CHAPTER 4

# Nuclear Magnetic Resonance and Chemometrics to Assess Geographical Origin and Quality of Traditional Food Products

**R. Consonni** and **L. R. Cagliani**

Contents


I. Introduction 88
A. Food characterization 89
B. NMR techniques 90
C. Chemometrics 92
II. Geographical Origin of Foods 96
A. Wine 98
B. Olive oil 106
C. Cheese 112
D. Cereals 116
E. Honey 119
F. Fish 121
G. Meat 122
H. Other foods 125
III. Quality and Authenticity of Foods 131
A. Wine and beer 131
B. Vegetables 138
C. Balsamic and Traditional Balsamic Vinegar of Modena 141
D. Fruit juice 143
E. Olive oil 145
F. Fish and meat 149
G. Dairy products 150
H. Honey 153


Institute for the Study of Macromolecules, ISMAC, National Council of Research, Lab. NMR, Milan, Italy

Advances in Food and Nutrition Research, Volume 59
ISSN 1043-4526, DOI: 10.1016/S1043-4526(10)59004-1

I. Coffee 154
J. Vanilla 154
IV. Conclusions 156
References 157

**Abstract** In this globalization era, the opening of the markets has put at almost everybody's disposal a wide variety of foods, allowing everybody to taste food flavors and aromas from different nations. Notwithstanding this opportunity, countries try to preserve their markets by developing protection policies. A few countries have adopted different denominations to label their "typical food" products in order to give them additional value. Besides, the term "typical food" is widely thought of as something anchored to the local traditions, with geographical meaning and made with typical raw materials. Then a "typical food" starts to be considered "traditional" when it is made following specific and old recipes. As a matter of fact, these products acquire particular organoleptic characteristics that are not reproducible when produced in different places. In this review, NMR studies coupled to multivariate statistical analysis are presented with the aim of determining geographical origin and key quality characteristics.

## I. INTRODUCTION

In the field of food consumption, particular attention from both EU and national government organizations is dedicated to the so-called ''typical'' products, whose principal characteristics are the documentation of historical methods, geographical location, and quality of raw materials. ''Traditional'' foods, like PDO (Protected Designation of Origin) and PGI (Protected Geographical Indication), as well as DOCG (Demoninazione di Origine Controllata e Garantita) and IGT (Indicazione Geografica Tipica) for wines with Italian origins, are included in these product types. With the term ''traditional,'' people often think about a typical food prepared with very old recipes, handed on from father to son, usually manufactured in a specific geographical area and with particular procedures of preservation and ripening. In Italy, the production of these foods is regulated by a law (D.M. 18 July 2000) and it involves a traditional making procedure, older than 25 years and homogeneous for the interested area. In other words, they are part of the culture, and they carry inside a piece of history. These products (more than 4000 in Italy) are sensorially and nutritionally enriched, thus being unique products; nowadays some of these foods are under protection because they must be anchored to their original making procedures and locations. For these

reasons, different denominations, like PDO, PGI, and TSG (Traditional Specialty Guaranteed), have been created for identifying and assuring their specific properties.

In this review, nuclear magnetic resonance (NMR) techniques and chemometrics applied to several foods are summarized according to both quality and geographical origin determination. The pivotal importance of NMR techniques in modern science is mirrored in three Nobel Prize awards since 1990. One of the main challenges in dealing with NMR is to obtain a very good spectrum in terms of quality and resolution. A second challenge is concerning the ''as much as possible'' resonance assignment procedure, which is the most time-consuming part of the data analysis, even though some general databases of spectra are available online. Anyway, the most widely used approach to confirming the proposed assignment is still the addition of standard compounds. The food matrix is in general a complex mixture to analyze, by NMR, from the chemical point of view. Several classes of compounds are present, in different amount and sometimes their signals are overlapped to other strong uninteresting ones, most likely residual protons of deuterated organic solvent or impurities. Nevertheless, the chemical composition analysis of complex matrices is one of the possible investigations that NMR could afford. The concentration and determination of different chemical species can be quantified in a single experiment and with the aid of chemometrics, the extraction of discriminating variables for samples differentiation could be achieved.

## A. Food characterization

Food characterization as well as adulteration detection is nowadays one of the most attractive challenges from both the technological and computational points of view. During the last decades, food characterization was obtained by using ''classical techniques'' or ''physical analysis'' making use of colorimetric, potentiometric, ash, and enzymatic determinations. Starting from the 1980s, chromatographic techniques opened new frontiers to food characterization especially involving LC (liquid chromatography) and HPLC (high-performance liquid chromatography) methods. MS (mass spectrometry) played a predominant role in the identification of chemical compounds. The introduction of combined techniques, such as HPLC–MS, lowered the detection limits; the strongest instrument in terms of sensitivity acts as a detector coupled to the strongest mixture analyzer, constituting the primer for the development of the so-called ''advanced'' techniques. This combined approach was named by Hirschfeld ''hyphenated'' in 1980 according to a chapter on ''Hyphenated Techniques'' in *Analytical Chemistry* (Wiley-VCH, 1998): ''... In general, the term *hyphenated* ... the marriage ... of two separated analytical techniques via

appropriate interfaces, usually with the backup of a computer tying everything together.'' This relatively new definition indicates an analytical online system coupling separation with detection methods. It has several advantages compared with the analog offline technique: it results in much better repeatability of analytical results, contamination is minimized and it can automatically be run. However, the best systems are those where the experimental conditions of the separation and detection technique fit one to the other.

In recent years, even though most of the classic analytical methods are still used because they are part of certified methodologies, new analytical techniques are emerging and progressing more and more. The majority of these constitute the so-called ''omics techniques,'' as they concern high-throughput identification and quantification of small molecules. As a matter of fact, this approach was revealed to be the most informative with the capability of identifying food constituents, using metabolite profiling allowing determination of the composition, adulteration, and quality of foods simultaneously. Cordella *et al.* (2002) reviewed the applications of eight selected techniques in food characterization and adulteration detection: the final judgment appeared to be the arrangement of several techniques to collect as much data as possible to obtain the better characterization of the food under investigation. In a recent review, Wishart (2008) presented different metabolomics technologies with their advantages and drawbacks in food science and nutrition research. Several efforts were also made in nutrition studies as well as in food consumption monitoring. From these studies, it appeared that metabolomics is very promising in food science and nutrition research, with two main limitations: technologies and databases turned into high costs and time-consuming procedures. Further improvements are obtained with the use of chemometric methods; statistical analysis of large data sets allows comparisons and the evaluation of differences or similarities present within samples.

## B. NMR techniques

NMR is a unique and versatile spectroscopic method for measuring samples in all phase states, providing detailed molecular information of the system under investigation. In complex matrices, like foods, several chemical compounds could be detected in a quantitative way with a single experiment. Only in the last few years, NMR has emerged as a practical alternative solution to quality assurance challenges, even though the effective cost of the NMR system slowed the diffusion of this technology.

NMR spectroscopy is essentially based on energy absorption by atomic nuclei with active spin (nonzero value of quantum number spin I) in the presence of a static magnetic field. This energy absorption is

obtained with the application of an appropriate radiofrequency pulse; after this pulse, the excited nuclei will relax and their emitted energy is detected as a time-dependent signal intensity. Fourier transformation of this signal will produce a frequency-dependent signal intensity proportional to the number of excited nuclei, which constitute the NMR signal. Each nucleus shows a specific position in the frequency spectrum, called chemical shift: this value is characteristic of the chemical environment surrounding the observed nuclei. The measurable parameters of the resulting spectral lines (line positions, intensities, line width, multiplicity, and transients in time-dependent experiments) can be interpreted in terms of molecular structure, conformation, molecular motions, exchanges, and other rate processes. Different NMR techniques and instruments give different information, always treating the sample in a nondestructive way. High resolution (HR), low resolution or low field (LF), and imaging (NMRI) are the possible NMR techniques. HR allows qualitative and quantitative analysis of samples as well as molecular structure determinations in solution. Samples in solid state and in semisolid state were analyzed by using magic angle spinning condition (HR-MAS spectroscopy). This technique, introduced by Andrew *et al.* (1958) and Lowe (1959) rapidly became widely accepted in food characterization (Bertocchi and Paci, 2008). It takes advantage of measuring chemical shifts and multiplicity of signals and with neglected preparation for solid samples. LF gives information about the relaxation time, strictly correlated with intra- and intermolecular motions, diffusion processes, and structural properties of liquids in porous systems or amorphous phase and systems consisting of different phases. It lacks chemical shift information and the acquired $^{1}H$ signal consists of a single absorption line containing information arising from all protons present in the specimen. Very interesting and recent applications are focused on the measurement of bound and free water in foods, whereas in the particular case of aqueous solution, the relaxation characteristics of the single-water absorption signal bring about a wealth of information about the solute molecules that could be highly relevant in the characterization of a specific sample. Moreover, the possibility of recording the $^{1}H$ relaxation rate over an extended range of magnetic fields allows the identification of different contributions that various components exercise in the foodstuff (field cycling NMR).

MRI gives a multidimensional picture of the object under investigation on the basis of contrast enhanced by NMR parameters, usually relaxation time, spin density, or chemical shift. It is a useful technique to get information about the physiological conditions of many tissues and is widely used in agricultural applications, medical diagnosis, and recently with growing interest, in food quality characterization. Imaging can be obtained monitoring different nuclei; the rule ''the most abundant, the more sensitive'' is almost always adopted and in this respect usually $^{1}H$,

$^{19}F$, and $^{23}Na$ are commonly used. MRI was used for the first time in 1973 (Lauterbur, 1973) and nowadays is successfully applied to food characterization (Falcone *et al*., 2006).

A particular mention is needed for describing one analytical technique capable of detecting the exact site-specific isotope ratio. This technique, called SNIF-NMR (site specific natural isotope fractionating technique), constitutes the most specific and sophisticated method used for food authenticity determination.

All of the briefly described NMR techniques have been adopted usefully in the characterization of food. The role of NMR in food chemistry has been growing progressively during the last years even though, as far as we know, NMR has never been adopted as an official analytical methodology. The only exception is the quality determination of olive oils for the Lazio region, obtained with a regional law by Segre in 2001. The main problem affecting all other techniques, even though more sophisticated, is that they are focused on compounds of only a specific chemical class. A choice has to be made on the basis of the expected parameters to be evaluated or the type of fraud to be detected. A solution to this problem is the use of a broad analytical screening performed with the detection of different compounds in a single experiment, as NMR can do.

## C. Chemometrics

In recent years, the increased specificity and sensitivity of the analytical instruments offered the feasibility of obtaining a wide range of information in one shot. This technological breakthrough became more and more attractive and thus a ''normal'' approach to studying foods, in terms of either quality or authenticity assessment. This approach, largely accepted as ''metabolic profiling'' or ''metabonomics,'' is also improperly denoted ''metabolomics,'' while in fact this term properly refers to the collection of small molecules that can be found in a cell, organ, or organism. In the past, the metabolomics approach created two different schools of thought: (a) the chemometric approach, in which the chemical compounds were not identified, but their spectral patterns were statistically analyzed to identify relevant spectral features that could differentiate samples; and (b) the targeted or comprehensive profiling, where the focus was to identify and quantify the chemical compounds as much as possible and then perform statistics to identify the most relevant biomarkers. Nowadays, the term chemometrics is largely used without any specific intent.

### 1. Monovariate statistical analysis

In chemistry, masses of data are obtained by measuring many variables on an ensemble of chemical samples or by recording many signals from an industrial process in order to track its behavior. A data collection task,

whether in science, business, or engineering, typically involves many measurements made on several samples. Such data variability has traditionally been analyzed by using one or two variables at a time. However, to discover the relationships among all samples and variables efficiently, all data must be processed simultaneously. Chemometrics is intended to extract information in multivariate chemical data, using the tools of statistics and mathematics. It is typically used for three primary purposes: to explore patterns of association in data, to track the properties of materials on a continuous basis, and to prepare and use multivariate classification models. In general, the algorithms applied have demonstrated significant capacity in analyzing and modeling a wide variety of data types for an even more diverse set of applications. In general, different mathematical methods can be used to explore experimental data, based on the different possible targets. The first general phase is the data exploration, which gives information about statistical parameters of each variable, correlation among variables, and so on. In particular, the first aim is the data dimensionality reduction. Among the possible systems, analysis of variance (Miller and Miller, 1993), ANOVA, is used to select the variables most significant in sample differentiation. It is a univariate statistical technique for testing the null hypothesis that two or more samples are drawn from the same population; high values of the *F*-test suggest that the null hypothesis can be discarded. This technique is no longer used for large data sets (especially in the case of spectroscopic data). The extension of ANOVA is called ''multivariate data analysis'' (MANOVA), and it is used whenever more than one correlated variable is concerned and they cannot be simply combined. MANOVA selects discriminant variables with high indices of reliability.

## 2. Multivariate statistical analysis

Unlike monovariate methods, where only one variable is considered, in multivariate statistical analysis, correlations among more variables are concerned. This approach is largely used nowadays especially because spectroscopic data are often adopted to explore the quality of foods and metabolic content. Multivariate data analysis is usually applied for addressing these aspects: (a) data overview, (b) classification and or discrimination among groups of observations, and (c) regression modeling between two blocks of data (X and Y). These applications reflect the main stages of multivariate analysis. One of the main aims of this technique is to reduce the dimensionality. Among the so-called ''compression techniques,'' principal component analysis, PCA (Geladi and Kowalski, 1986; Jackson, 1991) and the correlated methods is widely used and recognized as one of the main ''unsupervised'' compression technique for primarily analysis of data. This method finds linear combinations of the variables in the original data, called PCs, which are orthogonally

related and describe the major trends in the data. When the minimum meaningful number of PCs has been found, by means of *loadings* and *score*s matrices, the original data matrix can be rebuilt. Inspection of the loadings gives indications on how the PCs are obtained from the original variables and how much the variable has in common with that PC. Scores show how the observations are clustered together on the basis of their variables.

Another compression technique, being part of the so-called ''classification methods,'' is the cluster analysis (CA; Romesburg, 1984) applied to evaluate similarities and clusters among samples. This approach based on ''similarities'' or ''classification'' methods could also be split into *hierarchical* or *nonhierarchical* approaches. Commonly, two types of clustering are used: K-Mean and Tree Clustering, named TCA. These classification methods are without ''*a priori*'' hypothesis in finding meaningful groups, and the result is often used for further statistical analysis. Dendrograms are usually adopted as a graphical representation tool to visualize the data clustering.

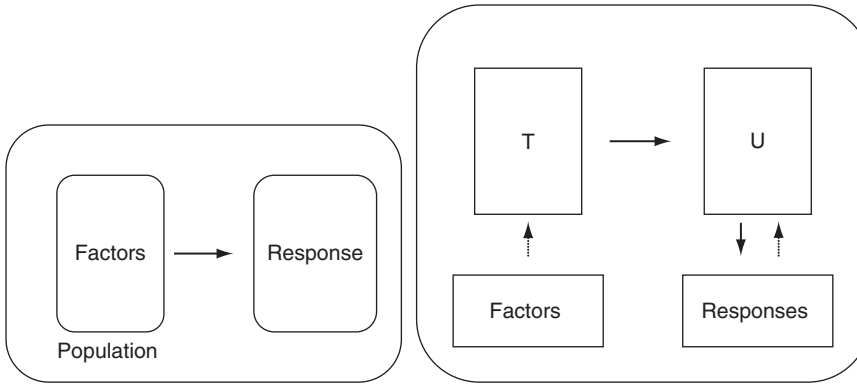
Discriminant analysis (DA) performs samples classification with an ''*a priori*'' hypothesis. This hypothesis is based on a previously determined TCA or other CA protocols. DA is also called ''discriminant function analysis'' and its natural extension is called MDA (multiple discriminant analysis), which sometimes is named ''discriminant factor analysis'' or CDA (canonical discriminant analysis). Among these type of analyses, linear discriminant analysis (LDA) has been largely used to enforce differences among samples classes. Another classification method is known as QDA (quadratic discriminant analysis) (Frank and Friedman, 1989) an extension of LDA and RDA (regularized discriminant analysis), which works better with various class distribution and in the case of high-dimensional data, being a compromise between LDA and QDA (Friedman, 1989).

More recently, ICA (independent component analysis) has been developed for the analysis of signals from complex mixtures (Comon, 1994). In this approach, the coefficients of the linear expansion of the data vectors must be mutually independent; this requires higher order statistics in determining the ICA expansion and also some nonlinearities must be used in the learning phase, thus resulting in a more meaningful data representation with respect to PCA. Generalized discriminant analysis (GDA; McLachland, 1992) is used to determine whether a given classification of samples into a group is appropriate. Therefore, each sample is assigned to a group and a model is searched and computed to maximize the classification. A general aim is to find out a mathematical model with high predictive capacity for a variable obtained from known values derived from the ensemble of independent variables; these types of protocols are called ''regression methods.'' The simplest model describes the $Y$ variable linearly dependent on the $X$ variable; this casual dependence is

a linear regression. Science often involves controllable variables (factor or predictor variables) to explain, to regulate, or to predict the behavior of other variables (response variables). When factors are few, not significantly redundant (collinear), and show a correct relationship to the responses, the multiple regression can be the right way to turn data into information. When spectroscopic data are concerned, factors (variables) can be hundreds and highly collinear; the responses are components that need to be predicted for future samples. In these cases, Partial Least Squares Projections to Latent Structures (Wold *et al.*, 1984), PLS, is used to create multivariate calibration models with predictive capacity. In principle, multiple linear regression can be used with a large number of factors. However, if this number is bigger than the number of observations, the model will fail to predict a new data set because of the overfitting problem. In such cases, there could be only a few underlying or latent factors that account for most of the variation in the response. The origin of PLS acronym can be explained by considering the general idea of PLS, which is to extract these latent factors accounting for the largest manifest factor variation possible, while optimally modeling the response. In Scheme 4.1, the general aim of PLS is summarized: factors are used to predict responses in the population. This is achieved indirectly by extracting the latent factors *T* and *U* from factors and responses respectively.

The extracted factors *T* (X scores) are used to predict the Y scores *U* and then to build predictions for the responses. In PLS, the X and Y scores are chosen so that the relationship between successive scores is as strong as possible.

Currently, several linear and nonlinear multivariate classification methods exist: the choice implies the evaluation of discriminatory power against the ability to interpret the meaning of class differences. In this respect, Soft Independent Modeling of Class Analogy (SIMCA;




**SCHEME 4.1** Graphical representation of PLS.

Wold, 1976) is a well-established method for multivariate classification; disjoint PCA is used for fitting each class and it is largely used even though it does not give easily accessible class differences information, thus hampering the quality of interpretation of the classification model. PLS discriminant analysis (PLS-DA) has largely been used for explaining differences among overall class properties that become progressively more complicated with an increasing number of classes. The relatively new orthogonal PLS-DA (OPLS-DA; Bylesjö *et al.*, 2006) approach has been demonstrated to be the most revealing of the generated models. OPLS-DA is obtained as an extension of the PLS method featuring an orthogonal signal correction (OSC) filter (Trygg and Wold, 2002). In other words, compared to PLS-DA, OPLS-DA effectively separates predictive from nonpredictive (orthogonal) loadings variation and this is particularly enforced when a two-class model is concerned.

### 3. Artificial neural networks

In recent years, a progressive increase in the number of articles and reviews has appeared in the literature concerning the use of chemometrics applied with different techniques in food analysis and quality determination and very recently, a review on artificial neural networks was also presented (Marini, 2009). The latter technique, based on mathematical methods, was originally born with the aim of mimicking human brain functionality. The introduction of two key concepts (Hopfield, 1982) caused a big expansion of interest and application across a large range of problems, up to the recent years when it has been applied in both exploratory and regression data analysis, especially when nonlinear trends in data are present. For any additional details, we suggest consulting the review by Marini (2009).

