Food Chemistry 170 (2015) 308–315

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem

Analytical Methods

# A DNA barcoding approach to identify plant species in multiflower honey

I. Bruni [a], A. Galimberti [a], L. Caridi [a], D. Scaccabarozzi [b], F. De Mattia [a], M. Casiraghi [a], M. Labra [a,*]

[a] Università degli Studi di Milano-Bicocca, ZooPlantLab, Dipartimento di Biotecnologie e Bioscienze, Piazza della Scienza 2, 20126 Milano, Italy
[b] Parco Regionale della Grigna Settentrionale, Via Fornace Merlo 2, 23816 Barzio, Italy

ARTICLE INFO

*Article history:*
Received 15 August 2013
Received in revised form 25 May 2014
Accepted 13 August 2014
Available online 23 August 2014

*Keywords:*
Food traceability
Honey
Molecular markers
Pollen identification
*rbcL*
*trnH-psbA*

ABSTRACT

The purpose of this study was to test the ability of DNA barcoding to identify the plant origins of processed honey. Four multifloral honeys produced at different sites in a floristically rich area in the northern Italian Alps were examined by using the *rbcL* and *trnH-psbA* plastid regions as barcode markers. An extensive reference database of barcode sequences was generated for the local flora to determine the taxonomic composition of honey. Thirty-nine plant species were identified in the four honey samples, each of which originated from a mix of common plants belonging to *Castanea*, *Quercus*, *Fagus* and several herbaceous taxa. Interestingly, at least one endemic plant was found in all four honey samples, providing a clear signature for the geographic identity of these products. DNA of the toxic plant *Atropa belladonna* was detected in one sample, illustrating the usefulness of DNA barcoding for evaluating the safety of honey.


## 1. Introduction

The European Union recognises the importance of food traceability, defined as the ability to track food through all stages of production, processing and distribution (Regulation 178/2002, Official Journal of the European Communities, 2002). Food traceability has a pivotal role in global trade markets, and its relevance is growing every year (Galimberti et al., 2013; Karlsen, Dreyer, Olsen, & Elvevoll, 2013). Concerns about food traceability are even more important for foods whose area of production strongly influences the quality of the final product.

For example, the quality of honey is closely related to the flower composition and the climatic and environmental conditions of the production area (Bogdanov, Haldimann, Luginbuhl, & Gallmann, 2007; Stolzenbach, Byrne, & Bredie, 2011; Pohl, Stecka, Sergiel, & Jamroz, 2012). The Commission of the European Communities in Council Directive DENLEG 2000/10 (Council Directive, 2000) established guidelines for the general and specific compositions of the main honey varieties that can be marketed in European countries. The principal labelling requirements that must be indicated are the floral origin, physicochemical properties, organoleptical characteristics and regional provenance (Council Directive, 2000). Physicochemical parameters (e.g., pH, conductivity, sugar, aminoacids, vitamins and minerals) can be easily determined by analytical methods, whereas organoleptical properties are usually determined by sensory (i.e., visual, taste and olfactory) approaches. Although more complicated to determine, geographic provenance can be inferred by analysing the spectrum of pollen residuals which indicates the plants visited by bees during honey production.

In recent years, several approaches have been set up to evaluate honey characteristics (Batista et al., 2012; Camina, Pellerano, & Marchevsky, 2012; Manyi-Loh, Ndip, & Clarke, 2011) and to assess the quality of different products. The composition of pollen is commonly analysed by microscopy to determine the botanical ingredients of honey. This technique, known as melissopalynology, has been the most common method for identifying and counting pollen grains in honey in the last 30 years (Bambara, 1991; Escriche, Kadar, Juan-Borrás, & Domenech, 2011; Louveaux, Maurizio, & Vorwohl, 1978). Melissopalynology can determine the frequency of pollen from different species and can be used with other approaches, to classify honey (Persano Oddo & Piro, 2004). European standards define honey as 'unifloral' when it is from a completely or partially botanical origin, including its pollen (>45%) and physicochemical and sensory characteristics corresponding to its species of origin (see Persano Oddo & Piro, 2004 for some examples).

However, melissopalynology is time-consuming, requires specialised knowledge and involves a laborious counting procedure, making it a challenge to interpret the results and to identify

* Corresponding author. Tel.: +39 02 64483472; fax: +39 02 64483450.
*E-mail address:* massimo.labra@unimib.it (M. Labra).

http://dx.doi.org/10.1016/j.foodchem.2014.08.060
0308-8146/

botanical origins. Furthermore, it may not be possible to recognise individual species from a micromorphological analysis of pollen (Khansari et al., 2012). Although some pollen (e.g., of *Eucalyptus* and *Castanea*) show recognisable morphological traits in honey, others (e.g., pollen of some Campanulaceae and Lamiaceae) are not well distinguishable by their micromorphological traits (Khansari et al., 2012; Salmaki, Jamzad, Zarre, & Bräuchler, 2008).

Recently, researchers have applied molecular tools to analyse the composition of honey by using primers and probes specifically designed to recognise local plant species in honey (Laube et al., 2010). DNA markers, such as nuclear 18S rDNA (Olivieri, Marota, Rollo, & Luciani, 2012) and the plastid *trnL* gene (Valentini, Miquel, & Taberlet, 2010), were used to test their ability to identify plant traces from different honey samples. This approach is based on 'DNA barcoding', in which the species composition of mixed matrices is determined by comparing sequences of the same DNA region with a reference database (Casiraghi, Labra, Ferri, Galimberti, & De Mattia, 2010; Galimberti et al., 2013). The selection of universal informative markers is very important (Casiraghi et al., 2010; Sandionigi et al., 2012) to identify the botanical composition of honey and to differentiate pollen belonging to closely related taxa.

The Plant Working Group of the Consortium for the Barcode of Life (CBOL; http://www.barcoding.si.edu/plant_working_group.html) suggested the plastid coding regions *rbcL* and *matK* as core barcodes for plant identification. Additional regions, such as *trnH-psbA* and ITS2, could be used to analyse closely related taxa (Hollingsworth, Graham, & Little, 2011). Although the *matK* gene is considered a good DNA marker because it evolves rapidly (Hilu & Liang, 1997), its amplification requires specific primer combinations for different angiosperm families (Dunning & Savolainen, 2010). Therefore, it is not suitable for the analysis of unknown complex matrices, such as honey. The goal of this study was to evaluate the usefulness of the *rbcL* region and the *trnH-psbA* spacer as DNA barcoding tools for identifying the botanical constituents of honey.

