Food Chemistry 214 (2017) 183–191

Contents lists available at ScienceDirect

Food Chemistry

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

Analytical Methods

Rapid identification of the botanical and entomological sources of honey using DNA metabarcoding

Sean W.J. Prosser *, Paul D.N. Hebert

Centre for Biodiversity Genomics, Biodiversity Institute of Ontario, University of Guelph, 50 Stone Road East, Guelph, Ontario N1G 2W1, Canada

ARTICLE INFO

Article history:
Received 2 March 2016
Received in revised form 7 July 2016
Accepted 11 July 2016
Available online 12 July 2016

Keywords:
Honey
Metabarcoding
Next generation sequencing
Food security
Fraud

ABSTRACT

Honey is generated by various bee species from diverse plants, and because the value of different types of honey varies more than 100-fold, it is a target for fraud. This paper describes a protocol that employs DNA metabarcoding of three gene regions (ITS2, rbcLa, and COI) to provide an inexpensive tool to simultaneously deliver information on the botanical and entomological origins of honey. This method was used to examine seven varieties of honey: light, medium, dark, blended, pasteurized, creamed, and meliponine. Plant and insect sources were identified in five samples, but only the botanical or insect source could be identified in the other two. Two samples were found to be misrepresented. Although this method was generally successful in determining both plant and insect sources, honeys rich in polyphenolic compounds or subject to crystallization were recalcitrant to analysis, so further research is required to combat honey adulteration and mislabeling.

1. Introduction

Derived from pollen and nectar, honey is made by several species of bees aside from the European honey bee, *Apis mellifera* (Jones, 2010). Polyfloral honey is ubiquitous, but monofloral honey can be produced by deploying hives where flowers of a particular plant species are dominant. Due to their rarity, unique flavours, and medicinal properties, monofloral honeys can be very expensive. For example, Manuka honey, which derives from two species of *Leptospermum*, retails at about $100/kg (Ahmed & Othman, 2013) due to its health benefits (Yaghoobi, Kazerouni, & Kazerouni, 2013). The value of honey is also influenced by its insect source; honey from stingless bees (e.g. *Melipona beecheii*) is highly valued, creating an incentive to dilute or substitute with honey from *A. mellifera* (Vit, Medina, & Enriquez, 2004).

While some effort has been made to develop protocols to ascertain the entomological sources of honey (Ramón-Sierra, Ruiz-Ruiz, & de la Luz Ortiz-Vázquez, 2015; Schnell, Fraser, Willerslev, & Gilbert, 2010), most work has focused on identifying its plant origins. Past studies have often relied upon diagnostic phytochemicals (e.g. Cotte et al. (2004), Tosun (2013)) or melissopalynology, the study of pollen in honey (e.g. de Franca Alves and de Assis Ribeiro dos Santos (2014)). Although the latter approach requires considerable expertise and cannot distinguish many plant species (Kaskoniene & Venskutonis, 2010), it is a powerful diagnostic tool, especially when used with other methods (Hawkins et al., 2015). However, melissopalynology is ineffective in cases where low-value honey is filtered to remove its source pollen and spiked with pollen from the 'desired' monofloral.

Although none of the above protocols can reliably determine the plant source of an 'unknown' honey (Kaskoniene & Venskutonis, 2010; Laube et al., 2010), such uncertainty can often be resolved through the genetic analysis of targeted gene regions isolated from honey. Termed metabarcoding, this approach overcomes many of the limitations of other analytical methods (Schnell et al., 2010) and is gaining power because of increased access to high-throughput sequencing platforms (Hawkins et al., 2015; Richardson et al., 2015; Valentini, Miquel, & Taberlet, 2010). Although honey includes pollen, nectar, and bee products, most prior genetic studies have only examined pollen (Hawkins et al., 2015; Richardson et al., 2015; Valentini et al., 2010). While such analysis reveals the major botanical components of honey, full authentication requires analysis of DNA isolated from liquid honey because it can reveal both the source bee and "hidden" floral components in cases where honey has been filtered to remove pollen.

This paper describes a cost-effective protocol to identify the botanical and entomological components of honey using metabarcoding to examine three gene regions. Pollen components are examined using nuclear ITS2, pollen-free plant material with the plastid gene rbcLa, and the source bee species using mitochondrial

* Corresponding author.
E-mail address: sprosser@uoguelph.ca (S.W.J. Prosser).

http://dx.doi.org/10.1016/j.foodchem.2016.07.077
0308-8146/

cytochrome *c* oxidase subunit I (COI). ITS2 (~350 bp) was chosen to identify the pollen signature of honey because it can discriminate most plant species (Yao et al., 2010). Additionally, ITS2 is found in all pollen grains, while plastid markers can be absent (Bell, Burgess, Okamoto, Aranda, & Brosi, 2016). A segment (162 bp) of the rbcLa gene was used to detect trace and/or degraded plant DNA in honey (but not necessarily from pollen). Finally, a segment (120 bp) of the COI gene was used to identify the entomological source of honey. After developing a protocol for the authentication of honey, its performance was tested on seven of the most common forms of honey to ascertain its advantages and limitations.

2. Materials and methods

2.1. Honey samples

Seven honey samples were examined (Table 1); six from *A. mellifera* obtained from retail outlets or apiarists in Ontario, and one from *M. beecheii* obtained from a producer in Mexico. The six samples from *A. mellifera* represented the major types of commercial honey. Light-coloured honey is the most common form and was likely the type analyzed in previous studies (e.g. Bruni et al. (2015), Jain, de Jesus, Marchioro, and de Araujo (2013), Valentini et al. (2010)). Three other liquid honeys (medium, dark, and blueberry-flavoured) were analyzed to ascertain if genetic analysis was impeded by higher levels of polyphenols or flavonoids (Pyrzynska & Biesaga, 2009). The last two samples included a liquid and a creamed honey that had undergone pasteurization and controlled-crystallization, respectively. They were included to ascertain if this additional processing affected genetic analysis. Finally, the most valuable honey produced in Mexico and Guatemala, meliponine, was analyzed to confirm that it derived from *M. beecheii* because it is exempt from the oversight imposed on honey producers (Vit et al., 2004) and is thus a target for fraud.

2.2. Preparation of honey for DNA extraction

When viewed from the perspective of DNA analysis, honey includes two components – pollen and liquid. The pollen component contains DNA from the major botanical sources, while the liquid may contain DNA from minor botanical sources and its source insect. DNA was extracted from a homogenized aliquot of liquid or semi-solid honey to examine the liquid component. For pollen analysis, DNA was extracted after concentrating the pollen grains by first heating 10 mL of honey and vortexing it with a 4:1 vol of sterile water at 56 °C. This mixture was centrifuged at 5000*g* for 30 min to pellet the pollen and the supernatant was decanted. The pollen pellet was then re-suspended in 25 mL of sterile 56 °C water and centrifuged at 5000*g* for 15 min. The supernatant was decanted and the pollen pellet was re-suspended in 1 mL of 50% ethanol, transferred to a 2 mL microfuge tube, and centrifuged at 12,000*g* for 15 min. The supernatant was decanted, and the pollen pellet was dried at 56 °C for 45 min. When DNA recovery was compared using three methods (unground, ground with 3 mm tungsten beads, and ground with sand), the first two methods produced DNA concentrations that were just 1% of those from pollen ground with sand (data not shown). Therefore, 100 mg of sterile sand (grains 0.3 mm in diameter) was added to the dried pellet, and the pollen/sand mixture was manually ground with a pestle to pulverize the pollen prior to DNA extraction.

