FEASR

REGIONE DEL VENETO

Fondo europeo agricolo per lo sviluppo rurale: l'Europa investe nelle zone rurali

Reg. (CE) n. 1698/2005 - PSR 2007 - 2013
DGR n. 199 del 12/02/08
MISURA 124 – "Cooperazione per lo sviluppo di nuovi prodotti, processi e tecnologie nel settore agricolo, alimentare e forestale"

Progetto UNIMIELE – Relazione finale

Sviluppo e messa a punto del metodo di indagine.

Sono state standardizzate le procedure di estrazione dei campioni, i protocolli di acquisizione degli spettri, e i sistemi di elaborazione degli stessi. In particolare, sono state definite le quantità di campione da analizzare e le quantità di solvente da impiegare. È stata valutata la riproducibilità del metodo di estrazione raccogliendo gli spettri NMR di diversi estratti dello stesso lotto.

Per quanto riguarda la procedura di estrazione, il protocollo messo a punto prevede i seguenti passaggi:

1 - Se il miele è cristallizzato, viene reso fluido attraverso un breve processo di riscaldamento a microonde (qualche secondo), ponendo particolare attenzione a non superare i 35 °C.

2 - 6 g di miele in una provetta di teflon da centrifuga e si aggiungono 15 mL di acqua deionizzata.

3 - Si aggiungono 15 mL di $CHCl_3$. Si agita vigorosamente con agitatore meccanico per 10 minuti.

4 - L'emulsione che si forma in seguito all'agitazione viene rotta mediante centrifugazione per 15 minuti a 20 °C con velocità di 10000 rpm.

5 - Si raccoglie la parte organica sottostante in provetta di vetro e si pone ad essiccare sotto flusso di azoto.

6 - Si aggiungono 600 mL di $CDCl_3$ e si procede con l'analisi ponendo questa soluzione in un tubo NMR da 5 mm.

Gli scopi che si volevano ottenere da questa estrazione erano due: in primo luogo eliminare gli zuccheri (semplici e complessi) che costituiscono la parte maggioritaria dei campioni; in secondo luogo la procedura doveva essere semplice, veloce, e riproducibile. La soluzione adottata, di sciogliere il miele in acqua e successivamente di estrarre in cloroformio le componenti di interesse soddisfa entrambi i requisiti. I campioni sono pronti per l'analisi in meno di un'ora, e molti campioni possono essere preparati contemporaneamente da un unico operatore, rendendo la procedura efficiente.

Oltre agli zuccheri, altre sostanze (polari) possono rimanere nella fase acquosa e quindi non rientrare nell'analisi successiva. Anche sostanze poco polari, ma molto volatili, vengono perse durante la fase di evaporazione del cloroformio. Questo non inficia il risultato, come abbiamo potuto dimostrare con una serie di controlli. Infatti non si cerca di mappare l'intero "metaboloma", inteso come l'insieme delle sostanze presenti nel miele originario, ma di ottenere un campione nel quale le sostanze presenti siano rappresentative dell'origine botanica, sia come tipo di sostanze, che come loro quantità. Tra i controlli effettuati in questo senso vi è in primo luogo la ripetibilità, ossia la verifica che il risultato rimanga lo stesso in diverse ripetizioni della stessa misura. Inoltre, dato sicuramente più significativo, campioni "incogniti" analizzati a distanza anche di molto tempo sono stati correttamente classificati, il che dimostra come il protocollo di estrazione sia adeguato.

Il protocollo di acquisizione degli spettri NMR protonici era stato in parte sviluppato prima dell'inizio del progetto. Durante il progetto è stato verificato e convalidato.

Gli spettri sono stati acquisiti con uno strumento Bruker AVANCE DMX-600 operante a 600.09 MHz per il protone. La sequenza di impulsi recentemente sviluppata e pubblicata dal nostro gruppo (Rastrelli, F.; Schievano, E.; Bagno, A.; Mammi, S. NMR quantification of trace

components in complex matrices by band-selective excitation with adiabatic pulses. Magn. Reson. Chem. 2009, **47**, 868–872) consente di eliminare i segnali più intensi nello spettro, corrispondenti a cere e acidi grassi che poco differenziano i campioni e che oscurano segnali più deboli. Questo permette l'ottenimento di spettri più sensibili in minor tempo, facilitando e migliorando l'analisi. Tipicamente, è possibile aumentare il guadagno del ricevitore di 32 volte ed ottenere spettri con lo stesso rapporto Segnale/Rumore in metà tempo. Nella figura sottostante è mostrato il confronto fra uno spettro acquisito in questo modo ed uno acquisito con la sequenza di impulsi standard.

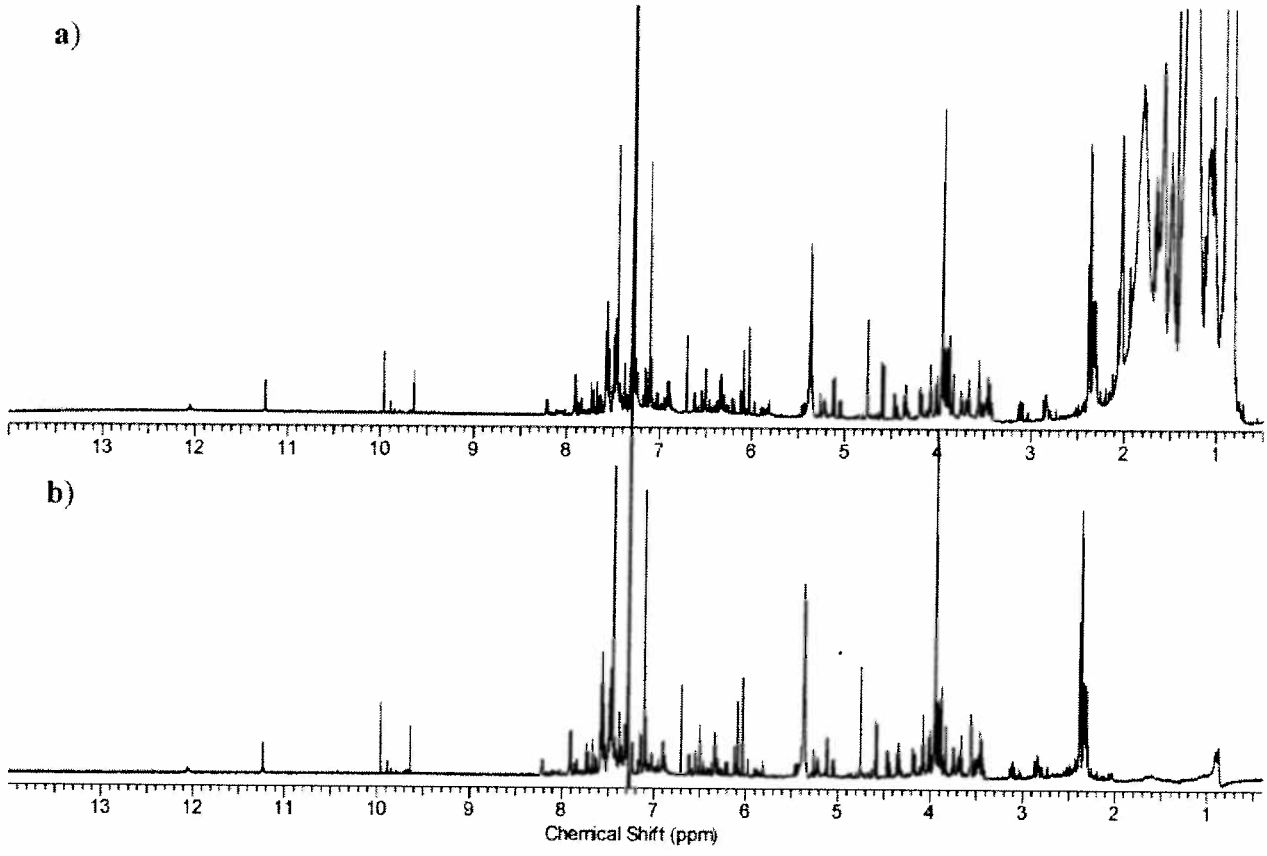


Figura 1. Confronto fra gli spettri di uno stesso campione di miele acquisiti con a) la sequenza standard (impulso e acquisizione) in 50 minuti e b) la sequenza da noi sviluppata in 25 minuti. Si noti la mancanza di segnale in b) fra 1 e 2 ppm.

Gli altri parametri di acquisizione, dopo ottimizzazione, sono i seguenti: temperatura, 298 K; tempo di recupero, 2 s; finestra spettrale, 8400 Hz; numero di scansioni, 256; punti acquisiti, 32K.

Successivamente, i dati sono stati elaborati utilizzando un software dedicato. L'ottimizzazione dei parametri di processamento ha portato all'utilizzo dei seguenti parametri: apodizzazione del FID mediante moltiplicazione per una funzione esponenziale con Line Broadening di 0.5 Hz; zero-filling dei dati fino a 128K punti; trasformata di Fourier; correzione di fase manuale; correzione di linea di base manuale. Come riferimento interno di chemical shift è stato utilizzato il segnale residuo del $CHCl_3$ a 7.27 ppm.

La funzione esponenziale utilizzata consente di migliorare il rapporto Segnale/Rumore senza incidere sulla risoluzione dello spettro. La scelta di effettuare la correzione di fase e di linea di base manualmente piuttosto che automaticamente deriva dalla necessità di analizzare l'intero spettro, e quindi di avere un'ottima qualità della linea di base in tutto lo spettro. Questo non è sempre garantito dalle operazioni automatiche, specialmente quando compiute su spettri acquisiti in tempi diversi.

Nella figura sottostante è mostrata una porzione degli spettri di mieli di origine botanica diversa. Da questi spettri è evidente la possibilità di differenziarli.

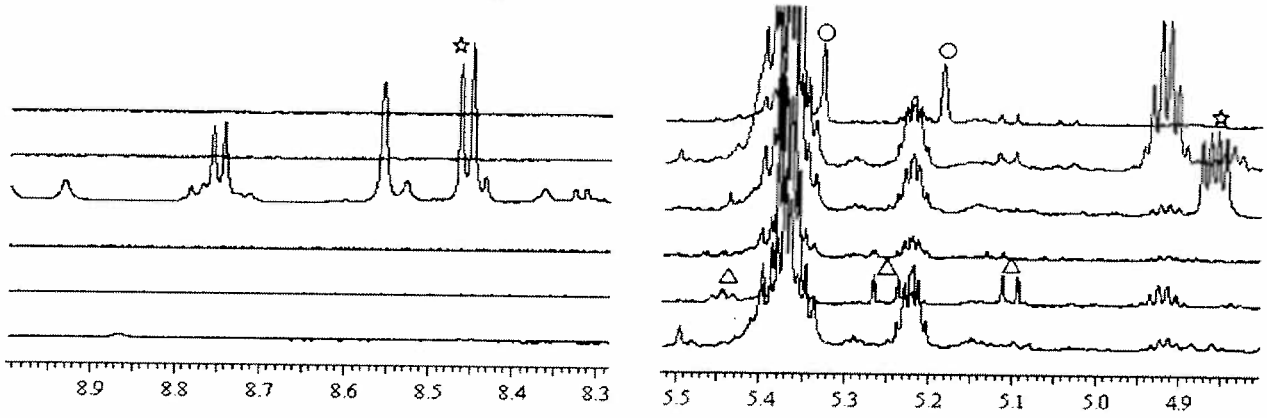


Figura 2. Confronto fra gli spettri di campioni di miele di diversa provenienza botanica. Dall'alto: tiglio, eucalipto, castagno, acacia, arancio, melata.

Gli spettri così elaborati sono stati suddivisi in intervalli di 0.04 ppm e sottoposti ad integrazione. I valori di integrazione così ottenuti sono le variabili utilizzate nell'analisi statistica. Si costruisce quindi una matrice formata da tutte le variabili misurate (circa 2000) per il numero di campioni studiati. Per poterle confrontare fra loro, le misure sui diversi campioni devono essere normalizzate scegliendo 1000 come somma del valore degli integrali su tutto lo spettro. Inoltre, le variabili vengono centrate rispetto alla media su tutti i campioni e scalate rispetto alla loro deviazione standard utilizzando l'algoritmo noto come "*Pareto scaling*".

L'analisi di un numero così elevato di variabili è possibile solamente utilizzando delle tecniche statistiche multivariate, ossia che sono in grado di evidenziare le differenze fra campioni diversi e di raggrupparli in base alle caratteristiche comuni. La più semplice di queste analisi è la PCA (*Principal Component Analysis*), che spiega le variabili secondo il principio di massima "varianza" dando peso maggiore a quelle che assumono valori diversi in campioni diversi, e trascurando invece quelle che sono simili in diversi campioni. Senza l'intervento dell'analista, questa tecnica evidenzia le variabili importanti per la caratterizzazione di campioni diversi. Tuttavia, non consente la costruzione di modelli di classi diverse di campioni.

Per classificare i campioni in gruppi omogenei occorre far uso di tecniche più sofisticate, quali la PLS-DA (*Projection to Latent Structures – Discriminant Analysis*), con la quale un certo numero di campioni viene utilizzato per definire le caratteristiche del modello di ciascuna classe (*training set*); i rimanenti campioni vengono utilizzati per validare i modelli ottenuti (*test set*). Con questa tecnica abbiamo inizialmente costruito modelli per quattro origini botaniche diverse, utilizzando 85 campioni per la costruzione del modello (26 acacia, 21 castagno, 20 tiglio, e 18 millefiori) e 33 per la sua validazione (9 acacia, 10 castagno, 6 tiglio, 7 millefiori, e 1 castagno-tiglio). Dalla figura sottostante è evidente come le quattro origini siano perfettamente distinguibili: sui 118 campioni analizzati, solo un campione di acacia ed uno di tiglio vengono classificati come millefiori.

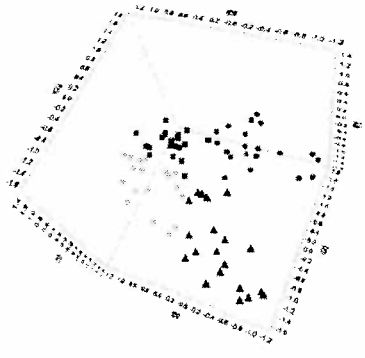


Figura 3. Identificazione mediante PLS-DA di campioni di miele di diversa provenienza botanica. Nero: tiglio; blu: castagno; rosso: acacia; verde: millefiori.

Un limite della PLS-DA consiste nel numero di classi modellizzabili contemporaneamente. Dato che i campioni che volevamo distinguere hanno molte origini botaniche diverse, questo metodo è risultato inadeguato. Nella seconda parte del progetto abbiamo quindi utilizzato un approccio gerarchico, nel quale PCA e PLS-DA (nella sua versione O2PLS-DA) sono state impiegate per riconoscere *pattern* simili nei vari campioni e per identificare particolari variabili, caratteristiche di ciascuna classe. Sono state quindi definite delle "super-variabili" utilizzate nel secondo stadio dell'analisi, basato su un approccio Bayesiano Naïve e sul metodo SVM (*Support Vector Machines*). Con questo approccio siamo riusciti a classificare tutte le origini botaniche (7) per le quali avevamo a disposizione un numero congruo di campioni. I risultati di questo metodo vengono presentati più oltre.

Campionatura.

Dopo la messa a punto della procedura di acquisizione dati e parallelamente all'affinamento dell'analisi statistica è stata acquisita una quantità elevata di dati. I campioni necessari per le analisi sono stati procurati dai partner Rigoni di Asiago e Miele Veneto. Nelle tabelle sottostanti è riassunta l'origine e la quantità dei campioni esaminati, suddivisa secondo i tempi nei quali sono stati acquisiti. A sinistra sono riportate le specie maggiori, a destra le minori.

Origine botanica	***totale campioni***	***raccolto 2008***	***raccolto 2009***	***raccolto 2010***
ACACIA	79	46	29	4
CASTAGNO	76	50	7	19
TIGLIO	81	22	33	26
MILLEFIORI	48	18	24	6
ARANCIO	55	22	17	16
EUCALIPTO	56	18	12	26
MELATA	56		30	26
TOTALE	**451**	**176**	**152**	**123**

Origine botanica	***totale campioni***
LIMONE	4
TRIFOGLIO	2
MELO	4
MONTAGNA	9
SULLA	4
TIMO	1
GIRASOLE	2
TOTALE	**26**

I campioni di miele sono provenienti da diverse regioni italiane (vedi tabella sottostante).

ORIGINE BOTANICA	**N° totale campioni**	**REGIONI ITALIANE**
Miele di acacia	79	Veneto – Piemonte- Lombardia – Emilia Romagna
Miele di castagno	76	Veneto – Piemonte- Lombardia – Emilia Romagna- Umbria - Toscana – Sicilia
Miele di tiglio	81	Veneto – Piemonte- Lombardia – Emilia Romagna
Miele di millefiori	48	Veneto – Piemonte- Lombardia – Emilia Romagna – Campania – Sicilia
Miele di arancio	54	Sicilia
Miele di eucalipto	56	Sicilia - Campania
Miele di melata	58	Veneto – Piemonte- Lombardia – Emilia Romagna – Sicilia
Miele di limone	4	Sicilia
Miele di trifoglio	1	Toscana
Miele di melo	4	Veneto- Trentino
Miele di montagna	9	Veneto- Piemonte
Miele di sulla	3	Sicilia
Miele di timo	1	Sicilia
Miele di girasole	1	Toscana

Il numero di campioni raccolti è risultato significativo solamente per le varietà: miele di acacia, di castagno, di tiglio, di arancio, di eucalipto, di millefiori e di melata. Per le altre varietà, non è stato possibile reperire un numero sufficiente di campioni.

Nella prima tabella, in rosso sono evidenziate le specie botaniche che erano state indicate nel progetto come oggetto primario di studio. Fra queste mancano tarassaco e rododendro, e i campioni di melo sono solamente quattro. Non ripetiamo qui le considerazioni su questa campionatura che facemmo in sede di relazione in itinere, che valgono anche per la campionatura finale. In particolare, l'inserimento di mieli diversi da quelli indicati originariamente si è rivelato importante in quanto, per esempio, alcuni mieli dichiarati e riconosciuti uniflorali con altri metodi vengono classificati come millefiori dal metodo NMR (vedi più oltre). Particolarmente interessante da questo punto di vista è pertanto il confronto, riportato più avanti, fra le interferenze riscontrate nell'analisi pollinica e quelle possibili nell'analisi NMR.

L'origine botanica dichiarata dagli apicoltori è stata confermata:

- Dalla valutazione dei marcatori organolettici/sensoriali (visivi, olfattivi e gustativi) delle varietà monoflorali (modulo 055 in allegato)
- Dall'analisi chimico-fisica di alcuni parametri caratterizzanti e specifici delle varietà botaniche (schede tecniche di caratterizzazione dei mieli STA in allegato).

Per tutti i campioni, le analisi sensoriali e chimico-fisiche hanno confermato l'origine botanica dichiarata dagli apicoltori. In alcuni campioni di Tiglio è stato messo in evidenza una "contaminazione" da miele di Castagno; questa contaminazione si spiega in quanto le due varietà si trovano nella stessa fascia di altitudine e hanno lo stesso periodo di fioritura.

Dalla "banca miele" sono stati selezionati vari campioni da sottoporre all'analisi pollinica classica; questa analisi è stata effettuata dal CRA-API (Consiglio per la Ricerca e la Sperimentazione in Agricoltura - Unità di Ricerca di Apicoltura) e ha dato i risultati riportati nella tabella sottostante:

Tipologia miele	**N° campioni**	**Esito**
Acacia	5	Origine botanica confermata
Arancio	5	Origine botanica confermata o spettro pollinico compatibile per 2 campioni; miele a prevalenza di sulla e agrumi per 2 campioni; miele a prevalenza di sulla per 1 campione
Castagno	5	Origine botanica confermata o spettro pollinico compatibile
Eucalipto	4	Origine botanica confermata o spettro pollinico compatibile
Tiglio	7	Spettro pollinico compatibile con l'origine botanica dichiarata per 6 campioni; un campione risulta "a prevalenza di castagno e tiglio"
Melata	5	Origine botanica confermata.

Un'altra considerazione riguarda la numerosità dei campioni (specie maggiori) studiati. Nel progetto si era indicato un numero di 100 come traguardo per ciascuna specie botanica. Nel corso dell'analisi statistica dei dati abbiamo riscontrato che i modelli predittivi che si possono costruire non cambiano molto quando il numero dei campioni analizzati supera le 50 unità per specie botanica. In altre parole, aumentare la campionatura di molto sopra le 50 unità per specie botanica non apporta alcun beneficio nell'individuazione dell'origine di mieli incogniti. Questo risultato non era prevedibile a priori, in quanto dipende fortemente dalla variabilità dei campioni analizzati: se tutti i campioni di una determinata origine botanica sono sufficientemente simili fra loro, non è

necessario utilizzarne molti per la costruzione del modello. Abbiamo cercato in ogni caso di favorire le specie botaniche originariamente indicate nel progetto, acquisendo dati per un numero più elevato di campioni di queste specie.

Costruzione del modello finale e suo utilizzo.

Come più sopra indicato, l'approccio statistico utilizzato è cambiato nel corso del progetto fino all'ottenimento di un risultato applicabile a tutti i campioni analizzati. Con i campioni acquisiti nel 2009 e 2010 è stato costruito un modello che è poi stato impiegato per verificare la provenienza botanica dei campioni acquisiti successivamente. Questo modello riguarda le sette tipologie di miele maggiormente rappresentate, ossia sei uniflorali e la tipologia millefiori. Non è possibile includere le altre tipologie nel modello, a causa della differente numerosità.

Nella fase 1 del modello è stata eseguita una PCA separatamente su campioni appartenenti a ciascuna tipologia di miele. Questa analisi serve per evidenziare eventuali *outliers* (ne sono stati individuati 4) e per selezionare da ciascuna PCA, mediante la tecnica di campionamento "*onion D-optimal design*", circa il 70% delle osservazioni (242 campioni) per costruire il cosiddetto *training set*, che viene utilizzato per costruire il modello di predizione. La restante parte dei campioni (107) è stata utilizzata come *test set*, ossia come validazione dei modelli ottenuti. Successivamente, i 100 ulteriori campioni studiati sono stati collocati nel modello, utilizzandoli ancora come validazione del metodo. Le numerosità delle varie origini usate sono indicate nella tabella sottostante.