## 2. Materials and methods

### 2.1. Study area and honey sampling

In this study, four honeys (ORT; CON, MON and BAI) produced in the regional park of Grigna Settentrionale (Northern Italy), were selected to investigate their botanical composition through a DNA barcoding approach. These honeys were ready to be sold in the markets as 'multifloral hone', produced during the period of June–July 2012 by amateur beekeepers from four different localities in the park (Fig. 1). For each sample an aliquot of 25 ml (40 g) was stored at −20 °C and used for DNA extraction.

The regional park of Grigna Settentrionale covers a territory of over 5000 hectares around the Grigna massif (Italian Alps). Besides the rather limited altitude (the highest summit reaches the 2409 m), the protected area is characterised by a great variety of habitats and climates ranging from the typical alpine to the submediterranean climate caused by the strong influence of Como Lake.

The total flora of the four honey production areas selected in this study consists of 593 vascular plants, including cultivated species (Rossi, 2005) and many rare and endemic taxa. The vegetation of the park varied according to the altitude: the lowest slopes up to 900 m are covered by forests (hornbeam, oak, chestnut, ash and linden trees) alternated by dry and pasture meadows. At higher altitudes, up to 1800–1900 m, there are forests of beech and conifers, especially larch. Near the summit, there are heaths with Rhododendron, mountain pine, juniper and green alder, which mark the limit of the trees. At the highest altitudes, next to the top of the massif, the area is dominated by the typical grasslands of calcareous substrates, characterised by annual species (e.g., *Carex*, *Sesleria*) and chasmophytic vegetation.

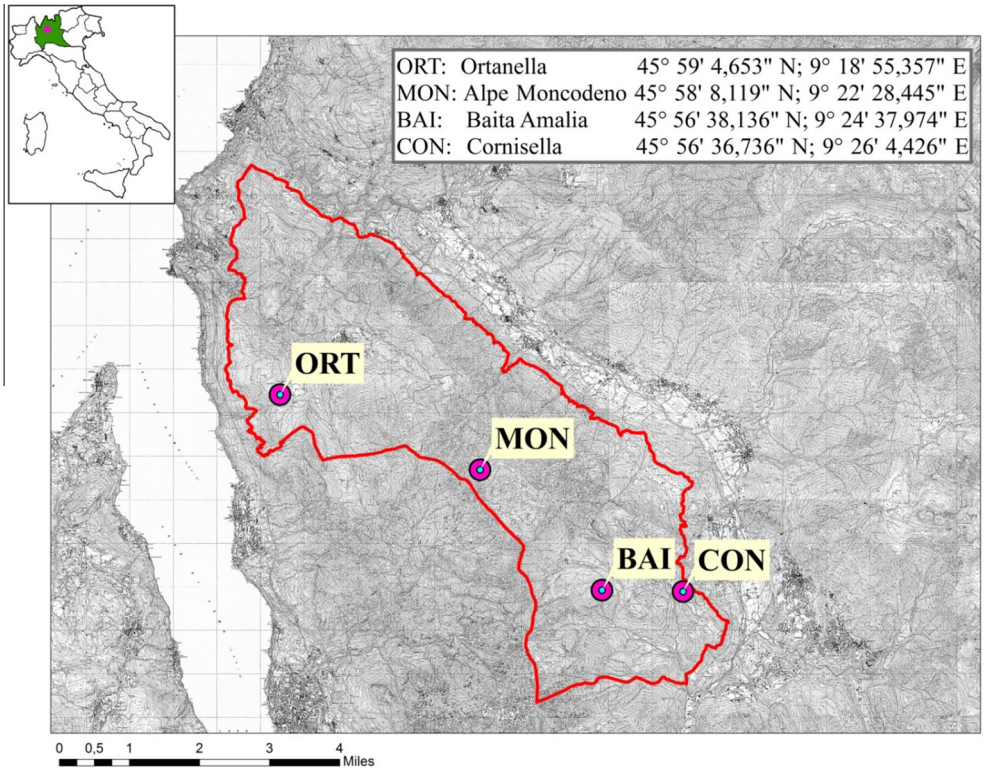


**Fig. 1.** Distribution map of honey production sites within the Grigna Settentrionale Regional Park (red line). The full names and geographic coordinates for the collection sites are provided. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The four selected honeys were produced at the medium altitudes between 700 and 1300 m, in the meadow area, surrounded by deciduous woods mainly dominated by oaks, beech trees and conifers (Fig. 1).

### 2.2. Reference DNA barcoding database

In this study a dedicated DNA barcoding reference database, consisting of 315 plant taxa, was assembled. Plants entries were selected from the floristic list related to the area of Grigna Settentrionale. The database included DNA barcoding sequences from all the most common species pollinated by bees and distributed in the honey production areas. Endemic and rare species were also included. A complete list of the species selected as reference for DNA barcoding analysis is provided as Supporting information (Table S1).

For each taxon, *rbcL* and *trnH-psbA* DNA barcode sequences were considered. A total of 104 taxa were newly characterised through DNA barcoding starting from fresh samples collected in the study area during the spring of 2012. For each individual, young leaves or buds were collected, and stored at −20 °C. All samples were vouchered as 'MIB:ZPL' following the protocol specified by the Global Registry of Biodiversity Repositories (http://grbio.org/), and the data standards for BARCODE Records (Hanner, 2009). Specimen and voucher codes are listed in Table S1 (Supporting information).

For the remaining 211 species, the *rbcL* and *trnH-psbA* sequences were retrieved from our private database (120 species, MIB:ZPL collection) and GenBank (91 species). Sequences of the former group are also available in GenBank (see Table S1), whilst records belonging to the latter category were chosen after a careful evaluation of accessions characteristics to avoid misidentification in the next bioinformatics analyses.

### 2.3. DNA extraction and purification

For each one of the four honey samples, a total of 25 ml were diluted with 25 ml of distilled water and heated to 45 °C for 5 min to permit easier handling and to decrease the honey viscosity. After 20 min of centrifugation at 13,000 rpm, the supernatant was discarded whilst the pellet was suspended in 20 ml of distilled water and dissolved by shaking. Samples were centrifuged again for 20 min at 13,000 rpm and the pellet (approximately 120 mg) was suspended in 200 μl of 1× TE buffer. One hundred microlitres of each processed sample were used for DNA extraction using DNeasy Isolation and Purification kit (Qiagen, Hilden, Germany).