2.3. DNA extraction

DNA was extracted using 400 μL of lysis buffer [700 mM guanidine thiocyanate (Sigma), 30 mM EDTA pH 8.0 (Fisher Scientific), 30 mM Tris-HCl pH 8.0 (Sigma), 0.5% Triton X-100 (Sigma), 5% Tween-20 (Fluka Analytical)] mixed with 2 mg/mL of Proteinase K (Promega) prior to use. This lysis buffer was added to 100 μL of liquid honey or to the ground pollen/sand mixture, followed by incubation for 18 h at 56 °C with gentle shaking.

DNA purification employed the method of Ivanova, Dewaard, and Hebert (2006) with slightly different volumes for the liquid and pollen lysates. The lysate was mixed with two volumes – 1 mL (liquid) or 800 μL (pollen) – of binding mix [3 M guanidine thiocyanate, 10 mM EDTA pH 8.0, 5 mM Tris-HCl pH 6.4, 2% Triton X-100, 50% ethanol] and applied to a silica membrane spin column (Epoch Biolabs), 700 μL at a time. The column was centrifuged at 6000*g* for 2 min after each application. The membrane was washed once with 500 μL (liquid) or 350 μL (pollen) of protein wash buffer [1.56 M guanidine thiocyanate, 5.2 mM EDTA pH 8.0, 2.6 mM Tris-HCl pH 6.4, 1.04% Triton X-100, 70% ethanol] and centrifuged at 6000*g* for 2 min. The column was washed a second time with 750 μL of wash buffer [50 mM NaCl (Fisher Scientific), 0.5 mM EDTA pH 8.0, 10 mM Tris-HCl pH 7.4, 60% ethanol] and centrifuged at 6000*g* for 4 min. The flow-through was discarded and the column was centrifuged at 10,000*g* for 4 min. The silica membrane insert was then transferred to a clean 1.5 mL microfuge tube and dried at 56 °C for 30 min.

To release DNA from the silica membrane, 40 μL (liquid) or 50 μL (pollen) of elution buffer [10 mM Tris-HCl, pH 8.0, prewarmed to 56 °C] was applied directly to the membrane and allowed to incubate at room temperature for 1 min. DNA was eluted from the column via centrifugation at 10,000*g* for 5 min. The pollen DNA was quantified using a Qubit 2.0 fluorometer (Life Technologies) and adjusted to approximately 0.5 ng/μL with elution buffer. The DNA from liquid honey was in such low abundance that it could not be quantified to allow normalization, so it was directly used for PCR.

2.4. PCR amplification for Ion Torrent sequencing

PCR amplicons require adapter sequences to enable their characterization on an Ion Torrent PGM (Life Technologies) and multiplex identifier (MID) tags to associate the sequence reads with a particular honey sample. This is typically accomplished with

Table 1
Types of honey analyzed in this study. All information regarding provenance, production, and advertised botanical and entomological sources are directly or indirectly extracted from labels on the container.

Honey type	Provenance	Produced in	Sold in	Advertised botanical source	Advertised entomological source
Light	Canada	Canada	Canada	Polyfloral	*A. mellifera*
Medium	Australia/Brazil	Australia	Canada	Orange blossom	*A. mellifera*
Dark	Canada	Canada	Canada	Buckwheat	*A. mellifera*
Blended	Canada	Canada	Canada	Polyfloral with blueberry extract	*A. mellifera*
Pasteurized	Canada	Canada	Canada	Polyfloral	*A. mellifera*
Creamed	France	France	Canada	Lavender	*A. mellifera*
Meliponine	Mexico	Mexico	Mexico	Polyfloral	*M. beecheii*

"fusion primers", gene-specific sequences tailed with adaptor/MID sequences, which are incorporated into amplicons during PCR. However, because fusion primers were less effective at amplifying target sequences than untailed primers (personal observation), a two-step PCR was employed. The first PCR used untailed primers to amplify DNA from liquid honey or pollen, while the second employed fusion primers to incorporate sequencing adapters and MID tags into the amplicons generated by the first round of PCR. PCR regimes were identical for the three markers unless otherwise stated. Table 2 provides primer details.

Initial PCR reactions consisted of 6.25 μL of 10% D-(+)-trehalose dihydrate (Fluka Analytical), 2.0 μL of Hyclone ultra-pure water (Thermo Scientific), 1.25 μL of 10X PlatinumTaq buffer (Invitrogen), 0.625 μL of 50 mM $MgCl_2$ (Invitrogen), 0.125 μL of each 10 μM primer, 0.0625 μL of 10 mM dNTP (KAPA Biosystems), 0.060 μL of 5 U/μL PlatinumTaq DNA Polymerase (Invitrogen), and 2 μL of template DNA for a total reaction volume of 12.5 μL. Thermocycling conditions were 94 °C for 5 min, 40–60 cycles (40 for ITS2, 60 for rbcLA and COI) of [94 °C for 30 s, 48–53 °C for 30 s (53 °C for ITS2, 55 °C for rbcLa, 48 °C for COI), 72 °C for 30–45 s (45 s for ITS2, 30 s for rbcLa and COI)], and a final extension of 72 °C for 10 min.

The products from the first PCR were used as template for the second PCR, which employed an identical PCR mix excepting the inclusion of fusion primers (Table 2). This PCR was run with eight replicates per reaction with a thermocycling regime of 94 °C for 4 min, 20–25 cycles (20 for ITS2, 25 for rbcLa and COI) of [94 °C for 40 s, 51–56 °C for 40 s (56 °C for ITS2, 55 °C for rbcLa, 51 °C for COI), 72 °C for 30–45 s (45 s for ITS2, 30 s for rbcLa and COI)], and a final extension of 72 °C for 5 min. Following amplification, products from the eight PCRs for each sample were pooled. For each of the seven honey samples and one negative control (n = 8), the rbcLa and COI products were pooled so they could be purified in two cleanup reactions (i.e. one for rbcLa and one for COI). The ITS2 samples were not pooled so that they could be normalized to permit an equal concentration after cleanup. This was done because ITS2 was expected to amplify differently among the seven honey samples, so the final data could be biased by over-representation of a particular product if normalization was not performed.