Origine botanica	***n. campioni***	***training set***	***test set***	***acquisiti nel 2011***
ACACIA	73	52	21	4
CASTAGNO	60	42	18	16
TIGLIO	61	43	18	19
MILLEFIORI	48	33	15	0
ARANCIO	39	27	12	15
EUCALIPTO	32	20	12	24
MELATA	36	25	11	22
TOTALE	**349**	**242**	**107**	**100**

Il *training set* è stato usato per costruire modelli del tipo "1 *vs* tutti" con un'analisi O2PLS-DA, dove ogni classe di miele viene confrontato con le altre classi, considerate tutte insieme in un'unica grande classe. Con questo metodo abbiamo ottenuto 7 modelli. Ciascuno di questi modelli contiene una componente predittiva che è stata usata per due scopi:

1-Predire la classe di campioni incogniti, nel nostro caso prima di tutto del *test set*, utilizzato quindi per verificare la bontà dei modelli costruiti, e poi di altri 100 campioni acquisiti successivamente. Anche di questi campioni era nota l'origine botanica, per cui questo secondo *set* di campioni fornisce una valida indicazione della trasportabilità del metodo a campioni prodotti in annate diverse.

2-Ricercare potenziali marcatori, ossia sostanze presenti solo o in una quantità maggiore in un certo tipo di miele, da utilizzare come molecole guida nella definizione e caratterizzazione della provenienza botanica.

I modelli ottenuti hanno mostrato un elevato grado di affidabilità, come dimostrato dalle seguenti tabelle. In queste tabelle, la *sensitivity* rappresenta, per ogni specie botanica, il rapporto percentuale tra i campioni correttamente predetti e il totale dei campioni appartenenti all'origine botanica considerata. La *precision* rappresenta invece il rapporto percentuale tra i campioni correttamente predetti di quella classe e il totale di tutti i campioni del *dataset* assegnati a quella classe.

	P_Ar	P_Cast	P_AC	P_Ti	P_EU	P_Me	P_Mil	TOTAL	SENSITIVITY
Arancio	11	0	0	0	0	0	1	12	92%
Castagno	0	16	1	1	0	0	0	18	89%
Acacia	0	0	21	0	0	0	0	21	100%
Tiglio	0	1	0	17	0	0	0	18	94%
Eucalipto	0	0	0	0	12	0	0	12	100%
Melata	0	0	0	0	0	10	1	11	91%
Millefiori	1	0	1	1	0	0	12	15	80%
PRECISION	92%	94%	91%	90%	100%	100%	86%		

Nella tabella qui sopra vengono riportati i risultati ottenuti per il *test set*. Su 107 campioni, 99 vengono predetti correttamente. I campioni che non vengono predetti correttamente rappresentano per la maggior parte reali plausibili errori di classificazione in origine. L'unico risultato inatteso è un miele di castagno assegnato ad acacia. In questo caso, un confronto diretto (PLS-DA) fra tutti i campioni di castagno e acacia predice il campione come castagno, sebbene il coefficente di predizione per l'acacia non risulti trascurabile. Complessivamente, la *sensitivity*, ossia **la correttezza della predizione di specie botaniche uniflorali è sempre intorno o superiore al 90%, con punte del 100%.**

	P_Ar	P_Cast	P_AC	P_Ti	P_EU	P_Me	P_Mil	TOTAL	SENSITIVITY
Arancio	14	0	1	0	0	0	0	15	93%
Castagno	0	15	0	0	0	0	1	16	94%
Acacia	0	0	1	0	0	0	3	4	25%
Tiglio	0	1	2	14	0	0	2	19	74%
Eucalipto	2	0	0	0	22	0	0	24	92%
Melata	0	0	0	0	1	19	2	22	86%
Millefiori									
PRECISION	88%	94%	25%	100%	96%	100%			

La tabella qui sopra illustra i risultati ottenuti per i 100 campioni analizzati successivamente. In questo caso la correttezza di predizione è inferiore al caso precedente. I maggiori problemi, sia come *sensitivity* che come *precision*, si riscontrano con i mieli di acacia. Uno dei motivi potrebbe risiedere nel fatto che i composti identificati come "marcatori" del miele di acacia dall'analisi statistica sono in realtà presenti in tutti i miele, ma sono contenuti in quantità maggiore nei mieli di acacia. È possibile che mieli di annate diverse possano contenere quantità diverse di tali composti.

Per ciascuno dei casi dubbi è stato effettuato un confronto puntuale fra le due tipologie di origine botanica (ad es., castagno *vs* tiglio). In molti casi, ma non tutti, l'origine dichiarata è stata confermata, anche se la presenza di elevate quantità dell'altro miele è indubbia dall'analisi NMR. Questo risultato indica che il metodo NMR ha ancora bisogno di essere rifinito. In particolare, il tentativo ambizioso di mettere a confronto un numero elevato di classi diverse di composti comporta una diminuzione della capacità predittiva, che però può essere recuperata mediante confronti diretti specifici. Il risultato di questi confronti, applicato all'ultima tabella, la corregge come segue:

	P_Ar	P_Cast	P_AC	P_Ti	P_EU	P_Me	P_Mil	TOTAL	SENSITIVITY
Arancio	15	0	0	0	0	0	0	15	100%
Castagno	0	16	0	0	0	0	0	16	100%
Acacia	0	0	1	0	0	0	3	4	25%
Tiglio	0	1	0	17	0	0	1	19	89%
Eucalipto	2	0	0	0	22	0	0	24	92%
Melata	0	0	0	0	1	20	1	22	91%
Millefiori									
PRECISION	88%	94%	100%	100%	96%	100%			

Ricerca di potenziali marcatori di origine botanica.

Nella costruzione di ciascun modello viene prodotto un grafico, chiamato "*S-plot*", che fornisce indicazioni su quali variabili sono importanti nel caratterizzare quella particolare specie botanica. In altre parole, vengono individuati dal modello i segnali nello spettro NMR che contraddistinguono i campioni appartenenti a quella classe. Come esempio, viene qui sotto riportato il grafico relativo al miele di melata.

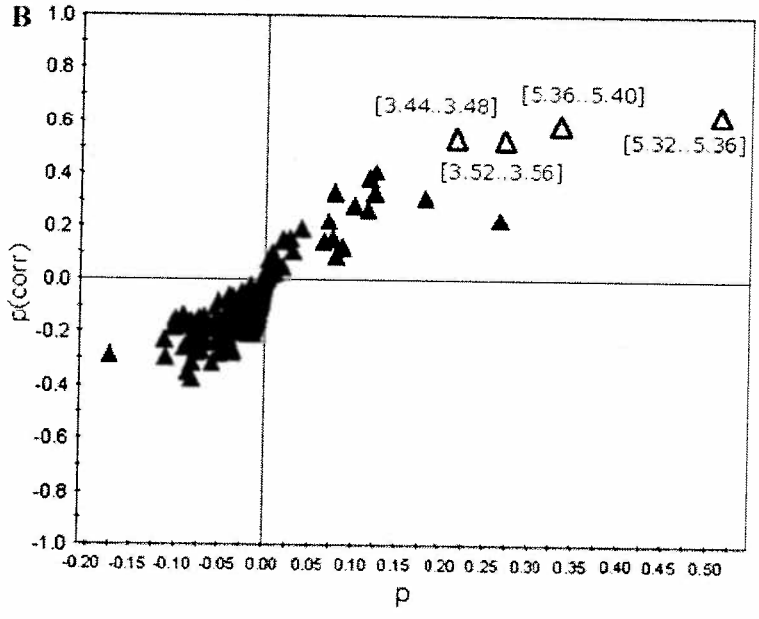


Figura 4. *S-plot* relativo ai campioni di miele di melata. Le risonanze attribuite al marcatore sono indicate con triangoli vuoti.

Da questo grafico è possibile quindi risalire alle molecole contenute nel miele di melata che lo differenziano dagli altri tipi di miele, grazie ai segnali indicati da triangoli vuoti. Per arrivare a questo risultato è stato necessario, in quasi tutti i casi, isolare la sostanza (o le sostanze) caratteristiche da tutte quelle presenti nell'estratto analizzato. Solamente nel caso dei marcatori del tiglio è stato possibile individuarne la struttura nella miscela, a causa della forte intensità dei loro segnali.

Gli estratti in cloroformio di una quantità di miele variabile fra i 36 e i 60 g sono stati trattati con una miscela 95:5 di acetonitrile e acqua per precipitare le sostanze meno polari. Nel caso dei mieli di melata, il marcatore è precipitato ed è stato purificato mediante cromatografia su gel di silice con eluizione isocratica con etere di petrolio:etile acetato 3:1. Per gli altri mieli, invece, i marcatori son rmasti in soluzione e sono stati purificati mediante cromatografia su gel di silice con eluizione a step di metanolo in cloroformio. I due marcatori per il miele d'acacia sono stati ritrovati nella frazione a 0.25% di metanolo; quello di miele di eucalipto nella frazione a 0.5% di metanolo; i due marcatori di miele d'arancio nella frazione a 1.25% di metanolo.

La struttura dei marcatori è stata determinata utilizzando principalmente la spettroscopia NMR e la spettrometria di massa. Le molecole individuate sono riportate nella figura sottostante.

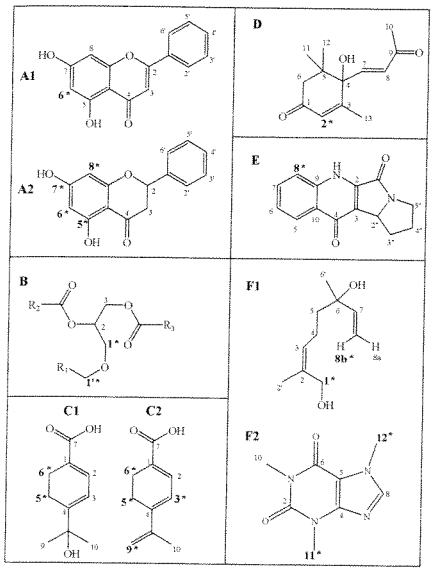


Figura 5. Marcatori identificati e caratterizzati per le sei classi di mieli uniflorali studiate.

A: Acacia
B: Melata
C: Tiglio
D: Eucalipto
E: Castagno
F: Arancio

A molte delle sostanze identificate vengono riconosciute specifiche proprietà farmaceutiche e questo è un fattore che può essere validamente utilizzato per la valorizzazione del prodotto.

Siamo ancora alla ricerca di composti che siano presenti unicamente nel miele di acacia, così come anche di altre sostanze da affiancare a quelle già identificate.

Confronto fra il metodo sviluppato ed i metodi tradizionali.

Il metodo sviluppato presenta degli indubbi vantaggi rispetto all'analisi melissopalinologica. A fronte di un costo simile, o possibilmente anche inferiore, l'analisi NMR è sicuramente più veloce (abbiamo dovuto attendere sei mesi per avere l'analisi di 30 campioni, mentre con l'NMR l'analisi di 30 campioni richiede una settimana al massimo). Rimangono alcuni punti da risolvere per rendere il metodo ancor più robusto. Soprattutto occorre stabilire un protocollo di trattamento dei dati nei casi dubbi.

Rispetto all'analisi sensoriale, che è la procedura universalmente adottata per la maggior parte dei campioni, il metodo NMR è più oggettivo e trasportabile, non richiede esperienza specifica nel campo, anche se il suo maggiore svantaggio è rappresentato dal costo dell'analisi.

In conclusione, il metodo da noi sviluppato può essere affiancato alle analisi correntemente impiegate e fornire un valido supporto nei casi dubbi o controversi.

Differenziazione geografica.

I tentativi fatti di classificare i campioni di miele in nostro possesso in base all'origine geografica non hanno per ora avuto successo. Ci siamo concentrati sulle 5 varietà per le quali abbiamo campioni provenienti da molte regioni. Una parziale differenziazione è possibile tra i mieli provenienti "dal nord" e quelli provenienti "dal sud" dell'Italia. La principale differenza consiste nella maggiore presenza di cere in questi ultimi. Questo risultato potrebbe riflettere una diversa procedura di estrazione e produzione del miele, e sembra interessante proseguire lo studio in questa direzione.

Divulgazione dei risultati.

I risultati del presente progetto sono stati divulgati usando i normali canali delle pubblicazioni dei lavori scientifici. In particolare, sono stati pubblicati due articoli su riviste internazionali di prestigio, che sono allegati alla presente relazione. I riferimenti bibliografici relativi sono riportati di seguito.

E. Schievano, E. Peggion, and S. Mammi
^{1}H Nuclear Magnetic Resonance Spectra of Chloroform Extracts of Honey for Chemometric Determination of its Botanical Origin. *J. Agr. Food Chem.*, **58,** 57–65 (2010).

E. Schievano, M. Stocchero, E. Morelato, C. Facchin, and S. Mammi
An NMR-Based Metabolomic Approach to Identify the Botanical Origin of Honey. *Metabolomics*, DOI 10.1007/s11306-011-0362-8 (2012, in press).

Oltre a questi articoli, il lavoro è stato presentato a congressi scientifici, mediante poster e comunicazioni orali. I riferimenti a tali eventi sono riportati di seguito.

E. Schievano e S. Mammi
Lo Spettro ^{1}H-NMR di Estratti in Cloroformio del Miele: un'Affidabile Impronta Digitale dell'Origine Botanica.
Presentato all'VIII Congresso Nazionale di Chimica degli Alimenti, Marsala (TP), 20-24 Settembre 2010, Abstract Book p. 183.

E. Schievano, C. Facchin, E. Morelato, M. Stocchero, and S. Mammi
Floral Markers in Honeys of Various Botanical Origin.
Presentato al XL Congresso Nazionale di Risonanze Magnetiche, Parma, 26-28 Settembre 2011, Abstract Book p. 85.

E. Morelato, C. Facchin, M. Stocchero, S. Mammi, and E. Schievano
NMR-Based Metabolite Fingerprinting to Identify the Botanical Origin of Honey.
Presentato al XL Congresso Nazionale di Risonanze Magnetiche, Parma, 26-28 Settembre 2011, Abstract Book p. 99.

S. Mammi
Rintracciabilità Alimentare: un Contributo dall'NMR.
Presentato al Workshop "Applicazione della Risonanza Magnetica Nucleare in Campo Agro-Alimentare", C.R.E.A. – Università di Siena, Colle Val d'Elsa (SI), 14 Gennaio 2012.

Infine, il lavoro è stato condotto anche come tirocinio di tesi, triennale o magistrale, di vari studenti. Gli elaborati finali sono documentazione pubblica. Le tesi finora scritte sono le seguenti:

Applicazione della Spettroscopia NMR alla Determinazione delle Origini Geografiche e Botaniche del Miele
Laurea Triennale in Chimica, Facoltà di Scienze MM. FF. NN., Università di Padova, A.A. 2008-09.
Laureando: De Filippo Christian Corrado

Applicazione della Spettroscopia di Risonanza Magnetica Nucleare e Analisi Multivariata per la Caratterizzazione dei Mieli d'Arancio
Laurea Triennale in Chimica Industriale, Facoltà di Scienze MM. FF. NN., Università di Padova, A.A. 2008-09.
Laureando: Pivato Antonio

Rilevazione del 3,5,5-trimetil-4-(3'-oxobutenil)-2-cicloesenone nel Miele di Eucalipto mediante HPLC e NMR
Laurea Triennale in Chimica, Facoltà di Scienze MM. FF. NN., Università di Padova, A.A. 2008-09.
Laureando: Bergamasco Johnny

Determinazione dell'Origine Botanica del Miele: un Nuovo ed Affidabile Metodo di Risonanza Magnetica Nucleare e Chemiometria
Laurea Specialistica in Chimica Industriale, Facoltà di Scienze MM. FF. NN., Università di Padova, A.A. 2008-09.
Laureando: Pergher Marco

Caratterizzazione del Miele di Melata e Identificazione del Marker del Miele di Arancio mediante Risonanza Magnetica Nucleare
Laurea Triennale in Chimica Industriale, Facoltà di Scienze MM. FF. NN., Università di Padova, A.A. 2009-10.
Laureando: Mazzaretto Ivan

L'Origine Botanica del Miele: un Approccio Metabolomico
Laurea Triennale in Chimica, Facoltà di Scienze MM. FF. NN., Università di Padova, A.A. 2009-10.
Laureando: Carollo Silene

Determinazione dell'Origine Botanica del Miele: Identificazione del Pinocembrin e Acido Abscissico nel Miele di Acacia
Laurea Triennale in Chimica, Facoltà di Scienze MM. FF. NN., Università di Padova, A.A. 2010-11.
Laureando: Rigodanza Francesco

Fingerprinting Metabolico Basato su NMR per Identificare l'Origine Botanica del Miele
Laurea Specialistica in Chimica, Facoltà di Scienze MM. FF. NN., Università di Padova, A.A. 2010-11.
Laureando: Morelato Elisa

Metodi di Indagine per il Controllo della Qualità e la Prevenzione di Frodi Alimentari
Laurea Triennale in Chimica, Facoltà di Scienze MM. FF. NN., Università di Padova, A.A. 2011-12.
Laureando: Carniato Giulia

Prospettive future.

Il metodo che abbiamo messo a punto è scientificamente robusto ed affidabile. È quindi pronto per essere divulgato presso i possibili beneficiari. Secondo quanto indicato nel progetto originario, questi sono i consumatori, i trasformatori, i produttori, e gli organismi di controllo.

Il primo passo in questa direzione consisterà nella pubblicazione del metodo in riviste italiane specializzate nel settore (quali Apitalia, Apoidea, L'Apicoltore italiano, L'Apis), in maniera da raggiungere i trasformatori e i produttori.

Parallelamente saranno contattati i NAS e la Guardia di Finanza per offrire questo metodo in alternativa o in aggiunta all'analisi melissopalinologica e a quella organolettica.

Solo mediante un utilizzo da parte dei soggetti su indicati si potrà giungere ad una valorizzazione del metodo. Se trasformatori, produttori, e organismi di controllo potranno verificare la validità e la convenienza del metodo, ciò potrebbe portare al suo inserimento in un disciplinare di produzione e alla predisposizione di un "certificato di origine", con indubbio vantaggio per i consumatori.

Una delle sfide che intendiamo raccogliere riguarda la caratterizzazione di provenienza geografica. Se finora i risultati in questa direzione non sono stati soddisfacenti per quanto riguarda il confronto fra diverse regioni italiane, è pur vero che le frodi maggiori riguardano mieli di

provenienza extraeuropea. Dovrebbe essere più facile distinguere mieli italiani da mieli di altra origine europea o addirittura extraeuropea, a causa delle maggiori differenze botaniche e climatiche.

La pubblicazione del nostro articolo su *J. Agr. Food Chem.* ci ha procurato una collaborazione internazionale. con la Prof.ssa Patricia Vit Olivier del Departmento Ciencia de Alimentos, Facultad de Farmacia, Universidad de Los Andes, Venezuela. La Prof.ssa Vit studia fra le altre cose delle particolari api sudamericane senza pungiglione. Abbiamo già studiato il miele prodotto da queste api, e il nostro metodo lo distingue con facilità dai mieli di *A. mellifera.*

Un'altra collaborazione internazionale sorta nel corso del presente progetto è stata con la Facultad de Ciencias Químicas, Universidad de Castilla-La Mancha, Ciudad Real, Spagna. Nel corso del 2010 la Dottoranda María Moreno Pérez ha trascorso un periodo di quattro mesi nel nostro laboratorio occupandosi anche del presente progetto.

ALLEGATI

Rigoni di Asiago	SCHEDA ANALISI SENSORIALE MIELE	Mod.055 rev.1 pag.1 di 1

Data **CAMPIONE** **FORNITORE** **Quantità**

Origine Botanica Geografica

ESAME VISIVO							
Fluido ☐	Cristallizzato ☐						
Limpido SI NO	Omogeneo SI NO			Pulito SI NO			
COLORE							
Bianco paglierino			1	2	3	4	5
Crema			1	2	3	4	5
Nocciola			1	2	3	4	5
Giallo			1	2	3	4	5
Ambra			1	2	3	4	5
Marrone			1	2	3	4	5
Nero			1	2	3	4	5

ESAME TATTILE							
Fluido			1	2	3	4	5
Cristallizzato			1	2	3	4	5
Tipo di cristalli	Fini ☐	Medi ☐	Grossi ☐	solubili ☐	Taglienti ☐	Tondeggianti ☐	Resistenti ☐

ESAME OLFATTIVO/RETROLFATTIVO						
Floreale	Fiori freshi	1	2	3	4	5
	Fiori secchi	1	2	3	4	5
Fruttato	Agrumi	1	2	3	4	5
	Frutta matura	1	2	3	4	5
	Frutta cotta/essicata	1	2	3	4	5
Vegetale	Fresco	1	2	3	4	5
	Secco	1	2	3	4	5
	Erbe aromatiche/balsamiche	1	2	3	4	5
	Bosco/resina	1	2	3	4	5
	Sottobosco	1	2	3	4	5
Caldo	Speziato	1	2	3	4	5
	Pasticceria	1	2	3	4	5
	Cuoio	1	2	3	4	5
Empireumatico	Caramellato	1	2	3	4	5
	Tostato	1	2	3	4	5
Odori anomali	Fermentato	1	2	3	4	5
	Ossidato	1	2	3	4	5
	Microbiologico	1	2	3	4	5

ESAME GUSTATIVO						
Dolce		1	2	3	4	5
Amaro		1	2	3	4	5
Acido		1	2	3	4	5
Astringente		1	2	3	4	5
Persistente		1	2	3	4	5

GIUDIZIO FINALE						
Franchezza		1	2	3	4	5
Ricchezza		1	2	3	4	5
Livello edonico		1	2	3	4	5
Corrispondenza all'origine botanica dichiarata		1	2	3	4	5

NOTE:

Rigoni di Asiago	SCHEDA TECNICA MATERIA PRIMA	Sigla documento:	STA-01
		Revisione:	0
		Data:	06/12/2011
		Pagina	1 di 1
		Filename:	STA-01 miele di acacia
Titolo della STA:	MIELE DI ACACIA		

TIPO DI MIELE: monoflora

PROPRIETA' ORGANOLETTICHE:

- Stato fisico: liquido
- Colore: da bianco acqua a giallo paglierino chiaro, trasparente
- Odore: molto leggero
- Sapore: dolce e delicato ricorda il profumo dei fiori

CARATTERISTICHE CHIMICO-FISICHE

➢ Colore mm pfund	0 - 15
➢ Gradi Brix	> 80.0°
➢ Umidità %	< 18.0
➢ HMF mg/Kg	< 5.0
➢ Diastasi	>5.0
➢ pH	3.8 – 4.2
➢ Acidità meq/Kg	<20.0
➢ Cond. Elettrica µS/cm	< 200
➢ Fruttosio %	40.0 – 44.0
➢ Glucosio %	26.0 – 30.0
➢ Saccarosio %	4.5 – 5.5

RESIDUI:

Le seguenti sostanze: pesticidi, sulfamidici, antibiotici, metalli pesanti, non fanno parte della composizione naturale del miele, devono dunque risultare non rilevabili.