The DNA extraction of the local species used to set up the reference DNA database was performed by using the same commercial kit stating from 100 mg of fresh plant materials (young leaves or buds). Purified DNA concentration of each sample was estimated both fluorometrically with a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific, USA) by measuring the absorbance (Abs) at 260 nm and by comparison of ethidium bromide-stained band intensities with λ DNA standard. DNA extracts were used as template for DNA barcoding analyses when they showed a minimum concentration of 10 ng/μl.

### 2.4. DNA barcoding analysis

DNA barcoding analysis was performed with the plastidial *rbcL* region and the *trnH-psbA* intergenic spacer. For PCR amplification and sequencing of *rbcL*, the primer combination was 1F: 5′-ATGTCACCACAAACAGAAAC-3′ and 724R: 5′-TCGCATGTACCTGCAGTAGC-3′ (Fay, Bayer, Alverson, de Bruijn, & Chase, 1998). The primer combination used for *trnH-psbA* was trnH: 5′-CGCGCATGGTGGATTCACAATCC-3′ and psbA: 5′-GTTATGCATGAACGTAATGCTC-3′ (Newmaster & Ragupathy, 2009). PCRs were performed starting from 10 ng of DNA by using puReTaq Ready-To-Go PCR beads (Amersham Bioscience, Freiburg, Germany) in a 25 μl reaction according to the manufacturer's instructions. PCR cycles consisted of an initial denaturation step for 7 min at 94 °C, 35 cycles of denaturation (45 s at 94 °C), annealing (30 s at 50 °C for *rbcL* and 53 °C for *trnH-psbA*), extension (1 min at 72 °C) and a final extension at 72 °C for 7 min.

PCR products obtained from the reference species were directly sequenced. The amplification products obtained from honeys samples were checked by electrophoresis on 1.5% (w/v) agarose. The PCR products were cloned using the pGEM-T Easy Vector System (Promega Corporation, Madison, WI, USA). Recombinant plasmids were isolated using Miniprep kit (Applied Biosystems, Foster City, CA) and the insert size and DNA concentration were assessed by gel electrophoresis on 2.0% (w/v) agarose stained with ethidium bromide.

For each one of the five honey samples, 100 clones were randomly selected to proceed with the insert sequencing.

All samples (reference species and clones) were bidirectionally sequenced for each barcode region with the same primer pairs used during the PCR step. Sequences were obtained by using an ABI 155 3730XL automated sequencer at Macrogen Inc., Korea. Manual editing of raw traces and subsequent alignments of forward and reverse sequences enabled us to amend sequencing errors. The 3′ and 5′ terminals were clipped to generate consensus sequences for each taxon. The identification of short inverted repeat regions in the *trnH-psbA* spacer was performed as reported by Whitlock, Hale, and Groff (2010), using the EMBOSS Software package (Rice, Longden, & Bleasby, 2000). The EINVERTED algorithm (Guindon & Gascuel, 2003) was used with default parameters to detect the occurrence of inversions in the *trnH-psbA* region.

To define honey composition, the 100 sequences were firstly aligned with Clustal W 2.1 (Larkin et al., 2007) and analysed with MEGA 5.1 (Tamura et al., 2011) to define MOTUs (Molecular Operational Taxonomic Units). The resulting sequences were used to identify the plant composition of the four honeys using a BLAST analysis (Altschul, Gish, Miller, Myers, & Lipman, 1990) conducted on the dedicated DNA barcoding reference database (Table S1, Supporting information). Each MOTU was assigned to the species showing the nearest matches (maximum identity) according to Bruni et al. (2012) and De Mattia et al. (2012). When the value of identity matches was lower than 99% the MOTU was considered as 'unidentifiable'. The analysis was performed separately for both the two tested markers and results were combined to identify the plant species in honey samples.

## 3. Results

DNA extracted from the four honey samples and the local reference taxa was of high quality (ratios of absorbance, $A_{260/280}$ and $A_{260/230}$ ~1.80 and >1.90, respectively) and provided good yields (10–25 ng/μl). Amplification was successful (i.e., non-specific bands were absent) and provided DNA concentration of purified amplicons >50 ng/μl when standard primer pairs for the amplification of *rbcL* and *trnH-psbA* were used. No short inverted repeat regions were found in the *trnH-psbA* sequences. Accession numbers of DNA barcode sequences for each species are provided in Table S1 (Supporting information).

For each honey sample, 100 clones were sequenced for both markers (Tables 1–4) and grouped into coherent molecular operational taxonomic units (MOTUs) according to their sequence similarity. Between 12 and 15 MOTUs were identified for each honey sample, with 1–26 sequences each. Identified MOTUs were compared with the reference database. For each sample, all MOTUs were associated with plant species (Tables 1–4). However, the

**Table 1**
Molecular identification of plant species detected in honey from Baita Amalia (BAI). The number of clones for each MOTU, the species match in reference database (RD) and the related identity values (ID%) obtained with the BLAST search are reported for the two barcode regions.

| Identified plants | *rbcL*-molecular identification | | | *trnH-psbA* molecular identification | | |
|---|---|---|---|---|---|---|
| | MOTUs (n. clones) | Species match in RD | ID (%) | MOTUs (n. clones) | Species match in RD | ID (%) |
| *Bromus erectus* | 3 | *Bromus erectus* | 100 | 1 | *Bromus erectus* | 100 |
| *Cirsium arvense* | 12 | *Cirsium arvense*<br>*Carlina acaulis*<br>*Cirsium vulgare*[b] | 100<br>100<br>100 | 9 | *Cirsium arvense* | 100 |
| *Centaurea jacea gaudini* | 0 | – | – | 5 | *Centaurea jacea gaudini* | 100 |
| *Fagus sylvatica* | 13 | *Fagus sylvatica* | 100 | 17 | *Fagus sylvatica* | 100 |
| *Juniperus communis* | 3 | *Juniperus communis* | 100 | 0 | – | – |
| *Leucanthemum vulgare* | 10 | *Leucanthemum vulgare*<br>*Artemisia campestris* | 100<br>99.3 | 9 | *Leucantheumum vulgare* | 100 |
| *Ostrya carpinifolia* | 11 | *Ostrya carpinifolia* | 100 | 13 | *Ostrya carpinifolia* | 100 |
| *Quercus pubescens/Q. petraea* | 16 | *Quercus pubescens*<br>*Quercus petraea* | 100<br>100 | 9 | *Quercus pubescens*<br>*Quercus petraea* | 100<br>99.1 |
| *Solanum nigrum/S. villosum* | 4 | *Solanum nigrum*<br>*Solanum villosum* | 100<br>100 | 8 | *Solanum nigrum*<br>*Solanum villosum* | 100<br>99.8 |
| *Trifolium montanum* | 12 | *Trifolium montanum*<br>*Trifolium repens* | 100<br>99.8 | 17 | *Trifolium montanum* | 100 |
| *Veronica officinalis* | 7 | *Veronica officinalis* | 100 | 0 | – | – |
| *Xerolekia speciosissima*[a] | 3 | *Xerolekia speciosissima*<br>*Buphtalmum salicifolium*<br>*Cyanus triumfetti*[b] | 100<br>99.6<br>99.4 | 8 | *Xerolekia speciosissima* | 100 |
| Unidentifiable | 6 | – | – | 4 | – | – |