2.5. PCR cleanup and Ion Torrent sequencing

Amplicons were purified using a magnetic bead protocol which permitted the selective retention of particular amplicon lengths by adjusting the ratio (v:v) of magnetic beads to amplicons. The same protocol was applied to PCR products for all three markers with minor differences for ITS2. Including primer and adapter sequences, plant ITS2 amplicons are less than 500 bp, while common fungal contaminants are above 500 bp (personal observation). ITS2 cleanup was therefore optimized to select DNA between 350 bp and 500 bp, aiding removal of both shorter primer dimers and longer fungal ITS2 amplicons. For the shorter rbcLa and COI amplicons, DNA ranging from 120 bp to 500 bp was selected. To remove DNA greater than 500 bp in length, 80 μL of PCR product was mixed with 40 μL of magnetic beads (Aline Biosciences, Woburn, Massachusetts), and incubated at room temperature for 8 min. The solution was placed on a magnetic rack for 2 min and 116 μL of the supernatant was transferred to a clean 1.5 mL tube. To remove residual primers and primer dimers, 23 μL of water and 36–71 μL of fresh beads (36 μL for ITS2, 71 μL for rbcLa and COI) were added to the supernatant and incubated at room temperature for 8 min. After incubation, the solution was placed on a magnetic rack for 2 min and the supernatant was discarded. The pellet was washed three times with 300 μL of freshly prepared 80% ethanol and then air dried for 12 min. PCR products were eluted from the beads with 36 μL of sterile water, and 32 μL was transferred to a clean tube. The purified PCR products were quantified using a Qubit 2.0 fluorometer and adjusted to 1 ng/μL. For ITS2, the eight 1 ng/μL samples were mixed in equal amounts and diluted 200X. For rbcLa and COI, the 1 ng/μL samples were mixed 2:1 (v:v, rbcLa:COI) and diluted 335X. To make the final library for sequencing, 15 μL of the diluted ITS2 sample was mixed with 10 μL of the diluted rbcLa/COI sample. The library was then sequenced on an Ion Torrent PGM using a 318 v.2 chip, following the manufacturer's instructions.

2.6. Data analysis

Following sequencing, each read was associated with a particular marker and honey sample via the MID tags. This resulted in one dataset per marker (ITS2, rbcLa, COI) for each sample (n = 8), creating 24 datasets which were each run through an analytical pipeline that: (1) identified and removed primer and adapter sequences (Martin, 2011); (2) removed reads shorter than 100 bp (e.g. primer dimers) and those with low quality scores (QV < 20) (github.com/ucdavis-bioinformatics/sickle); (3) de-replicated reads with 100% identity (http://hannonlab.cshl.edu/fastx_toolkit/index.html); (4) combined reads into operational taxonomic units (OTU) with 98% identity (Edgar, 2013); and (5) assigned a taxonomic name to each OTU by performing a BLAST search using custom reference databases. A custom reference database was prepared for each marker by downloading all ITS2 and rbcLa sequences from land plants and all COI sequences from the family Apidae from GenBank on January 15, 2016. The BLAST identification results were manually inspected by removing any OTUs that were contaminants or that were unreliable due to low coverage (<10 reads) or ambiguous identification results. Raw sequence reads were uploaded to the Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra/) under BioProject PRJNA313930.

Species-level identifications for plants were only accepted if they were based on ITS2 sequences. Given intraspecific and interspecific distances reported in other ITS2 studies on plants (Liu, Zeng, Yang, Chu, & et al., 2012; Liu, Zeng, Yang, Ren, & et al., 2012; Madesis, Ganopoulos, Ralli, & Tsaftaris, 2012), species level matches are reliable if queried sequences show 98%–100% identity with reference sequences and possess sufficiently low E-values (e.g. $<10^{-50}$). For OTUs based on short rbcLa sequences (162 bp), assignments were only made to a family level because even full-length rbcLa barcodes (ca. 552 bp) often do not provide reliable assignments beyond a genus (e.g. Hawkins et al. (2015)). The source insect was considered to be *A. mellifera* or *M. beecheii* if the COI OTU sequences matched 98–100% to the appropriate species.

3. Results

All of the negative controls except one were negative for DNA. The sole exception involved the recovery of one plant OTU (*Dioscorea*) from a negative control for ITS2. Accordingly, this species was excluded from analysis of the metabarcoding data.

A total of 72 botanical sources were detected in the light honey (Table 3), but only 16 could be assigned to a species (Table 4). ITS2-based analysis of the pollen extract detected 17 plant species from nine orders and nine families, while rbcLa-based analysis of the liquid extract detected 55 plant sources from 31 orders and 55 families, including 13 of the 17 plant OTUs found in pollen (Table 4). While most of the plant sources were consistent with a Canadian provenance, four species in the pollen extract and nine of the plant families in the liquid extract do not occur in North America (Table 4), indicating that this honey was diluted with product from

Table 2
Primers used in this study. Adapter sequences are shown in bold, multiplex identifier (MID) tag sequences are underlined, and gene-specific sequences are shown in regular font. A.K.A – "also known as", name of primer in original publication if different from this study.

Primer	A.K.A	Direction	Sequence (5′ → 3′)	Purpose	Target amplicon	References
ITS2-F	ITS-S2F	Forward	ATGCGATACTTGGTGTGAAT	Initial amplification	ITS2 (350 bp)	Chen et al. (2010)
ITS2-R	ITS4	Reverse	TCCTCCGCTTATTGATATGC	Initial amplification	ITS2 (350 bp)	White, Bruns, Lee, and Taylor (1990)
ITS2-F1	ITS-S2F-trP1	Forward	**CCTCTCTATGGGCAGTCGGTGAT**ATGCGATACTTGGTGTGAAT	Sequencing adapter fusion	ITS2 (350 bp)	Chen et al. (2010)
ITS2-R1	ITS4-ion29	Reverse	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**TCGACCACTCTCCTCCGCTTATTGATATGC	Sequencing adapter fusion	ITS2 (350 bp)	White et al. (1990)
ITS2-R2	ITS4-ion30	Reverse	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**CGAGGTTATCTCCTCCGCTTATTGATATGC	Sequencing adapter fusion	ITS2 (350 bp)	White et al. (1990)
ITS2-R3	ITS4-ion31	Reverse	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**TCCAAGCTGCTCCTCCGCTTATTGATATGC	Sequencing adapter fusion	ITS2 (350 bp)	White et al. (1990)
ITS2-R4	ITS4-ion32	Reverse	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**TCTTACACACTCCTCCGCTTATTGATATGC	Sequencing adapter fusion	ITS2 (350 bp)	White et al. (1990)
ITS2-R5	ITS4-ion33	Reverse	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**TTCTCATTGAACTCCTCCGCTTATTGATATGC	Sequencing adapter fusion	ITS2 (350 bp)	White et al. (1990)
ITS2-R6	ITS4-ion34	Reverse	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**TCGCATCGTTCTCCTCCGCTTATTGATATGC	Sequencing adapter fusion	ITS2 (350 bp)	White et al. (1990)
ITS2-R7	ITS4-ion36	Reverse	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**AAGGAATCGTCTCCTCCGCTTATTGATATGC	Sequencing adapter fusion	ITS2 (350 bp)	White et al. (1990)
ITS2-R8	ITS4-ion37	Reverse	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**CTTGAGAATGTCTCCTCCGCTTATTGATATGC	Sequencing adapter fusion	ITS2 (350 bp)	White et al. (1990)
rbcLa-F	N/A	Forward	ATGTCACCACAAACAGAGACTAAAGC	Initial amplification	rbcLa (163 bp)	Levin et al. (2003)
rbcLa-R	MrbcL163-R1	Reverse	CGGTCCAYACAGYBGTCCAKGTACC	Initial amplification	rbcLa (163 bp)	Ivanova, Kuzmina, Braukmann, Borisenko, and Zakharov (2016)
rbcLa-F1	rbcLa-F-ion11	Forward	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**TCCTCGAATCATGTCACCACAAACAGAGACTAAAGC	Sequencing adapter fusion	rbcLa (163 bp)	Levin et al. (2003)
rbcLa-F2	rbcLa-F-ion12	Forward	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**TAGGTGGTTCATGTCACCACAAACAGAGACTAAAGC	Sequencing adapter fusion	rbcLa (163 bp)	Levin et al. (2003)
rbcLa-F3	rbcLa-F-ion13	Forward	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**TCTAACGGACATGTCACCACAAACAGAGACTAAAGC	Sequencing adapter fusion	rbcLa (163 bp)	Levin et al. (2003)
rbcLa-F4	rbcLa-F-ion14	Forward	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**TTGGAGTGTCATGTCACCACAAACAGAGACTAAAGC	Sequencing adapter fusion	rbcLa (163 bp)	Levin et al. (2003)
rbcLa-F5	rbcLa-F-ion15	Forward	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**TCTAGAGGTCATGTCACCACAAACAGAGACTAAAGC	Sequencing adapter fusion	rbcLa (163 bp)	Levin et al. (2003)
rbcLa-F6	rbcLa-F-ion16	Forward	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**TCTGGATGACATGTCACCACAAACAGAGACTAAAGC	Sequencing adapter fusion	rbcLa (163 bp)	Levin et al. (2003)
rbcLa-F7	rbcLa-F-ion17	Forward	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**TCTATTCGTCATGTCACCACAAACAGAGACTAAAGC	Sequencing adapter fusion	rbcLa (163 bp)	Levin et al. (2003)
rbcLa-F8	rbcLa-F-ion18	Forward	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**AGGCAATTGCATGTCACCACAAACAGAGACTAAAGC	Sequencing adapter fusion	rbcLa (163 bp)	Levin et al. (2003)
rbcLa-R1	MrbcL163-R1-ion1-trP1	Reverse	**CCTCTCTATGGGCAGTCGGTGAT**CTAAGGTAACCGGTCCAYACAGYBGTCCAKGTACC	Sequencing adapter fusion	rbcLa (163 bp)	Ivanova et al. (2016)
COI-F	AncientLepF2	Forward	ATTRRWRATGATCAARTWTATAAT	Initial amplification	COI (120 bp)	Hebert et al. (2013)
COI-R	Apis238_R1	Reverse	TAATCAAAATCTAATATTATTTATTCG	Initial amplification	COI (120 bp)	This study