Rigoni di Asiago
Assicurazione Qualità
Marina Panozzo

Rigoni di Asiago	SCHEDA TECNICA MATERIA PRIMA	Sigla documento:	STA-02
		Revisione:	0
		Data:	06/12/2011
		Pagina	1 di 1
		Filename:	STA-02 miele di arancio
Titolo della STA:	MIELE DI ARANCIO		

TIPO DI MIELE: monoflora

PROPRIETA' ORGANOLETTICHE:

- Stato fisico: cristallizzazione ritardata generalmente a granulazione media o fine
- Colore: bianco, traslucido
- Odore: tipico del fiore
- Sapore: caratteristico e delicato, lievemente acidulo

CARATTERISTICHE CHIMICO-FISICHE

➢ Colore mm pfund	10-40
➢ Gradi Brix	> 80.0°
➢ Umidità %	< 19.0
➢ HMF mg/Kg	< 10.0
➢ Diastasi	>5.0
➢ pH	3.6 – 4.2
➢ Acidità meq/Kg	< 25.0
➢ Cond. Elettrica µS/cm	<400

RESIDUI:

Le seguenti sostanze: pesticidi, sulfamidici, antibiotici, metalli pesanti, non fanno parte della composizione naturale del miele, devono dunque risultare non rilevabili.

Rigoni di Asiago
Assicurazione Qualità
Marina Panozzo

Rigoni di Asiago	SCHEDA TECNICA MATERIA PRIMA	Sigla documento:	STA-03
		Revisione:	0
		Data:	06/12/2011
		Pagina	1 di 1
		Filename:	STA-03 miele di castagno
Titolo della STA:	MIELE DI CASTAGNO		

TIPO DI MIELE: monoflora

PROPRIETA' ORGANOLETTICHE:

- Stato fisico: generalmente liquido o a cristallizzazione molto ritardata
- Colore: da ambra ad ambra scuro con tonalità rossastra
- Odore: deciso, forte e penetrante, aromatico
- Sapore: forte e persistente con retrogusto amarognolo più o meno accentuata

CARATTERISTICHE CHIMICO-FISICHE

➢ Colore mm pfund	60 – 120
➢ Gradi Brix	> 80.0°
➢ Umidità %	< 18.0
➢ HMF mg/Kg	< 5.0
➢ Diastasi	>10.0
➢ pH	4.5-5.7
➢ Acidità meq/Kg	< 40
➢ Cond. Elettrica µS/cm	> 800

RESIDUI:

Le seguenti sostanze: pesticidi, sulfamidici, antibiotici, metalli pesanti, non fanno parte della composizione naturale del miele, devono dunque risultare non rilevabili.

Rigoni di Asiago
Assicurazione Qualità
Marina Panozzo

Rigoni di Asiago	SCHEDA TECNICA MATERIA PRIMA	Sigla documento:	STA-04
		Revisione:	0
		Data:	06/12/2011
		Pagina	1 di 1
		Filename:	STA-04 miele di eucalipto
Titolo della STA:	MIELE DI EUCALIPTO		

TIPO DI MIELE: monoflora

PROPRIETA' ORGANOLETTICHE:

- Stato fisico: cristallizzato a granulazione fine o media, compatta
- Colore: da ambra chiaro ad ambra, con tonalità grigiastre
- Odore: intenso e caratteristico
- Sapore: forte e persistente simile all'odore ma più gradevole

CARATTERISTICHE CHIMICO-FISICHE

➢ Colore mm pfund	40 – 70
➢ Gradi Brix	> 80.0°
➢ Umidità %	< 18.0
➢ HMF mg/Kg	< 10.0
➢ Diastasi	>10.0
➢ pH	3.7 – 4.1
➢ Acidità meq/Kg	< 40
➢ Cond. Elettrica µS/cm	< 900

RESIDUI:

Le seguenti sostanze: pesticidi, sulfamidici, antibiotici, metalli pesanti, non fanno parte della composizione naturale del miele, devono dunque risultare non rilevabili.

Rigoni di Asiago
Assicurazione Qualità
Marina Panozzo

Rigoni di Asiago	SCHEDA TECNICA MATERIA PRIMA	Sigla documento:	STA-05
		Revisione:	0
		Data:	06/12/2011
		Pagina	1 di 1
		Filename:	STA-05 miele di tiglio
Titolo della STA:	MIELE DI TIGLIO		

TIPO DI MIELE: monoflora

PROPRIETA' ORGANOLETTICHE:

- Stato fisico: liquido o spesso cristallizzato pastoso
- Colore: molto variabile: i mieli più puri sono di colore ambra chiaro con tonalità gialle, ma tendono a scurirsi in misura più o meno marcata in presenza di inquinamenti da castagno o melata.
- Odore: tipico, mentolato, balsamico, persistente
- Sapore: intensamente aromatico, simile all'odore.

CARATTERISTICHE CHIMICO-FISICHE

➢ Colore mm pfund	30 – 150
➢ Gradi Brix	> 80.0°
➢ Umidità %	< 19.0
➢ HMF mg/Kg	< 8.0
➢ Diastasi	>12.0
➢ pH	3.9 – 5.3
➢ Acidità meq/Kg	< 40
➢ Cond. Elettrica µS/cm	> 400

RESIDUI:

Le seguenti sostanze: pesticidi, sulfamidici, antibiotici, metalli pesanti, non fanno parte della composizione naturale del miele, devono dunque risultare non rilevabili.

Rigoni di Asiago
Assicurazione Qualità
Marina Panozzo

Rigoni di Asiago	SCHEDA TECNICA MATERIA PRIMA	Sigla documento:	STA-06
		Revisione:	0
		Data:	06/12/2011
		Pagina	1 di 1
		Filename:	STA-06 miele di melata
Titolo della STA:	MIELE DI MELATA		

TIPO DI MIELE: melata

PROPRIETA' ORGANOLETTICHE:

- Stato fisico: liquido o a cristallizzazione molto ritardata
- Colore: ambra scura fino a quasi nero
- Odore: deciso, vegetale-fruttato
- Sapore: forte e persistente, poco dolce, ricorda quello della melassa

CARATTERISTICHE CHIMICO-FISICHE

➢ Colore mm pfund	80-110
➢ Gradi Brix	> 80.0°
➢ Umidità %	< 17.0
➢ HMF mg/Kg	< 5.0
➢ Diastasi	>15.0
➢ pH	4.3 – 5.9
➢ Acidità meq/Kg	< 50
➢ Cond. Elettrica µS/cm	>800

RESIDUI:

Le seguenti sostanze: pesticidi, sulfamidici, antibiotici, metalli pesanti, non fanno parte della composizione naturale del miele, devono dunque risultare non rilevabili.

Rigoni di Asiago
Assicurazione Qualità
Marina Panozzo

Rigoni di Asiago	SCHEDA TECNICA MATERIA PRIMA	Sigla documento:	STA-07
		Revisione:	0
		Data:	06/12/2011
		Pagina	1 di 1
		Filename:	STA-07 miele di fiori
Titolo della STA:	MIELE DI FIORI		

TIPO DI MIELE: poliflora

PROPRIETA' ORGANOLETTICHE:

- Stato fisico: liquido o cristallizzato
- Colore: variabile da ambra chiaro ad ambra scuro
- Odore: variabile, caratteristico delle varietà botaniche che lo compongono
- Sapore: variabile , caratteristico delle varietà botaniche che lo compongono

CARATTERISTICHE CHIMICO-FISICHE

➢ Colore mm pfund	30-100
➢ Gradi Brix	> 80.0°
➢ Umidità %	< 18.0
➢ HMF mg/Kg	< 8.0
➢ Diastasi	>12.0
➢ pH	3.8 – 4.4
➢ Acidità meq/Kg	< 40.0

RESIDUI:

Le seguenti sostanze: pesticidi, sulfamidici, antibiotici, metalli pesanti, non fanno parte della composizione naturale del miele, devono dunque risultare non rilevabili.

Rigoni di Asiago
Assicurazione Qualità
Marina Panozzo

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY
ARTICLE

J. Agric. Food Chem. 2010, 58, 57–65 57
DOI:10.1021/jf9022977

^{1}H Nuclear Magnetic Resonance Spectra of Chloroform Extracts of Honey for Chemometric Determination of Its Botanical Origin

ELISABETTA SCHIEVANO,* EVARISTO PEGGION, AND STEFANO MAMMI

Dipartimento di Scienze Chimiche, Università di Padova, Via Marzolo, 1, Padova 35131, Italy

In this work, we present a new NMR study, coupled with chemometric analysis, on nonvolatile organic honey components. The extraction method is simple and reproducible. The ^{1}H NMR spectra of chloroform extracts acquired with a fast and new pulse sequence were used to characterize and differentiate by chemometric analysis 118 honey samples of four different botanical origins (chestnut, acacia, linden, and polyfloral). The spectra collection, processing, and analysis require only 30 min. The ^{1}H spectrum provides a fingerprint for each honey type, showing many characteristic peaks in all spectral regions. Principal component analysis (PCA) and projection to latent structures by partial least squares-discriminant analysis (PLS-DA) were performed on selected signals of the spectra to discriminate the different botanical types and to identify characteristic metabolites for each honey type. A distinct discrimination among samples was achieved. According to the distance to model criterion, there was no overlap between the four models, which proved to be specific for each honey type. The PLS-DA model obtained has a correlation coefficient R^2 of 0.67 and a validation correlation coefficient Q^2 of 0.77. The discriminant analysis allowed us to classify correctly 100% of the samples. A classification index can be calculated and used to determine the floral origin of honey as an alternative to the melissopalinology test and possibly to determine the percentage of various botanical species in polyfloral samples. Preliminary data on the identification of marker compounds for each botanical origin are presented.

KEYWORDS: Honey; floral origin; NMR; multivariate statistical analysis

INTRODUCTION

Honey is defined as "the natural sweet substance produced by *Apis mellifera* bees from the nectar of plants or from secretions of living parts of plants or excretions of plant-sucking insects on the living parts of plants, which the bees collect, transform by combining with specific substances of their own, deposit, dehydrate, store, and leave in honeycombs to ripen and mature" (*1*). There have been many reported beneficial effects from the use or consumption of honey, including antimicrobial properties (*2*, *3*), antioxidant effects (*4*, *5*), wound healing effects (*6*), and prevention of life-threatening pathologies such as diabetes, cardiovascular deseases (*7*), and cancer (*8*). Honeybees and their products can also be employed as potential bioindicators of environmental contamination (*9*).

The composition and properties of a particular honey sample depend strongly on the type of flowers visited by the bees, as well as on the climatic conditions in which the plants grow and on contributions of the beekeeper (*10*, *11*).

The Codex Alimentarius Standard (*12*) and the European Union Council Directive (*1*) specify that the term "honey" may be completed by a reference to the origin, whether blossom or plant, provided the product comes predominantly from the indicated source and has the appropriate organoleptic, physicochemical, and microscopic properties corresponding to that origin.

The interpretation of "predominantly" remains ambiguous, and the definition of unifloral or polyfloral nowadays is not based on physical–chemical parameters but on melissopalynological analysis. Usually, honey is considered unifloral when the pollen frequency of one plant is over 45% (*13*). For honey samples with under-represented pollen grains (e.g., lavender, citrus, and rosemary) botanical classification may be achieved with a percentage pollen frequency of only 10–20%. Melissopalynology, i.e., the identification and quantification of pollen grains contained in honey, has been traditionally used to ascertain the botanical origin of honeys, although with some limitations (*14*). Specifically, melissopalynology requires trained analysts and the previous knowledge of pollen morphology. In spite of these problems, palynological analysis remains the reference method nowadays.

In the last few decades, specific chemical and physical properties of honey have been used to determine its botanical origin (*14*–*16*) and new analytical techniques have been proposed toward this aim. They are based, for example, on the determination of total flavonoids, profile of volatiles, and amino acid and carbohydrate composition.

*To whom correspondence should be addressed. Tel: 39-0498275742. Fax: 39-0498275239. E-mail: elisabetta.schievano@unipd.it.

Published on Web 11/25/2009

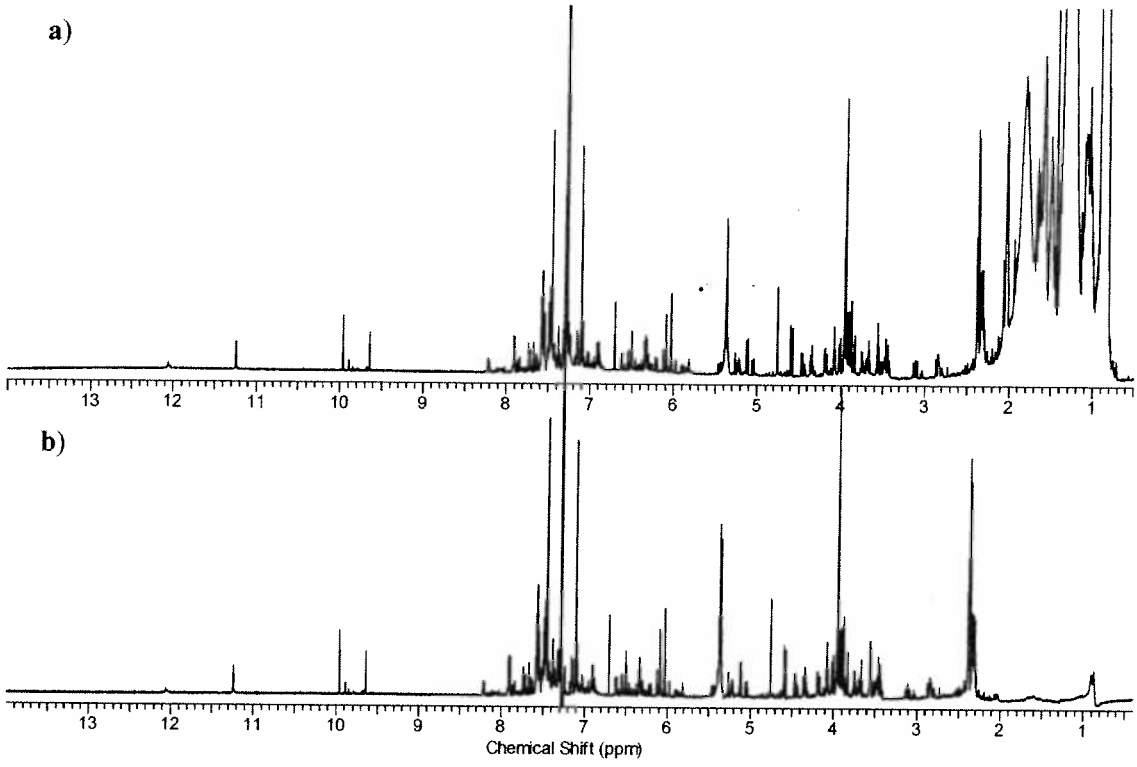


Figure 1. Comparison between complete 1D ^{1}H spectra of a honey sample obtained with (**a**) a single pulse sequence in 50 min and (**b**) the modified DPFGSE sequence in 25 min.

An improvement in the determination of the botanical origin can certainly be achieved by the application of the multivariate analytical approach. Several applications to honey classification have been reported (*17*−*21*). For example, attempts have been made to use some physical and chemical properties (*22*) or the mineral content (*23*) of honey. Although interesting results were obtained, most of these studies were based on a fairly limited number of honey samples or used a combination of parameters based on several independent measurements.

Recently, also nuclear magnetic resonance (NMR) techniques have been proposed to identify and classify honey of different floral sources (*24*, *25*) or geographic origins (*26*, *27*). One of the main advantages of this technique is that structural and quantitative information can be obtained on a wide range of chemical species in a single NMR experiment. NMR is frequently applied to food samples that can be directly examined as liquids (*28*), but very simple extraction or sample preparation procedures may also be used (*29*). In the case of honey, the botanical or geographical differentiations were based on different carbohydrate composition. Consonni and Cagliani (*27*) showed that this parameter can be used to distinguish honeys of different geographical origins and hinted at the possibility of separating acacia and polyfloral honeys. The ^{13}C spectrum was suggested to be the best probe of carbohydrate composition. Lolli et al. (*25*) classified samples of different botanical origins by using 2D HMBC experiments coupled with multivariate statistical analysis, by dissolving the samples in water or DMSO. In only one NMR study, conducted at 300 MHz, the solid-phase methanol extracts were analyzed (*24*). In this paper, we show that ^{1}H NMR spectra of organic extracts can be used as a "fingerprint" to differentiate the botanical origin of honey. The use of high fields also allows the identification and characterization of some of these compounds as botanical markers.

In this work, we present a new NMR study, coupled with chemometric analysis, on nonvolatile organic compounds. ^{1}H spectra of chloroform extracts of honey were analyzed with a fast and new pulse sequence (*30*). The extraction method is simple, and reproducible, and it yields an extract that has never been analyzed before. The advantage of this approach is to eliminate the compounds most present in the honey mixture, i.e., the carbohydrates, and to retain the aroma compounds and those hydrophobic substances that differ the most in honeys of various sources. Also, the extraction procedure yields a concentrated solution amenable to fast NMR analysis. A total of 118 samples of acacia, linden, chestnut, and polyfloral honeys were analyzed. By using principal component analysis (PCA) and supervised techniques (projection to latent structures by partial least squares-discriminant analysis, PLS-DA), a classification model according to floral origin was obtained, with high predictability power. A classification index can be calculated and used to determine floral origin of honey as an alternative to the melissopalinology test and possibly to determine the percentage of various botanical species in polyfloral samples.

During this analysis, we were able to isolate and identify some molecular markers of these botanical origins.

MATERIALS AND METHODS

Samples. A total of 118 honey samples of different botanical origins were analyzed. Among them, 93 were obtained directly from the producers with certified origin, while 25 were commercial products.

Samples of four different botanical origins were obtained from Veneto apiaries: 28 acacia (*Pseudoacacia robinia L.*), 23 chestnut (*Castanea sativa*), 22 linden (*Tilia spp.*), and 20 polyfloral honeys. Commercial samples included 7 acacia, 8 chestnut, 4 linden, 5 polyfloral, and 1 declared chestnut−linden honey.

Sample Preparation. Portions of samples (6 g) were weighted in a centrifuge tube and dissolved with 15 mL of deionized water. Fifteen milliliters of $CHCl_3$ were added, and the mixture was mechanically stirred for 10 min. The biphasic mixture was then centrifuged at 10 000 rpm for 15 min at 4 °C. The lower chloroform phase was collected, and the solvent

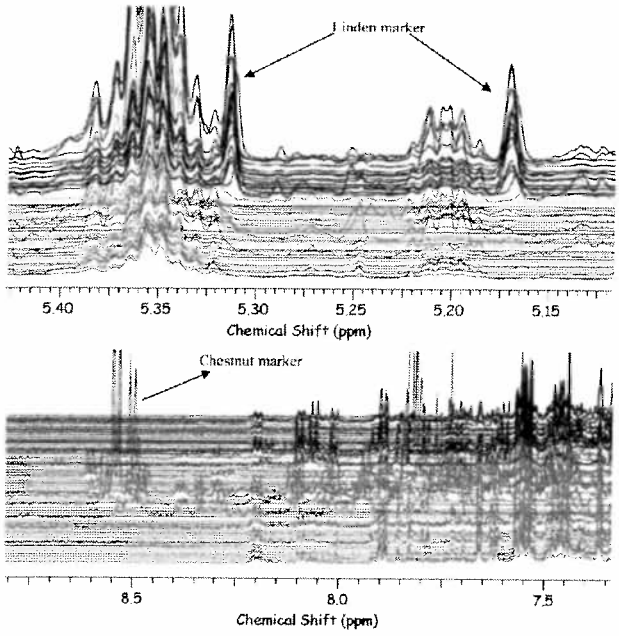


Figure 2. Tile plot of all spectra of chestnut, linden, and acacia honeys: region between 5.10 and 5.45 ppm (top) and between 7.2 and 9.0 ppm (bottom), with the arrows indicating proton signals of markers of linden and of chestnut. Red, blue, and black colors indicate acacia, chestnut, and linden honeys, respectively.

was evaporated under a gentle stream of nitrogen. The solid residue was dissolved in 600 μL of $CDCl_3$ and put in an NMR tube.

To identify markers from chestnut and acacia honey, separations were conducted using a silica gel column eluted with $CHCl_3$ and a gradient of 0−5% MeOH.

NMR Analysis. Spectra were recorded on a Bruker Avance 600 DMX instrument, operating at 600.09 MHz for ^{1}H and equipped with a 5 mm TXI *xyz* gradient inverse probe.