[a] Endemic species.
[b] More than 3 species showed Maximum identity values in BLAST higher than 99%.

**Table 2**
Molecular identification of plant species detected in honey from Alpe Moncodeno (MON). The number of clones for each MOTU, the species match in reference database (RD) and the related identity values (ID%) obtained with the BLAST search are reported for the two barcode regions.

| Identified plants | *rbcL*-molecular identification | | | *trnH-psbA* molecular identification | | |
|---|---|---|---|---|---|---|
| | MOTUs (n. clones) | Species match in RD | ID (%) | MOTUs (n. clones) | Species match in RD | ID (%) |
| *Acer platanoides/A. campestre* | 13 | *Acer platanoides*<br>*Acer pseudoplatanus*<br>*Acer campestre* | 100<br>99.8<br>99.8 | 9 | *Acer platanoides*<br>*Acer campestre* | 100<br>99.6 |
| *Aster amellus* | 5 | *Aster amellus*<br>*Solidago virgaurea*<br>*Cyanus triumfetti* [b] | 100<br>100<br>99.8 | 9 | Aster amellus | 100 |
| *Atropa belladonna* | 2 | *Atropa belladonna* | 100 | 0 | – | – |
| *Buplerum petraeum* | 4 | *Buplerum petraeum*<br>*Buplerum stellatum* | 100<br>100 | 2 | *Buplerum petraeum* | 100 |
| *Campanula trachelium* | 0 | – | – | 2 | *Campanula trachelium* | 100 |
| *Centaurea jacea/C. rhaetica* | 11 | *Centaurea jacea*<br>*Centaurea jacea gaudini*<br>*Carlina acaulis*[b] | 100<br>99.7<br>99.5 | 6 | *Centaurea jacea*<br>*Centaurea rhaetica* | 100<br>100 |
| *Centaurea triumfetti* | 0 | – | – | 8 | *Cyanus triumfetti* | 100 |
| *Fagus sylvatica* | 21 | *Fagus sylvatica* | 100 | 13 | *Fagus sylvatica* | 100 |
| *Genista tinctoria* | 7 | *Genista tinctoria*<br>*Laburnum anagyroides* | 100<br>99.4 | 5 | *Genista tinctoria* | 100 |
| *Laserpitium nitidum* | 5 | *Laserpitium nitidum* | 100 | 2 | *Laserpitium nitidum* | 100 |
| *Minuartia grignensis*[a] | 2 | *Minuartia grignensis* | 100 | 5 | *Minuartia grignensis* | 100 |
| *Primula grignensis*[a]*/P. glaucescens*[a] | 1 | *Primula grignensis*<br>*Primula glaucescens* | 100<br>100 | 0 | – | – |
| *Quercus pubescens/Q. petraea* | 14 | *Quercus pubescens*<br>*Quercus petraea* | 100<br>100 | 17 | *Quercus pubescens*<br>*Quercus petraea* | 100<br>99.1 |
| *Trifolium montanum* | 12 | *Trifolium montanum*<br>*Trifolium repens* | 100<br>99.8 | 15 | *Trifolium montanum* | 100 |
| *Trifolium pratense* | 0 | – | – | 3 | *Trifolium pratense* | 100 |
| Unidentifiable | 3 | – | – | 4 | – | – |

[a] Endemic species.
[b] More than 3 species showed Maximum identity values in BLAST higher than 99%.

**Table 3**
Molecular identification of plant species detected in honey from Cornisella (COR). The number of clones for each MOTU, the species match in reference database (RD) and the related identity values (ID%) obtained with the BLAST search are reported for the two barcode regions.

| Identified plants | *rbcL*-molecular identification | | | *trnH-psbA* molecular identification | | |
|---|---|---|---|---|---|---|
| | MOTUs (n. clones) | Species match in RD | ID (%) | MOTUs (n. clones) | Species match in RD | ID (%) |
| *Acer platanoides/A. campestre* | 7 | *Acer platanoides*<br>*Acer pseudoplatanus*<br>*Acer campestre* | 100<br>99.8<br>99.8 | 7 | *Acer platanoides*<br>*Acer campestre* | 100<br>99.6 |
| *Anthyllis vulneraria* | 0 | *Anthyllis vulneraria* | 100 | 0 | – | – |
| *Campanula raineri*[a]*/Physoplexis comosa* | 12 | *Campanula raineri*<br>*Campanula elatinoides*<br>*Campanula rotundifolia*[b] | 100<br>100<br>99.0 | 12 | *Campanula raineri*<br>*Physoplexis comosa* | 100<br>99.3 |
| *Castanea sativa* | 13 | *Castanea sativa*<br>*Quercus pubescens*<br>*Quercus petraea* | 100<br>99.8<br>99.8 | 13 | *Castanea sativa* | 100 |
| *Genista tinctoria* | 7 | *Genista tinctoria*<br>*Laburnum anagyroides* | 100<br>99.4 | 7 | *Genista tinctoria* | 100 |
| *Fagus sylvatica* | 11 | *Fagus sylvatica* | 100 | 11 | *Fagus sylvatica* | 100 |
| *Geranium robertianum* | 7 | *Geranium robertianum* | 100 | 7 | *Geranium robertianum* | 100 |
| *Laserpitium nitidum* | 9 | *Laserpitium nitidum* | 100 | 9 | *Laserpitium nitidum* | 100 |
| *Phyteuma scheuchzeri* | 1 | *Phyteuma scheuchzeri*<br>*Centaurea nigrescens*<br>*Campanula barbata* [b] | 100<br>99.4<br>99.4 | 1 | *Phyteuma scheuchzeri* | 100 |
| *Quercus pubescens/Q. petraea* | 3 | *Quercus pubescens*<br>*Quercus petraea* | 100<br>100 | 3 | *Quercus pubescens*<br>*Quercus petraea* | 100<br>99.1 |
| *Rubus idaeus* | 5 | *Rubus idaeus*<br>*Rubus caesius*<br>*Rubus ulmifolius* | 100<br>99.6<br>99.6 | 5 | *Rubus idaeus* | 100 |
| *Tanacetum corymbosum* | 13 | *Tanacetum corymbosum*<br>*Achillea millefolium* | 100<br>99.8 | 13 | *Tanacetum corymbosum* | 100 |
| *Tilia cordata* | 4 | *Tilia cordata* | 100 | 4 | *Tilia cordata* | 100 |
| *Thlaspi rotundifolium*[a] | 3 | – | – | 3 | *Thlaspi rotundifolium* | 100 |
| *Viola tricolor/V. hirta* | 4 | *Viola tricolor*<br>*Viola hirta* | 100<br>99.0 | 0 | – | – |
| Unidentifiable | 5 | – | – | 5 | – | – |

