanning with special barcode scanners will immediately make it possible to have all certain species' indications.

A big issue emerged from this study was the lack of reference sequences available for species and taxa comparison. This issue has determined the impossibility of discriminating some groups such as: *Livistona chinensis Jacq, Trachycarpus fortunei Hook., Phoenix dactylifera L.; Phoenix reclinata Jacq., Trithrinax campestris Burmeist., Anthemis cupaniana Tod. ex Nyman, Butia capitata (Mart.) Becc., Senecio candidus (C. Presl) DC, Aloe arborescens Mill., Aloe plicatilis L., Iris pseudopumila Tineo, Iris germanica L., Salvia officinalis L., Cistus salvifosius L., Cistus x pulverulentus Delilei, Cistus albidus L., Cistus skanbergii, Dianthus rupicola subsp. aeolicus Lojac., Dianthus busambrae Soldano & F. Conti.* This

Table 4. Levels of genetic divergence for larger families. Genetic divergences calculated with the parameter K2P% (Kimura 1980).

Family	rbcL			matK				trnH-psbA	
	N. seq	Variable sites	GD%	N. seq	Variable sites	GD%		N. seq	Variable sites
Arecaceae	107	24/533	0,7	115	101/770	1,5		36	85/676
Fabaceae	25	32/543	1,5	21	181/815	6,4		9	67/329
Euphorbiaceae	27	39/540	1,2	14	88/769	4		18	164/736
Asteraceae	67	54/563	2,1	83	157/797	4,4	Jacobaea	14	22/416
							Helichrysum	14	44/533
							Anthemis	20	8/348

evidence demonstrated the great importance of creating molecular databases that incorporate the widest possible biodiversity with universal markers. In addition, it highlighted the importance of creating a dedicated database of the main floricultural species of ornamental interest, which can support the practical application of the molecular protocol for the purposes of traceability and monitoring by control bodies (Giovino *et al.* 2014). Although DNA Barcoding can reach 80-90% of resolution levels, it can lack sufficient discrimination power in some families, including *Ericaceae, Lamiaaceae, Orchidaceae*. This is due to the modest evolutionary distance between closely related species evolved from recent divergence, as suggested before (Hollingsworth *et al.*, 2009). Although in some cases, the multi-locus approach can have a great success, the evaluation of additional barcoding regions in relation to the success of discrimination, requires the use of individual taxonomic groups with difficult discrimination (Hollingsworth *et al.*, 2011).

In conclusions, this work confirmed the high performances of rbcL and matK markers usin a total of 100 plant taxa, belonging to 20 different families. The taxa successfully sequenced for at least one of the considered markers were 98 and 61% of the total evaluated ones at level of species or subspecies. Considering that the failure of taxa is linked to particular genus, or species, with very low evolutionary divergence, this result confirms the potential of the barcoding approach for the rapid analysis of unknown samples. Cryptic groups found in this study highlighted the already well-known technical problems due to the low level of matK amplification and sequencing success. Anyway, this marker greater power of discrimination compared to rbcL.Therefore, we can conclude that although the adoption of core markers appeared to be a good compromise, in some cases the multi-locus approach and the addition of the third trnH-psbA marker can promote greater success, as demonstrated here.

The evaluation of additional barcoding regions can be useful for increasing the success of discrimination, but thisdepends on the individual taxonomic groups showing problems of PCR amplification and sequencing with core markers. However, it is worthy to notice that a large sample of references related to eachtaxa is necessary to validate the accuracy of the method.This study highlighted the great importance of creating molecular databases incorporating the widest possible biodiversity with universal markers, developing a dedicated database, especiallyfor floricultural species with ornamental interest to enhance their traceability and monitoring of commercial exchanges by control national authorities.

DISCLOSURE STATEMENT

No fincial interests or benefits has been arisen from the application of our research.

DATA AVAILABILITY STATEMENT

We declare that all data presented here have been included in the present work and related supplemental material.

DATA DEPOSITION

Most of the analysed species were submitted in the Bold database within the project: FMED: Manager Giovino Antonio (Tab. 5).

REFERENCES

AsmussenCB, Dransfield J, DeichmannV, BarfodAS, PintaudJC, Baker WJ. 2006. A new subfamily classification of the palm family (Arecacee): Evidence from plastid DNA phylogeny. Bot J Linn Soc. 151(1):15-38.

Aubriot X, Lowry PP 2nd, Cruaud C, Couloux A, Haevermans T. 2013. DNA barcoding in a biodiversity hot spot: potential value for the identification of Malagasy Euphorbia L. listed in CITES Appendices I and II. Mol EcolResour. 13(1):57-65. doi: 10.1111/1755-0998.12028.