The 1D spectra were acquired using a modified double pulsed field gradient spin echoes (DPFGSE) sequence (*30*). Specifically, the cluster $[G-S-G]$, where G represents a pulsed field gradient and S is a generic element (usually a 180° soft pulse), was modified by the addition of an inversion hard pulse after the first gradient (that is, $G-\pi-S-G$) and incorporating an inversion Reburp pulse of 2 kHz sweep width and 10 ms duration centered at 1 ppm. The introduction of a π pulse in the DPFGSE sequence allowed us to remove the highest signals present in the 0−2 ppm region. All gradient pulses were followed by a 100 μs recovery delay. The typical acquisition parameters of this experiment were as follows: temperature, 298 K; recycle time, 2 s; spectral window, 6000 Hz; number of scans, 256; data points, 32K; receiver gain, 8K. The parameters for the 1D spectra obtained with the standard single-pulse sequence were as follows: temperature, 298 K; recycle time, 2 s; spectral window, 6000 Hz; number of scans, 1024; data points, 32K; receiver gain, 256.

Data were processed using the ACD software (ACD/Specmanager 7.00 sofware, Advanced Chemistry Development Inc., 90 Adelaide Street West, Toronto, Ontario, Canada M5H 3V9). Fourier transformation was performed after zero-filling the FID data to 128K points and after apodization using a decreasing exponential with line broadening of 0.5 Hz. The spectra were phased and baseline-corrected using the ACD manual routine, and the ^{1}H NMR chemical shifts were referenced to the residual $CHCl_3$ signal at 7.27 ppm. Each ^{1}H spectrum was segmented into identical intervals ("buckets") of 0.04 ppm, and the signal intensity in each interval was integrated. The spectra were normalized to the total sum of integral covering the δ interval 13−1.8 and excluding the δ region 7.26−7.28, which contains the residual solvent peak. The resulting normalized integrals composed the data matrix that was submitted to multivariate analysis.

Identification of markers for each botanical origin was obtained through 1D and 2D spectra. The following parameters were used.

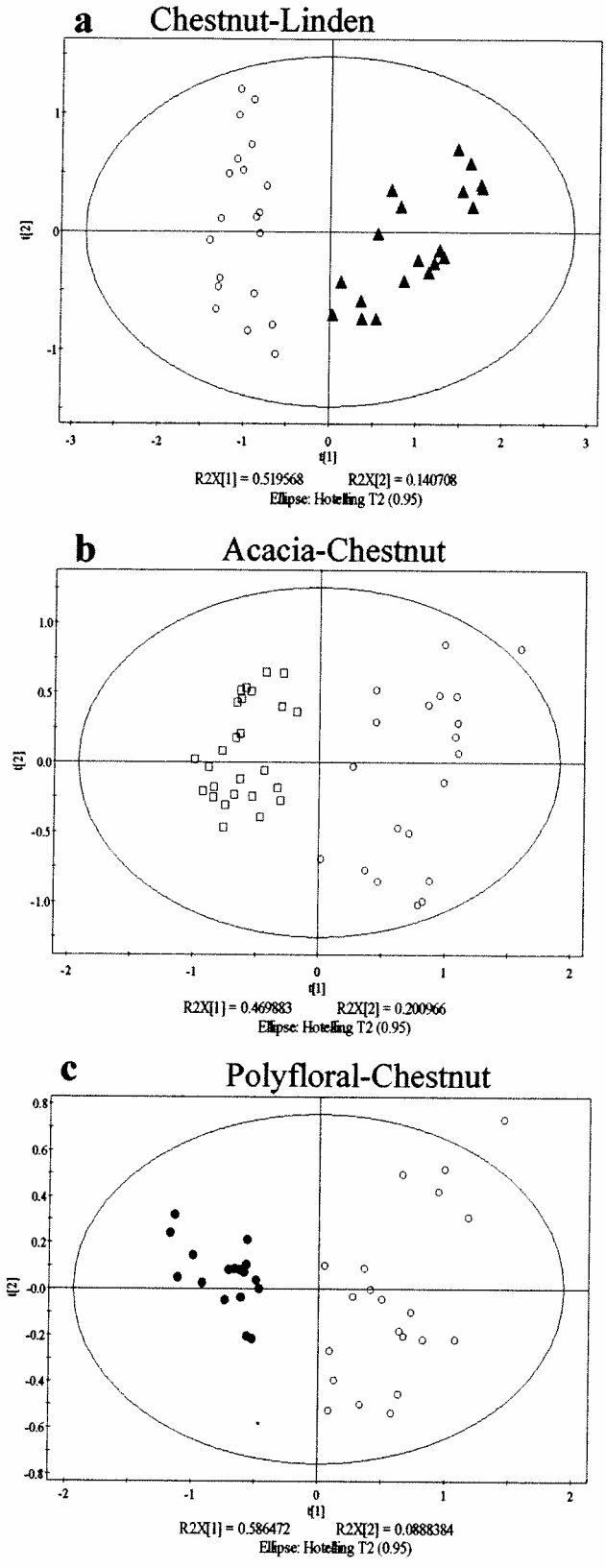


Figure 3. PCA score plots of the analysis performed on the 85 samples of the training set. Two botanical origins were compared at a time. As an example, the results obtained comparing chestnut honey with the other three types of honeys are reported: (**a**) chestnut−linden; (**b**) acacia−chestnut; (**c**) polyfloral−chestnut.

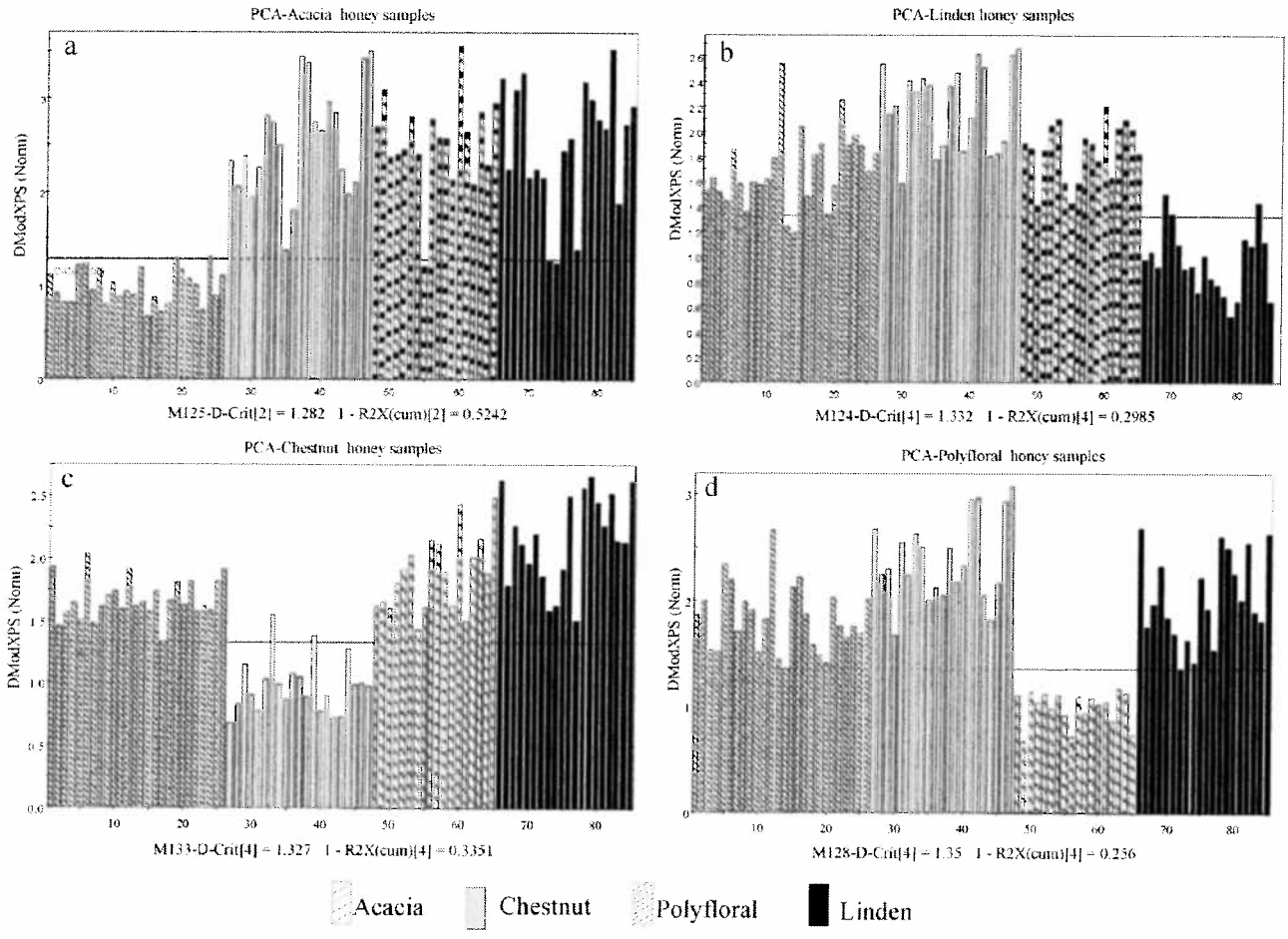


Figure 4. Distance to the model (DmodXPS) plots for the training set: (**a**) DModX plot taking acacia honey as a reference; (**b**) DModX plot taking linden honey as a reference; (**c**) DModX plot taking chestnut honey as a reference; (**d**) DModX plot taking polyfloral honey as a reference. The red lines represent the maximum tolerable distance (Dcrit) for the considered data set. Moderate outliers have DModX values larger than Dcrit.

(i) 1D selective TOCSY spectra were recorded with a 180° Gaussian shaped pulse of 79.8 ms, a TOCSY mixing time of 70 ms, 32 scans, and 16K data points (*31*).

(ii) ^{1}H−^{1}H TOCSY spectra were recorded in the time-proportional phase incrementation (TPPI) mode, with a spectral window of 10 ppm in both dimensions, 2048 × 512 data points, 2 s relaxation delay, 70 ms mixing time, and 32−128 scans.

(iii) ^{1}H−^{1}H COSY spectra were recorded in magnitude mode, with a spectral window of 10 ppm in both dimensions, 2048 × 512 data points, 2 s relaxation delay, and 16 scans.

(iv) ^{1}H−^{1}H NOESY spectra were recorded in the TPPI mode, with a spectral window of 10 ppm in both dimensions, 2048 × 512 data points, 2 s relaxation delay, 1.2 s mixing time, and 128−256 scans.

(v) HMQC spectra were recorded in the TPPI mode, with a spectral window of 10 ppm (^{1}H) and 220 ppm (^{13}C), 1 s relaxation delay, 1024 × 256 data points, and 128−256 scans.

(vi) HMBC spectra were recorded in the TPPI mode, with a spectral window of 10 ppm (^{1}H) and 220 ppm (^{13}C), 1 s relaxation delay, 1024 × 256 data points, and 600−800 scans.

Statistical Analysis. Principal component analysis (PCA) and PLS-DA using "mean centering" as data pretreatment (*32*) was conducted using the software SIMCA-P11 (Umetrics, Umea, Sweden).

Data were visualized by plotting either the PC scores, where each point in the score plot represents an individual sample, or the loading plot, which permits us to identify the spectral regions with the greatest influence on the separation and clustering of the samples and, therefore, to deduce which compounds are responsible for such clustering (markers).

To validate the robustness of the discrimination, the samples were divided into a training set and a test set. The latter was composed of two randomly selected samples from each botanical class and of the samples purchased in commercial stores for a total of 33 samples (i.e., 9 acacia, 10 chestnut, 6 linden, 7 polyfloral, and 1 declared chestnut−linden). The statistical analysis was performed on the training set composed of 85 samples: 26 acacia, 21 chestnut, 20 linden, and 18 polyfloral honeys.

SIMCA-P calculates normalized prediction distances of the samples to the models (DmodX) whose critical values (Dcrit) were computed with 0.95 confidence intervals. The distance of each sample to each of the PC models was computed and plotted in a DmodX plot (*33*). This approach was used to assess the classification performance of samples by predicting class membership and to evaluate the specificity of the models.

A PLS-DA model was also applied, separating the samples into classes according to their origin. The validity of the PLS-DA model was assessed using the correlation coefficient R^2 and the cross-validation correlation coefficient Q^2. The latter was derived using the default option of SIMCA-P. The same training set and test set were used as in the PCA.

RESULTS AND DISCUSSION

Spectral Analysis. Representative spectra of a honey chloroform extract are reported in **Figure 1**. The spectrum of **Figure 1a** was obtained with the standard single-pulse sequence in 50 min. The spectrum is dominated by signals in the 0−2 ppm region, which originate from protons of hydrocarbon chains belonging to high-molecular-weight *n*-alkanes or linear fatty acids already identified as honey constituents (*34*). These signals are not relevant for our study; they hide resonances of other molecules that may be markers

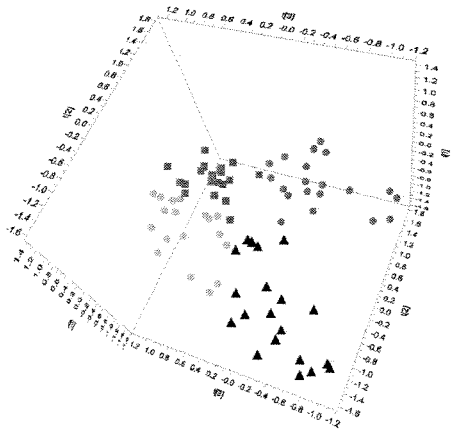


Figure 5. PLS-DA 3D score plot of the honey classification model. The axes of the plot are PLS-DA components 1−3. Red, blue, black, and green colors indicate acacia, chestnut, linden, and polyfloral honeys, respectively.

of botanical origin and, furthermore, limit the usable receiver gain value so that the very weak signals are inaccurately sampled and poor integrated intensities result. Removal of the strongest signals, by means of the sequence described in Materials and Methods (*30*), allowed an increase of the receiver gain, which resulted in shorter acquisition times, improved digitization of the small amplitude peaks, and lower integration errors.

A typical spectrum obtained with the modified DPFGSE sequence is reported in **Figure 1b**; it was obtained in only 25 min and has the same S/N ratio of the spectrum of **Figure 1a**.

The spectrum appears very crowded in the entire spectral range and provides global information about the complex extraction mixture, although deciphering the chemical content of such samples from the NMR data is less straightforward. In general, the advantage of NMR is that all types of compounds give rise to signals simultaneously, so that the NMR spectrum represents a fingerprint of the matrix under study. The choice of chloroform as a solvent offers great advantages compared to other solvents previously used in NMR studies of honey. The residual chloroform signal is very sharp and hides a very small region at 7.26 ppm which does not influence the analysis. On the other hand, solvents such as DMSO and MeOH are less suitable, since they present large signals in very important areas (around 3.4 ppm for MeOH and around 2.5 ppm for DMSO).

Superimposing the spectra of all the analyzed samples reveals the regions in which signals present only in one honey type are clearly visible. As an example, two expansions of the spectra are reported in **Figure 2**, in which signals exclusively from linden samples (top) and chestnut samples (bottom) appear.

To evaluate the differences between the various botanical origins, a chemometric analysis was carried out using SIMCA-P, as described in Materials and Methods.

Evaluation of the Discriminant Ability of the NMR Spectra. As a first step, an unsupervised approach by means of PCA was applied to the training set. PC analysis was performed, comparing two botanical origins at a time. The plots of the first two PCs of all the PCAs demonstrate that very good discrimination of honeys according to their botanical origins is reached. All R^2 and Q^2 values were over 70% and 60%, respectively. As an example, the comparisons of chestnut honey with the other honey types are reported in **Figure 3**.

This procedure provided the variables responsible for sample separation and produced a final data matrix with optimal sample clusterization.

Table 1. Classification List Reprojected onto the PLS-DA Models Performed by Considering All Training Set Samples.[a]

sample		probability of class membership			
botanical origin	geographic origin	acacia	chestnut	linden	polyfloral
		Training Set			
acacia	Veneto	**0.77**	−0.15	0.23	0.14
acacia	Veneto	**0.80**	0.09	0.08	0.03
acacia	Veneto	**0.97**	−0.03	−0.02	0.07
acacia	Veneto	**0.96**	−0.03	0.02	0.06
acacia	Veneto	**1.02**	0.30	0.12	−0.44
chestnut	Veneto	0.03	**1.20**	−0.15	−0.08
chestnut	Veneto	0.23	**0.91**	−0.10	−0.04
chestnut	Veneto	−0.08	**0.95**	−0.05	0.18
chestnut	Veneto	0.05	**0.61**	0.26	0.08
chestnut	Veneto	0.01	**1.16**	−0.12	−0.05
linden	Veneto	0.26	0.02	**0.90**	−0.18
linden	Veneto	0.02	0.08	**0.90**	−0.01
linden	Veneto	0.04	−0.06	**0.93**	0.08
linden	Veneto	−0.32	0.38	**0.97**	−0.03
linden	Veneto	−0.04	0.08	**0.74**	0.22
polyfloral	Veneto	0.13	0.11	−0.13	**0.90**
polyfloral	Veneto	−0.24	0.17	0.03	**1.05**
polyfloral	Veneto	−0.15	0.22	0.24	**0.69**
polyfloral	Veneto	0.07	0.05	0.00	**0.87**
polyfloral	Veneto	0.06	0.14	0.06	**0.74**
		Test Set			
acacia*	Veneto	**0.62**	−0.05	0.20	0.23
acacia*	Veneto	**1.04**	0.07	−0.16	0.05
acacia	Abruzzo	**0.86**	0.00	0.32	-0.18
acacia	France	**0.85**	−0.05	−0.24	0.44
acacia	Italy	**1.06**	−0.12	−0.12	0.18
acacia	France	**0.93**	−0.13	0.12	0.08
acacia	Veneto	**0.62**	−0.13	0.22	0.20
acacia	Italia	**0.53**	−0.11	0.10	0.49
acacia	Italy	0.43	−0.09	0.09	***0.57***
chestnut*	Veneto	−0.21	**0.87**	0.29	0.05
chestnut*	Veneto	−0.19	**0.71**	0.32	0.16
chestnut	Italy	−0.05	**1.47**	−0.43	0.02
chestnut	France	−0.08	**0.69**	0.12	0.27
chestnut	Italy	−0.17	**0.93**	−0.08	0.31
chestnut	Piedmont	−0.49	**1.10**	0.24	0.14
chestnut	Italy	−0.26	**1.01**	−0.03	0.28
chestnut	Veneto	0.04	**0.76**	0.22	-0.02
chestnut	Veneto	0.06	**1.14**	−0.27	0.08
chestnut	France	0.03	0.28	0.20	***0.48***
linden*	Italy	0.04	0.08	**0.70**	0.19
linden*	Veneto	0.54	−0.08	**0.99**	-0.45
linden	Italy	−0.04	0.05	**1.20**	-0.21
linden	Italy	−0.03	0.01	**0.80**	0.22
linden	France	0.30	0.35	**0.52**	0.42
linden	Italy	0.15	−0.1	0.40	***0.55***
polyfloral*	Veneto	0.00	0.02	0.20	**0.79**
polyfloral*	Veneto	0.09	0.12	0.05	**0.74**
polyfloral	Austria	0.33	0.08	0.16	**0.42**
polyfloral	Italy	0.48	0.00	−0.03	**0.55**
polyfloral	Italy	0.55	−0.12	−0.09	**0.66**
polyfloral	Lithuania	0.22	0.18	0.02	**0.62**
polyfloral	Veneto	0.24	0.25	0.12	**0.63**
chestnut−linden	Piedmont	−0.29	**0.89**	0.41	−0.01

[a] The results of only five training set samples per botanical origin are reported for clarity. The results of all test samples are reported. The eight test samples randomly selected from those obtained directly from the producers with certified origin are marked with an asterisk. The complete classification list is given in the Supporting Information.

By modeling each class separately through PCA, it became evident that each botanical origin is characterized by specific

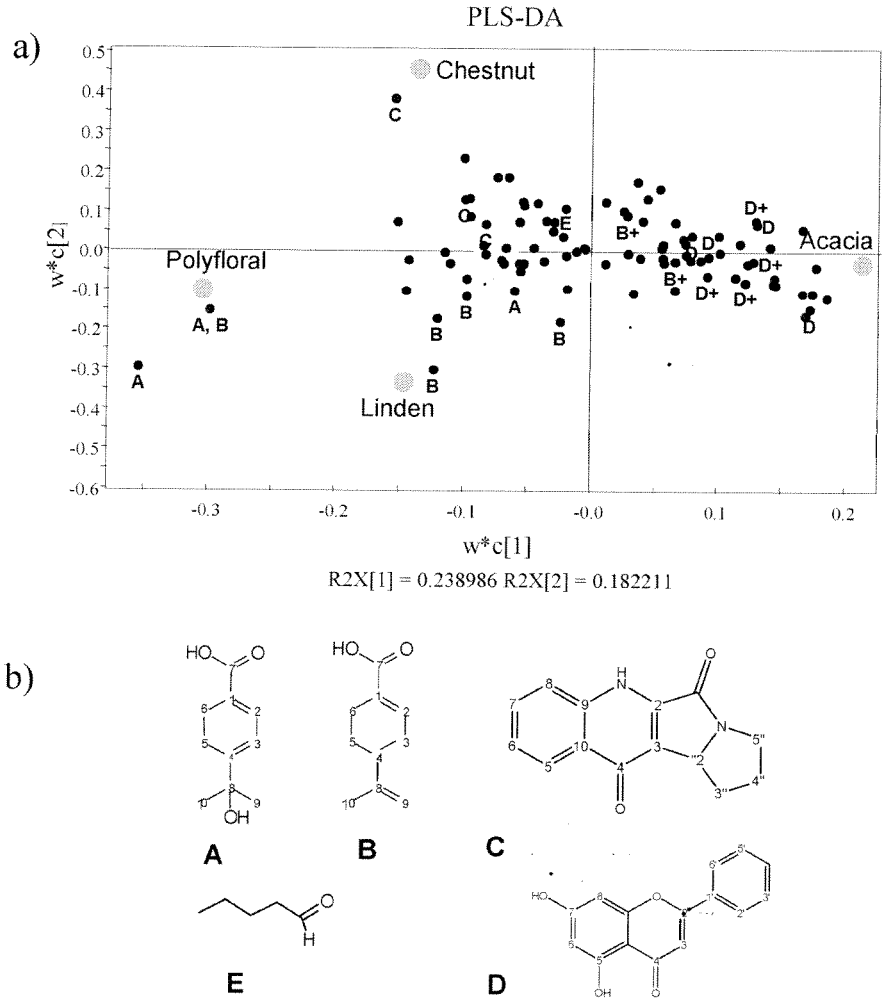


Figure 6. (**a**) PLS-DA weight plot for the 85 samples of the training set. The identified markers are indicated with letters. (**b**) Structures of identified markers.

resonances. The specificity of the models was demonstrated with the DmodX criterion and the critical value (Dcrit) with 95% confidence intervals. These parameters show the distance of each sample to the model in the X space (*33*).