[a] Endemic species.
[b] More than 3 species showed Maximum identity values in BLAST higher than 99%.

relationship between the number of clones and the composition of MOTUs did not account for the abundance of each species in the samples.

In several cases the BLAST analysis performed with *rbcL* was not able to identify a species with more than 99% similarity. For example, for Fagaceae, the *rbcL* did not distinguish *Castanea sativa* Mill. from *Quercus pubescens* Willd. and *Quercus petraea* (Mattuschka) Liebl. In contrast, DNA barcoding analysis performed with the *trnH-psbA* marker identified all plant species in the honey samples with rare exceptions for some congeneric species, namely *Acer platanoides* L., *Acer campestre* L., *Centaurea jacea* (Boiss. & Reut.) Gremli, *Centaurea rhaetica* Moritzi, *Q. pubescens*, *Q. petraea*, *Solanum nigrum* L. and *Solanum villosum* Mill., as previously shown by Pirredda, Simeone, Attimonelli, Bellarosa, and Schirone (2011).

The combination of data from using both *rbcL* and *trnH-psbA* allowed us to identify 12, 14, 14 and 15 plant species in the BAI, MON, ORT and COR samples respectively (Tables 1–4), with a minimum of 38 plant species overall (Fig. 2). Only a few MOTUs could not be identified by BLAST searches, which was probably due to the incomplete floristic coverage of the reference database. The four analysed honeys exhibited a prevalence of pollen types from a variety of floral sources widely distributed in Grigna Settentrionale Park, including: *A. platanoides.*, *A. campestre*, *C. sativa*, *Fagus sylvatica*, *Q. pubescens* and *Q. petraea*. Several endemic species were detected, such as *Xerolekia speciosissima* (L.) Anderb. in the BAI sample (Table 1); *Minuartia grignensis* (Rchb.) Mattf., *Primula grignensis* Moser and *Primula glaucescens* Moretti in the MON sample (Table 2); *Campanula raineri* Perp. and *Thlaspi rotundifolium* (L.) in the COR sample (Table 3); and *M. grignensis* in the ORT sample (Table 4).

## 4. Discussion

The goal of this study was to evaluate the effectiveness of the DNA barcoding approach for determining the botanical composition of honey. High-quality DNA was obtained from all samples. The *rbcL* and *trnH-psbA* markers were easily amplified and sequenced. However, *rbcL* had limited ability for identification, especially of congeneric taxa, because it identified most MOTUs only to the genus level. Using the *trnL* marker, Valentini et al. (2010) found that conserved genes were unable to distinguish closely related taxa co-occurring in complex matrices. In contrast, almost every MOTU detected in the honey samples was assigned to a species when *trnH-psbA* was used. Thus, the plastid spacer was the most suitable marker to determine the plant species composition of honey.

These findings support the combined use of the conventional core-barcode markers with the *trnH-psbA* spacer to differentiate congeneric taxa (Federici et al., 2013; Hollingsworth et al., 2011). The *trnH-psbA* is suitable for characterising honey from a limited geographic area with well-known flora because the region is not very well represented in public databases, which would be

**Table 4**
Molecular identification of plant species detected in honey from Ortanella (ORT). The number of clones for each MOTU, the species match in reference database (RD) and the related identity values (ID%) obtained with the BLAST search are reported for the two barcode regions.

| Identified plants | *rbcL*-molecular identification | | | *trnH-psbA* molecular identification | | |
|---|---|---|---|---|---|---|
| | MOTUs (n. clones) | Species match in RD | ID (%) | MOTUs (n. clones) | Species match in RD | ID (%) |
| *Aster amellus* | 12 | *Aster amellus*<br>*Solidago virgaurea*<br>*Cyanus triumfetti*[b] | 100<br>100<br>99.8 | 6 | *Aster amellus* | 100 |
| *Castanea sativa* | 13 | *Castanea sativa*<br>*Quercus pubescens*<br>*Quercus petraea* | 100<br>99.8<br>99.8 | 26 | *Castanea sativa* | 100 |
| *Carlina acaulis* | 0 | – | – | 5 | *Carlina acaulis* | 100 |
| *Carduus defloratus/Cirsium erisithales* | 13 | *Carduus defloratus*<br>*Cirsium vulgare*<br>*Carlina acaulis*[b] | 100<br>100<br>99.8 | 4 | *Carduus defloratus*<br>*Cirsium erisithales* | 100<br>99.4 |
| *Centaurea rhaetica/C. jacea* | 3 | *Centaurea jacea*<br>*Centaurea jacea gaudini*<br>*Carlina acaulis*[b] | 100<br>99.7<br>99.5 | 8 | *Centaurea jacea*<br>*Centaurea rhaetica* | 100<br>100 |
| *Fagus sylvatica* | 11 | *Fagus sylvatica* | 100 | 4 | *Fagus sylvatica* | 100 |
| *Geranium rotundifolium* | 3 | *Geranium rotundifolium*<br>*Geranium phaeum*<br>*Geranium sylvaticum*[b] | 100<br>99.4<br>99.0 | 4 | *Geranium rotundifolium* | 100 |
| *Melittis melissophyllum* | 6 | *Melittis melissophyllum* | 100 | 0 | – | – |
| *Minuartia grignensis*[a] | 4 | *Minuartia grignensis* | 100 | 7 | *Minuartia grignensis* | 100 |
| *Quercus pubescens/Q. petraea* | 9 | *Quercus pubescens*<br>*Quercus petraea* | 100<br>100 | 12 | *Quercus pubescens*<br>*Quercus petraea* | 100<br>99.1 |
| *Rubus idaeus* | 8 | *Rubus idaeus*<br>*Rubus caesius*<br>*Rubus ulmifolius* | 100<br>99.6<br>99.6 | 2 | *Rubus idaeus* | 100 |
| *Trifolium montanum* | 9 | *Trifolium montanum*<br>*Trifolium repens* | 100<br>99.8 | 3 | *Trifolium montanum* | 100 |
| *Trifolium pratense* | 8 | *Trifolium pratense*<br>*Trifolium rubens* | 100<br>99.5 | 11 | *Trifolium pratense* | 100 |
| *Veronica officinalis* | 0 | – | – | 6 | *Veronica officinalis* | 100 |
| Unidentifiable | 1 | – | – | 2 | – | – |