## II. GEOGRAPHICAL ORIGIN OF FOODS

In recent years, consumers all around the world have shown an increasing interest for both high-quality food products and reliable indicators of geographical origin. The reasons for this can be traced to patriotism, health benefits or specific organoleptic and culinary qualities associated with regional products, media attention, decreasing confidence in the quality and safety of products coming from outside the local region, country, or EU; concern about animal welfare; and an emphasis on environmentally friendly production methods (Luykx and van Ruth, 2008).

The EU plays a major role in enhancing high-quality attributes and in sustaining this wide range of cultures and culinary traditions. In 1992, EU regulation 2081/92 and 2082/92 introduced systems to promote and to protect geographical indications and designations of origin for

agricultural products and foodstuffs. Successively, these rules were replaced by 509/2006 and 510/2006 regulations respectively, and finally 1898/2006 EU regulation was added. The main aims of these regulations were the promotion of products with specific characteristics and particularly those coming from less-favored or rural areas, the improvement of the income of farmers in returning to a ''genuine effort to improve quality,'' the retention of population in rural areas and the provision of clear and succinct information to consumers regarding the product origin. Differentiation among food products can be attributed to the unique local features of the product, its history, or its distinctive character linked to natural or human factors such as soil, climate, local know-how, and traditions.

These regulations allowed the application of different geographical indications to food products and precisely PDO, PGI, and TSG.

PDO status could be used only for agricultural products or foodstuffs which are produced, processed, and prepared in a given specific geographical area, by using recognized know-how and whose quality or properties are significantly or exclusively determined by the geographical environment, including natural and human factors. Examples of PDO products are Saffron from Aquila and S. Gimignano and Traditional Balsamic Vinegar of Modena (TBVM) and Reggio Emilia (Italy), Royal potatoes (United Kingdom), Roquefort (cheese from France), and Kalamata (olive oil from Greece). PGI designation covers agricultural products and foodstuffs closely linked to a specific geographical area for quality, goodwill, or other characteristic properties and whose production, processing, or preparation takes place within the determined geographical area. Typical PGI food products are Calcot de Valls (vegetables from Spain), Scotch beef (Scotland), and Nurnberger rostbratwurste (sausage from Germany). Finally, TSG is a trademark applied to agricultural products or foodstuffs presenting features that distinguish them from other similar products belonging to the same category. These products must be obtained by using traditional raw materials or must present traditional compositions or must be manufactured by using production and/or processing reflecting traditional methods. The main aim of the TSG designation is to allow high-quality products that are not necessary linked to a geographic area to be differentiated from other products in order to obtain an economical advantage. TSG products are, for example, cow milk Mozzarella (cheese from Italy), Karjalanpiirakka (Finnish biscuits), Lambic, Gueze Lambic or Geuze (Belgian beers), and Jamón Serrano (Spanish ham).

These products have quality levels that could (a) enhance food security, inasmuch as they contribute to rural development and support small producers for accessing markets, (b) complying with food safety

requirements, assuring added-value attached to the product's specifications, (c) contribute to social welfare and food diversity and biodiversity preservation. Indeed, thanks to the link between such products and their areas of origin, they can help to preserve local resources, maintain traditions, strengthen the organization of local stakeholders, and to prevent delocalization and the rural exodus.

Procedures or ingredients adopted not in accordance with the established rules are not allowed the use of these designations. These procedures will assure a high-quality level for these selected food products, thus protecting both producers and consumers from possible fraudulent products or improper processing/production methods. Notwithstanding these rules, appropriate controls and analysis are quite often not available to enforce these approaches. Just one example is the requirement for geographical origin of several foods, which consists of a paper documentation rather than an instrumental determination. In this respect, several researchers focused their studies on applying different analytical techniques to the determination of the geographical origin of foods, as well as authenticity and quality investigations.

We have found that most of the selected papers focused on geographical food characterization by NMR techniques combined with chemometrics were dealing with wine and olive oil. Other characterizations of origin involved honey, cheese, fish and meat, and cereals foods; we found only a single paper dealing with mustard oil, green tea, chamomile, cod liver oil, propolis, concentrated tomato paste, and cocoa. A summarized table, concerning methodologies and applications potentially useful for geographical fraud detection of foods, is reported in Table 4.1.

## A. Wine

Wine is a fermented alcoholic beverage mainly composed of water, ethanol, glycerol, sugars, organic acids, and inorganic ions. Due to the ever increasing attention to the ''naturalness'' of wine by both consumers and controlling authorities, a large number of researchers focused their attention on the development of fast and efficient methods for the geographic origin determination and authenticity of wines. From the chemical point of view, wine is a mixture of amino acids, polyphenols, and sugars; from the NMR point of view, to obtain a good-quality proton spectrum of wine is not so simple. Suppression of the strong intensity signals of ethanol (with its satellite signals) and water, as well as obtaining flat baseline spectra, is quite challenging. Recently, NMR spectroscopy was applied to investigate samples of Aglianico DOC red wine from Basilicata and Campania (two regions in the south of Italy) (Fig. 4.1). Ten selected resonances from the $^{1}$H NMR spectra were integrated, quantified, and analyzed by PCA: four of them were found to be important for

**TABLE 4.1** Summary of papers dealing with NMR and chemometric for geographical origin determination

Geographical origin determination

| Food type | Type/region | Statistical method | Methodology | References |
|---|---|---|---|---|
| Wine | France | ANOVA, PCA | SNIF, IRMS | Day *et al.* (1995) |
| | Bordeaux | ANOVA, PCA | SNIF | Martin *et al.* (1999) |
| | Slovenia | PCA, LDA | SNIF, IRMS | Ogrinc *et al.* (2001) |
| | Apulia | PCA, HCA, DA | $^{1}H$ | Brescia *et al.* (2002a) |
| | Slovenia | HCA | $^{1}H$ | Košir and Kidrič (2002) |
| | Carbenet, Merlot/Bordeax | PCA | $^{1}H$ | Pereira *et al.* (2005) |
| | Spirit beverages/different nations | CDA, CBT | $^{1}H$ | Petrakis *et al.* (2005) |
| | Chinese | PCA | $^{1}H$ | Du *et al.* (2007) |
| | Cabernet, Campbell, Shiraz/different nations | PCA, PLS-DA | $^{1}H$ | Son *et al.* (2008) |
| | Aglianico DOC | PCA | $^{1}H$ | Viggiani and Castiglione Morelli (2008) |
| | Muscat/Korea | PCA, PLS-DA | $^{1}H$ | Son *et al.* (2009a) |
| Olive oil | Italy, Israel | PCA, PCR, PLS | $^{13}C$ | Shaw *et al.* (1997) |
| | Campania, Lazio, Sicily, Umbria | PCA | $^{1}H$ | Sacchi *et al.* (1998) |
| | Apulia | PCA, HCA, DA | $^{1}H$ | Sacco *et al.* (2000) |
| | Liguria, Tuscany, Lazio, Sicily, Apulia, Garda | ANOVA, LDA, TCA | $^{1}H$ | Mannina *et al.* (2001b) |
| | Tuscany | CA, DA | $^{1}H$ | Mannina *et al.* (2001a) |
| | Italy | PCA | $^{13}C$ | Vlahov *et al.* (2001) |
| | Lazio | PCA, LDA | $^{1}H$, $^{13}C$ | D'Imperio *et al.* (2007) |

(*continued*)

**TABLE 4.1** (*continued*)

| Geographical origin determination | | | | |
|---|---|---|---|---|
| Food type | Type/region | Statistical method | Methodology | References |
| | Apulia | LDA | $^{13}C$ | Vlahov *et al.* (2003) |
| | Garda, Veneto | PCA | $^{1}H$ | Mannina *et al.* (2005) |
| | Europe | PCA, LDA, PLS-DA, PNN | $^{1}H$ | Rezzi *et al.* (2005) |
| | Garda | PCA | $^{1}H$ | Schievano *et al.* (2006) |
| | Greek | CDA, CBT | $^{1}H$, $^{31}P$ | Petrakis *et al.* (2008) |
| Cheese | Buffalo mozzarella | PCA, HCA, DA | $^{1}H$, ICP, IRMS, HPIC | Brescia *et al.* (2005) |
| | Emmenthal | CA | MAS | Shintu *et al.* (2006) |
| | Parmigiano Reggiano, Grana type | PCA, OPLS | $^{1}H$ | Consonni and Cagliani (2008b) |
| | Cow milk | PCA | HPIC, ICP, $^{1}H$, IRMS | Sacco *et al.* (2009) |
| Cereal | Wheat | PCA | MAS | Sacco *et al.* (1998) |
| | Wheat | PCA, DA | MAS, IRMS | Brescia *et al.* (2002b) |
| | Bread | ANOVA, PCA, DA | MAS, Image, IRMS | Brescia *et al.* (2007) |
| Honey | Italy, Europe, non-Europe | PCA, PLS-DA | $^{1}H$ | Consonni and Cagliani (2008a) |
| | Corsican, non-Corsican | PLS-DA, PLS-GP | $^{1}H$ | Donarski *et al.* (2008) |
| Fish | Gilthed Sea Bream | LDA, PNN | $^{1}H$ | Rezzi *et al.* (2007) |
| | Salmon | PNN, SVM | $^{13}C$ | Aursand *et al.* (2009) |

| | | | | |
|---|---|---|---|---|
| Meat | Apulia | PCA, DA | MAS | Sacco *et al.* (2005) |
| | America, Canada, Australia | ANOVA, PCA, DA | MAS | Shintu *et al.* (2007) |
| Mustard oil | Canada, India | PCA | SNIF, IRMS | Remaud *et al.* (1997) |
| Green Tea | China, Japan, Vietnam, India, Indonesia, Bangladesh | PCA, CA | $^{1}$H | Le Gall *et al.* (2004) |
| Camomille | Egypt, Hungary, Slovakia | PCA | $^{1}$H, PLE | Wang *et al.* (2004) |
| Cod liver oil | Scotland, Norway | PCA, PLS-GA | $^{13}$C, GC | Standal *et al.* (2008) |
| Propolis | Asian, Africa, European | PCA | $^{1}$H | Watson *et al.* (2006) |
| Concentrated tomato paste | Italy, China | PCA, OPLS-DA | $^{1}$H | Consonni *et al.* (2009) |
| Cocoa | Africa, South America | ANOVA | LF | Hernandez and Rutledge (1994) |

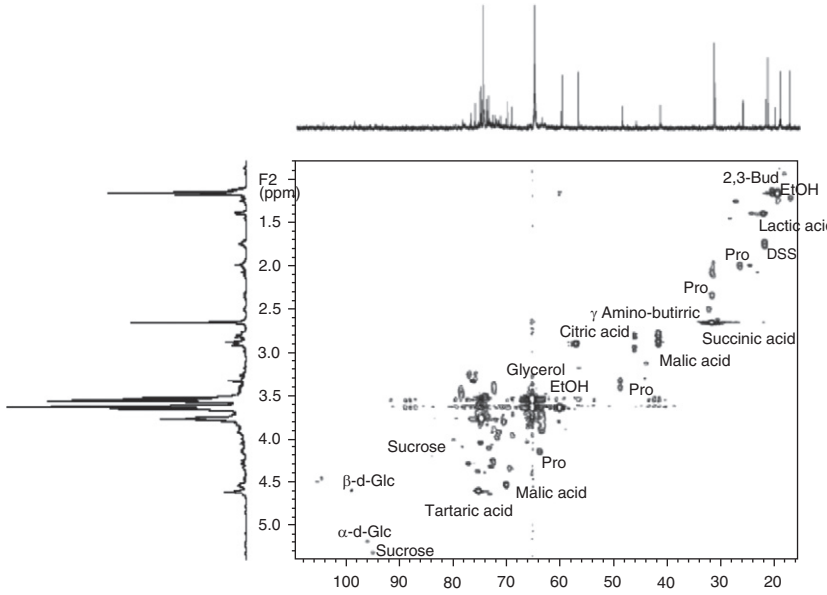



**FIGURE 4.1** Part of a 2D $^{1}H$–$^{13}C$ GHSQC spectrum of a wine from Venosa (in Basilicata region) acquired at 9.4 T (400 MHz). Wine was first lyophilized and then dissolved in $D_2O$. Abbreviations used are as follow: Glc, glucose; 2,3-Bud, 2,3-butanediol; and Pro, proline. (From Viggiani and Castiglione Morelli, 2008.)

sample differentiation, namely proline, 2,3-butanediol, succinic acid, and marginally glycerol. Only on the basis of the quantities of the first three metabolites (succinic acid, 2,3-butanediol, and proline), a good separation of the Aglianico DOC samples was obtained corresponding to both their geographical origin and vintages (Viggiani and Castiglione Morelli, 2008).

A combination of $^{1}H$ NMR and PCA was applied by Du *et al.* (2007) for the classification and the determination of the geographical origin of Chinese wines. On the basis of the total metabolite content, dry red, white, medium dry white, and blended wines were well separated by using PCA. Red wines were enriched in proline, lactate, and aromatic compounds compared to white ones, which had larger amounts of malate, citrate, and glycerol. The high-proline content in red wines seemed to be caused by the presence of grape skin during the fermentation. As expected, sweet wines were much more enriched in sugars and had significantly less hippuric acid compared to blended wines. A good separation corresponding to the production locations was also achieved by performing PCA on only dry red wines produced in three different regions from the north of China: Yantai and Changli (in the coastal region) and Shacheng (in the continental region). Organic acids, proline, glycerol,

ethyl esters, 2,3-butanediol, and 3-hydroxyl-2-butanone were to be the most important variables for sample separation which was believed to be mainly due to the fermentation process for different wines and environmental variations, like local climate, soil, underground water, sunlight, and rainfall.

A good separation among red wines from the north, center, and south of the Apulia (region of south Italy) was achieved by Brescia *et al.* (2002a) by applying PCA, DA, and HCA (hierarchical clustering analysis) protocols, by considering the combination of routine analysis (density, alcohol content, acidity, etc.) and $^{1}$H NMR data. In particular, statistical evaluation of the data indicated that the content of heavy metals and organic acids (Ba, Mn, Zn, Al, Fe, and citrate) characterized the wines coming from the center region, while Mg, Ca, K, Br, Cl, and tartrate characterized the wines coming from north and south of Apulia. Amino acids (isoleucine, valine, citrulline, and leucine) were also among the major substances responsible for geographical discrimination in the case of the analytical and NMR data sets, respectively. Slightly worse results were obtained in the case of wine origin prediction by using the NMR data, because some of the signals overlapped. The same research group compared the use of NMR spectroscopy, HPICE (high-performance ion chromatography exclusion) and ICP-OES (inductively coupled plasma optical emission spectroscopy) for the classification of Slovenia and Apulia wines. Also in this case, classical analytical and NMR data were considered separately for statistical analysis with PCA, HCA, and RDA in order to compare the discriminating potential of each methodological approach. Interestingly, both approaches led to a very good separation among wines from the two countries. In particular, NMR data allowed a better clustering of samples with PCA, while a better origin prediction was obtained for Apulian and Slovenian wines by using RDA. The content of heavy metals was most responsible for discrimination between the wines when analytical data were used alone, while the content of amino acids (isoleucine, proline, and citrulline) was the most discriminating when NMR data was evaluated (Brescia *et al.*, 2003b).

The variability in skin and pulp tissue composition for grape berries of Merlot Noir, Cabernet Franc, and Cabernet Sauvignon harvested at their mature stage in four appellations in the Bordeux area were investigated by $^{1}$H NMR in addition to conventional physicochemical analysis and PCA (Pereira *et al.*, 2005). A very good separation among wine samples coming from Bordeaux appellation, Saint-Emilion, Buzet, and Pessac-Léognan areas was achieved by considering physicochemical variables measured for 134 samples of grape berries: sugar content, mineral nitrogen, and total acidity resulted to be the most discriminant variables. $^{1}$H NMR data were less discriminant than the physicochemical ones, considering both grape berries and skin extracts, but allowed the

characterization of the metabolites involved in group separation: fructose, sucrose, and amino acids such as GABA, proline, and arginine.

In a recent article, metabolite differences in pulp, skin, seed of grapes, and wines coming from three different regions of South Korea (Yeoncheon, Yeongdong, and Chochiwon) were investigated by using $^{1}$H NMR with PCA and PLS-DA. PCA performed on NMR data for pulp and skin extracts led to a quite good separation of the Yeoncheon and Chochiwon samples corresponding to their geographical origin but better results were achieved when wines aged for 3 and 6 months were considered. PLS-DA loadings for Yeongcheon wines showed higher levels of lactate, proline, and glycerol by comparison to Chochiwon wines which were enriched in 2,3-butanediol, malate, tartarate, citrate, and succinate (Son *et al.*, 2009a).

By evaluating a few samples of white wine coming from three wine-growing regions in Slovenia, Košir and Kidrič (2002) found a good separation of wine varieties for Chardonnay, Welsch Riesling, Sauvignon, and one sample of Riesling, on the basis of the intensities of seven amino acids (arginine, citrulline, lysine, proline, isoleucine, alanine, and valine) signals in $^{1}$H NMR spectra and the Ward's hierarchical clustering method. Adding signals relative to glycerol, butylene glycol, and succinic acid, a differentiation between wine samples from Coastal and Continental regions (Drava and Sava regions) was achieved, even though it was not possible to distinguish wines coming from the two continental regions. Therefore, different pedoclimatic properties influenced much stronger glycerol, butylene glycol, and succinic acid content rather than the content of amino acids. Better results were achieved by Ogrinc *et al.* (2001) by investigating more than 100 wines from the same areas with SNIF-NMR and IRMS methods combined with PCA and LDA. By considering the deuterium/hydrogen isotopic ratio of the methylenic site in the ethanol molecule, $(D/H)_{II}$ and $\delta^{13}C$ values, a separation between wines coming from Coastal and Continental regions was achieved, but by adding $\delta^{18}O$ values, a differentiation between wines coming from the two continental regions, Drava and Sava, was also possible. The $\delta^{18}O$ values were modified by the meteorological events during grape ripening and harvesting.

The metabolite content of Cabernet Sauvignon samples coming from California, Australia, and France and Shiraz samples coming from Australia were evaluated by $^{1}$H NMR spectra and were well differentiated by performing PLS-DA (Fig. 4.2). Pairwise comparison of Australian, French, and Californian Cabernet Sauvignon wines indicated a higher content of succinic and tartaric acids in French wines by comparison with the Californian ones that had higher levels of proline. French Cabernet wines showed a higher content of glucose in comparison with Australian wines which had higher levels of proline, tartaric and gallic acids, and 2-phenylethanol. Finally, Californian wines had greater

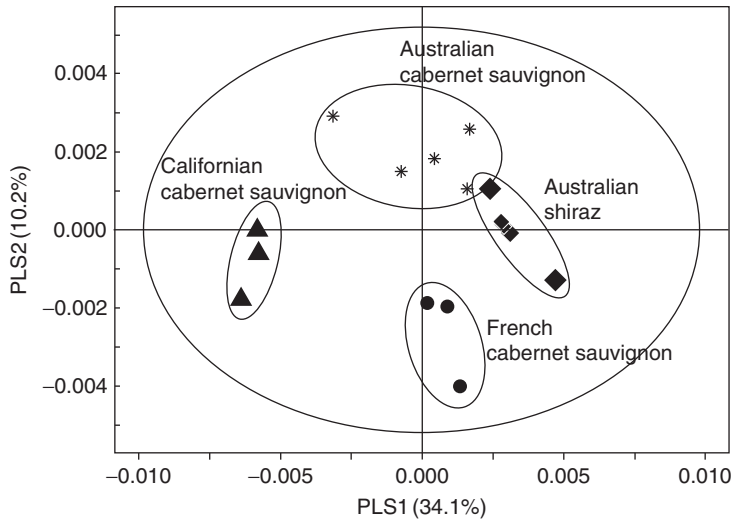



**FIGURE 4.2** PLS-DA score plot performed on the $^{1}$H NMR spectra of wines. The plot shows the clear discrimination among Australian Shiraz wines (square) and French (circle), Californian (triangle), and Australian (star) Cabernet Sauvignon wines. (From Son *et al.*, 2008.)

amounts of proline and glucose compared to Australian ones, which conversely presented with higher amounts of glycerol, tartaric, succinic and gallic acids, 2,3-butanediol, and 2-phenylethanol (Son *et al.*, 2008).