In **Figure 4**, all the samples were compared using the Dcrit generated for each of them. We found that almost all acacia observations of the prediction set are predicted within the critical distance. Moreover, the other three species are far outside the tolerance interval of the acacia model. This means that the PCA model trained on acacia recognizes honeys of this botanical origin in the classification phase, and it is specific for this species. Few ambiguities exist (five false positive and four false negative results). In each of these cases, though, the correct classification is obtained by comparing all the models.

The good discrimination obtained with the PC analysis prompted us to perform a PLS-DA classification on the training test to derive a model with high prediction ability. The PLS-DA model obtained shows high discrimination, with $R^2 = 0.67$, $Q^2 = 0.77$, and $K = 0.87$ (**Figure 5**). The robustness of the PLS-DA model was evaluated using the classification list shown in **Table 1**. Each sample was classified by means of a "probability of class membership" indicative of its representativeness. The results of correct and ambiguous classifications obtained for both training and test set are displayed in **Table 1**: when the "probability of class membership" is larger than 0.5 (highlighted in bold), the object is considered correctly predicted; the samples incorrectly predicted are indicated in bold italics. All the samples belonging to the training set and the eight test samples (labeled with an asterisk in **Table 1**) randomly selected from those directly obtained from the producers with certified origin were correctly classified (see Table SI-1 in the Supporting Information). Also, all the commercial test samples fit the model space defined by the training set by using Hotelling's T^2 test and the distance to the model test. Two samples, declared to be acacia and linden, respectively, were classified as polyfloral, although the "probability of class membership" for the origin specified in the label was higher than that of the other monofloral types. The honey sample declared to contain both linden and chestnut honeys is correctly classified; in fact, it is evident by the class index that both botanical origins are present.

The variables responsible for this discrimination are shown in the weight plot (**Figure 6a**) that graphs together the *x*-variable weights "*w*" and the *y*-variable weights "*c*" (*35*), showing the relationships between the selected variables and the four botanical origins. Some of these resonances are labeled with letters and correspond to the identified compounds reported in **Figure 6b**.

Characterization of Botanical Markers. The identification of the compounds that give rise to resonances characteristic of the different botanical origins was undertaken using both 1D and 2D NMR techniques.

To identify markers of linden honey, TOCSY, COSY, NOESY, HMQC, and HMBC spectra were acquired directly

Table 2. ^{1}H and ^{13}C NMR (ppm) Assignment of Compound **C**

position	integration	^{1}H	^{13}C	COSY	HMBC	NOESY
1			175.1			
2			126.3			
3						
4			139.8			
5	1	7.57	118.4		124.2, 132.8, 127.5	
6	1	7.42	124.2		118.4, 127.5, 132.8	7.57, 8.44
8	1	8.44	126.3		132.8, 139.8, 175.1	7.42
7	1	7.69	132.8		126.3,139.8	7.42, 7.57
9			127.5			
2′			167.0			
2″	1	4.80	63.0	2.63, 1.26	29.26 (weak), 126.3, 167.0	2.42, 2.45, 2.63, 3.52
3″a	1	1.26	29.26	2.42, 2.45, 2.63, 4.80	63, 126.3	
3″b	1	2.63	29.26	1.26, 2.42, 2.45, 4.80	41.52	4.80, 1.26
4″a	1	2.42	29.26	1.26, 2.63, 3.52, 3.66	63.0	2.62, 3.52, 4.80
4″b	1	2.45	29.26	1.26, 2.63, 3.52, 3.66	63.0	1.26, 3.68
5″a	1	3.52	41.52	2.42, 2.45, 3.66	29.26	2.42, 3.66
5″b	1	3.66	41.52	2.42, 2.45, 3.52	29.26, 167.0	2.45, 3.52
HN		9.74				7.57

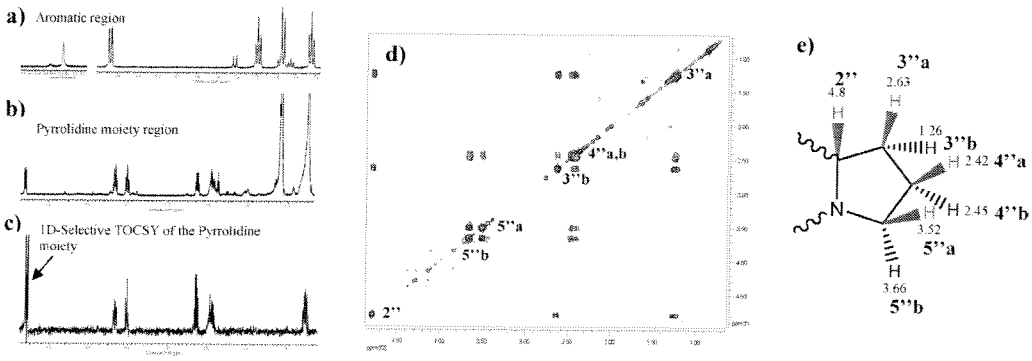


Figure 7. Expansions of the aromatic region (**a**) and of the pyrrolidine moiety region (**b**) of the purified chestnut marker. (**c**) Selective TOCSY spectrum, obtained with a 70 ms mixing time and a selective 180° Gaussian shaped pulse of 79.8 ms centered at 4.08 ppm. All the resonances belonging to the pyrrolidine moiety are present. (**d**) Portion of a COSY spectrum with the assignment of the pyrrolidine moiety. (**e**) Structure and chemical shift assignment of the pyrrolidine moiety.

on an organic extract. We identified two principal terpene acids from chemical shift information, mass spectroscopy data, and literature data (*36*): i.e., 4-(1-hydroxy-1-methylethyl)cyclohexa-1,3-dienecarboxylic acid (**A**) and 4-(1-methylethenyl)cyclohexa-1,3-dienecarboxylic acid (**B**). In all the analyzed linden honey samples, signals from compound **B** were more intense than those from compound **A**. The mass spectroscopy data and the ^{1}H and ^{13}C assignments are reported in the Supporting Information (Figure SI-1 and Table SI-2).

The identification of markers for chestnut and acacia honeys was conducted after purification of the organic extracts through silica gel column chromatography. Many fractions were analyzed, and some of them were identified. Here, we present preliminary results.

A fraction from acacia honey returned a HRESi(+)MS pseudomolecular ion of m/z 254.05 corresponding to the molecular formula $C_{15}H_{10}O_4$. The ^{1}H NMR resonances of this fraction (the assignment is reported in Table SI-3 of the Supporting Information) allowed us to identify this compound as crysin (**D**). This substance is present in higher quantity in this honey but is also present in the other honeys. The compound labeled with the letter **E** in **Figure 6a** corresponds to hexanal.

The chestnut marker (**C**) was isolated and identified as the compound recently detected by Beretta et al., who characterized it in DMSO (*37*). Our NMR spectra were obtained in $CDCl_3$, and the complete assignment is reported in **Table 2**. Our assignment agrees with that of Beretta et al., except for protons 3″ and 4″. This point is illustrated in **Figure 7**, where a selection of the COSY spectrum is shown. The NOESY correlations allowed us to obtain also the stereospecific assignment, shown in **Figure 7e**. In Figures SI-2 and SI-3 (Supporting Information), the UV spectrum and the mass spectrum are reported, respectively. Interestingly, the UV spectrum is identical with that of an unidentified marker of chestnut honey recently reported (*38*).

In conclusion, we presented a simple method to determine the botanical origin of honey, characterized by rapid sample preparation and short acquisition and processing time of the spectra. The PLS-DA approach permitted a discrimination of samples by calculating a specific model for each honey type with pratically no overlap. The results are promising in the perspective of developing similar models suitable to identify honeys of other botanical origins. The method we developed is able to distinguish also polyfloral honeys, allowing a complete classification of all the honey types analyzed. Each honey sample can be classified by means of the "probability of class membership", reported in **Table 1** and in Table SI-1 (Supporting Information). Specifically, we chose a value of 0.50 as the cutoff to assign a sample to that class. The possible presence of other botanical types is indicated by a high value of the corresponding class.

Many of the methods mentioned in the Introduction allow one to clearly discriminate between several types of unifloral honeys, but polyfloral honeys, which represent the majority of the honeys produced, are rarely considered. Only two methods applied discriminant analysis on a high number of samples (*20*, *22*). Devillers et al. (*22*) obtained 100% of correct classification on eight different monofloral honey types, but they used data from eight independent measurements (conductivity, pH, free acidity, and percentages of fructose, glucose, and raffinose). On the basis of front-face fluorescence spectroscopy, Ruoff et al. (*20*) concluded that the classification rates for the unifloral honeys were generally > 90%, whereas the classification rate for the polyfloral honeys ranged between 48 and 75%. This method seems a promising approach to discriminate between several unifloral as well as polyfloral honey samples using a single spectroscopic measurement. Our method is also based on a single spectroscopic measurement, and it shows two principal advantages with respect to that proposed by Ruoff et al.: i.e., reduced sample preparation time and better molecular characterization of the components.

To be used in practice, it would be necessary to extend the domain of application of the method we propose to other categories of unifloral honeys and to expand the database. Work is in progress in this direction, also with the aim to identify new chemical entities as botanical markers.

Supporting Information Available: Tables and figures giving additional characterization and test data. This material is available free of charge via the Internet at http://pubs.acs.org.

LITERATURE CITED

(1) Council Directive of the European Union Council Directive 2001/110/EC December 2001 relating to honey. *Off. J. Eur. Communities: Legis.* **2002**, *L10*, 47−52.

(2) Bogdanov, S. Nature and origin of the antibacterial substances in honey. *Lebensm.-Wiss.-Technol.* **1997**, *30*, 748–753.

(3) Weston, R. J. The contribution of catalase and other natural products to antibacterial activity of honey: a review. *Food Chem.* **2000**, *71*, 235–239.

(4) Buratti, S.; Benedetti, S.; Cosio, M. S. Evaluation of the antioxidant power of honey, propolis and royal jelly by amperometric flow injection analysis. *Talanta* **2007**, *71*, 1387–1392.

(5) Beretta, G.; Granata, P.; Ferrero, M.; Orioli, M.; Maffei Facino, R. Standardization of antioxidant properties of honey by a combination of spectrophotometric/fluorimetric assays and chemometrics. *Anal. Chim. Acta* **2005**, *533*, 185–191.

(6) Mphande, A. N. G.; Killowe, C.; Phailira, S.; Jones, H. W.; Harrison, W. J. Effect of honey and sugar dressings on wound healing. *J. Wound Care* **2007**, *16*, 317–319.

(7) Al-Waili, N. S. Natural honey lowers plasma glucose, C-reactive protein, homocysteine, and blood lipids in healthy, diabetic, and hyperlipidemic subjects: comparison with dextrose and sucrose. *J. Med. Food* **2004**, *7*, 100.

(8) Oršolić, N.; Terzić, S.; Šver, L.; Bašić, I. Honey-bee products in prevention and/or therapy of murine transplantable tumours. *J. Sci. Food Agric.* **2005**, *85*, 363–370.

(9) Conti, M. E.; Botrè, F. Honeybees and their products as potential biondicators of heavy metals contamination. *Environ. Monit. Assess.* **2001**, *69*, 267–282.

(10) Al, M. L.; Daniel, D.; Moise, A.; Bobis, O.; Laslo, L.; Bogdanov, S. Physico-chemical and bioactive properties of different floral origin honeys from Romania. *Food Chem.* **2009**, *112*, 863–867.

(11) Azeredo, L. da C.; Azeredo, M. A. A.; de Souza, S. R.; Dutra, V. M. L. Protein contents and physicochemical properties in honey samples of Apis mellifera of different floral origins. *Food Chem.* **2003**, *80*, 249–254.

(12) Codex Alimentarius, CODEX STAN 12, Revised Codex standard for honey, http://www.codexalimentarius.net/web/standard_list.jsp.

(13) Maurizio, A. Microscopy of honey. In *Honey: A Comprehensive Survey*; Crane, E., Ed.; Heinemann: London, 1979; pp 240−257.

(14) Anklam, E. A review of the analytical methods to determine the geographical and botanical origin of honey. *Food Chem.* **1998**, *63*, 549–562.

(15) Bogodanov, S.; Ruoff, K.; Persano Oddo, L. Physico-chemical methods for the characterisation of unifloral honeys: a review. *Apidologie* **2004**, *35*, S4–S17.

(16) Arvanitoyannis, I. S.; Chalhoub, C.; Gotsiou, P.; Lydakis-Simantiris, N.; Kefalas, P. Novel quality control methods in conjunction with chemometrics (multivariate analysis) for detecting honey authenticity. *Crit. Rev. Food Sci. Nutr.* **2005**, *45*, 193–203.

(17) Rebane, R.; Herodes, K. Evaluation of the botanical origin of estonian uni and polyfloral honeys by amino acid content. *J. Agric. Food Chem.* **2008**, *56*, 10716–10720.

(18) Cotte, J. F.; Casabianca, H.; Giroud, B.; Albert, M.; Lheritier, J.; Grenier-Loustalot, M. F. Characterization of honey amino acid profiles using high-pressure liquid chromatography to control authenticity. *Anal. Bioanal. Chem.* **2004**, *478*, 1342–1350.

(19) Goodacre, R.; Radovic, B. S.; Anklam, E. Progress toward the rapid non-destructive assessment of the floral origin of European honey using dispersive Raman spectroscopy. *Appl. Spectrosc.* **2002**, *556*, 521–527.

(20) Ruoff, K.; Luginbühl, W.; Künzli, R.; Bogdanov, S.; Bosset, O. J.; von der Ohe, K.; von der Ohe, W.; Amadò, R. Authentication of the botanical and geographical origin if honey by front-face fluorescence spectroscopy. *J. Agric. Food Chem.* **2006**, *54*, 6858–6866.

(21) Serrano, S.; Villarejo, M.; Espejo, R.; Jodral, M. Chemical and physical parameters of Andalusian honey: classification of citrus and eucalyptus honeys by discriminant analysis. *Food Chem.* **2004**, *87*, 619–625.

(22) Devillers, J.; Morlot, M.; Pham-Delegue, M. H.; Dore, J. C. Classification of monofloral honeys based on their quality control data. *Food Chem.* **2004**, *86*, 305–312.

(23) Fernández-Torres, R.; Pérez-Bernal, J. L.; Bello-López, M. Á.; Callejón-Mochón, M.; Jiménez-Sánchez, J. C.; Guirarúm-Pérez, A. Mineral content and botanical origin of Spanish honeys. *Talanta* **2005**, *65*, 686–691.

(24) Beretta, G.; Caneva, E.; Regazzoni, L.; Golbamaki Bakhtyari, N.; Maffei Facino, R. A solid-phase extraction procedure coupled to ^{1}H NMR, with chemometric analysis, to seek reliable markers of the botanical origin of honey. *Anal. Chim. Acta* **2008**, *620*, 176–182.

(25) Lolli, M.; Bertelli, D.; Plessi, M.; Sabatini, A. G.; Restani, C. Classification of Italian honeys by 2D HR-NMR. *J. Agric. Food Chem.* **2008**, *56*, 1298–1304.

(26) Donarski, J. A.; Jones, S. A.; Charlton, A. J. Application of cryoprobe ^{1}H nuclear magnetic resonance spectroscopy and multivariate analysis for the verification of Corsican honey. *J. Agric. Food Chem.* **2008**, *56*, 5451–5456.

(27) Consonni, R.; Cagliani, L. Geographical characterization of polyfloral and acacia honeys by nuclear magnetic resonance and chemometrics. *J. Agric. Food Chem.* **2008**, *56*, 6873–6880.

(28) Belton, P. S.; Delgadillo, I.; Gil, A. M.; Holmes, E.; Nicholls, A.; Nicholson, J. K.; Spraul, M. Use of high-field ^{1}H NMR spectroscopy for the analysis of liquid foods. *J. Agric. Food Chem.* **1996**, *44*, 1483–1487.

(29) Schievano, E.; Pasini, G.; Cozzi, G.; Mammi, S. Identification of the production chain of asiago d'allevo cheese by nuclear magnetic resonance spectroscopy and principal component analysis. *J. Agric. Food Chem.* **2008**, *56*, 7208–7214.

(30) Rastrelli, F.; Schievano, E.; Bagno, A.; Mammi, S. NMR quantification of trace components in complex matrices by band-selective excitation with adiabatic pulses. *Magn. Reson. Chem.* **2009**, *47*, 868–872.

(31) Sandusky, P.; Raftery, D. Use of selective TOCSY NMR experiments for quantifying ninor components in complex mixtures: application to the metabonomics of amino acids in honey. *Anal. Chem.* **2005**, *77*, 2455–2463.

(32) Wold, S.; Kettaneh, N.; Tiessem, K. J. Hierarchical multiblock PLS and PC models for easier model interepretation and as an alternative to variable selection. *J. Chemom.* **1996**, *10*, 463–482.

(33) *User's Guide SIMCA Edition*; Umetrics AB: Umeå, Sweden, 2006.

(34) Jerković, I.; Mastelić, J.; Marijanović, Z. A variety of volatile compounds as markers in uniflloral honey from Dalmatia sage (Salvia officinalis L.). *Chem. Biodivers.* **2006**, *3*, 1307–1316.
(35) Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: a basic tool of chemometrics. *L. Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130.
(36) Frérot, E.; Velluz, A.; Decorzant, E.; Naef, R. From linden flower to linden honey. Part 2: glicosidic precursors of cyclohexa-1,3-diene-1-carboxylic acids. *Chem. Biodivers.* **2006**, *3*, 94–100.
(37) Beretta, G.; Vistoli, G.; Caneva, E.; Anselmi, C.; Maffei Facino, R. Structure elucidation and NMR assignment of two pyrrolidinyl quinoline alkaloids from chestnut honey. *Magn. Reson. Chem.* **2009**, *47*, 456–459.
(38) Truchado, P.; Martos, I.; Bortolotti, A.; Sabatini, G.; Ferreres, F.; Tomas-Barberan, F. A. Use of quinoline alkaloids as markers of the floral origin of chestnut honey. *J. Agric. Food Chem.* **2009**, *57*, 5680–5686.

Received for review July 3, 2009. Revised manuscript received October 8, 2009. Accepted October 28, 2009. The financial support of the Veneto Region (UNIMIELE project, 2008) is gratefully acknowledged.

Table SI-1. Classification list reprojected onto the PLS_DA models performed by considering all training set samples. The results of all training set samples and all test samples are reported. The eight test samples randomly selected from those obtained directly from the producers with certified origin are marked with an asterisk.