Table 2 (*continued*)

Primer	A.K.A	Direction	Sequence (5′ → 3′)	Purpose	Target amplicon	References
COI-F1	AncientLepF2-ion1	Forward	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**CTAAGGTAACATTRRWRATGATCAARTWTATAAT	Sequencing adapter fusion	COI (120 bp)	Prosser, DeWaard, Miller, and Hebert (2016)
COI-F2	AncientLepF2-ion2	Forward	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**TAAGGAGAACATTRRWRATGATCAARTWTATAAT	Sequencing adapter fusion	COI (120 bp)	Prosser et al. (2016)
COI-F3	AncientLepF2-ion3	Forward	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**AAGAGGATTCATTRRWRATGATCAARTWTATAAT	Sequencing adapter fusion	COI (120 bp)	Prosser et al. (2016)
COI-F4	AncientLepF2-ion4	Forward	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**TACCAAGATCATTRRWRATGATCAARTWTATAAT	Sequencing adapter fusion	COI (120 bp)	Prosser et al. (2016)
COI-F5	AncientLepF2-ion5	Forward	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**CAGAAGGAACATTRRWRATGATCAARTWTATAAT	Sequencing adapter fusion	COI (120 bp)	Prosser et al. (2016)
COI-F6	AncientLepF2-ion6	Forward	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**CTGCAAGTTCATTRRWRATGATCAARTWTATAAT	Sequencing adapter fusion	COI (120 bp)	Prosser et al. (2016)
COI-F7	AncientLepF2-ion7	Forward	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**TTCGTGATTCATTRRWRATGATCAARTWTATAAT	Sequencing adapter fusion	COI (120 bp)	Prosser et al. (2016)
COI-F8	AncientLepF2-ion8	Forward	**CCATCTCATCCCTGCGTGTCTCCGACTCAG**TTCCGATAACATTRRWRATGATCAARTWTATAAT	Sequencing adapter fusion	COI (120 bp)	Prosser et al. (2016)
COI-R1	Apis238_R1-ion1-trP1	Reverse	**CCTCTCTATGGGCAGTCGGTGAT**CTAAGGTAACTAATCAAAATCTAATATTATTTATTCG	Sequencing adapter fusion	COI (120 bp)	This study

Table 3
The number of reads, before and after quality control (QC) filtering, and number of operational taxonomic units (OTU) detected for each honey sample and negative control. The number of reads and OTUs are shown separately for each marker. The honey fraction (pollen or liquid) to which each marker was applied is shown in parenthesis. OTUs were counted only if they were composed of ten or more reads and were not obvious contaminants.

Sample	No. reads (raw data)			No. reads (after QC vetting)			No. OTUs (coverage >10)			
	ITS2 (pollen)	rbcLa (liquid)	COI (liquid)	ITS2 (pollen)	rbcLa (liquid)	COI (liquid)	ITS2 (pollen)	rbcLa (liquid)	COI (liquid)	Total
Light	391877	379240	238257	65931	254410	94121	17	55	1	73
Medium	413045	288808	203093	227750	202464	77234	1	1	1	3
Dark	24	313821	176469	0	190385	61613	0	4	1	5
Blended	50	43454	189576	0	2	66930	0	0	1	1
Pasteurized	823034	232255	209331	103583	153496	81598	0	2	1	3
Creamed	268695	47862	0	20947	24426	0	2	1	0	3
Meliponine	132612	320599	183412	7084	196506	95317	4	59	1	64
Negative control	1668	45207	4	799	8	0	1	0	0	1

another region. Based on the foreign plant taxa detected, the adulterant honey likely originated from China. The entomological source was confirmed as *A. mellifera* (Table 4).

In contrast to the light honey, only two botanical sources were detected in the medium honey: one from the pollen extract and one from the liquid extract (Table 3). The pollen-derived ITS2 sequence matched several *Eucalyptus* species, reflecting the fact that many members of this genus are very closely related, a factor precluding species-level identification. The botanical source from the liquid extract matched the citrus family Rutaceae (Table 4), supporting this honey's origin from orange blossoms. All botanical sources are consistent with its purported Australian/Brazilian provenance. The entomological source was confirmed as *A. mellifera* (Table 4).

The pollen extract from the dark honey failed to generate any sequences, but DNA from the liquid extract revealed four botanical sources (Table 3) representing four orders (Table 4). These plants were consistent with a Canadian provenance, but did not include buckwheat (*Fagopyrum esculentum*), despite this honey having been sold as a monofloral of this species. The failure to detect *F. esculentum*, or to amplify any DNA from the pollen extract, might reflect inhibition by the high level of polyphenols in this honey. The entomological source was verified as *A. mellifera* (Table 4).