Differentiation between the traditional Cypriot spirit ''Zivania'' and other alcoholic beverages considered competitors, such as grappa, ouzo, rakea, vodka, gin, tsouika, tsipouro, tsikoudia, and eau-de-vie was performed. However, excellent prediction was not achieved by applying CDA and CBT (classification binary tree) to $^{1}$H NMR data (Petrakis *et al.*, 2005).

From all these papers, it appears that the geographical characterization of wine is possible when metabolite profiling is used and moreover, it seems that glycerol, alcoholic derivates, and in few cases organic acids could contribute to discrimination between samples when the considered areas are near each other. Conversely, when very different geographical locations are concerned, sugars play the dominant role.

SNIF-NMR and/or IRMS techniques were often combined with trace element analyses (ICP-MS, ICP-OES, FAAS, ETAAS, GFFA) and chemometrics for the geographical characterization of wines. In a relatively old paper, Day *et al.* (1995) analyzed 165 grape samples collected in 1990 in four different production areas of France (Alsace, Beaujolais, Burgundy, and the Loire Valley). The combined use of isotopic and trace element data allowed an excellent classification of wine samples corresponding to

their geographical origin by performing PCA and CDA. The isotopic parameters of the two main constituents of wine, water ($(D/H)^{Q}_{W}$, $\delta^{18}O^{Q}_{W}$) and ethanol ($(D/H)_{I}$, $(D/H)_{II}$, $\delta^{13}C$), were reliable indicators of the climate since they were related to meteorological conditions and to photosynthetic activity, while trace elements were more related to soil type. The same procedure was applied to the Burgundy wine region, and a good classification was obtained for samples coming from different appellations of the same region. Martin *et al.* (1999) investigated wines samples harvested during several years coming from a small production area in the region of Bordeaux (France). In this case, by considering only the stable isotopic analysis of ethanol and water of wines, a good discrimination corresponding to wine vintage was achieved. Combining isotopic with trace elements data, an improvement in the geographical origin classification for wines of different appellations of the Bordeaux region was obtained.

## B. Olive oil

Olive oil is a product with high nutritional values and health benefits largely produced in the Mediterranean area. It is well appreciated for its organoleptic characteristics and content of antioxidant compounds. There are different methods for extracting the oil from olives and these processes, as well as the pedoclimatic conditions, the agronomic factors, the cultivar and the processing techniques, have a direct influence on olive oil quality. The EU legislation classifies olive oil into different categories on the basis of their quality. Different grades of virgin olive oils are designated, where the extra virgin olive oil (EVOO) is considered the highest quality product on the basis of chemical parameters, essentially the total acidity. In the case of EVOO, this value must be less than 0.8%, in virgin olive oil less than 2%, in olive oil less than 1.5%, and in lampante olive oil larger than 2%. The superior organoleptic characteristics and nutritional properties of EVOO make it a highly valuable food in the market and, for this reason, several types of fraud based mainly on addition with less expensive oils are encountered. From the chemical point of view, olive oil consists mainly of glycerides with lesser amounts of organic acids, phenolic compounds, terpenes, and sterols. From the NMR point of view, the proton spectrum is dominated by fatty acids signals, while, for an estimation and assignment of minor components, such as aldehydes and terpenes, accurate recording conditions need to be established.

In 1998, the Sacchi group published the first article dealing with the geographical characterization of olive oils by using $^{1}H$ NMR spectroscopy and multivariate statistical analysis (Sacchi *et al.*, 1998). In this paper, 55 EVOO samples, obtained from different olive varieties and coming from four Italian regions (Campania, Lazio, Sicily, and Umbria), were

investigated. Quantitative data of selected resonances from $^{1}$H NMR spectra due to minor components (sterols, *n*-alkanals, *trans*-2-alkenals, and others volatile compounds) were analyzed by PCA and HCA. Both statistical approaches led to a very good classification of olive oil samples corresponding to their geographical origin. In particular with HCA, 96% of samples were correctly classified, while in the PCA model, observations clearly clustered in four different groups corresponding to the original four Italian regions (Fig. 4.3).

In 2001, the same research group investigated more than 200 samples of EVOO (Mannina *et al.*, 2001a,b) collected in 3 years (1996, 1997, and 1998) and in different Italian areas (Liguria, Tuscany, Lazio, Sicily, Apulia, Garda lake, and in a borderline district between Lazio and Tuscany). Different statistical protocols (ANOVA, LDA, TCA, and K-Means Clustering) were performed on the intensity of 11 normalized resonances from $^{1}$H NMR spectra. These resonances, as mentioned above, were due to minor components of olive oil: hexanal, *trans*-2-hexenal, two unknown unsaturated aldehydes, formaldehyde, three terpenes, squalene, cycloartenol, and β-sitosterol (Segre and Mannina, 1997). Because the variability of these resonances is due to the environment, the cultivar, olive oil defects, and the year of production, the authors considered initially EVOO for single vintage in statistical analysis, obtaining a good differentiation among samples corresponding to their geographical origin (Fig. 4.4).

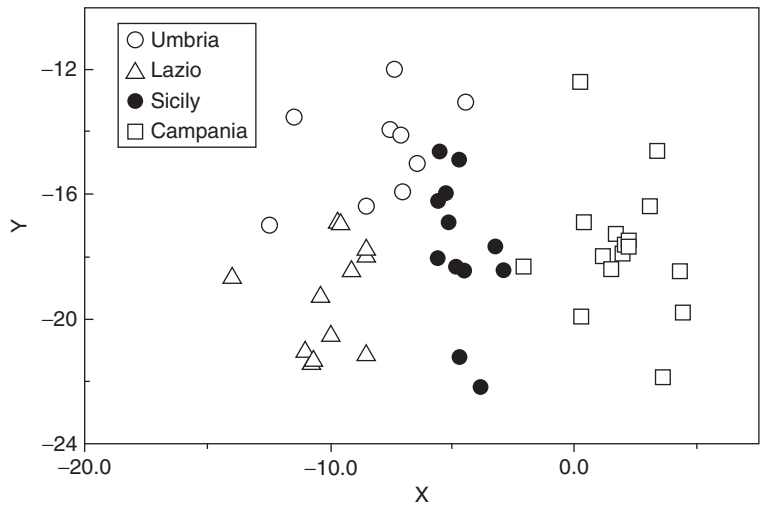



**FIGURE 4.3** PCA score plot performed considering Italian extra virgin oils from different geographical regions and selected NMR intensities. (From Sacchi *et al.*, 1998.)

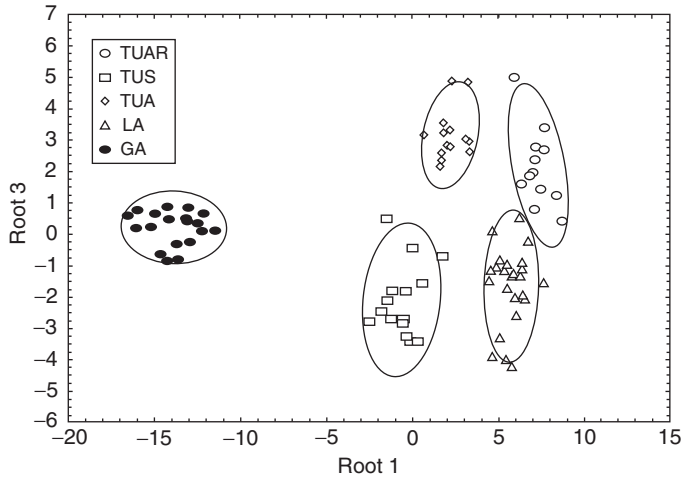



**FIGURE 4.4** LDA canonical score plot performed by considering extra virgin olive oils from different Italian areas in 1998 (Arezzo, TUAR; Seggianese, TUS; Lucca, TUA; Lazio, LA; and Lake Garda, GA). Ellipses represent the 95% confidence regions for each group. (From Mannina *et al.*, 2001b.)

By considering samples of different production years (1996 and 1997), a poorer, but adequate, geographic discrimination was achieved.

Multivarietal virgin and PDO olive oils harvested in 2003 in the north, the center, and the south area of Lazio were investigated by D'Imperio *et al.* (2007). PCA and LDA performed by considering the intensity of some selected resonances from both $^{1}H$ and $^{13}C$ NMR spectra (Fig. 4.5) led to a very good discrimination among samples from the three different pedoclimatic Lazio areas. Olive oils from the northern area had the lowest content of terpenes, while higher amounts were present in olive oils from the center, in addition to linolenic acid. Olive oils from the southern area showed a high level of squalene. The oleic acid content decreased slightly from the south to the north, and this might be related to different cultivars, agronomical practice, and rainfall amounts. A rather good separation among olive oils coming from the five provinces of Lazio was also achieved by using LDA (Fig. 4.6).

To confirm the accuracy of the NMR approach, the methodology used by this Segre research group was recognized in 2001 by the regional law no. 21 of August 3 as the official method to control the Lazio olive oil quality.

Italian EVOO from different cultivars and geographical areas of the Apulia region were investigated by Sacco *et al.* (2000), measuring analytical parameters (fatty acids in particular) and recording $^{1}H$ NMR spectra of phenolic extracts of olive oil. Their metabolite content was confirmed to

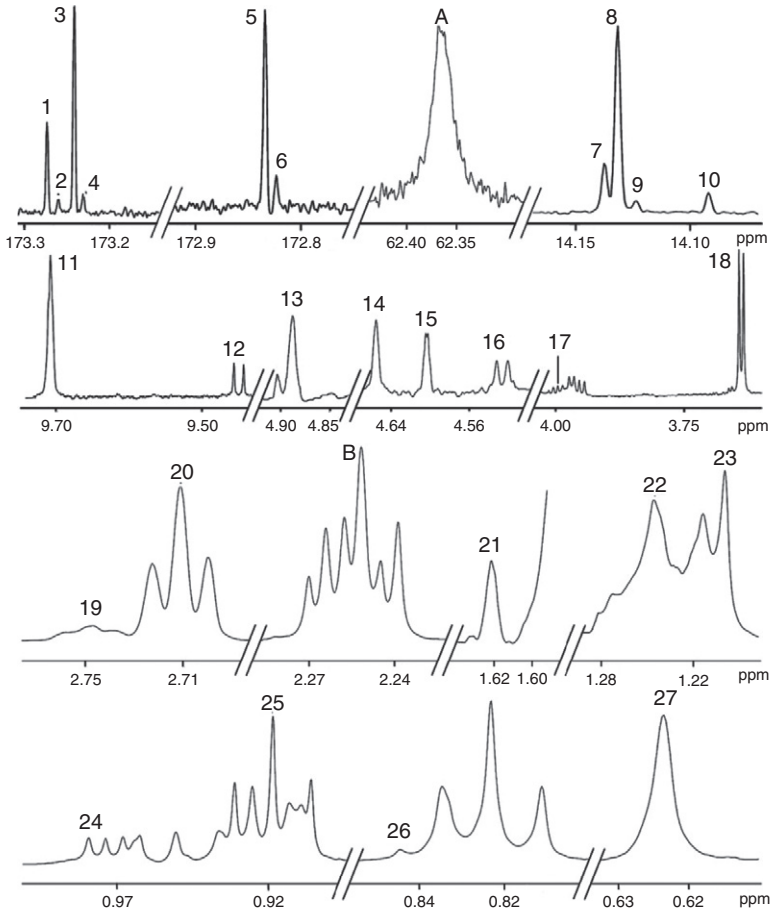



**FIGURE 4.5** $^{13}C$ (250 MHz) and $^{1}H$ NMR (600 MHz) signals used for the statistical analyses. 1, carbonyl signal of *sn*-1,3 saturated fatty chain; 2, carbonyl signal of *sn*-1,3-eicosen-11-oic and vaccenic fatty chains; 3, carbonyl signals of *sn*-1,3-oleic fatty chains; 4, carbonyl signals of *sn*-1,3-linoleic fatty chains; 5, carbonyl signals of *sn*-2-oleic fatty chains; 6, carbonyl signals of *sn*-2-linoleic fatty chains; "A": reference peak ($^{13}C$ NMR spectra) due to α-methylenic protons of glycerol moiety normalized to 100; 7, methyl of palmitic and stearic fatty chains; 8, methyl of oleic fatty chains; 9, methyl of eicosenoic and vaccenic fatty chains; 10, methyl of linoleic fatty chains; 11, hexanal; 12, *trans*-2-hexanal; 13, terpene 4; 14, terpene 3; 15, terpene 2; 16, terpene 1; 17, methylenic protons in α glycerol moiety of *sn*-1,3-diglycerides; 18, methylenic protons in α glycerol moiety of *sn*-1,2-diglycerides; 19, diallylic protons of linolenic fatty chains; 20, diallylic protons of linoleic fatty chains; "B": reference peak ($^{1}H$ NMR spectra) due to methylenic protons bound to C2 normalized to 1000; 21, squalene; 22, methylenic protons of all unsaturated fatty chains; 23, methylenic protons of palmitic and stearic fatty chains; 24, wax; 25, methyl of linolenic fatty chains; 26, methyl of linoleic fatty chains; 27, methyl-18 of β-sitosterol. (From D'Imperio *et al.*, 2007.)

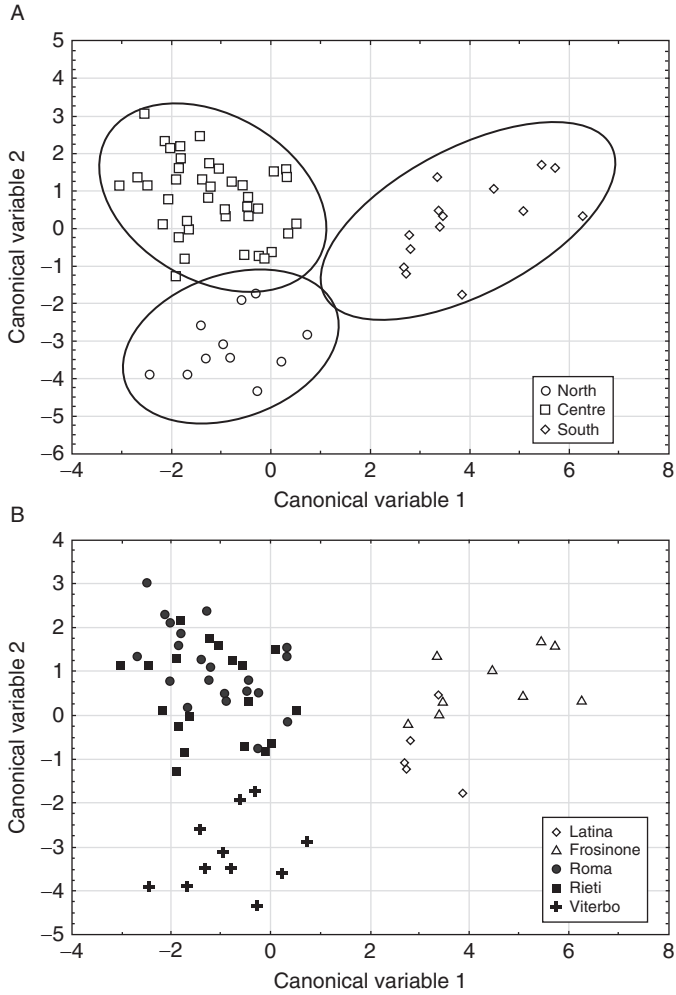



**FIGURE 4.6** LDA performed on the intensities of the variable 3, 4, 6, 8, 10, 12, 14, 15, 18, 19, 20, 21, 22, 23, 24, 26, and 27 (see Fig. 4.5). (A) Olive oil samples from the northern (11), the center (40), and the southern (14) areas of Lazio are differently labeled. The ellipse represents 95% confidence regions for each group. (B) Olive oil samples from five provinces of Lazio are differently labeled. (From D'Imperio *et al.*, 2007.)

be strongly affected by the variety, origin, and ripening stage of the olives (Amiot *et al.*, 1986; Montedoro and Garofolo, 1984; Solinas, 1987). Moreover, since phenolic compounds are strictly related to the typical bitter taste of olive oil and to the resistance of the oil to oxidation, the content of phenolic compounds is a useful index of olive oil quality. By performing DA on NMR data, a very good separation among samples coming from

the coast, north, and hinterland areas of Apulia was achieved with a prediction ability of 96% while no correlation between fatty acids composition and geographical origin was found; analytical data permitted only the discrimination of olive variety. By performing ANOVA and PCA on selected $^{1}$H NMR resonances (aldehydes, terpenes, squalene, linolenic acid, and *sn*-1,3-diglycerides), a differentiation among PDO EVOO from the Veneto region and the Garda Lake regions was achieved. By considering a very small number of samples, in a preliminary analysis, it was shown that a separation among EVOO from different pedoclimatic areas of the Veneto PDO area was still possible (Mannina *et al.*, 2005). With the same procedure, an excellent geographical origin differentiation between PDO EVOO coming from both the Veneto and Lombardia banks of Garda Lake (Fig. 4.7) was obtained (Schievano *et al.*, 2006). The PCA loading plot showed a higher amount of hexanal, *trans*-2-hexenal, and cycloartenol for EVOO samples from the Lombardia bank of Garda Lake.

For the first time, Shaw *et al.* (1997) investigated, by $^{13}$C NMR, EVOO samples of different cultivars coming from different Italian regions and from Israel. Thirty-nine carbon signals were integrated and used in statistical analysis, leading to a good sample separation corresponding to their variety. In addition, a good geographical origin prediction for samples of Toscana, Abruzzo, Puglia, and Israel was obtained by performing PLS, PCR, and PCA.

Olive oil samples coming from 13 PDO Italian areas of production were analyzed by $^{13}$C NMR DEPT (distortionless enhancement by polarization transfer), a particular pulse sequence used to improve the signal-to-noise ratio of $^{13}$C spectra (Vlahov *et al.*, 2001). Olive oils were dissolved

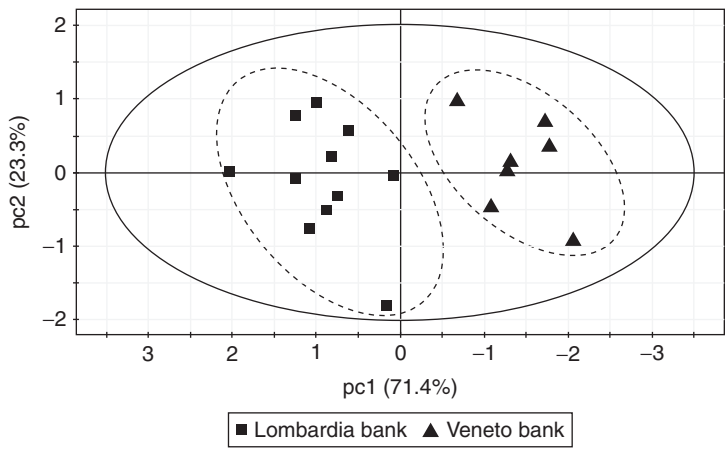



**FIGURE 4.7** PCA score plot performing considering all extra virgin olive oils coming from both Lombardia and Veneto bank of Garda lake and five selected NMR intensities.

in $CDCl_3$; 49 resonances were integrated and processed by PCA. The first two PCs explained 84% of the total variance leading to a grouping of PDO olive oils mainly corresponding to the cultivar type rather than the PDO area. The fatty acid composition of the triglyceride fraction was the most discriminant variable. Better results in geographical determination were obtained by investigating 173 olive oils samples from three PDO areas of Apulia by using the same NMR and statistical methodology (Vlahov *et al.*, 2003). The intensity data of triacylglycerol resonances were processed by LDA obtaining an excellent discrimination; olive oil samples from the PDO ''Colline di Brindisi'' and ''Terra di Bari'' were 90% correctly classified, whereas only 74% of ''Dauno'' PDO olive oils were correctly classified. In 2005, the Rezzi group investigated the profiles of olive oil samples harvested in 2002 and 2003 from different areas around the Mediterranean Sea such as Greece, Italy, Spain, Tunisia, and Turkey (Rezzi *et al.*, 2005). Among different statistical approaches used (LDA, PLS-DA, GDA BS, CC, and PNN), the better classification result, related to both geographical origin and production year, was achieved by applying PNN (from 58% to 100% on the external validation). It was shown that the variability due to production year is less determinant with respect to the country of origin.