[a] Endemic species.
[b] More than 3 species showed Maximum identity values in BLAST higher than 99%.

necessary for a wider variety of taxa. One of the reasons of this scarce representation was that stutter PCR products due to mononucleotide repeats had been frequently reported for *trnH-psbA* (Hollingsworth, 2008). However, the availability of *trnH-psbA* reference barcodes is improving, also due to the recent application of new technical advances to obtain high-quality sequences (e.g., appropriate polymerases and ideal PCR conditions; Fazekas, Steeves, & Newmaster, 2010).

Although melissopalynology and DNA barcoding perform best with a local reference database (morphological or molecular), DNA analysis is faster than melissopalynology. Moreover, DNA barcoding is the most standardised and universal DNA technique to be routinely used to analyse complex food matrices without any botanical knowledge (Galimberti et al., 2013).

An important result of this study is that DNA barcoding analysis can be used to infer the geographical origin of honey. Although the honey samples were comprised of a mix of common plants, such as *Castanea*, *Quercus* and *Fagus*, the presence of at least one endemic plant connected the honey samples to Grigna Settentrionale Regional Park. It would also be possible to determine the origin of honey from DNA data, but only with a detailed list of plants from the study area that have been studied at the molecular level. With next-generation sequencing, vast collections of samples, including complex food matrices, could be characterised at the molecular level for a relatively low price. Sampling is usually the limiting factor for this analysis, but it can be done by collecting plants at any life stage (a typical problem when only the morphological recognition is used) in collaboration with local amateurs or natural history museums.

Traces of DNA of *Atropa belladonna* L., a plant that is toxic for humans (Ashtiania & Sefidkonb, 2011), were detected in one of the honey samples. Although it does not necessarily mean that poisonous metabolites of *A. belladonna* were present in the honey sample, this result supports the use of DNA barcoding as an 'alarm bell' in the evaluation of food safety, as has been shown previously for several poisonous plants (e.g., Bruni et al., 2010). However, to date, any DNA barcoding study specifically aiming to assess the safety of honey is available despite some reports of honey with traces of dangerous plants such as *Rhododendron* spp. and *Crotalaria* spp. (Koca & Koca, 2007; Olivieri et al., 2012; Popescu & Kopp, 2013).

Even though the results of this study are based on only a few honey samples, they support the potential utility of DNA analysis in detecting fraudulent or mistaken labelling of honey (Camina et al., 2012) by comparing local flora species with those detected in the honey sample. DNA barcoding could act as reliable tool for honey traceability at different stages of production and distribution (Galimberti et al., 2013). The principal limitations of this approach are its inability to quantify the composition (in terms of biological units) of complex matrices and the lack of a general consensus on the selection of barcode regions, due to variable identification performances amongst different plant groups