No botanical sources were recovered from the honey blended with blueberry extract (Table 3), but the presence of *A. mellifera* DNA was confirmed (Table 4). The failure to detect botanical sources was due to a near failure of both ITS2 and rbcLa amplification. Although pollen grains were apparent during sample preparation, only 50 ITS2 reads were produced, none of which were retained after the quality control filter (Table 3). This result, and the fact that too few rbcLa reads were generated to produce reliable OTUs (Table 3), suggests that PCR amplification was inhibited, probably by blueberry-derived flavonoids that were co-purified with the DNA. The recovery of DNA from *A. mellifera* may mean that its concentration was higher in the honey.

The pasteurized honey contained very little pollen, suggesting that it had been filtered in addition to being pasteurized. As a consequence, no botanical sources were identified through analysis of the pollen extract (Table 3), but two plants representing two families from two orders (Table 4) were identified from the liquid extract. Both were consistent with Canadian origin and the entomological source was *A. mellifera* (Table 4).

Table 4
Identifications of the botanical and entomological sources detected in each honey. All identifications are based on unambiguous BLAST results of operational taxonomic units (OTUs) with >10 reads. All identifications based on ITS2 are from the analysis of pollen while those based on rbcLa or COI are from the analysis of liquid honey. The taxonomic rank and common names of each source are shown (if applicable), as well as the potential provenance of each source (potential provenance is shown in italics if a source is not local where the honey was reportedly produced).

Honey	Marker	Rank	Identification	Common name	Provenance
Light	COI	Species	*Apis mellifera*	European honey bee	Canada
Light	ITS2	Species	*Arctium lappa*	Greater Burdock	Canada
Light	ITS2	Species	*Barbarea orthoceras*	American yellowrocket	Canada
Light	ITS2	Genus	*Brassica* sp.	Cabbages et al.	Canada
Light	ITS2	Species	*Fritillaria verticillata*	Lily	*Asia*
Light	ITS2	Species	*Gliricidia sepium*	Quickstick	*Asia, South America*
Light	ITS2	Species	*Glycine max*	Soybean	Canada
Light	ITS2	Species	*Iris loczyi*	Iris	*Asia*
Light	ITS2	Species	*Lens culinaris*	Lentil	Canada
Light	ITS2	Species	*Lotus corniculatus*	Bird's-foot deervetch	Canada
Light	ITS2	Species	*Melilotus dentatus*	Sweet clover	Canada
Light	ITS2	Species	*Myosotis sylvatica*	Forget-me-not	Canada
Light	ITS2	Species	*Papaver rhoeas*	Common poppy	Canada
Light	ITS2	Species	*Rhus microphylla*	Sumac	Canada
Light	ITS2	Species	*Rosa chinensis*	China rose	*China*
Light	ITS2	Species	*Rubus allegheniensis*	Blackberry	Canada
Light	ITS2	Species	*Trifolium hybridum*	Clover	Canada
Light	ITS2	Species	*Vicia cracca*	Cow vetch	Canada
Light	rbcLa	Family	Acanthaceae	Acanthus	Canada
Light	rbcLa	Family	Aceraceae	Maple tree	Canada
Light	rbcLa	Family	Aizoaceae	Fig-marigold	*Asia, Africa*
Light	rbcLa	Family	Anacardiaceae	Sumac	Canada
Light	rbcLa	Family	Apiaceae	Parsnip	Canada
Light	rbcLa	Family	Apocynaceae	Dogbane	Canada
Light	rbcLa	Family	Araliaceae	Ginseng	Canada
Light	rbcLa	Family	Asparagaceae	Asparagus	Canada
Light	rbcLa	Family	Asteraceae	Aster	Canada
Light	rbcLa	Family	Brassicaceae	Mustard/crucifers	Canada
Light	rbcLa	Family	Caprifoliaceae	Honeysuckle	Canada
Light	rbcLa	Family	Caricaceae	Papaya	Canada
Light	rbcLa	Family	Celastraceae	Bittersweet/staff vine	Canada
Light	rbcLa	Family	Comandraceae	Bastard Toadflax	Canada
Light	rbcLa	Family	Cornaceae	Dogwood	Canada
Light	rbcLa	Family	Corynocarpaceae	Karaka	*Tropics*
Light	rbcLa	Family	Crassulaceae	Stonecrop	Canada
Light	rbcLa	Family	Cucurbitaceae	Gourds	Canada
Light	rbcLa	Family	Cyclanthaceae	Panama hat	Canada
Light	rbcLa	Family	Dilleniaceae	Dillenia	*Tropics*
Light	rbcLa	Family	Ebenaceae	Ebony	Canada
Light	rbcLa	Family	Elaeocarpaceae	Elaeocarpus	Canada
Light	rbcLa	Family	Euphorbiaceae	Spurge	Canada
Light	rbcLa	Family	Fabaceae	Legumes	Canada
Light	rbcLa	Family	Goodeniaceae	Goodenia	Canada
Light	rbcLa	Family	Grossulariaceae	Currant	Canada
Light	rbcLa	Family	Icacinaceae	Icacina	*Tropics*
Light	rbcLa	Family	Lamiaceae	Mint	Canada
Light	rbcLa	Family	Lauraceae	Laurel	Canada
Light	rbcLa	Family	Malvaceae	Mallow	Canada
Light	rbcLa	Family	Meliaceae	Mahogany	Canada
Light	rbcLa	Family	Moraceae	Mulberry	Canada
Light	rbcLa	Family	Moringaceae	Horseradish	Canada
Light	rbcLa	Family	Myrtaceae	Myrtle	Canada
Light	rbcLa	Family	Nelumbonaceae	Lotus lily	Canada
Light	rbcLa	Family	Oleaceae	Olive	Canada
Light	rbcLa	Family	Pandanaceae	Screw-pine/pandan	*Asia*
Light	rbcLa	Family	Phyllanthaceae		*Tropics*
Light	rbcLa	Family	Platanaceae	Plane tree	Canada
Light	rbcLa	Family	Poaceae	Grass	Canada
Light	rbcLa	Family	Potamogetonaceae	Pondweed	Canada
Light	rbcLa	Family	Rhamnaceae	Buckthorn	Canada
Light	rbcLa	Family	Ripogonaceae	Ripogonum	*Asia, Australia*
Light	rbcLa	Family	Rosaceae	Rose	Canada
Light	rbcLa	Family	Rubiaceae	Coffee	Canada
Light	rbcLa	Family	Rutaceae	Citrus	Canada
Light	rbcLa	Family	Salicaceae	Willow	Canada
Light	rbcLa	Family	Sapindaceae	Soapberry	Canada
Light	rbcLa	Family	Solanaceae	Tomato, potato, nightshade	Canada
Light	rbcLa	Family	Tetramelaceae		*Asia*
Light	rbcLa	Family	Theaceae	Tea	*Asia, South America*
Light	rbcLa	Family	Thymelaeaceae	Mezereum	Canada
Light	rbcLa	Family	Ulmaceae	Elm	Canada

Table 4 (*continued*)