In 2008, the Petrakis group investigated 131 EVOO (cv. Koroneiki) collected during 5 years (2001–2006) in three regions of Crete, two regions of Pelopponnesus and from the Zakynthos island in Greece, by using $^{1}H$ and $^{31}P$ NMR spectroscopy coupled with CDA and CBT (Petrakis *et al.*, 2008). The authors obtained a quite good sample discrimination corresponding to the geographical origin on the basis of linoleic and oleic acids, pinoresinol, 1,2-diacylglycerols, free acidity, free hydroxytyrosol, and total tyrosol. The inclusion of the harvesting year improved the classification of the samples. A discrimination among EVOO by considering six sites in Greece was also achieved even if the geographical predictions were 74% versus 87% obtained by considering the three macro regions.

## C. Cheese

In recent years, large varieties of cheese obtained the PDO trademark, such as Parmigiano Reggiano, buffalo milk Mozzarella and Gorgonzola (Italy), Camembert and Cantal (France), Sfela and Feta (Greece). For this reason, the geographical origin determination of cheese is increasing progressively, and is needed to protect consumers and producers from frauds and to assure the quality of these products.

The environmental conditions of a geographical area induce specific characteristics in the product, becoming a factor of primary importance in determining its typical nature. A PDO cheese in particular must be obtained from milk of animals bred in the PDO area and its organoleptic

characteristics are not reproducible in another geographical environment. There are very few papers in the literature dealing with the geographical origin determination of cheeses by using NMR spectroscopy and multivariate statistical protocols. In the first published article with this aim, buffalo mozzarella samples and the milk used for its production were investigated (Brescia *et al.*, 2005). The authors combined different analytical techniques such as HPIC (high-performance ion chromatography), ICP-AES (inductively coupled plasma atomic emission spectroscopy), $^{1}$H NMR, and IRMS (isotope ratio mass spectrometry) with chemometrics for the geographical differentiation between buffalo mozzarella cheeses and the corresponding milk used from the Caserta and Foggia areas (in the Campania and Apulia regions respectively, in the south of Italy). ''Mozzarella di Bufala Campana'' obtained the PDO recognition in 1996 and the PDO territory includes currently some areas within the Campania, Lazio, and Apulia regions of Italy. When only milk was analyzed, the authors obtained a very good separation between the Caserta and Foggia samples performing PCA, HCA, and DA with analytical and isotopic data. Considering only analytical data, mozzarella samples differentiation could not be obtained in accordance with their geographical origin. Only by combining isotopic parameters with NMR data, determined on aqueous mozzarella extracts, good results were achieved by carrying out PCA, HCA, and DA. This is a clear indication that the mozzarella production process plays a predominant role in the metabolite content, resulting in determinants that can be used for the geographical characterization. In particular, PCA has shown the $^{13}C/^{12}C$ ratio, acetate and tyrosine as the most discriminant variables. In a very recent paper, Sacco *et al.* (2009) investigated the geographical origin of cow milk used for the production of cow mozzarella with different analytical techniques. PCA of combined NMR and ICRMS data led to a better discrimination for water extracts of milk from central Europe and Apulia with respect to that obtained when chromatographic and emission spectroscopy data were considered.

Emmental cheese samples from seven different Europe regions were analyzed by Shintu and Caldarelli (2006). In this preliminary investigation, MAS spectroscopy has been successfully applied, enabling direct determination on the intact sample, avoiding derivatization or extraction that could alter the chemical content. By using resonances of fatty acids, organic and amino acids in PCA, a discrimination among different regions was achieved scoring with the sixth versus the fourth principal components, explaining totally only 13% of the total variance. Interestingly, Swiss samples were not particularly well differentiated, while the percentage of the correctly reclassified samples reached 89%. The use of canonical analysis performed on 10 selected variables allowed the sample separation according to their geographical origin (Fig. 4.8). Furthermore, the suitable spectral domains for geographical discrimination were

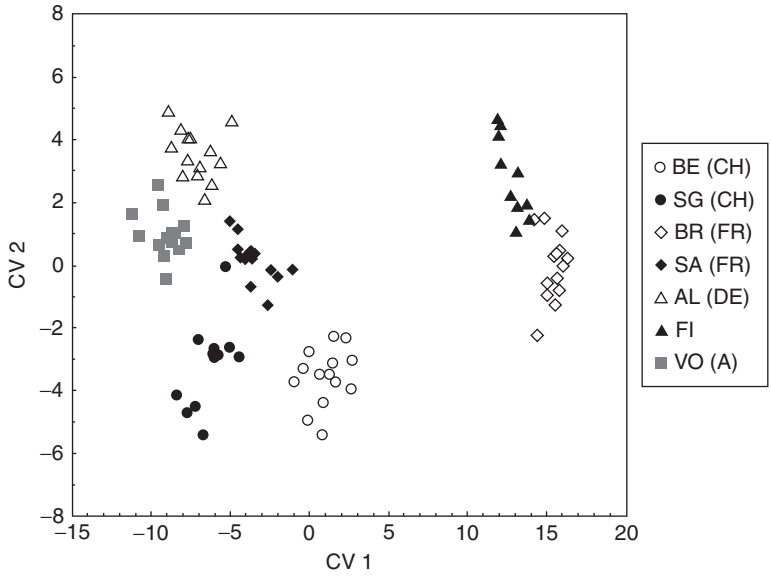



**FIGURE 4.8** Canonical analysis of Emmental cheeses from different countries. (From Shintu and Caldarelli, 2006.)

suggested to be unsaturated fatty acids, aspartic, the olefinic protons of lipids, serine and asparagine signals.

Recently, Consonni and Cagliani (2008b) obtained very good results for ripening (see Section III) and geographical characterization of Parmigiano Reggiano cheese, which received a PDO trademark in 1996. In this study, the authors analyzed the water-soluble components of Italian Parmigiano Reggiano cheeses (probably the most appreciated Italian cheese in the world) at different ripening stages and ''Grana-type'' cheeses from East Europe, the more common present in the Italian market. A typical $^{1}$H NMR spectrum of Parmigiano Reggiano cheese is reported in Fig. 4.9. By performing a PLS-DA protocol, a very clear differentiation between the Italian and foreign samples was achieved with the 72.5% of the total variance explained by the first two components (score plot in Fig. 4.10). The corresponding loading plot highlighted that foreign samples were characterized by a larger amount of leucine and isoleucine and also of lactic acid, butanoate, and acetic acid. Conversely, Italian cheeses were characterized by higher amounts of all other compounds, in particular threonine (which typified Parmigiano Reggiano samples aged for 30 months), valine, proline, glutammic acid, lysine, alanine, serine, arginine, and citrulline. In this case, the leucine content was a good marker for the ripening stage of Parmigiano Reggiano, thus suggesting that foreign samples are even less aged than 14-month Italian cheeses.

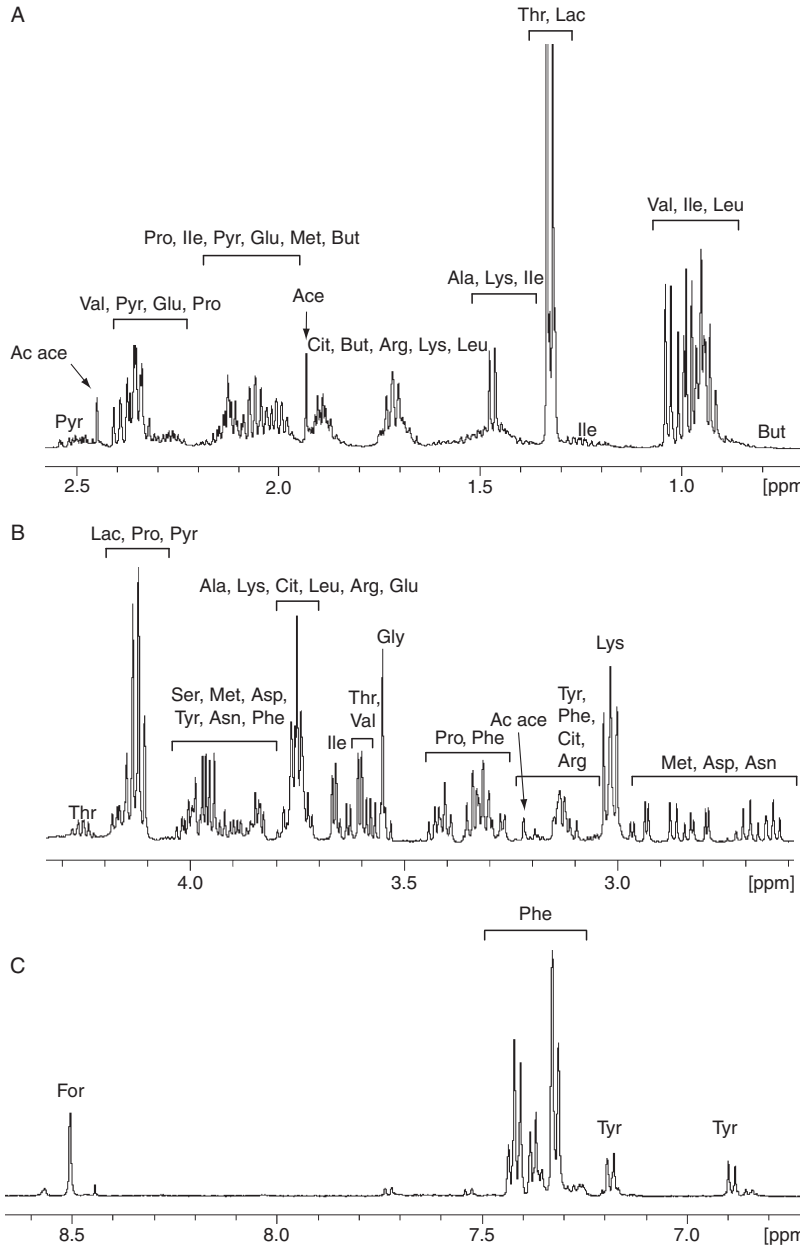



**FIGURE 4.9** $^{1}$H NMR spectrum of "Parmigiano Reggiano" aqueous extract sample. Principal spin system assignments are indicated. A, B and C are expanded regions of the NMR spectrum. (From Consonni and Cagliani, 2008b.)

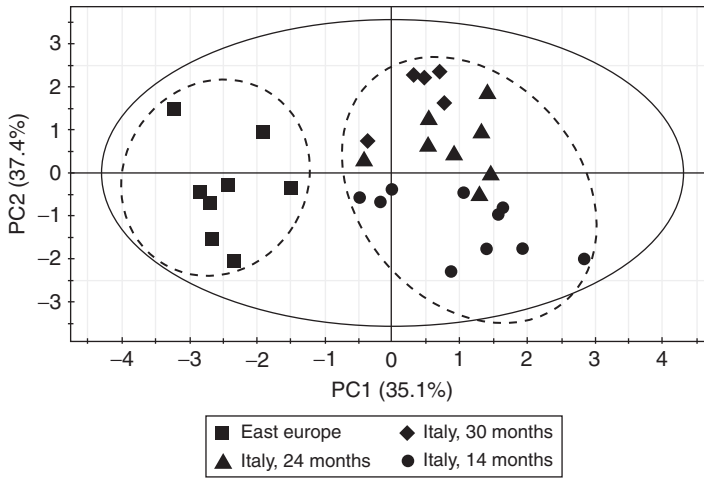



**FIGURE 4.10** PLS-DA score plot obtained by considering 23 Italian Parmigiano Reggiano samples and all east Europe countries "Grana-type" samples. In the score plot, filled symbols represent samples of 14 (circle), 24 (triangle), and 30 (diamond) months of ripening. (From Consonni and Cagliani, 2008b.)

## D. Cereals

The quality of dry pasta and therefore of durum wheat flour is strongly related to the protein content, gluten strength, color, and mycotoxin and chemical levels; all of these parameters depend on plant type, geographical origin, and production technology. It is therefore necessary to find analytical techniques that allow the characterization of these flour features. Consumers are interested in the geographical origin determination on these products. Authenticity markers need to be found to characterize typical products from specific Italian regions. This supports the promotion of a PDO stamp certifying durum wheat bread authenticity as strongly requested by the producers.

The first article dealing with the characterization of durum wheat flour in terms of geographical origin appeared in the literature in 1998. In this preliminary study, Sacco *et al.* (1998) investigated with MAS spectroscopy some durum wheat flour samples of different varieties, geographical origins (different Apulia areas), kinds of soil and times of harvest. $^{1}$H MAS NMR spectra was dominated by lipid and polysaccharide signals in particular due to α(1–6) and β(1–4) glucopolysaccharides. Spectroscopic data obtained using NOESY pulse sequence (Fig. 4.11) was digitalized and used for PCA. By considering both the lipid and saccharide regions, a good differentiation among samples was achieved. The same NMR technique coupled with IRMS was applied by Brescia *et al.* (2002b) to analyze samples of two durum wheat cultivars (Simeto and Colosseo), then converted to semolina, grown in 13 different locations in Italy.

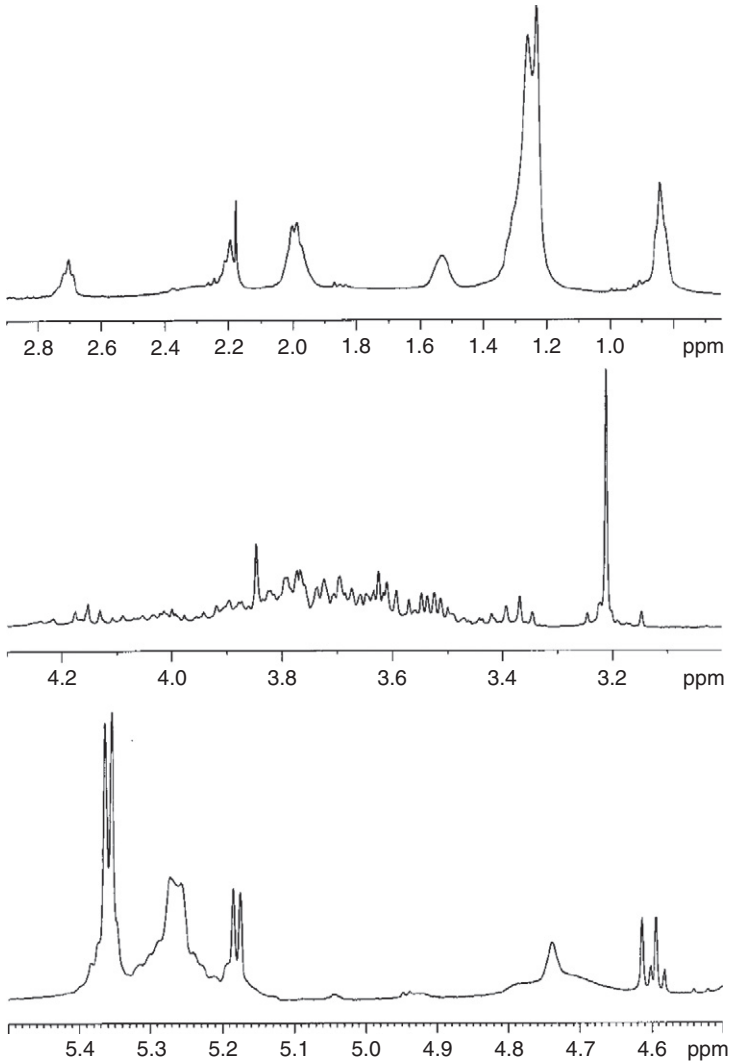



**FIGURE 4.11** $^{1}H$ NOESY spectrum (recorded at 400 MHz) of durum wheat flour dissolved in $D_2O$. (From Sacco *et al.*, 1998.)

The combined NMR and isotopic data were also compared to analytical/classical determinations. The latter did not allow achievement of a good geographical discrimination among samples in PCA, and also the prediction ability in DA was quite low (54%). Better results were obtained by considering the combination of isotopic ($^{13}C/^{12}C$, $^{18}O/^{16}O$, $^{15}N/^{14}N$) and NMR data (mainly by considering signals related to polysaccharides and triacylglycerols). In this case, a good geographical discrimination was

achieved in both PCA and DA protocols (prediction ability of 100%), as well as a good classification corresponding to cultivar. In 2007, the Brescia group applied the combined use of $^{1}$H HR-MAS NMR and IRMS to characterize different typical Italian breads coming from Altamura (PDO) and Laterza in Apulia, from Matera (which is going to obtain the PGI stamp from the European Community) in Basilicata and from Bonorva in the Sardinia region (Brescia *et al.*, 2007). In this study, not only the final product (bread) but also the raw material (semolina and the intermediate dough) were considered for evaluating the possibility of geographical differentiation. Performing PCA on isotopic ($\delta^{13}$C and $\delta^{15}$N ratios and average values) and NMR (intensities of lipids and polysaccharides) data, a significant differentiation between Bornova flour and the others was found, mainly based on $\delta^{15}$N values. Better results were achieved by considering dough samples. In this case, a very good differentiation among the four regions was achieved (Fig. 4.12), highlighting NMR and $\delta^{13}$C as the most discriminant variables. Sample discrimination was improved using DA, resulting in 87% of classification ability and 76% of predictive capacity. Interestingly, differences among samples were reduced when bread samples were concerned. In this case performing DA, only 60% of predictive capacity was reached. The authors attributed this behavior to the chemical modifications that occurred during the processing and in particular after cooking. Nevertheless, Bornova bread samples were well differentiated from the others on the basis of $\delta^{15}$N values.

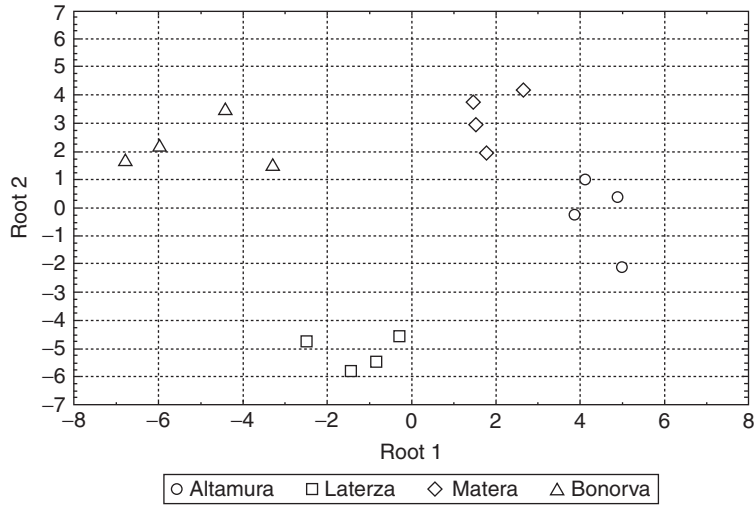



**FIGURE 4.12** Discriminant analysis score plot performed considering dough samples of different geographical origins considering both NMR and IRMS data (From Brescia *et al.*, 2007.)