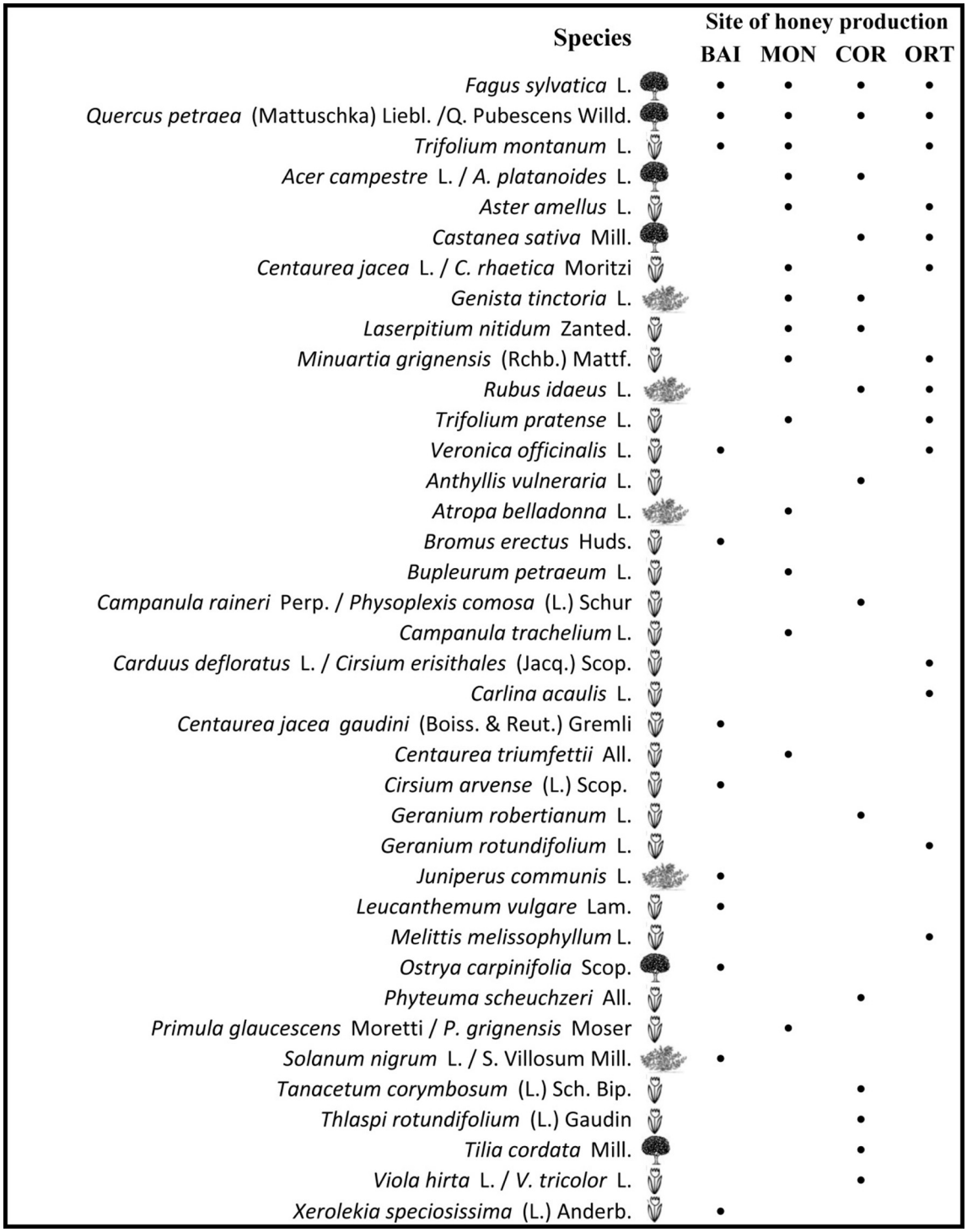


**Fig. 2.** Dot plot distribution of plant species in the four tested honeys. The plant typology (tree, flower, and shrub) and the collection sites are indicated.

(Casiraghi et al., 2010). European guidelines dictate that an accurate quantitative composition must be provided to certify the quality and to name the honey. Additional molecular techniques, such as real-time PCR based on SCAR with DNA barcoding markers, could be used to obtain relative abundances for plant species in honey samples or to detect possible contamination easily and rapidly (Jaakola, Suokas, & Häggman, 2010). The second limitation (i.e., universality of this approach) is also relevant but can be resolved with a well-populated database of DNA barcoding reference sequences for the local flora surrounding the beehives, as was the case in this study.

## 5. Conclusions

The origin, quality and safety of honey can be evaluated with DNA barcoding. High performance standards suggest that DNA barcoding can be considered a valid alternative to mellisopalynological analyses. However, two additional aspects should be considered before adopting DNA barcoding as a standard approach for ensuring the traceability of honey. First, an exhaustive analysis of the botanical composition of honey using DNA barcoding should be combined with next generation sequencing. A large number of DNA fragments could be sequenced without cloning the plasmid

vector, and more species present in trace amounts could be identified (Park et al., 2012). Second, DNA barcoding is already used as an identification technique in a legal context for fish traceability (e.g., by the Food and Drug Administration, USA and by the Philippines government, Galimberti et al., 2013). Further tests are necessary, but all stakeholders involved in the honey supply chain should seriously consider this opportunity.

## Acknowledgements

The authors would like to thank "Gruppo Botanofilo Comense" for their technical support during plant sampling. They are also indebted to Davide Nespoli for generating the figures. This work was supported by 'Fondazione Cariplo' – 'Italy' Grant C91H09000010003 with the project entitled "Insetti Pronubi: mezzi di connessione e diffusione di specie vegetali rare ed endemiche del parco regionale della Grigna Settentrionale". The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.foodchem.2014.08.060.

## References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology, 215*, 403–410.

Ashtiania, F., & Sefidkonb, F. (2011). Tropane alkaloids of *Atropa belladonna* L. and *Atropa acuminate* Royle ex Miers plants. *Journal of Medicinal Plants Research, 5*, 6515–6522.

Bambara, S. B. (1991). Using pollen to identify honey. *American Bee Journal, 131*, 242–243.

Batista, B. L., da Silva, L. R. S., Rocha, B. A., Rodrigues, J. L., Berretta-Silva, A. A., Bonates, T. O., Gomes, V. S. D., Barbosa, R. M., & Barbosa, F. (2012). Multielement determination in Brazilian honey samples by inductively coupled plasma mass spectrometry and estimation of geographic origin with data mining techniques. *Food Research International, 49*, 209–215.

Bogdanov, S., Haldimann, M., Luginbuhl, W., & Gallmann, P. (2007). Minerals in honey: Environmental geographical and botanical aspects. *Journal of Apicultural Research, 46*, 269–275.

Bruni, I., De Mattia, F., Galimberti, A., Galasso, G., Banfi, E., Casiraghi, M., & Labra, M. (2010). Identification of poisonous plants by DNA barcoding approach. *International Journal of Legal Medicine, 124*(6), 595–603.

Bruni, I., De Mattia, F., Martellos, S., Galimberti, A., Savadori, P., Casiraghi, M., Nimis, P. L., & Labra, M. (2012). DNA barcoding as an effective tool in improving a digital plant identification system: A case study for the area of Mt. Valerio, Trieste (NE Italy). *PLoS One, 7*, e43256.

Camina, J. M., Pellerano, R. G., & Marchevsky, E. J. (2012). Geographical and botanical classification of honeys and apicultural products by chemometric methods. A review. *Current Analytical Chemistry, 8*, 408–425.

Casiraghi, M., Labra, M., Ferri, E., Galimberti, A., & De Mattia, F. (2010). DNA barcoding: A six-question tour to improve users' awareness about the method. *Briefings in Bioinformatics, 11*, 440–453.

Council Directive. (2000). *Proposal for a Council Directive relating to honey—DENLEG/2000/10*. Brussels, 18-5-2000.

De Mattia, F., Gentili, R., Bruni, I., Galimberti, A., Sgorbati, S., Casiraghi, M., & Labra, M. (2012). A multi-marker DNA barcoding approach to save time and resources in vegetation surveys. *Botanical Journal of the Linnean Society, 169*, 518–529.

Dunning, L. T., & Savolainen, V. (2010). Broad-scale amplification of matK for DNA barcoding plants, a technical note. *Botanical Journal of the Linnean Society, 164*, 1–9.

Escriche, I., Kadar, M., Juan-Borrás, M., & Domenech, E. (2011). Using flavonoids, phenolic compounds and headspace volatile profile for botanical authentication of lemon and orange honeys. *Food Research International, 44*, 1504–1513.

Fay, M. F., Bayer, C., Alverson, W. S., de Bruijn, A. Y., & Chase, M. W. (1998). Plastid rbcL sequence data indicate a close affinity between *Diegodendron* and *Bixa*. *Taxon, 47*, 43–50.