Barrett R, Hebert PDN. 2005. Identifying spiders through DNA barcodes. Can J Zool. 83:481–491.

Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R,Winker K, Ingram KK, Das I 2007. Crypticspeciesas awindow on diversity and conservation. Trends in Ecologyand Evolution 22: 148–155.

Chase MW, CowanRS, Hollingsworth PM, van den Berg C, Madriñán S, Petersen G, Seberg O,Jørgsensen T, Cameron KM, Carine M, Pedersen N, Hedderson TAJ, Conrad F, Salazar GA, Richardson JE, Hollingsworth ML, Barraclough TG, Kelly L, Wilkinson M. 2007. A proposal for a standard protocol to barcode all land plants. Taxon. 56(2):295–299.

Chen SL, Yao H, Han JP, Liu C, Song JY, Shi LC, Zhu YJ, Ma XY, Gao T, Pang XH, Luo K, Li Y, Li XW, Jia XC, Lin YL, Leon C. 2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS ONE. 5:e8613. doi:10.1371/journal.pone.0008613.

De Vere N, Rich TC, Ford CR, *et al.* DNA barcoding the native flowering plants and conifers of Wales. *PLoS One*. 2012;7(6):e37945. doi:10.1371/journal.pone.0037945

Table 5. Species included in the Barcode of Life Data Systems (BOLD - www.barcodinglife.org).

FMED010-12 – **Dianthus busambrae [rbcLa:587]**
Taxonomy: Magnoliophyta, Magnoliopsida, Caryophyllales, Caryophyllaceae, Dianthus
Identifiers: D2-1[sampleid], PAL96708[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Palermo

FMED011-12 - **Dianthus rupicola subsp rupicola [matK:810,rbcLa:586,trnH-psbA:192]**
Taxonomy: Magnoliophyta, Magnoliopsida, Caryophyllales, Caryophyllaceae, Dianthus
Identifiers: D3.C[sampleid], PAL96722[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Palermo

FMED012-12 - **Dianthus rupicola subsp lopadusanum [matK:801,rbcLa:562,trnH-psbA:246]**
Taxonomy: Magnoliophyta, Magnoliopsida, Caryophyllales, Caryophyllaceae, Dianthus
Identifiers: D4.C[sampleid], PAL96723[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Isole Pelagie

FMED013-12 - **Genista madoniensis [rbcLa:582]**
Taxonomy: Magnoliophyta, Magnoliopsida, Fabales, Fabaceae, Genista
Identifiers: G2u[sampleid], PAL96710[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Palermo

FMED014-12 - **Genista demarcoi [matK:805,rbcLa:577]**
Taxonomy: Magnoliophyta, Magnoliopsida, Fabales, Fabaceae, Genista
Identifiers: G4u[sampleid], PAL96713[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Palermo

FMED015-12 – **Hieracium cophanense [matK:817,rbcLa:590]**
Taxonomy: Magnoliophyta, Magnoliopsida, Asterales, Asteraceae, Hieracium
Identifiers: H2.C[sampleid], PAL96873[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Palermo

FMED016-12 – **Helichrysum hyblaeum [matK:809,rbcLa:596]**
Taxonomy: Magnoliophyta, Magnoliopsida, Asterales, Asteraceae, Helichrysum
Identifiers: H6.C[sampleid], PAL96719[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Siracusa

FMED023-12 – **Ptilostemon greuteri [matK:793,rbcLa:577]**
Taxonomy: Magnoliophyta, Magnoliopsida, Asterales, Asteraceae, Ptilostemon
Identifiers: P1.C[sampleid], PAL96705[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Trapani

FMED027-13 - **Centaurea [matK:805,rbcLa:581]**
Taxonomy: Magnoliophyta, Magnoliopsida, Asterales, Asteraceae, Centaurea
Identifiers: C3.C[sampleid], PAL96729[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Palermo

FMED028-13 - **Brassica villosa subsp. bivoniana [rbcLa:568]**
Taxonomy: Magnoliophyta, Magnoliopsida, Brassicales, Brassicaceae, Brassica
Identifiers: B3.C[sampleid], PAL96874[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Palermo

FMED029-13 - **Centaurea [matK:836,rbcLa:588]**
Taxonomy: Magnoliophyta, Magnoliopsida, Asterales, Asteraceae, Centaurea
Identifiers: C1-3[sampleid], PAL86908[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Messina

FMED031-14 - **Brassica villosa [matK:798,rbcLa:557,trnH-psbA:350]**
Taxonomy: Magnoliophyta, Magnoliopsida, Brassicales, Brassicaceae, Brassica
Identifiers: B4u[sampleid], PAL96698[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Palermo