## E. Honey

Honey is a food product appreciated all around the world because it is a readily available source of energy and also because of its antibacterial and antioxidant activity (Bogdanov, 1997; Perez *et al.*, 2006). A large variety of monofloral and polyfloral honey types are available in the market, presenting large differences in physical, chemical, and organoleptic characteristics. Unfortunately, honey adulteration is widespread and both the quality and authenticity of honey need to be controlled to preserve the production area, to develop particular standards of quality and to protect consumers from commercial fraud (Bogdanov and Martin, 2002). The European Union Commission is encouraging the development of new analytical methods to verify and to assess the quality requirements for different honeys and to characterize their geographical origin. Following the Codex Alimentarius Standard for honey (Codex Alimentarius, 2002) and the Council directives (Council Directive, 2002), the use of the geographical origin is allowed when honey is produced exclusively within the area declared on the label. Actually, the geographical origin is evaluated by melissopalynology (pollen analysis) even though this methodology presents some limitations (Molan, 1998) such as the consolidated knowledge in pollen morphology and the need for highly specialized analyzers. In a recent study, Consonni and Cagliani (2008a) suggested the combined use of high-resolution $^{1}$H NMR and chemometrics as an alternative approach to the geographical characterization of honey. They analyzed the water-soluble content of polyfloral and acacia honey samples coming from different EU and non-EU countries. The PCA model based on NMR data resulted in a clear cut differentiation between polyfloral and acacia honey samples. The acacia ones were characterized by a higher amount of sucrose and fructose compared to the polyfloral samples. Interestingly, unsupervised PCA performed only with acacia samples led to a clear separation between Hungarian and Italian honeys. Italian samples were generally enriched with all water-soluble compounds. The geographical discrimination among Italian, Hungarian, and Argentinean polyfloral honeys was achieved as well, by performing a hierarchical PLS-DA. In this case, Argentinean honeys were enriched with phenylalanine and threonine when compared to the others. After hierarchical PLS-DA model validation with a test set, all samples were reevaluated, obtaining a correct prediction for all of them (Fig. 4.13). Preliminary $^{13}$C NMR data were also reported for analyzed samples; the spectra recorded in organic solvent revealed that the dominant signals were fructose and glucose in their different tautomeric isoforms. The quantitative analysis of these signals confirmed a higher F/G ratio for acacia than polyfloral honeys. The glucose and fructose isoforms analysis also allowed the identification of possible markers for polyfloral

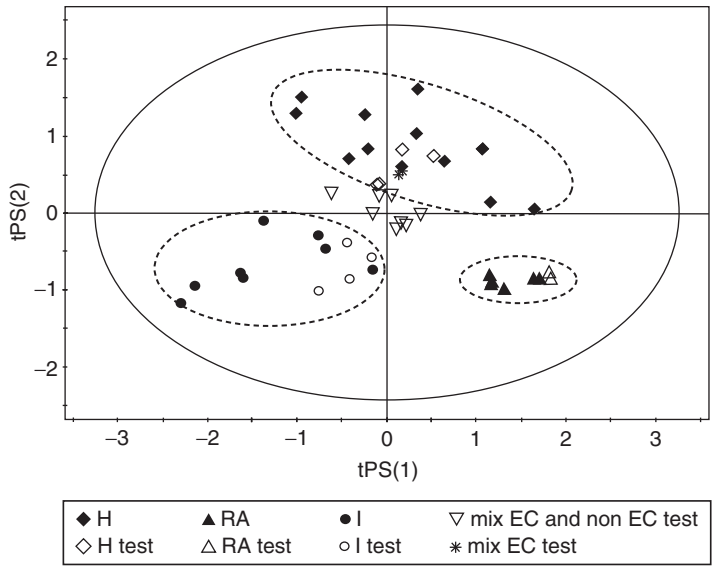



**FIGURE 4.13** Hierarchical PLS-DA performed considering 13 polyfloral honeys of certain origins (training set) with reprojection of polyfloral test set sample scores (10 samples). Filled symbols represent training set honey samples from Hungary (H, diamond), Italy (I, circle), and Argentina (RA, triangle), while open symbols represent test set honey samples from Hungary (H, diamond), Italy (I, circle), and Argentina (RA, triangle) from different EC countries (star) and from different EC and non-EC countries (inverted triangle). (From Consonni and Cagliani, 2008a.)

Argentinean and polyfloral and acacia Hungarian honeys. The βFP/βFF ratio was equal to one when compared to other samples, while for the second discrimination, shift deviations of a few carbons of both glucose and fructose isoforms were detected.

PDO Corsican honey samples and others from different regions of five countries (Austria, France, Germany, Ireland, and Italy) for a total of 182 samples of different varieties were investigated by Donarski *et al*. (2008) by using an NMR spectrometer equipped with a cryoprobe, to improve the sensitivity. $^{1}$H spectra of the water-soluble components of honey were employed for modeling different statistical approaches, like PLS-DA, two-stage genetic programming (GP; Davis *et al.*, 2006) and a new combination of PLS and GP (PLS-GP) by considering Corsican and non-Corsican honey samples. Among all, PLS-GP resulted to be the most accurate method, leading to a correct classification for Corsican honeys of 96.2% compared to 94.5% and 75.8% for two-stage GP and PLS-DA, respectively. PLS-GP used PLS as a variable selection step to determine input variables for GP. This resulted in a more easily interpretable model and was therefore used to identify the useful variables for classification. Among

these, trigonelline was identified for the first time in honey, by means of TOCSY spectrum and standards compounds, being a useful biomarker for geographical origin (saline habitat) or growth conditions (dry habitat).

## F. Fish

While in 1970 only 4% of the world's seafood came from the aquaculture, today the percentage has increased to 32% (FAO, 2007). Progress in aquaculture techniques has led to year-round availability of farmed fish and lower prices for consumers. To regulate labeling, packaging, and traceability and therefore to provide consumers with basic information on the characteristics of fish products, Commission Regulation (EC) 2065/2001 was introduced in 2002. According to this regulation, fish must be labeled with a specific commercial designation and scientific name, production method and fishing area. Accordingly, new analytical methods to assess both wild and farmed fish and to characterize their geographical origin need to be developed. Only two articles have appeared in the literature dealing with this aim by coupling the NMR technique and statistical methods. In the first one, Rezzi *et al.* (2007) investigated the lipid extracts dissolved in $CDCl_3$ of farmed and wild sea breams. One group came from markets in Italy, Greece, Croatia, and Turkey, while the second group was fished from the Mediterranean Sea. The PCA performed on the digitized NMR spectra of all samples led to a clear difference between the wild and farmed samples on the basis of methyl and methylene protons together with methylene and methyne protons in unsaturated fatty acids. Conversely, the LDA carried out on both PCA scores and NMR data allowed the classification of wild and farmed sea breams corresponding to their geographical origins. A similar study was carried out by Aursand *et al.* (2009) on Atlantic salmon. The authors analyzed more than 230 salmon samples, and among them 195 were of known Atlantic origins (wild salmon from Norway, Scotland, Canada, Iceland, Ireland and farmed salmon from Norway, Scotland, Canada, Iceland, Ireland, Faroes, and Tasmania), while the others were purchased on Italian, English, and Norwegian markets. By performing PNN (probabilistic neural networks) and SVM (support vector machines) on $^{13}C$ NMR data recorded on heterogeneous lipid extracts (Fig. 4.14) of wild and farmed salmon, they achieved an excellent discrimination between wild and farmed salmon samples (95.5% and 100%, respectively). The geographical origin classification was less feasible but the authors nevertheless achieved 82.2% and 99.3% of correct classification by PNN and SVM, respectively. Analyzing salmon samples purchased from the market, five of them were farmed instead of wild as indicated on the label. In addition, there were some dissimilarities between the geographical origin as stated on the label and resulted by the statistical classification. In this example, NMR spectroscopy

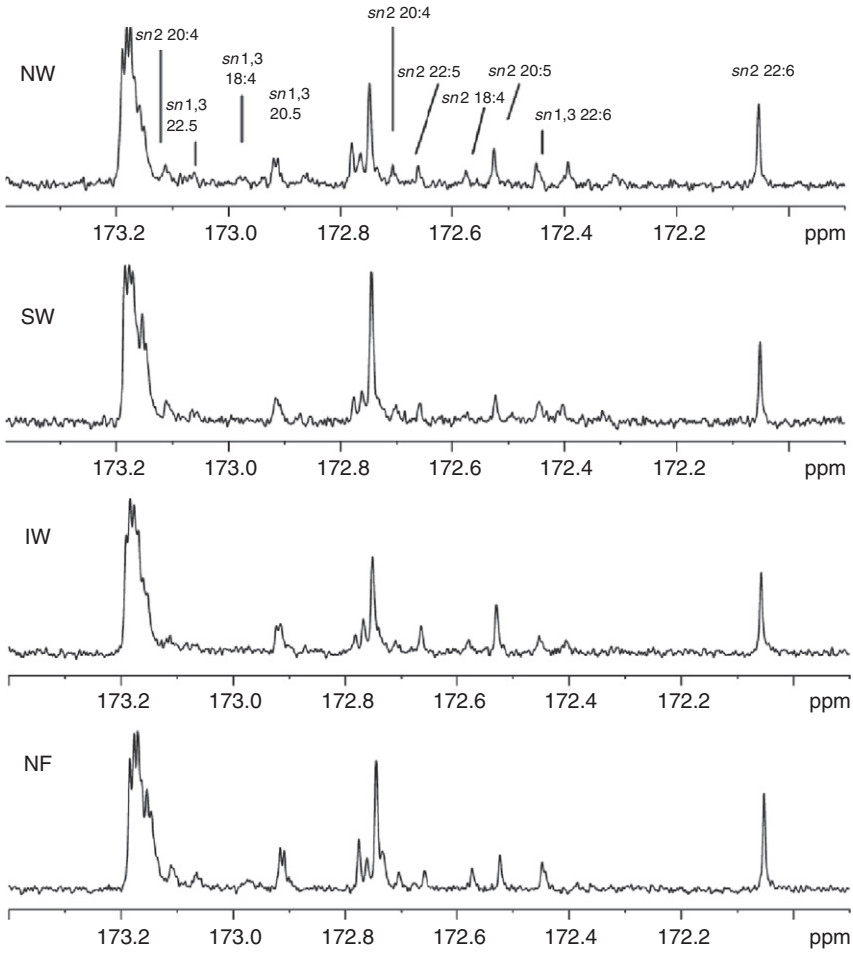



**FIGURE 4.14** $^{13}$C NMR (500 MHz) carbonyl region (173.4–172.0 ppm) of lipids extracted from salmon muscle of four different origins: (from top) wild salmon from Norway (NW), Scotland (SW), and Ireland (IW) and farmed salmon from Norway (NF). The position of fatty acids in triacylglycerols is designated (*sn*1,3 or *sn*2). (From Aursand *et al.*, 2009.)

combined with multivariate statistical methods was a very useful tool for traceability and for fraud identification.

## G. Meat

The health concerns such as Bovine Spongiform Encephalopathy and Foot and Mouth diseases in the beef meat industry have attracted more attention to the authenticity of meat products. One of the main authenticity

aspects related to fresh meat in Europe is strictly connected to the geographical origin. In 2000, the European Community introduced a regulation (1760/2000) (European Communities, 2000) obliging beef producers to use labels indicating the origin of meat. Meat products belonging to selected breeds and produced in particular area acquire added value in the market, mainly for PGI products. Origin identification of both raw materials and final products is primarily aimed to prevent fraud. In particular, meat is linked to the soil because of the production site and the animal diet.

The first article dealing with geographical characterization of meat by using NMR and chemometrics appeared in 2005. In this study, the Sacco group investigated 25 lamb meat samples from different breeds from three areas of the Apulia region (Comisana in the center, Comisana, Gentile and Merinizzata in the north and Sarda in the south) by using MAS spectroscopy, classical analysis, metal and isotopic ratio determinations, and multivariate statistical approaches (Sacco *et al.*, 2005). NMR samples were prepared as a semisolid pulp obtained by mixing meat and $D_2O$. In the MAS spectrum, fatty acids of triacylglycerol and carnosine signals together with amino acids, organic acids, and sugars were assigned. By performing PCA on NMR and isotopic determinations, poor results were achieved in geographical characterization in comparison with PCA carried out on routine and metal determinations. Nevertheless, the application of DA to the two data sets revealed a much better predictive capability for spectroscopic data (96% vs. 60%). NMR data showed that α-glucose and creatine were enriched in Merinizzata samples, Gentile samples were higher in triacylglycerols, while Sarda e Comisana samples from central Apulia were higher in unsaturated acids, which correlated with the lamb feeding regime. 1D MAS spectra were used by Shintu *et al.* (2007) to evaluate potential markers of one specific quality or geographic origin of dried beef samples of Australia, Brazil, Canada, Switzerland, and USA (Fig. 4.15). Buffered solutions were used for dissolving samples, thus preventing possible pH variations. The PCA did not lead to a good separation of samples corresponding to their geographical origin; only American and Canadian samples grouped fairly. A further investigation for evaluating differences in the dried meat's chemical composition according to its geographical origin was carried out by using DA. Excellent discriminations among the five groups were achieved (Fig. 4.16), and the first two canonical variables explained 88.5% of the total variance. American, Swiss, and Canadian samples had lower relative amounts of carnitine than the other two groups and higher relative concentrations of succinate and/or a compound not identified. Swiss samples were characterized by higher concentrations of phenylalanine and lower amounts of alanine and/or methionine and carnosine and/or tyrosine than the other analyzed samples.

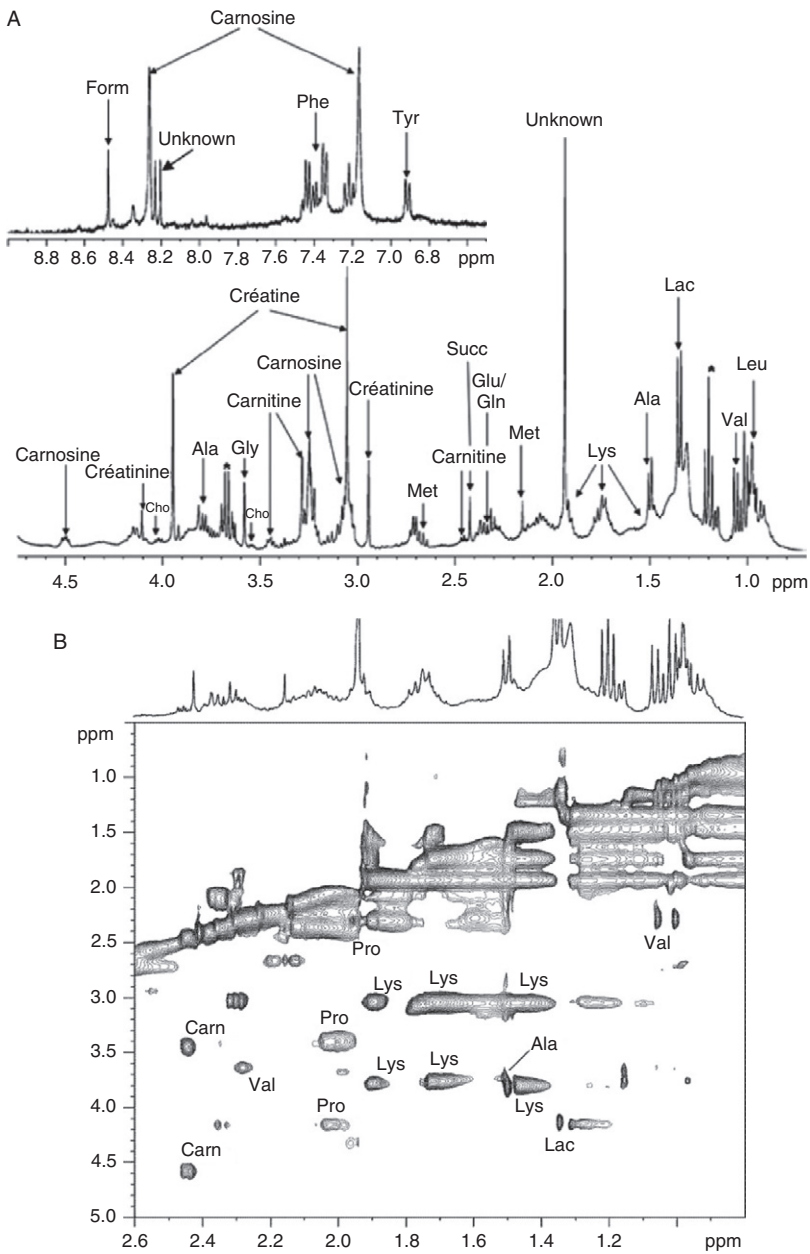



**FIGURE 4.15** (A) Portion of a $^{1}$H HRMAS spectrum (400 MHz) of a Swiss dried meat sample, with the labeling of some signals. Stars indicate residual ethanol from rotor washing. (B) Portion of the high-field region of a TOCSY spectrum of a dried meat sample, with some assignments. (From Shintu *et al.*, 2007.)

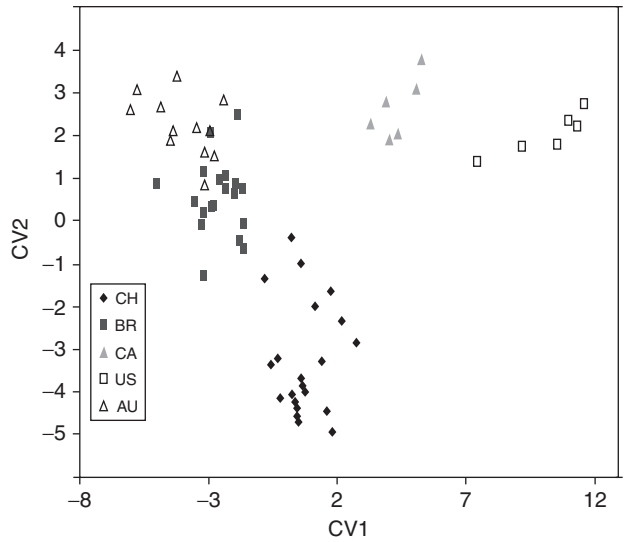



**FIGURE 4.16** Canonical analysis score plot performed considering 66 dried meat samples coming from different countries (CH, Switzerland; BR, Brazil; CA, Canada; US, USA; AU, Australia). (From Shintu *et al.*, 2007.)

## H. Other foods

Other geographical characterization studies performed on different food products by using NMR techniques and chemometrics are also found in the literature and reviewed here.