Honey	Marker	Rank	Identification	Common name	Provenance
Light	rbcLa	Family	Verbenaceae	Verbena	Canada
Light	rbcLa	Family	Vitaceae	Grape	Canada
Medium	COI	Species	*Apis mellifera*	European honey bee	Australia/Brazil
Medium	ITS2	Genus	*Eucalyptus* sp.		Australia
Medium	rbcLa	Family	Rutaceae	Citrus	Australia/Brazil
Dark	COI	Species	*Apis mellifera*	European honey bee	Global
Dark	rbcLa	Family	Araliaceae	Ivy	Canada
Dark	rbcLa	Family	Fabaceae	Legumes	Canada
Dark	rbcLa	Family	Lamiaceae	Mint	Canada
Dark	rbcLa	Family	Solanaceae	Tomato, potato, nightshade	Canada
Blended	COI	Species	*Apis mellifera*	European honey bee	Global
Pasteurized	COI	Species	*Apis mellifera*	European honey bee	Global
Pasteurized	rbcLa	Family	Asteraceae	Aster	Canada
Pasteurized	rbcLa	Family	Rosaceae	Rose	Canada
Creamed	ITS2	Genus	*Prunus* sp.	Cherry	France
Creamed	ITS2	Genus	*Rhus* sp.	Sumac	France
Creamed	rbcLa	Family	Anacardiaceae		France
Meliponine	COI	Species	*Melipona beecheii*	Stingless bee	Mexico
Meliponine	ITS2	Genus	*Avena* sp.		Mexico
Meliponine	ITS2	Genus	*Elymus* sp.		Mexico
Meliponine	ITS2	Species	*Glycine max*	Soybean	Mexico
Meliponine	ITS2	Species	*Pueraria montana*	Kudzu	Mexico
Meliponine	rbcLa	Family	Aceraceae	Maple tree	Mexico
Meliponine	rbcLa	Family	Amaranthaceae	Amaranth	Mexico
Meliponine	rbcLa	Family	Anacardiaceae	Sumac	Mexico
Meliponine	rbcLa	Family	Annonaceae	Custard apple	Mexico
Meliponine	rbcLa	Family	Apiaceae	Parsnip	Mexico
Meliponine	rbcLa	Family	Apocynaceae	Dogbane	Mexico
Meliponine	rbcLa	Family	Arecaceae	Palm tree	Mexico
Meliponine	rbcLa	Family	Asteraceae	Aster	Mexico
Meliponine	rbcLa	Family	Brassicaceae	Mustard/crucifers	Mexico
Meliponine	rbcLa	Family	Bromeliaceae	Bromeliads	Mexico
Meliponine	rbcLa	Family	Burseraceae	Torchwood	Mexico
Meliponine	rbcLa	Family	Cannabaceae	Hemp	Mexico
Meliponine	rbcLa	Family	Caprifoliaceae	Honeysuckle	Mexico
Meliponine	rbcLa	Family	Caricaceae	Papaya	Mexico
Meliponine	rbcLa	Family	Connaraceae	Cannarus	Mexico
Meliponine	rbcLa	Family	Cornaceae	Dogwood	Mexico
Meliponine	rbcLa	Family	Corynocarpaceae	Karaka	Mexico
Meliponine	rbcLa	Family	Crassulaceae	Stonecrop	Mexico
Meliponine	rbcLa	Family	Cucurbitaceae	Gourds	Mexico
Meliponine	rbcLa	Family	Cyclanthaceae	Panama hat	Mexico
Meliponine	rbcLa	Family	Ebenaceae	Ebony	Mexico
Meliponine	rbcLa	Family	Elaeocarpaceae	Elaeocarpus	Mexico
Meliponine	rbcLa	Family	Erythroxylaceae	Coca	Mexico
Meliponine	rbcLa	Family	Euphorbiaceae	Spurge	Mexico
Meliponine	rbcLa	Family	Fabaceae	Legumes	Mexico
Meliponine	rbcLa	Family	Gesneriaceae	Gesneriad	Mexico
Meliponine	rbcLa	Family	Goodeniaceae	Goodenia	Mexico
Meliponine	rbcLa	Family	Haemodoraceae	Bloodwort	Mexico
Meliponine	rbcLa	Family	Icacinaceae	Icacina	Mexico
Meliponine	rbcLa	Family	Loganiaceae	Logania	Mexico
Meliponine	rbcLa	Family	Loranthaceae	Mistletoe	Mexico
Meliponine	rbcLa	Family	Lythraceae	Loosestrife	Mexico
Meliponine	rbcLa	Family	Magnoliaceae	Magnolia	Mexico
Meliponine	rbcLa	Family	Melanthiaceae	Trillium	Mexico
Meliponine	rbcLa	Family	Melastomataceae	Melastome	Mexico
Meliponine	rbcLa	Family	Meliaceae	Mahogany	Mexico
Meliponine	rbcLa	Family	Moringaceae	Horseradish	Mexico
Meliponine	rbcLa	Family	Myrtaceae	Myrtle	Mexico
Meliponine	rbcLa	Family	Nothofagaceae	Southern beeches	Mexico
Meliponine	rbcLa	Family	Papaveraceae	Poppy	Mexico
Meliponine	rbcLa	Family	Peraceae		Mexico
Meliponine	rbcLa	Family	Phyllanthaceae		Mexico
Meliponine	rbcLa	Family	Poaceae	Grass	Mexico
Meliponine	rbcLa	Family	Potamogetonaceae	Pondweed	Mexico
Meliponine	rbcLa	Family	Rhizophoraceae	Mangroves	Mexico
Meliponine	rbcLa	Family	Rosaceae	Rose	Mexico
Meliponine	rbcLa	Family	Rubiaceae	Madder	Mexico
Meliponine	rbcLa	Family	Rutaceae	Citrus	Mexico
Meliponine	rbcLa	Family	Salicaceae	Willow	Mexico
Meliponine	rbcLa	Family	Sapindaceae	Soapberry	Mexico
Meliponine	rbcLa	Family	Sapotaceae	Sapodilla	Mexico
Meliponine	rbcLa	Family	Saxifragaceae	Saxifrage	Mexico
Meliponine	rbcLa	Family	Solanaceae	Tomato, potato, nightshade	Mexico

(*continued on next page*)

Table 4 (*continued*)

Honey	Marker	Rank	Identification	Common name	Provenance
Meliponine	rbcLa	Family	Stemonuraceae		Mexico
Meliponine	rbcLa	Family	Thymelaeaceae	Mezereum	Mexico
Meliponine	rbcLa	Family	Ulmaceae	Elm	Mexico
Meliponine	rbcLa	Family	Urticaceae	Nettle	Mexico
Meliponine	rbcLa	Family	Violaceae	Violet	Mexico
Meliponine	rbcLa	Family	Winteraceae	Wintera	Mexico

Three botanical sources were detected in the creamed honey, two from the pollen extract and one from the liquid extract (Table 3). One ITS2 (i.e. pollen) sequence matched closely to three species of cherry (*Prunus*), while the other matched several closely allied species of sumac (*Rhus*) (Table 4). The single botanical source detected from the liquid extract derived from the same family as *Rhus* and likely also reflects the detection of sumac. All botanical sources are consistent with French provenance, but no DNA sequences were recovered from *Lavandula* (family Lamiaceae) despite this honey's supposed monofloral origin from lavender. Furthermore, the entomological source could not be determined, challenging the authenticity of this honey.