Samples			**Probability of class membership**			
Training set	***Botanic origin***	***Geografic origin***	***Acacia***	***Chestnut***	***Linden***	***Polyfloral***
	Acacia	Veneto	**0.77**	-0.15	0.23	0.14
	Acacia	Veneto	**0.80**	0.09	0.08	0.03
	Acacia	Veneto	**0.97**	-0.03	-0.02	0.07
	Acacia	Veneto	**0.96**	-0.03	0.02	0.06
	Acacia	Veneto	**1.02**	0.30	0.12	-0.44
	Acacia	Veneto	**1.03**	-0.11	0.22	-0.14
	Acacia	Veneto	**0.82**	0.10	0.14	-0.05
	Acacia	Veneto	**1.01**	0.02	0.09	-0.12
	Acacia	Veneto	**0.98**	-0.02	0.12	-0.09
	Acacia	Veneto	**0.90**	-0.15	0.10	0.15
	Acacia	Veneto	**0.85**	0.00	-0.09	0.24
	Acacia	Veneto	**1.12**	0.00	-0.07	-0.05
	Acacia	Veneto	**0.66**	-0.01	0.19	0.16
	Acacia	Veneto	**0.62**	0.02	0.20	0.16
	Acacia	Veneto	**1.11**	0.07	-0.13	-0.05
	Acacia	Veneto	**0.88**	0.10	0.23	-0.21
	Acacia	Veneto	**0.78**	0.17	0.00	0.06
	Acacia	Veneto	**1.03**	-0.05	-0.13	0.15
	Acacia	Veneto	**0.60**	-0.04	0.51	-0.06
	Acacia	Veneto	**0.85**	0.04	0.04	0.08
	Acacia	Veneto	**1.08**	0.10	-0.23	0.04
	Acacia	Veneto	**0.94**	0.09	0.03	-0.07
	Acacia	Veneto	**0.97**	0.03	-0.07	0.07
	Acacia	Veneto	**0.75**	0.14	-0.19	0.30
	Acacia	Veneto	**1.11**	-0.03	-0.08	0.01
	Acacia	Veneto	**1.14**	-0.05	-0.11	0.02
	Chestnut	Veneto	0.03	**1.20**	-0.15	-0.08
	Chestnut	Veneto	0.23	**0.91**	-0.10	-0.04
	Chestnut	Veneto	-0.08	**0.95**	-0.05	0.18
	Chestnut	Veneto	0.05	**0.61**	0.26	0.08
	Chestnut	Veneto	0.01	**1.16**	-0.12	-0.05
	Chestnut	Veneto	0.05	**0.90**	-0.03	0.08
	Chestnut	Veneto	0.05	**1.07**	-0.04	-0.08
	Chestnut	Veneto	0.10	**1.14**	-0.16	-0.08
	Chestnut	Veneto	0.24	**0.62**	0.08	0.06
	Chestnut	Veneto	0.10	**0.77**	0.12	0.01
	Chestnut	Veneto	-0.12	**0.77**	0.03	0.32
	Chestnut	Veneto	0.13	**0.91**	-0.11	0.07
	Chestnut	Veneto	-0.10	**0.80**	0.13	0.17
	Chestnut	Veneto	-0.08	**0.98**	0.08	0.03
	Chestnut	Veneto	0.02	**1.12**	-0.05	-0.10
	Chestnut	Veneto	0.08	**0.98**	-0.03	-0.03
	Chestnut	Veneto	0.04	**0.82**	0.00	0.15
	Chestnut	Veneto	0.30	**0.64**	0.05	0.00
	Chestnut	Veneto	0.16	**0.64**	0.05	0.15
	Chestnut	Veneto	-0.40	**1.19**	0.26	-0.05
	Chestnut	Veneto	-0.19	**1.13**	0.11	-0.05
	Linden	Veneto	0.26	0.02	**0.90**	-0.18
	Linden	Veneto	0.02	0.08	**0.90**	-0.01
	Linden	Veneto	0.04	-0.06	**0.93**	0.08
	Linden	Veneto	-0.32	0.38	**0.97**	-0.03
	Linden	Veneto	-0.04	0.08	**0.74**	0.22
	Linden	Veneto	-0.02	-0.04	**0.86**	0.20
	Linden	Veneto	0.04	-0.03	**0.75**	0.23
	Linden	Veneto	0.22	0.10	**0.68**	0.01
	Linden	Veneto	0.21	0.02	**0.67**	0.09
	Linden	Veneto	0.17	0.20	**0.68**	-0.05
	Linden	Veneto	0.01	-0.04	**0.94**	0.10
	Linden	Veneto	0.17	0.17	**0.67**	-0.01

	Linden	Italy	-0.10	0.16	**1.10**	-0.16
	Linden	Italy	0.00	-0.06	**1.13**	-0.07
	Linden	Italy	-0.09	0.00	**1.11**	-0.02
	Linden	Veneto	-0.02	0.06	**0.98**	-0.01
	Linden	Veneto	-0.02	-0.14	**1.02**	0.13
	Linden	Veneto	0.41	-0.10	**0.84**	-0.15
	Linden	Veneto	0.00	-0.13	**0.86**	0.27
	Linden	Veneto	-0.05	0.01	**1.13**	-0.09
	Polyfloral	Veneto	0.13	0.11	-0.13	**0.90**
	Polyfloral	Veneto	-0.24	0.17	0.03	**1.05**
	Polyfloral	Veneto	-0.15	0.22	0.24	**0.69**
	Polyfloral	Veneto	0.07	0.05	0.00	**0.87**
	Polyfloral	Veneto	0.06	0.14	0.06	**0.74**
	Polyfloral	Veneto	0.02	-0.04	0.14	**0.89**
	Polyfloral	Veneto	-0.06	0.23	0.04	**0.79**
	Polyfloral	Veneto	0.40	-0.03	0.09	**0.53**
	Polyfloral	Veneto	0.02	0.05	0.16	**0.77**
	Polyfloral	Veneto	-0.07	-0.15	0.24	**0.98**
	Polyfloral	Veneto	-0.12	-0.06	0.15	**1.02**
	Polyfloral	Veneto	0.08	0.09	-0.17	**1.00**
	Polyfloral	Veneto	-0.03	0.07	-0.17	**1.13**
	Polyfloral	Veneto	-0.33	0.28	0.21	**0.84**
	Polyfloral	Veneto	0.14	-0.08	-0.13	**1.07**
	Polyfloral	Veneto	0.26	-0.15	-0.21	**1.10**
	Polyfloral	Veneto	0.17	-0.06	-0.05	**0.95**
	Polyfloral	Veneto	0.12	-0.07	0.07	**0.88**
Test set						
	Acacia*	Veneto	**0.62**	-0.05	0.20	0.23
	Acacia*	Veneto	**1.04**	0.07	-0.16	0.05
	Acacia	Abruzzo	**0.86**	0.00	0.32	-0.18
	Acacia	France	**0.85**	-0.05	-0.24	0.44
	Acacia	Italy	**1.06**	-0.12	-0.12	0.18
	Acacia	France	**0.93**	-0.13	0.12	0.08
	Acacia	Veneto	**0.62**	-0.13	0.22	0.20
	Acacia	Italia	**0.53**	-0.11	0.10	0.49
	Acacia	Italy	0.43	-0.09	0.09	***0.57***
	Chestnut*	Veneto	-0.21	**0.87**	0.29	0.05
	Chestnut*	Veneto	-0.19	**0.71**	0.32	0.16
	Chestnut	Italy	-0.05	**1.47**	-0.43	0.02
	Chestnut	France	-0.08	**0.69**	0.12	0.27
	Chestnut	Italy	-0.17	**0.93**	-0.08	0.31
	Chestnut	Piedmont	-0.49	**1.10**	0.24	0.14
	Chestnut	Italy	-0.26	**1.01**	-0.03	0.28
	Chestnut	Veneto	0.04	**0.76**	0.22	-0.02
	Chestnut	Veneto	0.06	**1.14**	-0.27	0.08
	Chestnut	France	0.03	0.28	0.20	***0.48***
	Linden*	Italy	0.04	0.08	**0.70**	0.19
	Linden*	Veneto	0.54	-0.08	**0.99**	-0.45
	Linden	Italy	-0.04	0.05	**1.20**	-0.21
	Linden	Italy	-0.03	0.01	**0.80**	0.22
	Linden	France	0.30	0.35	**0.52**	0.42
	Linden	Italy	0.15	-0.1	0.40	***0.55***
	Polyfloral*	Veneto	0.00	0.02	0.20	**0.79**
	Polyfloral*	Veneto	0.09	0.12	0.05	**0.74**
	Polyfloral	Austria	0.33	0.08	0.16	**0.42**
	Polyfloral	Italy	0.48	0.00	-0.03	**0.55**
	Polyfloral	Italy	0.55	-0.12	-0.09	**0.66**
	Polyfloral	Lithuania	0.22	0.18	0.02	**0.62**
	Polyfloral	Veneto	0.24	0.25	0.12	**0.63**
	Chestnut-linden	Piedmont	-0.29	**0.89**	**0.41**	-0.01

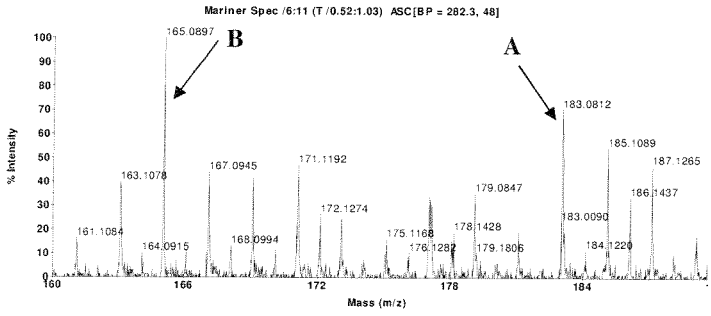
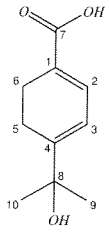


Figure SI-1. ESI-Mass spectroscopy of an extract of linden honey. The HRESi(+)MS pseudomolecular ions, corresponding to 4-(1-hydroxy-1-methylethyl)cyclohexa-1,3-diene-carboxylic acid (**A**) and 4-(1-methylethenyl)cyclohexa-1,3-diene-carboxylic acid (**B**), appear as the most intense peaks in the chromatogram.

A



B

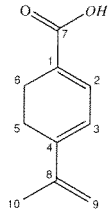
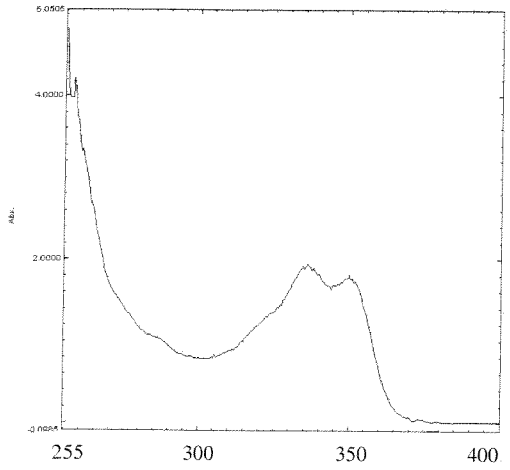


Table SI-2. ^{1}H, ^{13}C assignment of compounds **A** and **B**

	Compound **A**		Compound **B**	
Position	^{1}H	^{13}C	^{1}H	^{13}C
2	7.15	135.6	7.24	136.8
3	6.13	117.9	6.19	120.4
4				
5	2.50	23.9	2.52	23.9
6	2.50	21.7	2.52	21.7
7				
8				
9	1.79		5.18, 5.31	115.8
10	1.79		1.97	



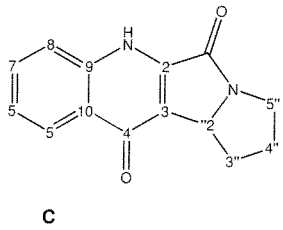


Figure SI-2. UV spectrum of the chestnut honey marker (**C**), a kynurenic acid derivative.

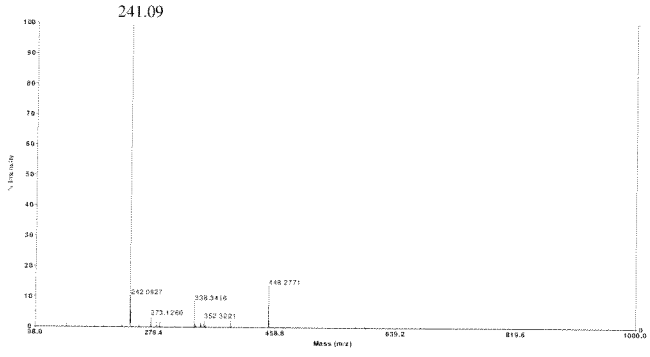
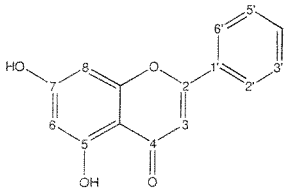


Figure SI-3. Mass spectroscopy of the chestnut honey marker. The HRESi(+)MS pseudomolecular ions corresponds to compound (**C**).

Table SI-3. ^{1}H assignment of Chrysin (**D**)

Position	^{1}H
6', 2'	7.54
3',4',5'	7.89
3	6.68
6	6.31
8	6.48



D

Deleted: ¶
Ta

Deleted: ¶
<sp>

Metabolomics
DOI 10.1007/s11306-011-0362-8

ORIGINAL ARTICLE

An NMR-based metabolomic approach to identify the botanical origin of honey

Elisabetta Schievano · Matteo Stocchero · Elisa Morelato · Chiara Facchin · Stefano Mammi

Received: 22 July 2011 / Accepted: 26 August 2011

Abstract NMR can be used in food analysis for origin discrimination and biomarker discovery using a metabolomic approach. Here, we present an example of this strategy to discriminate honey samples of different botanical origins. The NMR spectra of 353 chloroform extracts of selected honey samples were analyzed to detect possible markers of their floral origin. Six monofloral Italian honey types (acacia, linden, orange, eucalyptus, chestnut, and honeydew) were analyzed together with polyfloral samples. Specific markers were identified for each monofloral origin: two markers for acacia (chrysin and pinocembrin), one for chestnut (γ-LACT-3-PKA), two for orange (8-hydroxylinalool and caffeine), one for eucalyptus (dehydrovomifoliol), one for honeydew (a diacylglycerilether) and two for linden (4-(1-hydroxy-1-methylethyl)cyclohexa-1,3-diene-carboxylic acid and 4-(1-methylethenyl)cyclohexa-1,3-diene-carboxylic acid). An NMR-based metabolomic approach that used O2PLS-DA multivariate data analysis allowed us to discriminate the different types of honey. Two different classifiers were built based on different multivariate techniques. The high precision of the classification obtained suggests that this approach could be useful to develop generally applicable metabolomic tools to discriminate the origin of honey samples.

E. Schievano (✉) · E. Morelato · C. Facchin · S. Mammi
Dipartimento di Scienze Chimiche, Università di Padova,
Via Marzolo, 1, Padua 35131, Italy
e-mail: elisabetta.schievano@unipd.it

M. Stocchero
S-IN Soluzioni Informatiche, Via Ferrari 14,
36100 Vicenza, Italy

Keywords Honey · Botanical markers · Metabolomics · NMR (nuclear magnetic resonance) · O2PLS-DA multivariate analysis

1 Introduction

Humans have consumed honey as a food for at least 6000 years. Honey is the only sweetening agent that can be used without processing, and it is appreciated not only for its taste and aroma but also for its various nutritional and medicinal properties (Hennessy et al. 2010).

Honey is a very complex food, produced by bees from the nectar of blossoming plants or exudates secreted by another insect such as an aphid or other plant sap-sucking insects. Consequently, its properties and composition depend on bee species, species of the nectar-providing plant, geographic area, weather conditions, mode of storage and even harvesting technology and conditions (Castro-Vázquez et al. 2007).

Consumer choice is linked to unique organoleptic and aromatic properties of honey that depend principally on the botanical and geographical origins of the product. For this reason, origin assessment is important in quality control of honey as well as of many other foods. Honey regulation in the EU (Codex Alimentarius Commission 2001; European Commission 2002) states that the botanical and geographical origins of the product must be printed on the label. The control of commercially sold honey therefore requires the determination of parameters that unambiguously establish the origin as part of the overall quality of honey.

The classical approach to verify the botanical origin of honey is to use several complementary analytical methods. The number of physicochemical parameters necessary for a

Published online: 13 September 2011

complete, unambiguous characterization is very high, as they include viscosity, color, hygroscopicity, electrical conductivity, specific rotation, crystallization, density, pH and refractive index (Persano Oddo and Piro 2004). For this reason, melissopalinological properties are more commonly used. However, melissopalynology is not an easy technique because it requires very experienced analysts and it is very time consuming. According to some authors, melissopalynology is valid only for the determination of geographical origin of honey while it is less valid to determine its botanical origin (Stephens et al. 2010). As a result, faster characterization methods, which would be suitable for fast routine analysis of honey, are now being sought with the help of novel instrumental techniques (Kaškoniene and Venskutonis 2010).

Recently, many studies, based on the analysis of the compositional data of honey volatile compounds, phenolic acids, flavonoids, aroma compounds, carbohydrates, amino acids, and other constituents, have been performed with the aim to identify specific chemical marker compounds for honey samples of various origins. In a recent review (Kaškoniene and Venskutonis 2010), the studies carried out during the last two decades are summarized. The data reviewed on floral markers suggest that a better classification of various types of honey can be achieved by more than a single class of compounds, preferably in combination with statistical analysis of the data.

Spectroscopic techniques such as IR (Etzold and Lichtenberg-Kraag 2008; Bertelli et al. 2007), fluorescence (Ruoff et al. 2006), and NMR (Schievano et al. 2010) can be applied for metabolite fingerprinting. All these techniques can provide quantitative information without complex pre-treatment of the samples. The advantage of NMR spectroscopy is that it is able to provide structural information as well, and on a wide range of chemical species in a single experiment.

The objective of this study was to create an NMR spectra library of authentic monofloral and polyfloral Italian honeys collected over 2 years. The NMR spectra of chloroform extracts were acquired as recently described (Schievano et al. 2010). These spectra, containing high numbers of organic compounds including volatile and less volatile ones, can be considered fingerprints of each floral origin. We employed an NMR-based metabolomic approach (Wishart 2008) that uses O2PLS-DA multivariate data analysis to discriminate the different types of honey. A model was created with 242 samples from seven different botanical origins and then it was applied to 107 unknown samples to estimate its predictive success. The encouraging results obtained suggest that this approach could be useful to develop generally applicable metabolomic tools to discriminate the origin of honey samples.

The statistical approach singled out specific resonances belonging to molecular markers for each type of honey. One or two marker compounds for each botanical origin were identified and characterized.

2 Materials and methods

2.1 Chemicals

$CDCl_3$ (99.96%-d, H_2O <0.005% Euriso-Top, Gif sur Yvette, France), standard compounds (chrysin, 97% pure; pinocembrin, 95% pure; kynurenic acid, 98% pure) were purchased from Sigma Aldrich (Milan, Italy).

2.2 Honey samples

We analyzed 353 honey samples of seven different floral origins, obtained either from Rigoni S.p.A. or from trusted Veneto apiaries: 75 acacia, 60 chestnut, 62 linden, 40 orange, 32 eucalyptus, 36 honeydew, 48 polyfloral. Their botanical origin was ascertained by means of sensorial analysis. Generally, honey is classified by the floral source of the nectar from which it was made; however, honeydew is a type of honey produced by honeybees collecting a liquid secreted by another insect such as an aphid or other plant sap-sucking insects. Sometimes it is not properly considered "honey", primarily for its origin but also for its particular dark brown color and because it is not as sweet as honeys produced from nectar. Nevertheless, we decided to include it in the classification of the honey types analyzed in this work.

2.3 Sample preparation

The data base was built on ^{1}H-NMR spectra of chloroform extracts. A liquid/liquid extraction was carried out as reported by Schievano et al. (2010). Briefly, 6 g of honey were treated with 15 ml of water and 15 ml of chloroform and the mixture was mechanically stirred for 10 min after adjusting the pH of the aqueous solutions to 2.5. After centrifugation at 10000 rpm for 15 min, the lower chloroform phase was evaporated under a gentle stream of nitrogen. The solid residue was dissolved in 600 µl of $CDCl_3$ and put in an NMR tube.

The procedures to identify different marker compounds were specifically optimized for each botanical origin.

The markers of linden honey were characterized, with the aid of literature data, directly in the organic honey extract, because of their strong and isolated signals. The identification of markers for chestnut, acacia, orange and eucalyptus honeys were achieved after purification of the organic extracts through silica gel column chromatography.

Preliminary data on chestnut and acacia markers have been reported elsewhere (Schievano et al. 2010).

About 36–60 g of each type of honey were used to identify marker molecules. The chloroform extracts were initially treated with an acetonitrile:water = 95:5 solution to precipitate hydrophobic compounds, such as waxes. The solutions were dried and purified through a 25 cm × 2 cm silica gel column, equilibrated in chloroform. Eluent solutions were prepared with a 0–5% MeOH step gradient in $CHCl_3$.

The orange honey markers were both found in the 1.25% fraction; the eucalyptus honey marker was found in the 0.5% fraction while the acacia honey markers were both found in the 0.25% fraction. Different from the other markers, the honeydew marker precipitated from an acetonitrile:water = 95:5 solution and was purified on a silica gel column with petroleum ether:ethyl acetate = 3:1 as eluent.

In general, the marker compounds were obtained in sufficient purity, suitable for their spectroscopic identification.

2.4 NMR spectroscopy

One dimensional spectra were obtained with a 600-MHz Bruker Avance spectrometer. For each sample, 256 transient were collected as 64K points with a spectral width of 14 ppm, relaxation time of 2 s, at 298 K, using a modified double pulsed field gradient spin echoes sequence (Rastrelli et al. 2009). The total acquisition time for each NMR spectrum was 25 min. The FID (free induction decay) signal was zero filled to 128K and exponentially multiplied with a line broadening of 0.5 Hz before Fourier transformation. Phase and baseline correction were performed manually.

For the structural elucidation of the marker compounds, two dimensional homo and heteronuclear correlation NMR spectra (DQF-COSY, TOCSY, NOESY, HMQC, and HMBC) were collected. The heteronuclear experiments were acquired with gradient coherence selection while homonuclear experiments employed phase cycling for the same purpose. TOCSY spectra were acquired with a spectral window (SW) of 6400 Hz in both dimensions, 2K data points in the F2 dimension, 400–512 t_1 increments, 16 scans and 100–120 ms mixing times. COSY spectra were recorded in magnitude mode, with a SW of 10 ppm in both dimensions, 2K data points, 1 s relaxation delay, 16 scans and 400–512 t_1 increments. NOESY spectra were recorded in the TPPI mode, with a SW of 10 ppm in both dimensions, 2K data points, 2 s relaxation delay, 1.2 s mixing time, 128–256 scans and 512 t_1 increments. ^1H-^{13}C-HMQC spectra were recorded in the TPPI mode, with a SW of 10 ppm (^1H) and 220 ppm (^{13}C), 1K data points, 1 s relaxation delay, 128–256 scans and 180–200 experiments. ^1H-^{13}C HMBC spectra were recorded in the TPPI mode, with a SW of 10 ppm (^1H) and 220 ppm (^{13}C), 1K data points, 1 s relaxation delay, 600-800 scans and 180–200 experiments.

2.5 Mass spectrometry

Molecular masses of compounds were determined on a Mariner (PerSeptive Biosystems, Foster City, CA) mass spectrometer. The mass was calibrated using a mixture of neurotensin, angiotensin and bradykinin, at a concentration of 1 pmol/μl, as external standard.

2.6 Data preparation and statistical analysis

The spectral region from 2.16 to 13.00 ppm was bucketed and integrated using a constant interval of 0.04 ppm. The intervals containing the signals of the solvent were removed and the sum of the remaining integrals normalized for each spectrum. A data set of 353 samples and 265 variables was obtained.

In our study, we considered seven different types of honey. Usually, projection methods for classification, such as PLS-DA or O2PLS-DA, are able to produce efficient classification models for not more than 4–5 classes of samples. On the other hand, the SIMCA approach is not suitable for complex problems where Principal Component Analysis (PCA) is not able to extract the information useful for discrimination. For these reasons, we adopted a hierarchical approach in modeling our data. At the first level of modeling, PCA and O2PLS-DA (Trygg and Wold 2003; Paris et al. 2010) were applied to recognize patterns and to extract the hidden information useful for classification. In this stage, outliers were identified, single variables able to distinguish a particular type of honey from the others were highlighted, and a new set of super-variables was defined or calculated. These super-variables were used at the second level of modeling to build different classifiers based on a Naïve Bayes approach and Support Vector Machines (SVM) (Ivanciuc 2007).

PCA and O2PLS-DA were performed using the SIMCA P+ 11.0 package (Umetrics, Umeå, Sweden) while the Naïve Bayes classifier and C-SVC (Ivanciuc 2007) were implemented by WEKA 3.6.4 (University of Waikato, New Zealand) using the WLSVM tool for Support Vector Machines (EL-Manzalawy and Honavar 2005). The Response Surface approach and the designs for sampling were applied using MODDE 9.0 (Umetrics, Umeå, Sweden).