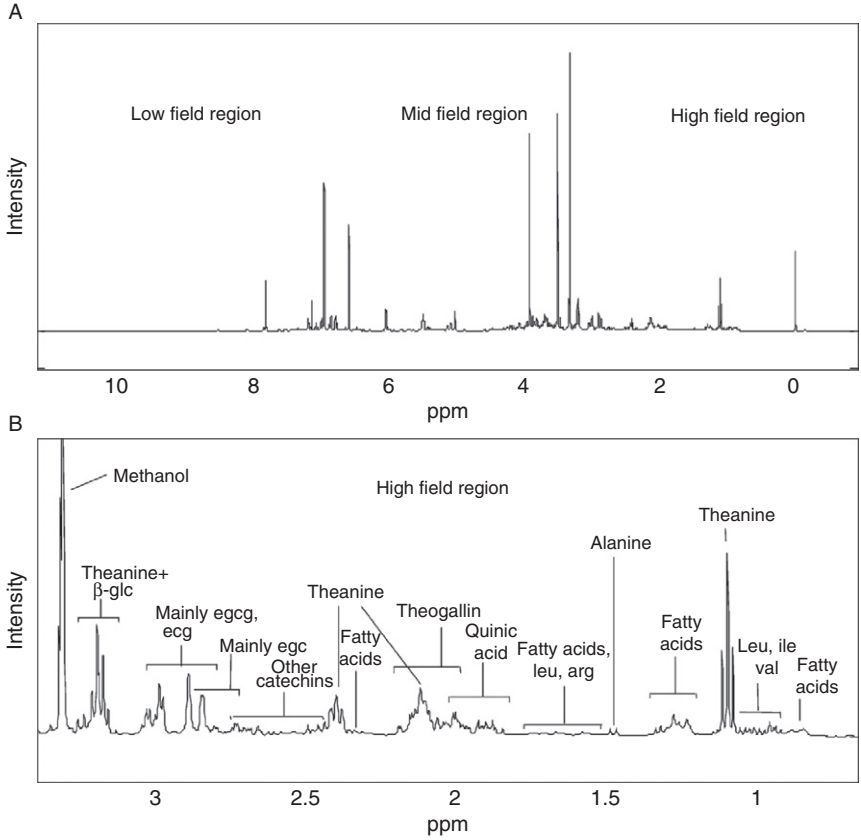
### 1. Mustard oil

Remaud *et al.* (1997) applied the combined use of isotopic analysis (SNIF-NMR and IRMS) and chemometrics to verify the authenticity of mustard oils. The major component of this oil was represented by allyl-isothiocyanate, which can be synthesized more conveniently than extracted from mustard seeds. For this reason, adulteration of natural mustard oil is very profitable. By performing PCA on the analytical parameters $\delta^{13}C$, $\delta^{34}S$, $\delta^{15}N$ (measured by IRSM) and $R_{2/1}$ and $R_{3/1}$ (relative isotope ratios obtained by $(D/H)_{I,II,III}$ measurement with SNIF-NMR) on both synthetic and natural allyl-isothiocyanates coming from the two major producers in the world, Canada and India, a very good differentiation among natural and synthetic samples was achieved. Moreover, a clear geographical differentiation between Canadian and Indian samples was obtained.

## 2. Green tea

In a more recent study, Le Gall *et al.* (2004) analyzed the methanol extracts of more than 190 green teas (Fig. 4.17) coming from China, Japan, Vietnam, India, Indonesia, and Bangladesh. Green tea has been studied during the last 10 years because of its health-related properties and for quality evaluation even thought its antioxidant, anticarcinogenic, antitumorigenic properties, and the cardiovascular disease protection are still to be demonstrated *in vivo* (Higdon and Frei, 2003). In any case, the distinctive characters of each green tea are affected by the geographical origin and therefore an objective analytical determination should support the actual sensory evaluation made by tea testers.

In the present paper, both PCA and CA applied to $^{1}$H NMR spectra indicated some separation between Chinese and non-Chinese samples even though it was not possible to group samples according to their



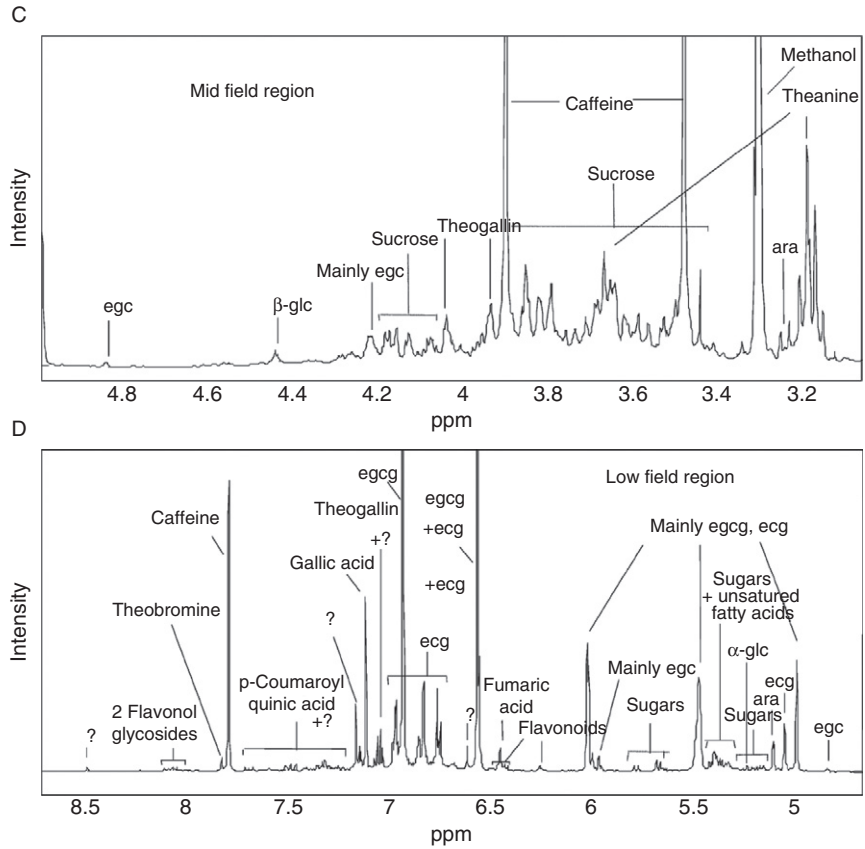



**FIGURE 4.17** Details of $^{1}$H NMR spectrum (recorded at 400 MHz) of a high grade Longjing green tea extract with assignments. Key: leu, leucine; ile, isoleucine; val, valine; arg, arginine; glc, glucose; and ara, 2-*O*-(β-L-arabinopyranosyl)-myo-inositol. (From Le Gall *et al.*, 2004.)

origins. The CA results suggested that a classification according to single countries might be possible by considering a larger non-Chinese sample data set.

### 3. Chamomile flowers (*Matricaria recutita* L.)

Samples from Egypt, Hungary, and Slovakia were studied by Wang *et al.* (2004) by $^{1}$H NMR spectroscopy and PCA. Samples were prepared by using two different extractions: simple water infusion and PSE (pressurized solvent extraction) using aqueous ethanol and water. PCA led to a very good discrimination of samples corresponding to their geographical origin for both extracts. In particular, for the boiling water infusion, the

loading plot indicated that sugars and glutamate/glutamine were the most discriminant variables. In particular, the second PC was affected positively by the differing percentage of stalk used in the extracts.

## 4. Cod liver oil

In a recent study, Standal *et al.* (2008) investigated 38 cod liver oils extracted from wild and farmed cods from Scotland and Norway by comparing $^{13}C$ NMR and GC for quality and geographical characterization. The PCA performed on the NMR data (123 chemical shift intensities were used as variables) led to an excellent differentiation between the Norway and Scottish samples. A discrimination between farm and wild samples and also among farms themselves was possible as well. To find the most discriminant variables, the authors carried out a PLS-GA (partial least squares-genetic algorithm) which indicated only 10 intensities out of the 123 initially used, as necessary for sample discrimination, even though not completely assigned. Among those assigned, such as the ω3 carbon atom of n-6 FA, the ω2 carbon atom of n-6 FA, and the n-6 fatty acids, the latter was the most important peak in wild/farmed differentiation. LDA achieved 100% and 96% of correct classification for wild/farmed and geographical origin determination, respectively. By using GC data, these values decreased to 97% and 63%, respectively. NMR data allowed an even better result in PCA. Conversely, use of GC data (fatty acids) led to a correct differentiation of samples corresponding to wild and farmed origins but discrimination among samples according to their geographical origins was not possible. $^{13}C$ NMR provided a more detailed lipid profile than GC since the positional distribution of fatty acids in triacylglycerols and information about lipid classes are specified.

## 5. Propolis

Watson *et al*. (2006) analyzed an organic solution of propolis samples from different countries all around the world. The PCA performed on NMR data led to a clear distinction, mainly due to sugars content, between Asian and European samples; African ones were more scattered even if barely grouped. Few propolis samples representative of Brazil, Pollen, and Solomon Islands were well differentiated.

## 6. Concentrated tomato paste

In a very recent study, Consonni *et al.* (2009) evaluated the geographical discrimination of Chinese and Italian triple concentrate tomato paste samples by $^{1}H$ NMR (a typical $^{1}H$ NMR spectrum with expansion and assignments is reported in Fig. 4.18) and multivariate statistical analysis. Despite a good market for imported triple concentrated tomato paste from China, the biggest world tomato producer, very few indications suggest the possibility of identifying the origin and the quality of different

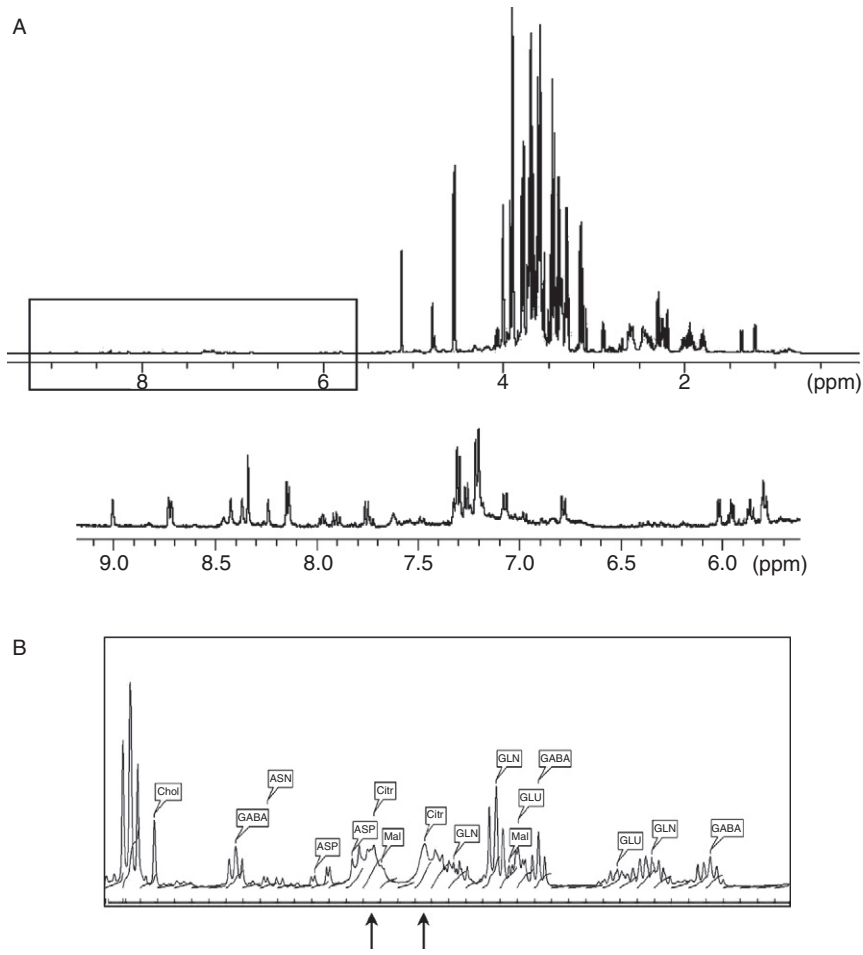



**FIGURE 4.18** (A) $^{1}$H NMR spectrum (recorded at 500 MHz) of a triple concentrated tomato paste sample lyophilized and dissolved in water. Expansion of the aromatic region is also shown. (B) Expansion of the aliphatic region with assigned resonances. Buckets involving citrate resonances, omitted in the second OPLS-DA model, are indicated by arrows. (From Consonni *et al.*, 2009.)

tomato products (Clement *et al.*, 2008; Sequi *et al.*, 2007). The Chinese product is cheaper than any others in the market and from a technological point of view, the equipment available in China is very modern coming mostly from Italy considered worldwide as the best manufacturer of tomato processing equipment.

Notwithstanding that different tomato cultivars and ripening stages were usually used to obtain the final product, the authors achieved an excellent discrimination of triple-concentrated tomato paste samples

corresponding to their geographical origin by performing the unsupervised PCA protocol on the water-soluble metabolite content of lyophilized samples (Fig. 4.19). To build a robust classification model, PCA scores were used for the OPLS-DA model, by considering a training and test sets obtained with D-Optimal Onion Design protocol. A correct classification of both the training and test set samples was achieved. Citrate was the most discriminant variable characterizing Chinese samples, while sugars (glucose and fructose) characterized the Italian samples. Because citrate could be added to triple-concentrated tomato paste for pH correction (even if not allowed by law), the discrimination of samples related to this variable could be biased. By excluding the citrate contribution (see Fig. 4.18 B), a new OPLS-DA was performed obtaining again a good classification corresponding to the geographical origin for all samples. Italian triple-concentrated tomato paste samples again showed a higher sugar content by comparison with Chinese ones which were now characterized by higher levels of aspartic acid and glutamine.

## 7. Cocoa

Hernandez and Rutledge (1994) investigated, by low resolution pulse NMR, the evolution of solid fat content (SFC) at 27.5 °C of cocoa masses of various geographical origin. The ANOVA analysis of a few quantitative parameters correlated with solid content and the speed of its transition from decomposition of fusion curves which indicated that the

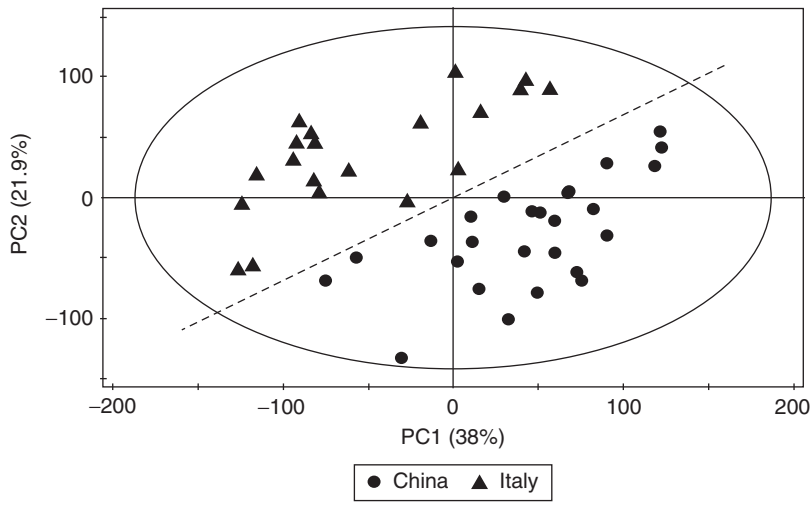



**FIGURE 4.19** PCA score plot performed by considering all triple concentrated tomato paste samples: filled triangles and dots represent Italian and Chinese samples, respectively. PC1 = 38.0%, PC2 = 21.9%. $R^2$ = 79.6%, and $Q^2$ = 56.0%. (From Consonni *et al.*, 2009.)

geographical origin plays a significant role in sample differentiation, based on the fluidification process. Moreover, the analysis suggested that both storage conditions and technological processes play a key role in the discrimination of sample quality (Hernandez and Rutledge, 1994).

## III. QUALITY AND AUTHENTICITY OF FOODS

With the use of the ''quality'' term, several equally important aspects must be considered. The quality determination is, as a matter of fact, a very complex investigation which covers ingredients, additives, fraud, production processes, etc. Furthermore, the term quality has a broad significance, consisting of both safety and healthy aspects. Consumer and producer demands, nutritional and sensorial aspects, and economical and ecological issues associated with food quality are also to be considered. With the recent opening of the market, quality and prices are in general offering different possibilities. The producers themselves have to be aware that their principal aim of reducing the prices of raw materials or finished products often requires a compromise in quality, and this, in general, is not always convenient. It is even more dramatic when this compromise is made with authenticity. Several examples nowadays have shown the dangers of the consumption of fraudulent foods, as well as poor-quality foods. For both quality and authenticity determination, a simple chemical analysis was traditionally adopted, often in combination with sensory analysis. Instrumental analysis has been used in very few cases. The requirement of quality determination was not completely addressed by these approaches. Fortunately, an increasing number of new determinations involve the use of advanced analytical methods and techniques, most of them coupled with chemometric methods. Simple chromatographic methods, for example, are almost useless nowadays for detecting the trace quantities allowed for some contaminants or to assess the geographical food origin. Thus, a significant discrepancy exists between the higher standards of quality and the inappropriate analytical approaches used to address the problem. If the new high quality standards need to be certified, better analytical methods are needed. Very interesting analytical tools are nowadays present in the literature, and most of them remain simply as excellent scientific exercises. A table presenting a summary of methodologies and applications that are potentially useful for quality/fraud detection of foods is reported in Table 4.2.

### A. Wine and beer

Wines consist of several hundred compounds in different concentrations; the dominant derivatives are ethanol, glycerol, sugars, and organic acids, while amino acids and flavonoids are present to a much lesser extent.

**TABLE 4.2** Summary of papers dealing with NMR and chemometric for quality characterization

Quality determination

| Food type | Type/region | Statistical method | Methodology | References |
|---|---|---|---|---|
| Wine, beer | Ale, Lager | PCA | $^{1}H$ | Duarte *et al.* (2002) |
| | Ale, Lager | PCA | $^{1}H$, FTIR | Duarte *et al.* (2004) |
| | Ale, Lager, Pilsner, Stark, Boch, Alt, Export | PCA, PLS | $^{1}H$ | Lachenmeier *et al.* (2005) |
| | Merlot, Cabernet | ANOVA, PCA | $^{1}H$, $^{13}C$ | Pereira *et al.* (2005) |
| | Merlot | PLS-DA | $^{1}H$ | Pereira *et al.* (2006) |
| | Shiraz | PCA | $^{1}H$ | Clark *et al.* (2006) |
| | Tempranillo, Viura, Garnacha | PLS | $^{1}H$ | Avenoza *et al.* (2006) |
| | Rioja | PCA | $^{1}H$ | Lòpez-Rituerto *et al.* (2009) |
| | Muscat Bailey | PCA, PLS-DA | $^{1}H$ | Son *et al.* (2009b) |
| | Korean Meoru | PCA | $^{1}H$ | Son *et al.* (2009c) |
| Vegetables | Potato | PLSR | LF | Povlsen *et al.* (2003) |
| | Potato | PCA | LF | Thybo *et al.* (2003) |
| | Watermelons | PLS-DA | $^{1}H$ | Tarachiwin *et al.* (2008) |
| | Wheat | PCA | MAS | Winning *et al.* (2009) |
| | Rice | PCA | $^{1}H$ | Fumagalli *et al.* (2009) |
| Vinegar | BVM, TBVM | PCA | $^{1}H$ | Consonni and Gatti (2004) |
| | BVM | PLS | $^{1}H$ | Consonni and Cagliani (2007) |
| | BVM, TBVM | ANOVA, PCA | $^{1}H$ | Caligiani *et al.* (2007) |
| | BVM, TBVM | PCA, PLS-DA | $^{1}H$ | Consonni *et al.* (2008c) |

| | | | | |
|---|---|---|---|---|
| Juices | Orange juice, pulp wash | PCA, DA | $^{1}H$ | Vogels *et al.* (1996) |
| | Orange juice, pulp wash | LDA, PCA | $^{1}H$ | Le Gall *et al.* (2001) |
| | Juices, rediluted, frauds | PLS | $^{1}H$ | Humpfer *et al.* (2008) |
| | Orange, grapefruit | ICA | $^{1}H$ | Cuny *et al.* (2008) |
| Olive oil | Leccino, Moraiolo, Dritta | PCA, PLS | $^{13}C$ | Vlahov *et al.* (1999) |
| | Coratina, Oliarola, Leccino, Peranzana | ANOVA, PCA, DA | $^{13}C$ | Brescia *et al.* (2003a) |
| | Nocellara, Biancolilla, Cerasuola,Tonda Iblea | MANOVA, PCA, TCA, MDS, LDA | $^{13}C$ | Mannina *et al.* (2003) |
| | Addition of hazelnut oil | ANN | $^{1}H$, $^{13}C$ | García-Gonzáles *et al.* (2004) |
| | Addition of lampante and refined oils | ANOVA | $^{31}P$ | Fragaki *et al.* (2005) |
| Fish, meat | Cod, salmon | PLS | LF | Jepsen *et al.* (1999) |
| | Pork | ANOVA | LF | Brown *et al.* (2000) |
| | Cod, salmon | PCA | LF | Aursand *et al.* (2008) |
| Dairy | Cow and buffalo milk | PCA, HCA, DA | $^{1}H$ | Brescia *et al.* (2004) |
| | Soymilk | PCA | $^{1}H$ | Yang *et al.* (2009) |
| Honey | American honeys | PCA, ANOVA | $^{1}H$ | Sandusky and Raftery (2005) |
| | Robinia, citrus, chestnut, eucalyptus, polyfloral | PCA, GDA | $^{1}H$, $^{13}C$ | Lolli *et al.* (2008) |
| Coffee | America, Canada, Australia | PCA, LDA | $^{1}H$ | Charlton *et al.* (2002) |
| Additives | Natural/synthetic vanilla | PCA | $^{13}C$ | Tenailleau *et al.* (2004) |

In a recent paper, 17 amino acids commonly present in Sauvignon wine from the Coastal region of Croatia were determined by complete assignment of their resonances by the use of $^{1}H$ and $^{13}C$ NMR spectroscopy (Košir and Kidrič, 2001). In 2003, the metabolite content of the skin and pulp of mature berry extracts for four wines from the Bordeaux area were analyzed (Pereira *et al.*, 2005). Differences readily observed were due to an absence of polyphenol content in pulp extracts (Fig. 4.20). PCA highlighted that sugars were mainly responsible for cluster separation among samples in both types of extracts.