The meliponine honey contained 63 detectable botanical sources (Table 3), although only two could be identified to species by ITS2-based pollen analysis (Table 4). Pollen analysis yielded four botanical sources from two families – oat (*Avena* sp.), soybean (*Glycine max*), kudzu (*Pueraria montana*), and wheatgrass (*Elymus* sp.) – although the oat and wheatgrass sources could only be identified to genus because their ITS2 sequences showed a >98% match to multiple species in their respective genera. Analysis of the liquid extract detected 59 botanical sources from 32 plant orders (Table 4), including the family to which soybean and kudzu belong. All botanical sources were consistent with a Mexican provenance and the entomological source was confirmed as *M. beecheii* (Table 4).

4. Discussion

Authentication of honey has traditionally relied upon melissopalynology and/or the biochemical analysis of carbohydrates (Tosun, 2013), proteins (Cotte et al., 2004), polyphenols (Campone et al., 2014), volatile organics (Spanik, Pazitna, Siska, & Szolcsanyi, 2014), and minerals (Arvanitoyannis, Chalhoub, Gotsiou, Lydakis-Simantiris, & Kefalas, 2005). However, melissopalynology cannot discriminate many plant species, while biochemical analysis is generally limited to determining the presence/absence of a specific plant species via a diagnostic chemical marker (Kaskoniene & Venskutonis, 2010).

Genetic analysis, specifically metabarcoding, combines the broad applicability of melissopalynology with the specificity of biochemical analysis, and is consequently gaining adoption as a standard tool for honey authentication. However, like chemical markers, the choice of genetic marker can significantly affect the taxonomic resolution of the assay. Early metabarcoding studies (Valentini et al., 2010) which used the trnL-UAA intron marker were unable to identify many plants beyond a family level. Hawkins et al. (2015) applied a different marker (rbcLa) to improve resolution, but it only delivered a species-level identification for one third of the taxa. Galimberti et al. (2014) and Bruni et al. (2015) increased taxonomic resolution by analyzing both rbcLa and the trnH-psbA spacer region, but the latter marker required some *a priori* knowledge of the botanical composition of the honey to generate the required reference library. Richardson et al. (2015) compared the results from melissopalynology with those obtained by metabarcoding honey with another gene region (ITS2), and concluded that genetic analysis was superior at identifying plants below the family level.

While several studies have demonstrated the usefulness of metabarcoding in ascertaining the botanical sources of honey, they only documented species recovered from pollen extracted from fresh honey. By comparison, the method presented here successfully analyzed both the botanical (pollen and pollen-free) and entomological sources of DNA in five of seven types of honey. The present results indicate that sequence recovery is more difficult from dark or flavoured honeys, likely reflecting inhibition by plant secondary compounds on PCR. However, since only three polyphenol-rich honeys were analyzed, further work is required to confirm that such honeys are always recalcitrant. If so, the inclusion of CTAB or PVP in the DNA extraction buffer might overcome the problem.

Pasteurization did not impede sequence recovery, as demonstrated by the detection of botanical and entomological DNA from pasteurized liquid honey. However, when honey is filtered to remove impurities, pollen is also removed, which could explain the recovery of just one plant source in the pollen fraction. Conversely, creamed honey only delivered sequences from the pollen fraction, suggesting that DNA does not persist in the crystal matrix of creamed honeys. Consequently, identification of the entomological source of creamed honeys may require protein analysis (Ramón-Sierra et al., 2015). Nevertheless, for non-creamed honeys, genetic analysis of the liquid fraction is the simplest method to determine the entomological source as well as botanical components not represented, or not detectable, in the pollen fraction (e.g. the "orange blossom" honey included in this study was only validated by analysis of the liquid fraction). Highly processed honeys dominate the retail market (Everstine, Spink, & Kennedy, 2013), and the present study has revealed the difficulty in recovering sequence information from them.

Most of the plant sources identified in this study were consistent with nectar collected during foraging, but the recovery of DNA from some wind-pollinated plants (e.g. members of Poaceae) mirrors results from other studies (Bruni et al., 2015; Hawkins et al., 2015). Such occurrences likely reflect non-specific incorporation by the bees (i.e. wind-blown pollen adheres to bees as they forage (Corbet, Beament, & Eisikowitch, 1982) and is subsequently incorporated into honey (Jones, 2010)).

While metabarcoding analysis enables the identification of the plant sources of honey, the number of sequence reads for a particular species is weakly correlated with its contribution to the honey. For example, Richardson et al. (2015) found no correlation between the number of reads for a species and its representation in the pollen fraction. The number of sequence reads from a particular species is heavily influenced by the number of original DNA templates. Both the number of ITS2 copies per nuclear genome and the number of copies of plastid genomes per cell varies among species, reflecting different forms of parental inheritance (Bell et al., 2016). Therefore, a scenario where the amount of amplification accurately reflects the degree to which all plant species contribute to a honey is unlikely. As well, even slight primer-template mismatches – an inevitable situation when using universal primers – can introduce significant amplification bias,

further affecting final read counts. Because of these factors, conclusions cannot be drawn about the abundance of each plant species in honey. Melissopalynology remains the best approach to quantify pollen in honey.

Although the protocol described here aids honey authentication, it has the same limitations as any method of genetic analysis involving pollen: it can be partially circumvented by replacing source pollen with a small amount of pollen from the desired plant. However, because pollen-free DNA in the liquid fraction of honey cannot be removed by filtration, evidence of the original plant and insect sources will likely be detectable using the present method. Indeed, analysis of the liquid extract from the light honey in this study revealed DNA from both temperate and tropical plants, indicating a blend of temperate and tropical honeys. Even without the analysis of pollen, the present method would have detected this case of adulteration. While the source components of honey cannot be fully masked, DNA can be added to ensure that a fraudulent honey has the genetic signatures of desired plants and insects. As a consequence, DNA metabarcoding needs to be coupled with biochemical analysis and palynology to maximize the prospects for authentication.

Conflict of interest

The authors declare no conflicts of interest.

Acknowledgments

Funding for this project was provided by the Ontario Ministry of Research and Innovation. We thank Thomas Braukmann for developing the bioinformatics platform used in this study, and Kylee Ronnenberg for assistance with data analysis and interpretation.

References

Ahmed, S., & Othman, N. H. (2013). Review of the medicinal effects of Tualang honey and a comparison with Manuka honey. *The Malaysian Journal of Medical Sciences, 20*(3), 6–13.

Arvanitoyannis, S., Chalhoub, C., Gotsiou, P., Lydakis-Simantiris, N., & Kefalas, P. (2005). Novel quality control methods in conjunction with chemometrics (multivariate analysis) for detecting honey authenticity. *Critical Reviews in Food Science and Nutrition, 45*(3), 193–203.

Bell, K. L., Burgess, K. S., Okamoto, K. C., Aranda, R., & Brosi, B. J. (2016). Review and future prospects for DNA barcoding methods in forensic palynology. *Forensic Science International: Genetics, 21*, 110–116.

Bruni, I., Galimberti, A., Caridi, L., Scaccabarozzi, D., De Mattia, F., Casiraghi, M., & Labra, M. (2015). A DNA barcoding approach to identify plant species in multiflower honey. *Food Chemistry, 170*, 308–315.