3 Results and discussions

3.1 Spectral analysis

Figure 1 shows representative NMR spectra of each different botanical origin. These ^1H NMR spectra were

obtained mixing the chloroform extracts of six different honey samples of the same botanical origin. To select these representative samples, single PCA models were obtained for each botanical group; D-optimal design with one central point was then applied to sample the space described by the PCA scores. In general, the highest signals are in the aliphatic region while signals of comparable intensity are found in the other regions. All the regions appear very crowded. Specifically, many peaks are present in the aromatic zone (6.0–9.0 ppm), but also signals from aldehydes, phenols and carboxylic acid are present.

Because of the important differences in the NMR spectra of the different honeys, these spectra can be considered as fingerprints characterizing each botanical origin. Although the information of the entire spectrum is necessary for a correct identification of the sample, some distinct differences in the chemical composition of metabolites among the various samples are already apparent by close

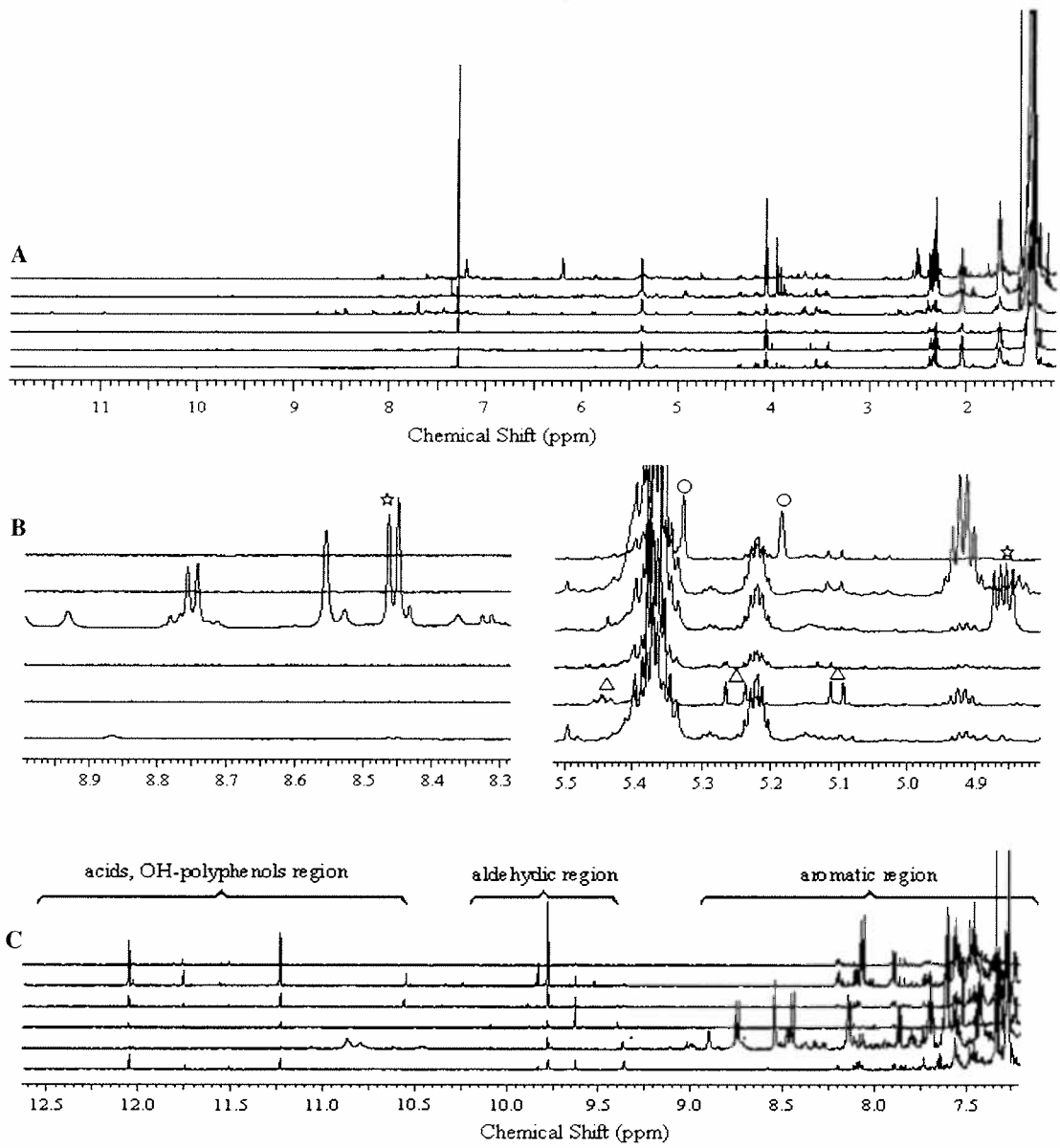


Fig. 1 **a** ^{1}H spectra of representative samples of six different monofloral honeys. From *top to bottom*: Linden, Eucalyptus, Chestnut, Acacia, Citrus and Honeydew. **b** Expansions of two regions in which resonances of some protons of the chestnut, orange and linden metabolites are highlighted. *Stars, open circles* and *open triangles* are on resonances of chestnut, linden, and orange markers, respectively. **c** Expansion of the 12.6–7.2 ppm region of the spectrum

inspection of specific portions of the spectra. These differences are visible in the zooms of two regions of the spectra (Fig. 1b, c), in which some resonances characterizing chestnut, orange and linden honeys are highlighted.

The statistical value of the differences observed in these spectra was assessed using the techniques described below, analyzing the spectra of the 353 samples individually. These spectra were recorded with the sequence reported in a previous work (Rastrelli et al. 2009), optimized to eliminate the strongest aliphatic signals and therefore to use a higher receiver gain and to improve the digitalization of weaker signals.

3.2 Statistical analysis

Pareto scaling and mean centering were applied before multivariate data analysis (Eriksson et al. 2006). The data set of 353 samples was investigated by performing PCA for each class of samples. On the basis of the Hotelling's T2 test at the confidence level of 95%, we identified four strong outliers, which were excluded from the dataset.

Balanced training (242 samples) and test sets (107 samples) were extracted from the remaining 349 samples by applying onion D-optimal design (Olsson et al. 2004) on the PCA scores performed considering each individual class. The PCA models for each class were also used to define the applicability domain for the classification models described in the following. We considered a given sample to belong to the applicability domain of the classification model if it belonged to at least one PCA model on the basis of its DModX+ parameter (Eriksson et al. 2006).

The training set was used to build *one-versus-all* O2PLS-DA models where each class of honey was compared with the other classes considered together as a unique class. In Table 1, the parameters describing the O2PLS-DA models obtained are reported.

The predictive component of each O2PLS-DA model was used both to highlight the variables corresponding to the signals of the putative markers for the individual classes of honey and to obtain a new set of descriptive variables useful to support a Naïve Bayes classifier. Indeed, inspecting the so called S-plot (Wiklund et al. 2008) for the *one-versus-all* discriminant models, we were able to select a total of 35 variables, which were interpreted to suggest the nature of putative markers for each class of honey, except for the class polyfloral. Twenty-two of these variables were assigned to chemical entities. All the selected variables were significant on the basis both of the *t*-test for the means and of the Mann–Whitney test for the medians (*P*-value less than 0.001). Specifically, we found 6 variables for the class orange (4 of these corresponding to signals attributed to the two putative marker compounds reported in the following), 7 for the class chestnut (1 of these corresponding to signals belonging to the putative marker), 8 for the class acacia (5 variables associated to signals produced by the two putative markers), 9 for the class linden (5 variables attributed to signals of the two putative markers), 1 for the putative marker of the class eucalyptus and 4 for that of the class honeydew. In Fig. 2, the S-plots used to characterize the class linden and the class honeydew are reported as examples.

The second use of the scores of the predictive component of the individual O2PLS-DA models was as descriptive variables to build a Naïve Bayes classifier useful to transform the scores of the discriminant models into the probability to belong to a specific class. For the test set, the combination of the O2PLS-DA and the Naïve Bayes classifier allowed us to obtain the predictions reported in Table 5 (Cohen's kappa coefficient equal to 0.91 and 7.5% of incorrectly classified samples).

A different classifier was built by considering SVM for classification. In this case, the 35 variables selected after inspection of the S-plots were used as descriptive variables. The training set was autoscaled and the space of the parameters defining the C-SVC investigated by applying a Response Surface approach based on Central Composite Face design. The performance of each C-SVC was quantified by the calculation of the Cohen's kappa coefficient during cross-validation with seven groups. Different kernel functions were tested. The best C-SVC was obtained using the radial basis function kernel (Cohen's kappa coefficient in cross-validation equal to 0.84, cost parameter equal to 200 and gamma equal to 0.05). The predictions obtained for the test set are summarised in the confusion matrix of Table 2 (Cohen's kappa coefficient equal to 0.90 and 8.4% of incorrectly classified samples).

The predictions of the two classifiers were compared in order to build a new classifier based on the *consensus* method. The rule used to obtain a class prediction for a given sample was the following: if the two classifiers predicted the same class, we accepted this prediction while, in the case of different class predictions, we considered the sample as unpredictable. Table 3 summarizes the predictions for the test set obtained by applying the *consensus*

Table 1 O2PLS-DA models for the *one-versus-all* strategy

Model	A	R^2	Q^2
Orange	1 + 2	0.72	0.66
Chestnut	1 + 2	0.73	0.66
Acacia	1 + 2	0.72	0.65
Linden	1 + 3	0.61	0.50
Eucalyptus	1 + 3	0.63	0.40
Honeydew	1 + 4	0.64	0.42
Polyfloral	1 + 3	0.45	0.20

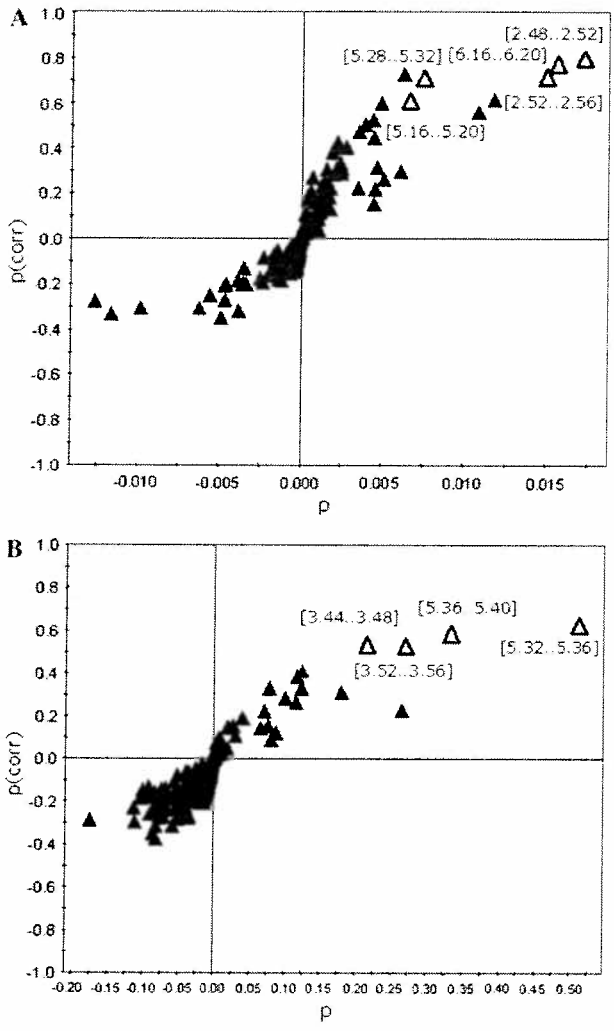


Fig. 2 S-plots for the O2PLS-DA models used to characterize the class linden (**a**) and the class honeydew (**b**). The signals of the putative markers are reported in *open triangles*

method. In general, the precision of the predictions was higher than the individual classifiers (Cohen's kappa coefficient equal to 0.95 and 4.1% of incorrectly classified samples) while the sensitivity was reduced as an effect of the increased number of unpredictable samples (column U in Table 4).

3.3 Marker compounds identification

Having achieved efficient separation and statistical validation, we explored the O2PLS-DA models to identify the marker signals underlying the separation of origins as previously described (Fig. 2).

To assign the important resonances to specific markers, the honey extracts were treated as described in paragraph 2.3. The NMR spectra of the different chromatographic fractions were analyzed to identify the resonances of interest. Unequivocal structural identification of the marker compounds was obtained using high-field (600 MHz) homo and hetero-correlation 2D-NMR experiments and MS analysis. The complete proton and carbon resonance assignment of all the compounds was obtained. In Fig. 3, the molecular structures of the markers are reported. The protons corresponding to relevant signals present in the S-plots are labelled in bold and marked with a star.

The marker compounds of acacia honey are chrysin (A1) and pinocembrin (A2), and they are responsible for four of the eight important signals found in the predictive component of the O2PLS-DA model for acacia. The other four signals were not assigned.

The O2PLS-DA model of orange honey highlighted six resonances, all arising from caffeine (F2) and 8-hydroxy-linalool (F1). The signal identifying the eucalyptus honey comes from dehydrovomifoliol, compound D.

All the seven signals characterizing the chestnut honey fall in the 7.90-8.50 ppm region. We assigned the signal at 8.45 ppm to H8 of γ-LACT-3PKA. The other signals probably arise from closely structurally related metabolites. It has been recently described that chestnut honey is rich in quinoline alkaloids (Beretta et al. 2009a, b). Further studies are needed to identify these molecules.

Table 2 Confusion matrix summarizing the predictions on the test set for the one-versus-all strategy

	P_OR	P_CH	P_AC	P_LI	P_EU	P_HO	P_PO	Total	Sensitivity (%)
OR	11	0	0	0	0	0	1	12	91.7
CH	0	16	1	1	0	0	0	18	88.9
AC	0	0	21	0	0	0	0	21	100.0
LI	0	1	0	17	0	0	0	18	94.1
EU	0	0	0	0	12	0	0	12	100.0
HO	0	0	0	0	0	10	1	11	90.9
PO	1	0	1	1	0	0	12	15	80.0
Precision (%)	91.7	94.1	91.3	89.5	100.0	100.0	85.7		

P prediction of the specified botanical origin, *OR* orange, *CH* chestnut, *AC* acacia, *EU* eucalyptus, *HO* honeydew, *PO* polyfloral

Table 3 Confusion matrix summarizing the predictions on the test set for the C-SVC model

	P_OR	P_CH	P_AC	P_LI	P_EU	P_HO	P_PO	Total	Sensitivity (%)
OR	11	0	0	0	0	0	1	12	91.7
CH	0	16	1	1	0	0	0	18	88.9
AC	0	0	21	0	0	0	0	21	100.0
LI	0	0	0	18	0	0	0	18	100.0
EU	0	0	1	0	10	0	1	12	83.3
HO	0	0	0	0	1	10	0	11	90.9
PO	0	0	2	0	0	1	12	15	80.0
Precision (%)	100.0	100.0	84.0	94.7	90.9	90.9	85.7		

P prediction of the specified botanical origin, *OR* orange, *CH* chestnut, *AC* acacia, *EU* eucalyptus, *HO* honeydew, *PO* polyfloral

Table 4 Confusion matrix summarizing the predictions on the test set for the consensus approach

	P_OR	P_CH	P_AC	P_LI	P_EU	P_HO	P_PO	U	N	Sensitivity (%)
OR	11	0	0	0	0	0	1	0	12	91.7
CH	0	16	1	1	0	0	0	0	18	88.9
AC	0	0	21	0	0	0	0	0	21	100.0
LI	0	0	0	17	0	0	0	1	18	94.4
EU	0	0	0	0	10	0	0	2	12	83.3
HO	0	0	0	0	0	9	0	2	11	81.8
PO	0	0	1	0	0	0	10	4	15	66.7
Precision (%)	100.0	100.0	91.3	94.4	100.0	100.0	90.9			

P prediction of the specified botanical origin, *OR* orange, *CH* chestnut, *AC* acacia, *EU* eucalyptus, *HO* honeydew, *PO* polyfloral

The S-plot, for the O2PLS-DA models used to characterize the class linden (Fig. 2a) highlights five signals belonging to the monoterpene acids, C1 and C2 reported in Fig. 3.

For honeydew, all four resonances (Fig. 2b) are assigned to a diacylglycerylether (compound B in Fig. 3).

The complete assignment of the resonances of the identified markers is reported in Table 5. As an example, Fig. 4 shows portions of ^{1}H-^{1}H TOCSY, NOESY and ^{1}H-^{13}C HSQC and HMBC spectra of the marker isolated from eucalyptus honey.

3.4 Discussion

Two different classifiers were built based on different multivariate techniques. Both classifiers were very robust in classification. The consensus approach allowed us to increase the precision of the classification obtained by considering the two classifiers separately (higher than 91%). A deep insight in the role played by the different metabolites was achieved analyzing the first level of our hierarchical model. An accurate analysis on fractions from silica gel column of honey chloroform extracts allowed us to identify a significant number of markers that discriminate the different types of honey. We found specific chemical biomarkers for several botanical origins. Other compounds, although present in all samples, identify specific botanical origins because of their relatively higher concentration.

The specific marker for chestnut honey is γ-LACT-3PKA, a molecule recently found (Beretta et al. 2009a, b; Schievano et al. 2010). This compound is specific for this type of honey and it was not present in other types of honey.

The biomarker useful to discriminate eucalyptus from the other honeys investigated here is the norisoprenoid dehydrovomifoliol (4-hydroxy-3,5,5-trimethyl-4-(3-oxo-1-butenyl)-2-cyclohexen-1-one). Other works confirm the presence of this compound in eucalyptus honey (Alissandrakis et al. 2011); however, dehydrovomifoliol should be considered a nonspecific chemical biomarker because it is found also in heather honey (Plutowska et al. 2011).

The orange blossom honey is characterized by two molecules, 8-hydroxylinalool and caffeine. 8-hydroxylinalool was found as one of the most abundant terpenes of citrus honeys (Escriche et al. 2011). Both compounds were recently proposed as fingerprint markers for the description of citrus honey (Melliou and Chinou 2011).

In linden honey, we found two monoterpene acids, cyclohexa-1,3-diene-1-carboxylic acid and 4-(1-hydroxy-1-methylethenyl)cyclohexa-1,3-diene-carboxylic acid. The occurrence of cyclohexa-1,3-diene monoterpenes is rare

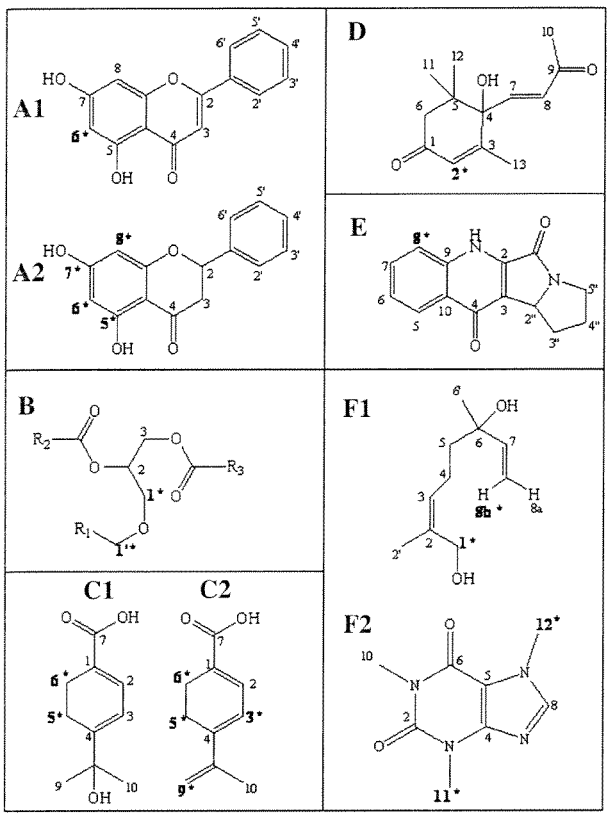


Fig. 3 Structures of the marker compounds identified. A1 = Acacia (Chrysin), A2 = Acacia (Pinocembrin), B = Honeydew (Diacylglycerylether), C1 = Linden (4-(1-hydroxy-1-methylethyl)cyclohexa-1,3-diene-carboxylic acid), C2 = Linden (4-(1-methylethenyl)cyclohexa-1,3-diene-carboxylic acid), D = Eucalyptus (dehydrovomifoliol), E = Chestnut (γ-LACT-3-PKA), F1 = Orange (8-hydroxylinalool), F2 = Orange (caffeine). The protons that originate the signals extracted from the S-Plots are labelled in bold and marked with a star

and some authors (Frérot et al. 2006) identified these acids in the form of their 1-O-β-gentobiosyl esters that are stored both in linden honey and linden nectar. They defined these compounds as specific biomarkers of this type of honey. A more recent work reports the glycosidic ester of C1 (Fig. 3) and in lower concentration the free form, also in chestnut honey (Truchado et al. 2009a, b). Our findings, however, indicate that the simultaneous presence of the two monoterpene acids is typical of linden honey and the occurrence of C1 in some chestnut honey samples likely arises from blending of chestnut nectar with linden nectar. The problem of discriminating linden and chestnut monofloral honeys is particularly acute because the two plants bloom at the same time and in the same habitat. The statistical method presented here, which makes use of the whole metabolic profile for discrimination, supports the classification provided from the sensorial analysis. In fact, only one chestnut sample is classified as linden while in several chestnut and linden samples the identified markers for both types of honey are present. The honeydew honey is distinguished from the other types by a diacylglycerylether, a compound present in all the other types but in relatively lower percentage. The three positions of glycerol are occupied by long chain fatty acids. To the best of our knowledge, the occurrence of this compound has not been reported yet, although the presence of fatty acids in honey has been studied (Alissandrakis et al. 2011, Isidorov et al. 2011). This compound probably is related to the composition of the combs and cuticular waxes (Tulloch and Hoffman 1972; Graddon et al. 1979; Bonaga et al. 1986; Tan et al. 1988).