The quality and metabolite content of grape berries are strongly influenced by microclimate conditions: the amount and distribution of light affects the photosynthetic capacity of the whole plant acting directly on enzyme activity. In another paper, the same research group (Pereira *et al.*, 2006) monitored the metabolite profile changes that occurred in grape berries upon sun and shadow exposure. Skin and pulp extracts of 60 samples for 20 berries (22 exposed and 38 shaded) collected at the mature stage in vineyards close to Bordeaux were analyzed by $^{1}H$ NMR spectroscopy and chemometrics. Amino acids, aromatic compounds, sugars, and organic acids were quantified and analyzed with PLS-DA in order to differentiate the effect of light on the samples. The most significant effect of light exposure was found to be the flavonols content, while amino acids were affected by the temperature changes caused by light exposure. In 2006, commercial wine fermentation was monitored by $^{1}H$ NMR spectroscopy (Clark *et al.*, 2006). The lactic acid concentration was found to increase, while those of malic and tartaric acids decreased, during the fermentation period. Most likely, this occurred because the ethanol concentration increased, thus forcing the precipitation of the latter two from the solution. Succinic and acetic acids, as well as methanol, rose to a plateau value. The same group showed in a recent paper how the use of two-dimensional correlated experiments could be a useful tool in the metabolite concentration determination during a red wine fermentation process (Kirwan *et al.*, 2008). The time course evolution of five types of red wine from Spain during both alcoholic and malolactic fermentation (MLF) was monitored by qHNMR spectroscopy (Avenoza *et al.*, 2006).

The use of this quantitative NMR technique, allowed identification, quantification, and characterization of bioactive natural compounds, while the use of a regression method, like PLS, was proposed for the quantification of partially overlapped NMR signals (Fig. 4.21).

This research group applied the same approach to another red wine type, Rioja (López-Rituerto *et al.*, 2009). In this study, PCA, performed on the entire fermentation time course of 207 days, demonstrated increases in ethanol, succinic, lactic, and acetic acids, while the alanine and malic acid concentrations decreased. Metabolite changes occurring during alcoholic fermentation were evaluated by performing PCA of the first 7 days of the

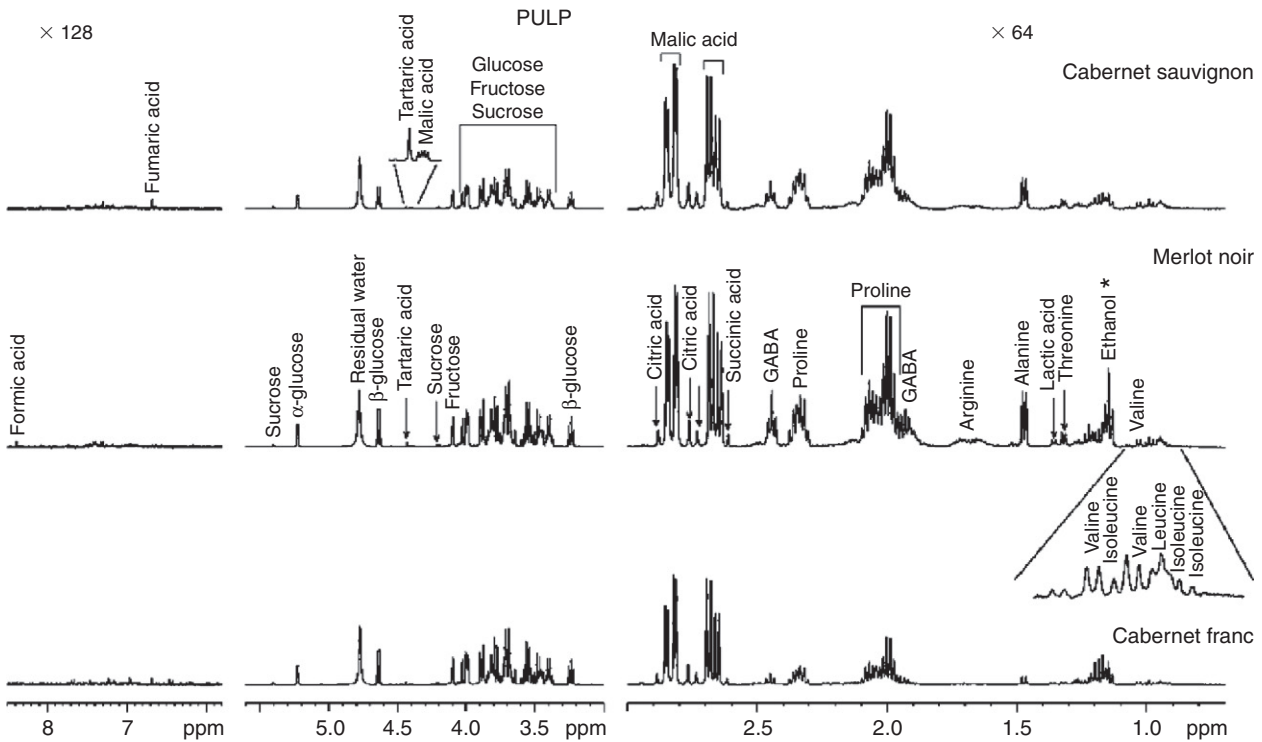



**FIGURE 4.20** $^{1}$H NMR spectra of freeze-dried pulp extracts of cv. Merlot noir, Cabernet-Sauvignon, and Cabernet Franc, acquired at 500 MHz with 64 scans and an acquisition time of 29 min. (From Pereira *et al.*, 2005.)

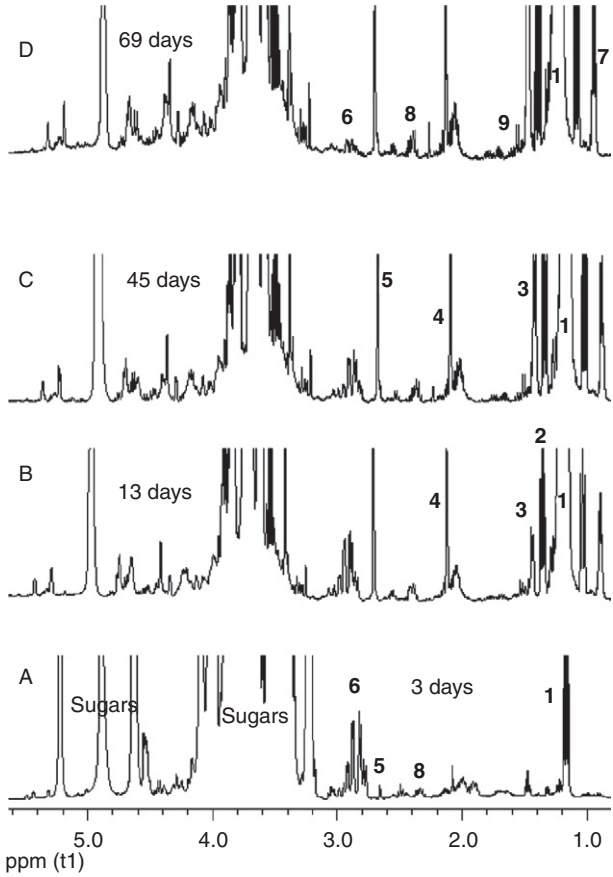



**FIGURE 4.21** $^{1}$H NMR spectra (400 MHz) of time course evolution of red wine in alcoholic and malolactic fermentations for grape red must (pH 3). Peaks: 1, ethanol; 2, ethanol satellites; 3, lactic acid; 4, acetic acid; 5, succinic acid; 6, malic acid; 7, 2,3-butanediol; 8, proline; 9, alanine. (From Avenoza *et al.*, 2006.)

fermentation process, which revealed an increase in ethanol, proline, succinic and lactic acids while conversely a decrease occurred in the level of alanine. Changes due to MLF could instead be evaluated during the rest of the fermentation process, which showed almost the same result as the previous PCA with the exception that the proline content was kept constant.

In another very recent paper (Son *et al.*, 2009b), the fermentative performances of yeast strains used for grape must fermentation were monitored by NMR and multivariate statistical methods. Characterization of the properties of wine yeasts is important because they affect wine quality. In this paper, the changes of metabolites in must during alcoholic

fermentation performed with different strains of *Saccharomyces cerevisiae* revealed different metabolite contents in the wines analyzed. In particular, glucose, organic, and amino acids were detected and evaluated during both fermentation and wine aging, thus highlighting the potential of the metabolite approach to address fermentation behavior and yeast strain effects. In another paper, the same group applied the same approach to the metabolite characterization of the MLF of the Korean wine Meoru (Son *et al.*, 2009c) with the aid of PCA. In this study, they found that wine fermented with *Saccharomyces bayanus* PC strain did not spontaneously induce the MLF, as had been observed in wines fermented with *S. cerevisiae* CDB and KUBY-501 strains. In addition, *S. cerevisiae* CDB was found to have the lowest fermentative behavior, resulting in a lower succinate content, while the KUBY-501 strain resulted in a lower amino acid content.

Among all possible wine components, antioxidants play an important role in health, due to their free radical trapping property. Their presence also enhances the aging of wine in the bottle with corresponding economic advantages. These compounds are phenolic, polyphenolic, and flavonoids and in general their content is quite high in wines. Taking advantage of the potential of NMR in characterizing chemical compounds, the aqueous and organic fractions of different aged white wines from Portuguese regions were analyzed. The NMR identified the presence of four different antioxidant compounds: two were tyrosol-like structures, while the other two were furan rings bonded to the anomeric carbon of sugar moieties (Oliveira *et al.*, 2008).

Another complex beverage is beer, consisting of several classes of chemical compounds like carbohydrates, amino acids, organic acids, and flavonols. Beer is essentially obtained from malted grains, hops, yeast, and water. A range of different analytical methods were applied to investigate the quality of beer. Again $^{1}$H NMR was successfully applied in combination with chemometrics to establish different quality of beer. Several different characters of beer can be due to the presence of different spices, herbs, and fruits, while the process or the methods used cause large differences among the beer samples. Ale and Lager beer samples were analyzed by means of NMR and PCA by Duarte *et al.* (2002). Differentiation between these two beer types has been achieved by considering the aromatic part of the $^{1}$H NMR spectrum. In particular, polyaromatic species, like polyphenols, have been suggested to play a role in sample differentiation. The same group developed few years later an NMR/FTIR combined method for investigating a possible correlation between type and sample composition (Duarte *et al.*, 2004). Obviously, $^{1}$H NMR provided more information regarding the chemical composition and thus sample differentiation was improved. In particular with PCA, the carbohydrate spectral region was discriminant for the differentiation of beers correlating with their composition of maltose, dextrins, and

glucose. When only the NMR aromatic region is considered, PCA was found to be particularly sensitive in distinguishing different beer types. The authors proposed their method as potentially applicable to the beer industry for a rapid (10 min) determination in quality control procedures. A further improvement of this method was proposed by the Lachenmeier group (Lachenmeier *et al.*, 2005) by using flow injection technology to change samples and speed up the data acquisition. In combination with PCA, NMR data of aromatic region could discriminate between beers made of wheat or barley malt, while a PLS model, verified through cross-validation procedure, could predict the lactic acid concentration as well as original gravity, an important legislation parameter used in Germany for beer classification.

## B. Vegetables

### 1. Potato

A relatively novel analytical tool, called SLICING, was employed to decompose the relaxation profiles obtained at LF for evaluating potato quality (Povlsen *et al.*, 2003). Potatoes from five cultivars at two storage times were evaluated by T2 measurements to investigate their structure and water distribution (Fig. 4.22). Furthermore, the relationship between NMR relaxation times on raw potatoes and sensory attributes on cooked samples was investigated by PLS models. The SLICING method decomposes the relaxation curves into four exponential components leading to a superior data analysis with respect to biexponential fitting, distribution

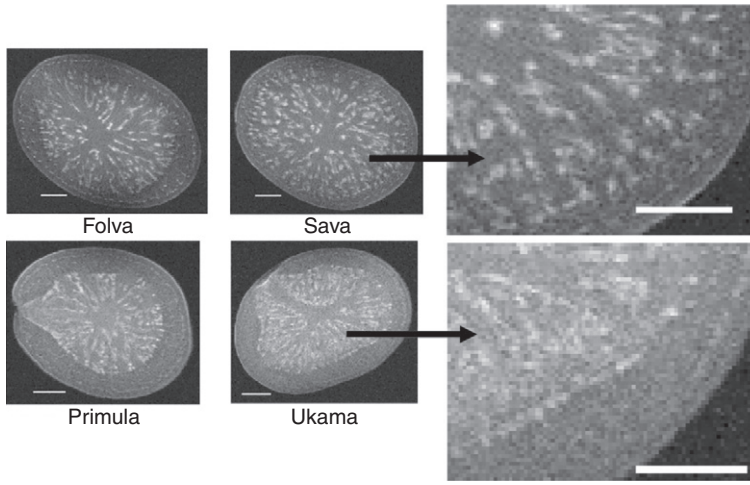



**FIGURE 4.22** Spin echo images of four raw potato varieties, scale bar: 1 cm. (From Thybo *et al.*, 2003.)

analysis, and PLS, in distinguishing the five potatoes cultivars. The prediction of six sensory attributes on the basis of NMR relaxation data was similar when PLS or SLICING algorithms were used.

The same group published in the same year another work on potatoes, by using in addition an NMR-imaging technique (Thybo *et al.*, 2003). In this work, the correlation between dry matter and NMR relaxation times (T1 and T2) and weighted images in four potato varieties were studied. PCA revealed variety differentiations based on dry matter content, showing, by PLS analysis, a high correlation between dry matter content and T2 NMR data. This correlation was very poor when NMR images were concerned. Furthermore, the authors found that the spatial distribution of water in the material determined by NMRI was not fully informative with respect to dry matter content.

A high-throughput analytical technique such as NMR has promise of adoption to investigate the metabolite content of vegetables. This information combined with multivariate statistical analysis is expected to offer a conclusive and exhaustive idea of quality discrimination and prediction with high reliability for new samples. In the field of fresh cut fruits and vegetables, demands for quality determination are increasing over the past decade.

### 2. Watermelons

Several spectroscopic and chromatographic techniques were used in order to assess the quality of watermelons, and once more, NMR provided a good tool. In a recent paper, Tarachiwin *et al.* (2008) investigated the metabolite content of watermelons (Fig. 4.23). The chemical constituents of seven graded watermelons grafted on gourd and three grafted on pumpkin were identified and differentiated by a PLS-DA protocol. Loadings revealed sucrose as the most characterizing variable for watermelons grafted on pumpkin compared to glucose and fructose for those grafted on gourd. These data were also correlated with sensory evaluation, largely influenced by the sugar content. Furthermore, the predictive results were also good for comparing the central flesh tissues from tissues collected near the outer edge.

### 3. Wheat

Plant tissues, and in particular, asynchronous protein metabolism in kernels of wheat, were investigated by NMR and chemometrics to address the problem of abiotic stress. Adverse climate events influence crop growth and yield, acting on the gene expression, protein synthesis, and metabolic pathways. The use of NMR spectroscopy could easily provide information about the metabolite content of all proton-bearing compounds, either in liquid extracts or in solid samples. Winning *et al.* (2009) investigated the metabolite content of wheat by high-resolution

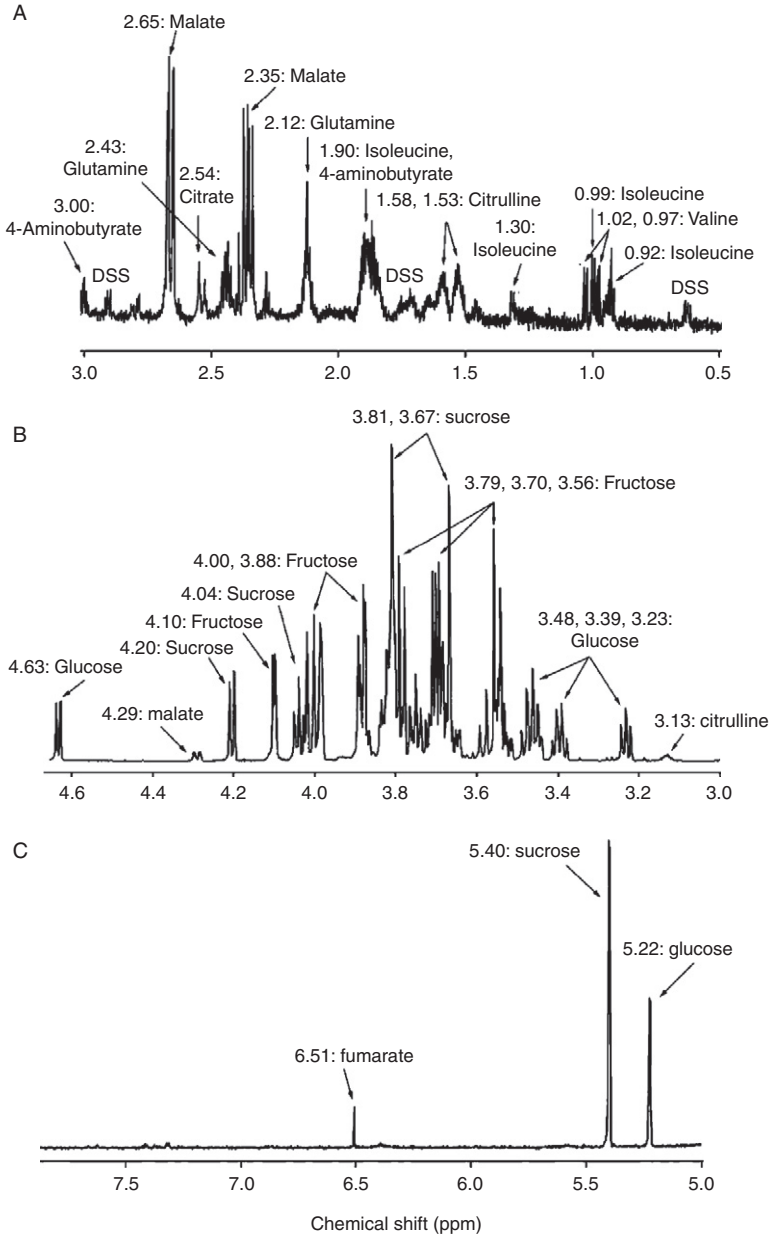



**FIGURE 4.23** $^{1}$H NMR spectra (750 MHz, $D_2O$, 25 °C) of watermelon extracted from the highest quality sample; expansions of different regions are reported. (From Tarachiwin *et al.*, 2008.)

and solid-state NMR of liquid extracts and flour, under different periods of drought stress. Proteins, carbohydrates, and amino acids were determined by MAS and HR spectroscopies in combination with protein content as determined by the standard Kjeldahl and Popov methods. PCA trajectories for the drought treatment indicated that two periods of drought will not have much effect on protein development during grain filling, while late drought did. Methanol extracts of wheat flour samples indicated fumaric acid as potential marker for drought conditions in mature kernels (Winning *et al.*, 2009).