Campone, L., Piccinelli, A. L., Pagano, I., Carabetta, S., Di Sanzo, R., Russo, M., & Rastrelli, L. (2014). Determination of phenolic compounds in honey using dispersive liquid-liquid microextraction. *Journal of Chromatography A, 1334*, 9–15.

Chen, S., Yao, H., Han, J., Liu, C., Song, J., Zhu, Y., ... Leon, C. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS ONE, 5*(1), e8613.

Corbet, S. A., Beament, J., & Eisikowitch, D. (1982). Are electrostatic forces involved in pollen transfer? *Plant, Cell and Environment, 5*, 125–129.

Cotte, J. F., Casabianca, H., Giroud, B., Albert, M., Lheritier, J., & Grenier-Loustalot, M. F. (2004). Characterization of honey amino acid profiles using high-pressure liquid chromatography to control authenticity. *Analytical and Bioanalytical Chemistry, 378*, 1342–1350.

de Franca Alves, R., & de Assis Ribeiro dos Santos, F. (2014). Plant sources for bee pollen load production in Sergipe, northeast Brazil. *Palynology, 38*(1), 90–100.

Edgar, R. C. (2013). UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nature Methods, 10*(10), 996–998.

Everstine, K., Spink, J., & Kennedy, S. (2013). Economically motivated adulteration (EMA) of food: common characteristics of EMA incidents. *Journal of Food Protection, 76*(4), 723–735.

Galimberti, A., De Mattia, F., Bruni, I., Scaccabarozzi, D., Sandionigi, A., Barbuto, M., ... Labra, M. (2014). A DNA barcoding approach to characterize pollen collected by honeybees. *PLoS ONE, 9*(10), e109363. http://dx.doi.org/10.1371/journal.pone.0109363.

Hawkins, J., de Vere, N., Griffith, A., Ford, C. R., Allainguillaume, J., Hegarty, M. J., ... Adams-Groom, B. (2015). Using DNA metabarcoding to identify the floral composition of honey: A new tool for investigating honey bee foraging preferences. *PLoS ONE, 10*(8), e0134735. http://dx.doi.org/10.1371/journal.pone.0134735.

Hebert, P. D. N., deWaard, J. R., Zakharov, E. V., Prosser, S. W. J., Sones, J. E., McKeown, J. T. A., ... La Salle, J. (2013). A DNA 'BarcodeBlitz': Rapid digitization and sequencing of a natural history collection. *PLoS ONE, 8*, e68535.

Ivanova, N. V., Kuzmina, M. L., Braukmann, T. W. A., Borisenko, A. V., & Zakharov, E. V. (2016). Authentication of herbal supplements using next-generation sequencing. *PLoS ONE, 11*(5), e0156426.

Ivanova, N. V., Dewaard, J. R., & Hebert, P. D. N. (2006). An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Molecular Ecology Notes, 6*, 998–1002.

Jain, S. A., de Jesus, F. T., Marchioro, G. M., & de Araujo, E. D. (2013). Extraction of DNA from honey and its amplification by PCR for botanical identification. *Food Science and Technology, 33*(4), 753–756.

Jones, R. (Ed.). (2010). *The beekeeper's bible: Bees, recipes & other home uses*. Richard Jones & Sharon Sweeney-Lynch.

Kaskoniene, V., & Venskutonis, P. R. (2010). Floral markers in honey of various botanical and geographic origins: A review. *Comprehensive Reviews in Food Science and Food Safety, 9*, 620–634.

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.

Levin, R. A., Wagner, W. L., Hoch, P. C., Nepokroeff, M., Pires, J. C., Zimmer, E. A., & Sytsma, K. J. (2003). Family-level relationships of Onagraceae based on chloroplast rbcL and ndhF data. *American Journal of Botany, 90*, 107–115.

Liu, Z., Zeng, X., Yang, D., Chu, G., Yuan, Z., & Chen, S. (2012). Applying DNA barcodes for identification of plant species in the family Araliaceae. *Gene, 499*, 76–80.

Liu, Z., Zeng, X., Yang, D., Ren, G., Chu, G., Yuan, Z., ... Chen, S. (2012). Identification of medicinal vines by ITS2 using complementary discrimination methods. *Journal of Ethnopharmacology, 141*, 242–249.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet Journal, 17*(1), 10–12.

Madesis, P., Ganopoulos, I., Ralli, P., & Tsaftaris, A. (2012). Barcoding the major Mediterranean leguminous crops by combining universal chloroplast and nuclear DNA sequence targets. *Genetics and Molecular Research, 11*(3), 2548–2558.

Prosser, S. W. J., DeWaard, J. R., Miller, S. E., & Hebert, P. D. N. (2016). DNA barcodes from century-old type specimens using next-generation sequencing. *Molecular Ecology Resources, 16*, 487–497.

Pyrzynska, K., & Biesaga, M. (2009). Analysis of phenolic acids and flavonoids in honey. *Trends in Analytical Chemistry, 28*(7), 893–902.

Ramón-Sierra, J. M., Ruiz-Ruiz, J. C., & de la Luz Ortiz-Vázquez, E. (2015). Electrophoresis characterisation of protein as a method to establish the entomological origin of stingless bee honeys. *Food Chemistry, 183*, 43–48.

Richardson, R. T., Lin, C. H., Sponsler, D. B., Quijia, J. O., Goodell, K., & Johnson, R. M. (2015). Application of ITS2 metabarcoding to determine the provenance of pollen collected by honey bees in an agroecosystem. *Applications in Plant Sciences, 3*(1), 1400066.

Schnell, I. B., Fraser, M., Willerslev, E., & Gilbert, M. T. P. (2010). Characterisation of insect and plant origins using DNA extracted from small volumes of bee honey. *Arthropod-Plant Interactions, 4*(2), 107–116.

Spanik, I., Pazitna, A., Siska, P., & Szolcsanyi, P. (2014). The determination of botanical origin of honeys based on enantiomer distribution of chiral volatile organic compounds. *Food Chemistry, 158*, 497–503.

Tosun, M. (2013). Detection of adulteration in honey samples added various sugar syrups with 13C/12C isotope ratio analysis method. *Food Chemistry, 138*, 1629–1632.

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

Vit, P., Medina, M., & Enriquez, M. E. (2004). Quality standards for medicinal uses of Meliponinae honey in Guatemala, Mexico and Venezuela. *Bee World, 85*(1), 2–5.

White, T., Bruns, T., Lee, S., & Taylor, J. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In M. A. Innis, D. H. Gelfand, J. J. Sninsky, & T. J. White (Eds.), *PCR protocols: A guide to methods and applications* (pp. 315–322). New York: Academic Press.

Yaghoobi, R., Kazerouni, A., & Kazerouni, O. (2013). Evidence for clinical use of honey in wound healing as an anti-bacterial, anti-inflammatory, anti-oxidant and anti-viral agent: A review. *Jundishapur Journal of Natural Pharmaceutical Products, 8*(3), 100–104.

Yao, H., Song, J., Liu, C., Luo, K., Han, J., Li, Y., ... Chen, S. (2010). Use of ITS2 region as the universal DNA barcode for plants and animals. *PLoS ONE, 5*(10), e13102.