Pinocembrin and chrysin are the two flavonoids found as markers for acacia honey. In general, flavonoids are found in almost all honey types (Alvarez-Suarez et al. 2010; Truchado et al. 2009a, b; Marghitas et al. 2010). With our extraction method, most flavonoids might be dissolved in the water fraction. Both pinocembrin and chrysin are rather hydrophobic and partition preferably in the organic phase. Although these two compounds were found in many other kinds of honey (Bertoncelj et al. 2011), we can identify them as acacia markers because of the higher limit of quantitation (LOQ) of our method. For chrysin and pinocembrin, the LOQ was estimated to be 0.64 mg/kg and this is higher than the concentration found by Bertoncelj et al. (2011) in all the honey types analyzed, except for acacia.

Interestingly, almost all the marker compounds presented in this paper are characterized by a specific biological activity.

γ-LACT-3PKA, for example, is a quinoline alkaloid derivative endowed with potential antinociceptive activity. It seems to play a role in the honeybee's tryptophan metabolism (Beretta et al. 2009a, b).

Dehydrovomifoliol (4-hydroxy-3,5,5-trimethyl-4-(3-oxo-1-butenyl)-2-cyclohexen-1-one), a derivative of abscisic acid, is a growth phytohormone (Guyot et al. 1999).

8-hydroxylinalool was demonstrated to have a special anti-microbial activity against Gram (±) bacteria strains and against pathogenic fungi, confirming the traditional reputation of honey as an antimicrobial agent (Melliou and Chinou 2011). This compound is also one of the important components of the essential oils of the flowers and fruits of the citrus genus.

Beside the known beneficial effects of caffeine, a recent article has also demonstrated that caffeine intake is associated with reduced hepatic fibrosis (Modi et al. 2010).

Also pinocembrin and chrysin have a demonstrated biological activity. Chrysin has antioxidant properties (Socha et al. 2011) and seems to play a really important action in reducing the metabolic activity of melanoma cells (Pichichero et al. 2010). Pinocembrin has been described as a good protector against oxidation and apoptosis induced by ischemia–reperfusion in rat brain (Liu et al. 2008).

Table 5 Assignments of the markers

Acacia						Honeydew		
Compound **A1**			Compound **A2**			Compound **B**		
Chrysin			Pinocembrin			Diacylglycerylether		
Position	δ H (ppm)	δ C (ppm)	Position	δ H (ppm)	δ C (ppm)	Position	δ H (ppm)	δ C (ppm)
2	–	163.1	2	5.44	78.8	**1***	3.56	68.2
3	6.68	103.0	3a	3.10	43.8	**1′***	3.45	71.4
4	–	181.8	3b	2.84	43.3	1R2	–	173.3
5-OH	12.81	161.4	4	–	197.3	1R3	–	173.7
6*	6.31	98.1	**5*-OH**	12.05	165.4	2	5.21	69.5
7-OH	12.81	164.3	**6***	6.01	95.6	2R1	1.56	29.8
8	6.48	94.1	**7*-OH**	12.05	168.3	2a(R2,R3)	2.30	34.0
9	–	157.0	**8***	6.01	97.2	2b(R2,R3)	2.33	
10	–	105.1	9	–	164.6	3a	4.18	61.8
1′	–	130.6	10	–	103.3	3b	4.35	
2′	7.89	126.3	1′	–	–	3R1	1.30	22.5
3′	7.54	129.1	2′	7.44	127.3	3 (R2,R3)	1.64	25.2
4′	7.54	132.0	3′	7.44	129.6	$-(CH_2)_n-$(R2,R3)	1.30	
5′	7.54	129.1	4′	7.44	129.7	$-CH_2-CR=$(R2,R3)	2.04	27.8
6′	7.89	126.3	5′	7.44	129.6	CH=CH (R2,R3)	5.36	129.9
			6′	7.44	127.3			

Linden						Eucalyptus		
Compound **C1**			Compound **C2**			Compound **D**		
4-(1-hydroxy-1-methylethyl) cyclohexa-1,3-diene-carboxylic acid			4-(1-methylethenyl) cyclohexa-1,3-diene-carboxylic acid			Dehydrovomifoliol		
Position	δ H(ppm)	δ C(ppm)	Position	δ H(ppm)	δ C(ppm)	Position	δ H(ppm)	δ C(ppm)
1	–		1	–		1	–	196.8
2	7.15	135.6	2	7.24	136.8	**2***	5.97	128.2
3	6.13	117.9	**3***	6.19	120.4	3	–	160.8
4	–		4	–		4	–	79.3
5*	2.50	23.9	**5***	2.52	23.9	5	–	42.5
6*	2.50	21.7	**6***	2.52	21.7	6a	2.51	49.8
7	–		7	–		6b	2.34	
8	–		8	–		7	6.47	130.3
9	1.79		**9a***	5.18	115.8	8	6.84	145
10	1.79		**9b***	5.31		9	–	198

Table 5 continued

Linden						Eucalyptus		
Compound **C1**			Compound **C2**			Compound **D**		
4-(1-hydroxy-1-methylethyl) cyclohexa-1,3-diene-carboxylic acid			4-(1-methylethenyl) cyclohexa-1,3-diene-carboxylic acid			Dehydrovomifoliol		
Position	δ H(ppm)	δ C(ppm)	Position	δ H(ppm)	δ C(ppm)	Position	δ H(ppm)	δ C(ppm)
			10	1.97		10	2.31	28.4
						11	1.03	23.8
						12	1.11	24.7
						13	1.89	18.8

Chestnut			Orange					
Compound **E**			Compound **F1**			Compound **F2**		
γ-LACT-3-PKA			8-Hydroxylinalool			Caffeine		
Position	δ H (ppm)	δ C (ppm)	Position	δ H (ppm)	δ C (ppm)	Position	δ H (ppm)	δ C (ppm)
2	–	126.3	**1***	4.00	67.9	1	–	–
3	–		2	–	136.8	2	–	151.8
4	–	139.8	3	5.43	125.7	3	–	–
		118.4						
5	7.57		4	2.08	23.2	4	–	148.0
6	7.42	124.2	5	1.61	42.4	5	–	107.8
7	7.69	132.8	6	–	73.1	6	–	155.5
8*	8.44	126.3	7	5.93	145.8	7	–	–
		127.5						
9	–		8a	5.09	111.2	8	7.55	141.2
10	–	139.8	**8b***	5.24	111.2	9	–	–
2′	–	167.0	2′	1.68	13.3	10	3.42	27.7
2″	4.80	63.0	6′	1.30	28.3	**11***	3.61	29.8
3″a	1.26	29.3				**12***	4.01	33.2
3″b	2.63							
4″a	2.42	29.3						
4″b	2.45							
5″a	3.52	41.5						
5″b	3.66							

The protons that originate the signals extracted from the S-Plots are labelled in bold and marked with a star

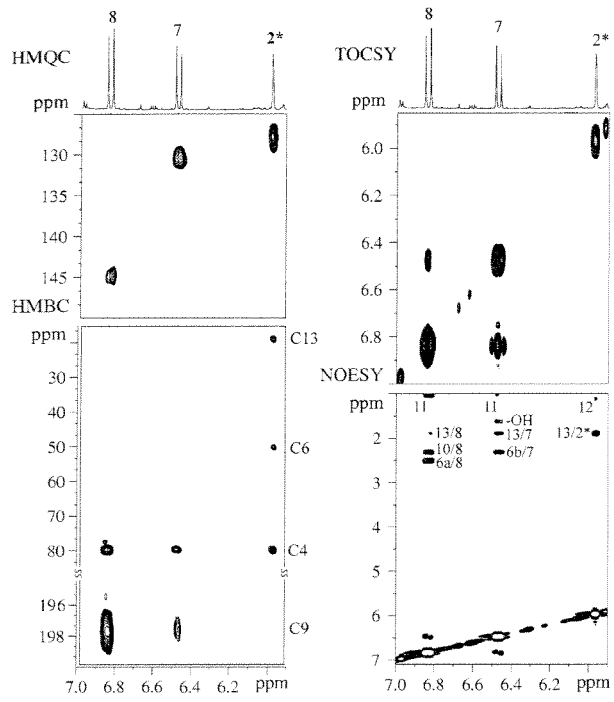


Fig. 4 Selected portions of 600 MHz two-dimensional NMR spectra and peak assignments of the eucalyptus marker (compound D in Fig. 3)

4 Concluding remarks

In summary, the present study demonstrates that a metabolomic approch based on ^{1}H-NMR and associated with O2PLS-DA is a very suitable tool to distinguish the botanical origin of honeys and to identify biomarker compounds. The proposed method requires very little sample preparation, is fast, reproducible, allows one to obtain a fingerprint of different classes of compounds at the same time, and is more objective than melissopalynological analysis.

Acknowledgments The financial support of the Veneto Region (UNIMIELE project, 2008) is gratefully acknowledged.

References

Alissandrakis, E., Tarantilis, P., Pappas, C., Harizanis, P. C., & Polossiou, M. (2011). Investigation of organic extractives from unifloral chestnut (*Castanea sativa L.*) and eucalyptus (*Eucalyptus Globulus Labill.*) honeys and flowers to identification of botanical marker compounds. *Food Science and Technology, 44*, 1042–1051.

Alvarez-Suarez, J. M., Tulipani, S., Díaz, D., Estevez, Y., Romandini, S., Giampieri, F., et al. (2010). Antioxidant and antimicrobial capacity of several monofloral Cuban honeys and their correlation with color, polyphenol content and other chemical compounds. *Food and Chemical Toxicology, 48*, 2490–2499.

Beretta, G., Artali, R., Caneva, E., Orlandini, S., Centini, M., & Maffei Facino, R. (2009a). Quinoline alkaloids in honey: Further analytical (HPLC-DAD-ESI-MS, multidimensional diffusion-ordered NMR spectroscopy), theoretical and chemometric studies. *Journal of Pharmaceutical and Biomedical Analysis, 50*, 432–439.

Beretta, G., Vistoli, G., Caneva, E., Anselmi, C., & Maffei Facino, R. (2009b). Structure elucidation and NMR assignments of two new pyrrolidinyl quinoline alkaloids from chestnut honey. *Magnetic Resonance in Chemistry, 47*, 456–459.

Bertelli, D., Plessi, M., Sabatini, A. G., Lolli, M., & Grillenzoni, F. (2007). Classification of Italian honeys by mid-infrared diffuse reflectance spectroscopy (DRIFTS). *Food Chemistry, 101*, 1565–1570.

Bertoncelj, J., Polak, T., Kropf, U., Korošec, M., & Golob, T. (2011). LC-DAD-ESI/MS analysis of flavonoids and abscisic acid with chemometric approach for the classification of Slovenian honey. *Food Chemistry, 127*, 296–302.

Bonaga, G., Giumanini, A. G., & Gliozzi, G. (1986). Chemical composition of chesnut honey: Analysis of the hydrocarbon fraction. *Journal of Agriculture and Food Chemistry, 34*, 319–326.

Castro-Vázquez, L., Díaz-Maroto, M. C., & Pérez-Coello, M. S. (2007). Aroma composition and new chemical markers of Spanish citrus honeys. *Food Chemistry, 103*, 601–606.

Codex Alimentarius Commission. (2001). Revised codex standard for honey. Codex Standard 12-1981. Rev.1, 1987. Rev.2, 2001.

EL-Manzalawy, Y., Honavar, V. (2005). WLSVM: Integrating LibSVM into Weka environment. http://www.cs.iastate.edu/~yasser/wlsvm.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikström, C., & Wold, S. (2006). *Multi- and megavariate data analysis. Basic principles and applications. Appendix II.* Umeå: Umetrics AB.

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.

Etzold, E., & Lichtenberg-Kraag, B. (2008). Determination of the botanical origin of honey by Fourier-transformed infrared spectroscopy: An approach for routine analysis. *European Food Research and Technology, 227*, 586–597.

European Commission. (2002). Council directive 2001/110/EC of 20 December 2001 relating to honey. *Official Journal of the European Communities, L10*, 47–52.

Frérot, E., Velluz, A., Decorzant, E., & Naef, R. (2006). From linden flower to linden honey. Glycosidic precursors of cyclohexa-1,3-diene-1-carboxylic acids. *Chemistry & Biodiversity, 3*, 94–100.

Graddon, A. D., Morrison, J. D., & Smith, J. F. (1979). Volatile constituents of some unifloral Australian honeys. *Journal of Agriculture and Food Chemistry, 27*, 832–837.

Guyot, C., Scheirman, V., & Collin, S. (1999). Floral origin markers of heather honeys: *Calluna vulgaris* and *Erica arborea*. *Food Chemistry, 64*, 3–11.

Hennessy, S., Downey, G., & O'Donnell, C. P. (2010). Attempted confirmation of the provenance of Corsican PDO honey using FT-IR spectroscopy and multivariate data analysis. *Journal of Agricultural and Food Chemistry, 58*, 9401–9406.

Isidorov, V. A., Czyżewska, U., Jankowska, E., & Bakier, S. (2011). Determination of royal jelly acids in honey. *Food Chemistry, 124*, 387–391.

Ivanciuc, O. (2007). Applications of support vector machines in chemistry. *Reviews in Computational Chemistry, 23*, 291–400.

Kaškonienė, 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.

Liu, R., Gao, M., Yang, Z., & Du, G. (2008). Pinocembrin protects rat brain against oxidation and apoptosis induced by

ischemia–reperfusion both in vivo and in vitro. *Brain Research, 1216*, 104–115.

Marghitas, L. A., Dezmirean, D. S., Pocol, C. B., Ilea, M., Bobis, O., & Gergen, I. (2010). The development of a biochemical profile of acacia honey by identifying biochemical determinants of its quality. *Notulae Botanicae Horti Agrobotanici Cluj, 38*, 84–90.

Melliou, E., Chinou, I. (2011, in press). Chemical constituents of selected unifloral Greek bee-honeys with antimicrobial activity. *Food Chemistry*, doi:10.1016/j.foodchem.2011.04.047.

Modi, A. A., Feld, J. J., Park, Y., Kleiner, D. E., Everhart, J. E., Liang, T. J., et al. (2010). Increased caffeine consumption is associated with reduced hepatic fibrosis. *Hepatology, 51*, 201–209.

Olsson, I.-M., Gottfries, J., & Wold, S. (2004). D-optimal onion designs in statistical molecular design. *Chemometrics Intelligent Laboratory Systems, 73*, 37–46.

Paris, D., Melck, D., Stocchero, M., D'Apolito, O., Calemma, R., Castello, G., et al. (2010). Monitoring liver alterations during hepatic tumorigenesis by NMR profiling and pattern recognition. *Metabolomics, 6*, 405–416.

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

Pichichero, E., Cicconi, R., Mattei, M., Gallinella Muzi, M., & Canini, A. (2010). Acacia honey and chrysin reduce proliferation of melanoma cells through alterations in cell cycle progression. *International Journal of Oncology, 37*, 973–981.

Plutowska, B., Chmiel, T., Dymerski, T., & Wardencki, W. (2011). A headspace solid-phase microextraction method development and its application in the determination of volatiles in honeys by gas chromatography. *Food Chemistry, 126*, 1288–1298.

Rastrelli, F., Schievano, E., Bagno, A., & Mammi, S. (2009). NMR quantification of trace components in complex matrices by band-selective excitation with adiabatic pulses. *Magnetic Resonance in Chemistry, 47*, 868–872.

Ruoff, K., Luginbühl, W., Künzli, R., Bogdanov, S., Bosset, J. O., Von der Ohe, K., et al. (2006). Authentication of the botanical and geographical origin of honey by front-face fluorescence spectroscopy. *J. Agr. Food Chem., 54*, 6858–6866.

Schievano, E., Peggion, E., & Mammi, S. (2010). ^{1}H nuclear magnetic resonance spectra of chloroform extracts of honey for chemometric determination of its botanical origin. *Journal of Agricultural and Food Chemistry, 58*, 57–65.

Socha, R., Juszczak, L., Pietrzyk, S., Gałkowska, D., Fortuna, T., & Witczak, T. (2011). Phenolic profile and antioxidant properties of Polish honeys. *International Journal of Food Science & Technology, 46*, 528–534.

Stephens, J. M., Schlothauer, R. C., Morris, B. D., Yang, D., Fearnley, L., Greenwood, D. R., et al. (2010). Phenolic compounds and methylglyoxal in some New Zealand manuka and kanuka honeys. *Food Chemistry, 120*, 78–86.

Tan, S.-T., Holland, P. T., Wilkins, A. L., & Molan, P. C. (1988). Extractives from New Zeland honeys. 1. White clover, Manuka and Kanuka unifloral honeys. *Journal of Agriculture and Food Chemistry, 36*, 453–460.

Truchado, P., Ferreres, F., & Tomas-Barberan, F. A. (2009a). Liquid chromatography–tandem mass spectrometry reveals the widespread occurrence of flavonoid glycosides in honey, and their potential as floral origin markers. *Journal of Chromatography A, 1216*, 7241–7248.

Truchado, P., Martos, I., Bortolotti, L., Sabatini, A. G., Ferreres, F., & Tomas-Barberan, F. A. (2009b). Use of quinoline alkaloids as markers of the floral origin of chestnut honey. *Journal of Agriculture and Food Chemistry, 57*, 5680–5686.

Trygg, J., & Wold, S. (2003). O2-PLS, a two-block (X–Y) latent variable regression (LVR) method with an integral OSC filter. *Journal of Chemometrics, 17*, 53–64.

Tulloch, A. P., & Hoffman, L. L. (1972). Canadian beeswax: Analytical values and composition of hydrocarbons, free acids and long chain esters. *The Journal of the American Oil Chemists' Society, 49*, 696–699.

Wiklund, S., Johansson, E., Sjöström, L., Mellerowicz, E. J., Edlund, U., Shockcor, J. P., et al. (2008). Visualization of GC/TOF-MS-based metabolomics data for identification of biochemically interesting compounds using OPLS class models. *Analytical Chemistry, 80*, 115–122.

Wishart, D. S. (2008). Metabolomics: Applications to food science and nutrition research. *Trends in Food Science & Technology, 19*, 482–493.

ABSTRACT

FEASR

REGIONE DEL VENETO

Fondo europeo agricolo per lo sviluppo rurale: l'Europa investe nelle zone rurali

Reg. (CE) n. 1698/2005 - PSR 2007 - 2013
DGR n. 199 del 12/02/08
MISURA 124 – "Cooperazione per lo sviluppo di nuovi prodotti, processi e tecnologie nel settore agricolo, alimentare e forestale"

Progetto UNIMIELE – Relazione finale

Riassunto

Nel corso di questo progetto è stato messo a punto un metodo oggettivo e veloce per la determinazione dell'origine botanica di mieli uniflorali.

Abbiamo messo a punto una procedura di estrazione delle componenti minoritarie poco polari utilizzando una miscela acqua/cloroformio. Lo spettro di Risonanza Magnetica Nucleare (NMR) di questo estratto fornisce un'impronta digitale, caratteristica per ciascun tipo di miele indagato.

Sono stati studiati circa 500 campioni di miele di 7 origini floreali diverse e provenienti da diverse regioni italiane, costruendo un database utile anche per studi futuri. Gli spettri NMR acquisiti sono stati analizzati utilizzando tecniche di analisi statistica multivariata. Sfruttando una parte dei campioni sono stati costruiti dei modelli di classificazione per le varie origini botaniche, che sono stati validati utilizzando i campioni rimanenti. Questi modelli sono in grado di predire efficacemente l'origine botanica di un miele incognito, soprattutto se usati in diversi gradi di complessità (confronto con tutte le classi, oppure a coppie di classi).

Il metodo da noi messo a punto è sicuramente più veloce rispetto all'analisi melissopalinologica. Rispetto all'analisi sensoriale, è più oggettivo e trasportabile e non richiede esperienza specifica nel campo. Il metodo sviluppato può essere affiancato alle analisi correntemente impiegate e fornire un valido supporto nei casi dubbi o controversi.

Per ciascuna specie botanica è stato possibile riconoscere dei metaboliti nel miele che caratterizzano il nettare di origine. Molte di queste molecole sono dotate di specifiche proprietà farmaceutiche e questo è un fattore che può essere validamente utilizzato per la valorizzazione del prodotto.

Il progetto è stato reso possibile grazie al finanziamento della Regione Veneto, DGR n. 199 del 12/02/08, MISURA 124 – "Cooperazione per lo sviluppo di nuovi prodotti, processi e tecnologie nel settore agricolo, alimentare e forestale", nell'ambito del Reg. (CE) n. 1698/2005 - PSR 2007–2013.

FEASR

REGIONE DEL VENETO

Fondo europeo agricolo per lo sviluppo rurale: l'Europa investe nelle zone rurali

Reg. (CE) n. 1698/2005 - PSR 2007 - 2013
DGR n. 199 del 12/02/08
MISURA 124 – "Cooperazione per lo sviluppo di nuovi prodotti, processi e tecnologie nel settore agricolo, alimentare e forestale"

Progetto UNIMIELE – Relazione finale

Abstract

During this project, an objective and rapid method to determine the botanic origin of unifloral honeys was developed.

We optimized a procedure to extract minor apolar components using a water/chloroform mixture. The Nuclear Magnetic Resonance (NMR) spectrum of this extract is a fingerprint of the type of honey under study.

Around 500 honey samples of seven different botanic origins were studied. The samples were produced in different Italian regions. A database was built, which can be used also for future studies. The NMR spectra were analyzed using multivariate statistical analysis techniques. A portion of the samples was used to build a classification model for each of the botanic origins. The remaining samples were used to validate these models. The models we produced can effectively predict the botanic origin of an unknown sample, especially if used in various degrees of complexity (comparison with all the possible classes, or pairwise).

The method we developed is certainly faster than the melissopalinological analysis. Relative to the sensorial analysis, it is more objective and transportable, and does not require competence in the specific field. This method can flank the analyses currently used and provide valid support in doubtful or controversial cases.

For each botanic species, specific metabolites were identified, which characterize the original nectar. Several of these molecules are endowed with specific pharmaceutical properties. This feature can be fruitfully exploited to improve the value of the product.

The present project was made possible through a financial contribution from the Regione Veneto, DGR n. 199 del 12/02/08, MISURA 124 – "Cooperazione per lo sviluppo di nuovi prodotti, processi e tecnologie nel settore agricolo, alimentare e forestale", within the Reg. (CE) n. 1698/2005 - PSR 2007–2013.