Fazekas, A. J., Steeves, R., & Newmaster, S. G. (2010). Improving sequencing quality from PCR products containing long mononucleotide repeats. *Biotechniques, 48*, 277–285.

Federici, S., Galimberti, A., Bartolucci, F., Bruni, I., De Mattia, F., Cortis, P., & Labra, M. (2013). DNA barcoding to analyse taxonomically complex groups in plants: The case of Thymus (Lamiaceae). *Botanical Journal of the Linnean Society, 171*, 687–699.

Galimberti, A., De Mattia, F., Losa, A., Bruni, I., Federici, S., Casiraghi, M., Martellos, S., & Labra, M. (2013). DNA barcoding as a new tool for food traceability. *Food Research International, 50*, 55–63.

Guindon, S., & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology, 52*(5), 696–704.

Hanner, R. (2009). *Data Standards for BARCODE Records in INSDC (BRIs)*. URL <http://barcoding.si.edu/PDF/DWG_data_standards-Final.pdf>. Accessed 15.04.14.

Hilu, K. W., & Liang, H. (1997). The matK gene: Sequence variation and application in plant systematics. *American Journal of Botany, 84*, 830–839.

Hollingsworth, P. M. (2008). DNA barcoding plants in biodiversity hot spots: Progress and outstanding questions. *Heredity, 101*, 1–2.

Hollingsworth, P. M., Graham, S. W., & Little, D. P. (2011). Choosing and using a plant DNA barcode. *PLoS One, 6*, e19254.

Jaakola, L., Suokas, M., & Häggman, H. (2010). Novel approaches based on DNA barcoding and high-resolution melting of amplicons for authenticity analyses of berry species. *Food Chemistry, 123*, 494–500.

Karlsen, K. M., Dreyer, B., Olsen, P., & Elvevoll, E. O. (2013). Literature review: Does a common theoretical framework to implement food traceability exist? *Food Control, 32*, 409–417.

Khansari, E., Zarre, S., Alizadeh, K., Attara, F., Aghabeigic, F., & Salmakia, Y. (2012). Pollen morphology of *Campanula* (Campanulaceae) and allied genera in Iran with special focus on its systematic implication. *Flora, 207*, 203–211.

Koca, I., & Koca, A. F. (2007). Poisoning by mad honey: A brief review. *Food and Chemical Toxicology, 45*, 1315–1318.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., & Higgins, D. G. (2007). Clustal W and clustal X version 2.0. *Bioinformatics, 23*, 2947–2948.

Laube, I., Hird, H., Brodmann, P., Ullmann, S., Schone-Michling, M., Chisholm, J., & Broll, H. (2010). Development of primer and probe sets for the detection of plant species in honey. *Food Chemistry, 118*, 979–986.

Louveaux, J., Maurizio, A., & Vorwohl, G. (1978). Methods of melissopalynology. *Bee World, 59*, 139–157.

Manyi-Loh, C. E., Ndip, R. N., & Clarke, A. M. (2011). Volatile compounds in honey: A review on their involvement in aroma, botanical origin determination and potential biomedical activities. *International Journal of Molecular Sciences, 12*, 9514–9532.

Newmaster, S. G., & Ragupathy, S. (2009). Testing plant barcoding in a sister species complex of pantropical *Acacia* (Mimosoideae, Fabaceae). *Molecular Ecology Resources, 9*, 172–180.

Official Journal of the European Communities. (2002). Regulation 178/2002. URL <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2002:031:0001:0024:EN:PDF>. Accessed 14.08.2013.

Olivieri, C., Marota, I., Rollo, F., & Luciani, S. (2012). Tracking plant, fungal, and bacterial DNA in honey specimens. *Journal of Forensic Sciences, 57*, 222–227.

Park, J. Y., Lee, S. Y., An, C. M., Kang, J. H., Kim, J. H., Chai, J. C., Chen, J. Y., Kang, J. S., Ahn, J. J., Lee, Y. S., & Hwang, S. Y. (2012). Comparative study between next generation sequencing technique and identification of microarray for species identification within blended food products. *Biochip Journal, 6*, 354–361.

Persano Oddo, L., & Piro, R. (2004). Main European unifloral honeys: Descriptive sheets. *Apidologie, 35*, S38–S81.

Pirredda, R., Simeone, M. C., Attimonelli, M., Bellarosa, R., & Schirone, B. (2011). Prospects of barcoding the Italian wild dendroflora: Oaks reveal severe limitations to tracking species identity. *Molecular Ecology Resources, 11*, 72–83.

Pohl, P., Stecka, H., Sergiel, I., & Jamroz, P. (2012). Different aspects of the elemental analysis of honey by flame atomic absorption and emission spectrometry: A review. *Food Analytical Methods, 5*, 737–751.

Popescu, R., & Kopp, B. (2013). The genus *Rhododendron*: An ethnopharmacological and toxicological review. *Journal of Ethnopharmacology, 147*, 42–62.

Rice, P., Longden, I., & Bleasby, A. (2000). EMBOSS: The European molecular biology open software suite. *Trends in Genetics, 16*(6), 276–277.

Rossi, P. (2005). *La flora del gruppo delle grigne*. Lecco: Comunità Montana.

Salmaki, Y., Jamzad, Z., Zarre, S., & Bräuchler, C. (2008). Pollen morphology of *Stachys* (Lamiaceae) in Iran and its systematic implication. *Flora, Morphology, Distribution, Functional Ecology of Plants, 203*, 627–639.

Sandionigi, A., Galimberti, A., Labra, M., Ferri, E., Panunzi, E., De Mattia, F., & Casiraghi, M. (2012). Analytical approaches for DNA barcoding data – How to find a way for plants? *Plant Biosystems, 146*, 805–813.

Stolzenbach, S., Byrne, D. V., & Bredie, W. L. P. (2011). Sensory local uniqueness of Danish honeys. *Food Research International, 44*, 2766–2774.

Tamura, K., Peterson, D., Peterson, N., Steker, G., Nei, M., & Kumar, S. (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution, 28*, 2731–2739.

Valentini, A., Miquel, C., & Taberlet, P. (2010). DNA barcoding for honey biodiversity. *Diversity, 2*, 610–617.

Whitlock, B. A., Hale, A. M., & Groff, P. A. (2010). Intraspecific inversions pose a challenge for the *trnH-psbA* plant DNA barcode. *PLoS One, 5*(7), e11533.