FMED039-16 - **Dianthus rupicola subsp rupicola [matK:810,trnH-psbA:192]**
Taxonomy: Magnoliophyta, Magnoliopsida, Caryophyllales, Caryophyllaceae, Dianthus
Identifiers: D3b[sampleid], FI18813[fieldid], FI18813[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Campania

FMED040-16 - **Dianthus rupicola subsp rupicola [matK:810,trnH-psbA:192]**
Taxonomy: Magnoliophyta, Magnoliopsida, Caryophyllales, Caryophyllaceae, Dianthus
Identifiers: D3c[sampleid], PAL72352[fieldid], PAL72352[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Trapani

FMED041-16 - **Dianthus rupicola [matK:790,trnH-psbA:188]**
Taxonomy: Magnoliophyta, Magnoliopsida, Caryophyllales, Caryophyllaceae, Dianthus
Identifiers: D6p[sampleid], PAL108619[fieldid], PAL108619[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Tunisia, Zembraisland

FMED042-16 - **Dianthus rupicola [matK:780,trnH-psbA:247]**
Taxonomy: Magnoliophyta, Magnoliopsida, Caryophyllales, Caryophyllaceae, Dianthus
Identifiers: D9p[sampleid], PAL108620[fieldid], PAL108620[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Spain, BalearicIslands, Majorca

FMED001-12 - Rosa sempervirens [matK:824,rbcLa:590]
Taxonomy: Magnoliophyta, Magnoliopsida, Rosales, Rosaceae, Rosa
Identifiers: R3-1[sampleid], SV Term[fieldid]
Depository: Research Unit for Mediterranean Flower Species
Collected in: Italy, Sicily, Palermo

FMED003-12 - Asteraceae [rbcLa:577]
Taxonomy: Magnoliophyta, Magnoliopsida, Asterales, Asteraceae
Identifiers: A4u[sampleid], PAL[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Palermo

FMED004-12 - Anthemis [matK:806,rbcLa:580]
Taxonomy: Magnoliophyta, Magnoliopsida, Asterales, Asteraceae, Anthemis
Identifiers: A7.C[sampleid], PAL[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Palermo

FMED005-12 - Brassica insularis [rbcLa:594]
Taxonomy: Magnoliophyta, Magnoliopsida, Brassicales, Brassicaceae, Brassica
Identifiers: B2.23[sampleid], PAL[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Isola di Pantelleria

FMED006-12 - Brassica [matK:779,rbcLa:582]
Taxonomy: Magnoliophyta, Magnoliopsida, Brassicales, Brassicaceae, Brassica
Identifiers: B5.C[sampleid], PAL[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Palermo

FMED009-12 - Dianthus rupicola subsp aeolicus [matK:623,rbcLa:578,trnH-psbA:256]
Taxonomy: Magnoliophyta, Magnoliopsida, Caryophyllales, Caryophyllaceae, Dianthus
Identifiers: D1-2[sampleid], PAL96703[museumid]
Depository: Palermo Botanical Garden, HerbariumMediterraneum
Collected in: Italy, Sicily, Messina

Domina G, Scibetta S, Scafidi F, Giovino A. 2017. Contribution to the identification of Dianthus rupicola (Caryophyllaceae) subspecies using morphological and molecular approaches. Phytotaxa. 291(1): 017–032. doi.org/10.11646/phytotaxa.291.1.2.

Gao T, Chen SL. 2009. Authentication of the medicinal plants in Fabaceae by DNA barcoding technique. Planta Med.; 75 - P-13DOI: 10.1055/s-2009-1216451

Gao T, Sun Z, Yao H, Song J, Zhu Y, Ma X, Chen S. 2011. Identification of Fabaceae plants using the DNA barcode matK. Planta Med. 77(1): 92-94.

Giovino A, Marino P, Domina G, Scialabba A, Schicchi R, Diliberto G, Rizza C, Scibetta S. Evaluation of the DNA barcoding approach to develop a reference data-set for the threatened flora of Sicily. 2014. Plant Biosyst. 150(4):631-640. DOI:10.1080/11263504.2014.989285.

Giovino A, Bertolini E, Fileccia V, Al hassan M, Labra M, Martinelli F. 2015a. Transcriptome Analysis of Phoenix canariensisChabaud in Response to Rhynchophorusferrugineus Olivier Attacks. Frontiers Plant Sci. 6:817. doi:10.3389/fpls.2015.00817.

Giovino A, Marino P, Domina G, Rapisarda P, Rizza G, Saia S. 2015b. Fatty acid composition of the seed lipids of Chamaerops humilis L. natural populations and its relation with the environment.Plant Biosystems. 149(4):767-776.

Hebert PDN, Cywinska A, Ball SL, deWaard JR. 2003. Biological identifications through DNA barcodes. Proc Biol Sci. 270(1512):313-2.

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W. 2004.Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly Astraptesfulgerator. Proc Natl Acad Sci U S A. 101(41):14812-7.

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM. 2004. Identification of birds through DNA barcodes. PLoS Biol. 2(10):e312.

Hollingsworth PM, Graham SW, Little DP. 2011. Choosing and Using a Plant DNA Barcode. PLoS ONE. 6(5): e19254.

Hollingsworth P.M. *et al.* CBOL Plant Working Group. 2009. A DNA barcode for land plants. Proc Nat Acad Sci United States Amer. 106:12794–12797.

Kimura M. 1980. A simple method for estimating evolutionary rates of base substitution through comparative studies of nucleotide sequences. J Mol Evol. 16(2):111-120

Meyer CP, Paulay G. 2005. DNA Barcoding: Error Rates Based on Comprehensive Sampling. PLoS Biol. 3(12): 422.

Saunders GW, Kucera H. 2010. An evaluation of rbcL, tufA, UPA, LSU and ITS as DNA barcode markers for the marine green macroalgae. Cryptogamie Algol. 31(4):487-528.

Savolainen Vincent, Cowan Robyn S, Vogler Alfried P, Roderick George K and Lane Richard Towards writing the encyclopaedia of life: an introduction to DNA barcoding360Phil. Trans. R. Soc. Bhttp://doi.org/10.1098/rstb.2005.1730.

Simeone MC, Piredda R, Papini A, Vessella F, Schirone B (2013) Application of plastid and nuclear markers to DNA barcoding of Euro – Mediterranean oaks

(Quercus, Fagaceae): problems, prospects and phylogenetic implications. Botanical Journal of the Linnean Society, 172(4): 478-499.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. Mol Biol Evol. 30(12): 2725-2729.

Tautz D, Arctander P, Minelli A, Thomas RH, Vogler AP. 2003. A plea for DNA taxonomy. Trends Ecol. Evol. 18(2): 70–74.

Varshney RK, Graner A, Sorrells ME. 2005. Genic microsatellite markers in plants: features and applications. Trends Biotechnol. 23(1):48-55.

Ward RD, Zemlak TS, Innes BH, Last PR, Hebert PDN. 2005. DNA barcoding Australia's fish species. Philos Trans R Soc Lond B Biol Sci. 360(1462): 1847–1857.

SUPPLEMENTARY FIGURES

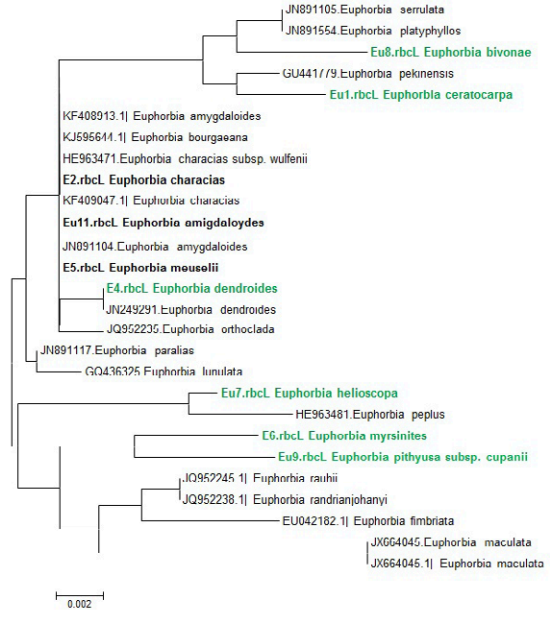

Fig. 4. Phylogenetic tree of Euphorbiaceae family with Neighbor Joining for *rbc*L marker.

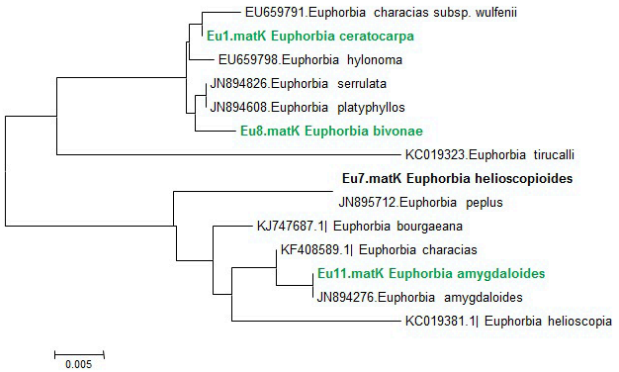

Fig. 5. Phylogenetic tree of Euphorbiaceae family with Neighbor Joining for *mat*K.