#### 4. Rice

The metabolite content of shoots and roots for Nippombare and Arborio cultivars of *Oryza sativa* under both biotic (fungus infection) and abiotic stress (drought and salt) conditions were investigated in a recent paper by Fumagalli *et al.* (2009). Water-soluble extracts were analyzed by $^{1}$H NMR spectroscopy, while MAS was applied to intact shoots and roots tissues. The PCA of abiotic stress-treated samples revealed the role played by sugars in sample differentiation. In both cultivars, sucrose and glucose increased upon stress conditions, thus suggesting a different sugar metabolism for the two cultivars in response to stress. The PCA score plot of biotic stress-treated samples revealed a general increase in metabolite content for the Nippombare cultivar by comparison with the Arborio cultivar. The infection time with fungi (24 or 48 h) was also clearly differentiated for the two cultivars.

#### 5. Green tea

In the already mentioned paper of Le Gall *et al.* (2004), the quality of green tea was investigated for 38 high-quality teas compared to 77 other Chinese teas (lower quality samples). PCA indicated that theogallin, theanine, monir sugars, epicatechin gallate, gallic acid, caffeine, and theobromine are largely present in high-quality samples, while quinic acid, sucrose, and epigallocatechin were accumulated in the other samples, thus confirming previous determinations.

### C. Balsamic and Traditional Balsamic Vinegar of Modena

Vinegars have been extensively studied by different groups with different analytical methods during the last 10 years. In recent years, a few research groups focused their attention on Balsamic and Traditional Balsamic Vinegars of Modena (BVM and TBVM respectively), the latter being the most famous and appreciated vinegar all over the world. TBVM obtained the PDO trademark in 2000 (Reg. CE no. 812/2000, GUCE L. 100 del 20.04.2000). This product is made of cooked must which is left to age into wooden barrels, of decreasing size and different wood type, for

at least 12 years before being sold but it can reach more than 25 years, giving rise to the ''extra old'' product. The chemical modifications experienced by the cooked must during this aging process are very complicated: sugar degradation, acetylated derivatives formation, enrichment of aroma from the barrels, etc. Also for this expensive food product, several different frauds, essentially set rules violations, have been encountered such as falsification of aging process and sugar and must addition. Again for this product, the analytical control is often based on sensory analysis and simple chemical–physical determinations, that, in our opinion, cannot fully address the fraud problem. For this reason, NMR was applied to this product in order to evaluate the possibility of a deeper characterization, also in terms of fraud detection. The primary interest was the aging determination. The first NMR study appeared in 2004 (Consonni and Gatti, 2004) showing that the quantification of five selected metabolites, being ethanol, acetic acid, malic acid, glucose, and HMF, could lead to a BVM and TVBM aging evaluation, simply by using PCA. Samples of vinegars resulted distributed in the PCA score plot in accordance to their aging process. In a further study, the same group approached relaxation analysis of BVM samples by measuring the spin-lattice relaxation time (T1) of acetic acid and β-glucose (Consonni and Cagliani, 2007). A PLS model based on these measurements combined with quantitative determination of the five selected metabolites previously discussed resulted in a very good BVM aging process determination (Fig. 4.24). In a recent paper, the same group presented a characterization of 72 BVM and TBVM samples with different aging processes (Consonni *et al.*, 2008c). A hierarchical PLS-DA model resulted in a high-predictable capability in terms of the aging process whose validation was checked on both training and test sets and further confirmed by accurate prediction of 41 unknown samples.

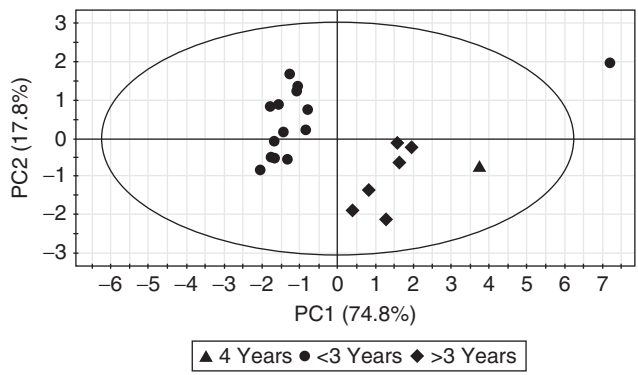



**FIGURE 4.24** Score plot of PCA for 21 samples of BVM and TBVM obtained by considering five selected compounds and T1 values for β-glucose and acetic acid.

In this model, the loading plot suggested that acetate, ethanol, and 3-hydroxy-2-butanone were the variables positively associated with ''young'' samples compared to sugars and HMF for the ''not young'' samples. This approach could be usefully employed as an analytical test for the balsamic vinegar aging process, based also on the high reproducibility of the NMR measurements. Chemical components of BVM were also presented in the paper of Caligiani *et al.* (2007) (Fig. 4.25).

A further improvement in the knowledge about BVM and TVBM arose from another paper from the Consonni group (Consonni *et al.*, 2008d). In this paper, $^{13}C$ NMR spectroscopy was successfully applied in TBVM authenticity determination. As already mentioned, this product, like other PDO foods, is produced under the application of set rules and frauds are perpetrated due to the large potential economic gain. In this latter chapter, the different isoforms of glucose and fructose were investigated. TBVM samples dissolved into water showed the ''natural ratio'' of two isoforms for glucose (α,β-pyranosidic) and three isoforms for fructose (α,β-furanosidic and β-pyranosidic). When organic solvent was used, a differentiation in the isoform ratios with respect to the natural one for both glucose and fructose was detected. In particular, the fructose isoform ratio shows the preferential degradation of the most abundant β-pyranosidic isoform. This selective fructose isoform degradation is due to sugar degradation that took place during the must cooking process. Three different factors suggested the presence of a fraud: among them, the analysis of the chemical shift deviations of C2 and C3 fructose carbon atoms of α- and β-furanosidic isoforms respectively with respect to reference values of highly trusted samples, revealed the authenticity of TBVM samples. This procedure has been patented (*MI2007 A001489*) and was successfully verified with some fraud samples. In a recent paper, focused on the chemical modification experienced during the vinegar maturation process, acetated sugar derivatives formation was monitored (Cirlini *et al.*, 2009).

## D. Fruit juice

Fruit juices are studied for several reasons, but mainly those applications are focused on fraud prevention as well as adulteration detection. The most common fraud is the addition of different and cheaper fruit, juice dilution, and sugar addition. In 1986, the first pioneering NMR study applied to orange and apple juices appeared in the literature (Eads and Bryant, 1986). At that time, the main problem was the water suppression peak in such matrices, but a small addition of manganese salt (relaxing agent) and the use of a spin echo sequence, resulted in a waterless spectra, avoiding baseline and phase distortions. In a following work, Martin *et al.* (1990) showed for the first time, the application of stable isotope techniques for adulteration detection in fruit juices. The method was developed

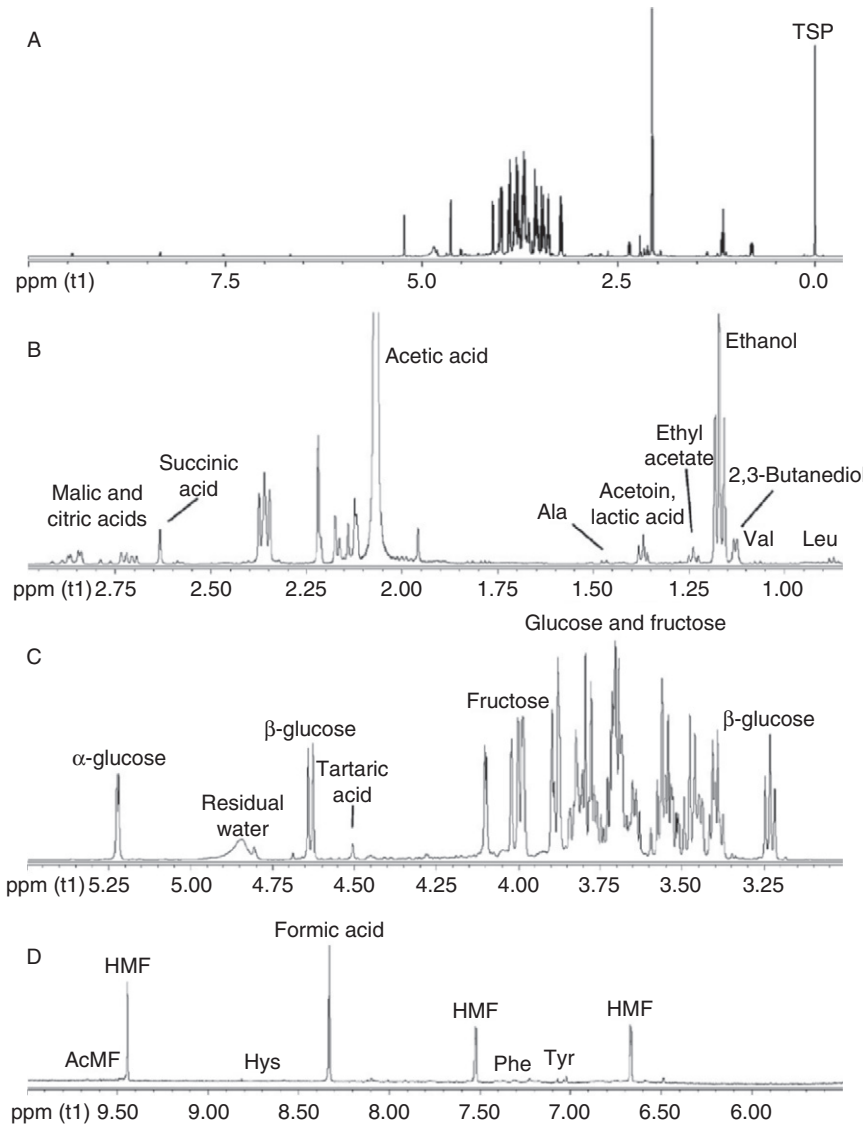



**FIGURE 4.25** $^{1}$H NMR spectrum of a BVM recorded at 600 MHz with different expansions and assignments. (Caligiani *et al.*, 2007.)

1 year later (Martin and Martin, 1991) under the name of SNIF-NMR and successively applied widely to the authenticity determination of several foods. A new screening method for fraud detection was proposed only 5 years later (Vogels *et al.*, 1996). In this study, the authors approached authenticity determinations by PCA. They found adulterations due to

detection of (a) naringenine or hesperidine (the principal flavonoids present in fruit peel of mandarins and oranges and grapefruits, respectively) indicative of pulp wash juices, (b) esters of butyric acid, and (c) sugars addition. The proposed NMR/PCA method, at that time, was quite promising especially for the easier sample preparation with respect to the other techniques. Later, in 1999, the same group presented an intercomparison study of $^{13}C$ content determination of organic acids and sugars in fruit juices, by using SNIF-NMR and IRMS determinations from different laboratories (Guillou *et al.*, 1999). The chemical shift difference between acids and sugars measured for pure juices and adulterated juices was large enough to detect the significant adulterations caused by the addition of organic acids. During recent years, with the development of high-field spectrometers and the use of more sophisticated multivariate statistical protocols, the use of the complete spectrum has been generally adopted for extracting all of the possible information to be used in sample characterization. In a recent paper (Le Gall *et al.*, 2001), more than 300 juice samples from different countries were analyzed with the aim of discriminating orange juices from pulp washed juices (Fig. 4.26).

A PCA model, built on a training set of samples and afterward validated with test set samples, was able to correctly predict 84 out of 88 authentic samples and 13 out of 17 pulp wash samples. The corresponding loading plot revealed that dimethylproline and an unknown compound were the most representative variables for pulp wash samples. Other not yet characterized resonances were also identified as potential markers for orange juice adulteration. In a very recent paper (Cuny *et al.*, 2008), the same group showed how ICA could be applied to NMR data in a fruit juice authentication protocol for orange/grapefruit discrimination. This ICA approach gave better results with respect to PCA because it is based on the idea of demixing spectra into a sum of pure signals instead of finding vectors indicating the direction of maximum dispersion. Moreover, they found that the supervised variable extraction would improve the determination. A standardized hyphenated protocol was proposed very recently by Humpfer *et al.* (2008). NMR analysis was proposed in conjunction with pattern recognition to (a) distinguish real juices from rediluted concentrated juices, (b) identify the geographical origin of the fruit used, (c) identify mixing of juices and concentrates, and (d) detect addition of artificial sugars. Then, a PLS model built on training set samples can predict new samples.

## E. Olive oil

The first NMR study of olive oil components appeared in 1993 by the Montedoro group (Montedoro *et al.*, 1993), in which three extracted polyphenolic compounds (one isomer of oleuropein aglycone, the dialdehydic

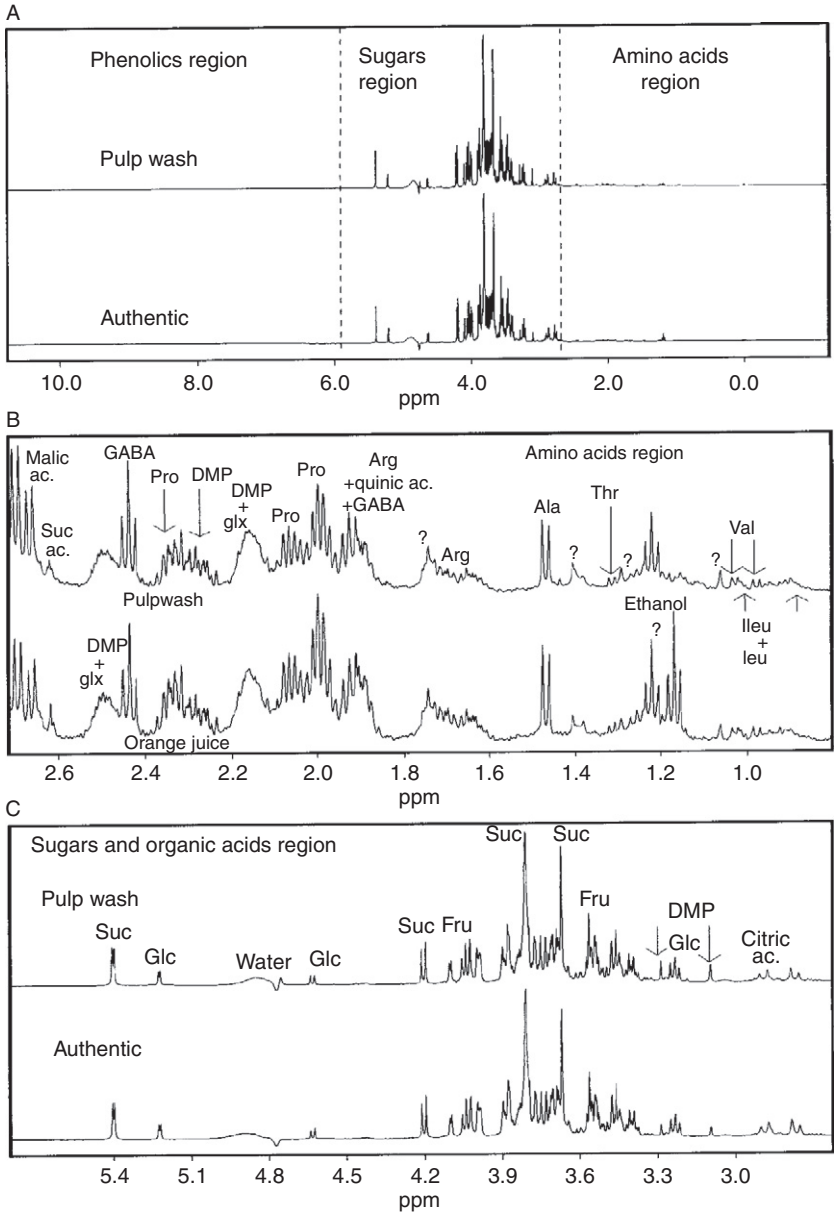



**FIGURE 4.26** (A) A typical orange juice and pulp wash samples $^{1}$H NMR spectra recorded at 500 MHz. (B) and (C) with expansions with assignments. (From Le Gall *et al.*, 2001.)

form of elenoic acid linked to 3,4-dihydroxylphenyl ethanol, and the dialdehydic form of elenoic acid linked to *p*-hydroxylphenyl ethanol) were characterized by $^{1}H$ and $^{13}C$ NMR spectroscopy. After several other studies were applied to olive oils with the underlying and growing interest around virgin olive oil, a review by Sacchi *et al.* (1997) on $^{1}H$ and $^{13}C$ NMR studies summarized the potential of NMR for virgin olive oil characterization and fraud detection. Concerning the authentication problem, like the analysis of acylglycerols and free fatty acids, the determination of *trans* fatty acids, the positional distribution of fatty acids, and the analysis of the unsaponificable fraction can be easily achieved by $^{13}C$ NMR spectroscopy. For quality assessment, the determination of free acidity, the profile of diacylglycerols, the evaluation of oxidant products, the analysis of the phenolic fraction as well as the determination of volatile compounds could be obtained by a combination of $^{1}H$ and $^{13}C$ NMR spectroscopy (Sacchi *et al.*, 1997). Later on, in 1999 $^{13}C$ NMR spectroscopy combined with multivariate analysis was applied to discriminate olive oil cultivars, like Leccino, Moraiolo, I-77, and Dritta (Vlahov *et al.*, 1999). In this approach, the integrated $^{13}C$ resonances obtained by a DEPT sequence were used to build a PLS model based on a training set and validated with test set samples. Furthermore, the PCA model indicated sample separation even though it was not always perfect. A few years later, another group (Brescia *et al.*, 2003a) approached the same aim by studying four different cultivars: Coratina, Oliarola, Leccino, and Peranzana. Also in this case, the $^{13}C$ resonances of 37 samples were integrated and by an ANOVA analysis, 67 original variables were reduced to 12, essentially due to saturated and unsaturated fatty acids. PCA and DA resulted in very good samples differentiation. Furthermore, the predictability of the DA model was evaluated by using 25 samples as a training set and 12 as a test set, reaching a predictability value of 95%. Finally, the NMR model was compared with the model based on chromatographic data, showing a very good agreement, thus indicating the capability of $^{13}C$ NMR analysis in cultivar determination. The composition of Nocellara, Biancolilla, Cerasuola, and Tonda Iblea cultivars was addressed in 2003 by the Mannina group by means of $^{13}C$ NMR spectroscopy and GC techniques combined with statistical analysis. In this chapter, the two techniques were evaluated by using different statistical methods like MANOVA, PCA, TCA, MDS, and LDA (Fig. 4.27). The obtained results for both analytical approaches were in good agreement and proved the usefulness of fatty acids in discriminating among the analyzed cultivars. In particular, 36 $^{13}C$ NMR peaks and 10 GC peaks were used as the most discriminating factors (Mannina *et al.*, 2003).

A difficult fraud determination is the addition of hazelnut oil to olive oil, and no official methods are described at present. In a relatively recent paper (García-Gonzáles *et al.*, 2004), ANN applied to $^{1}H$ and $^{13}C$ NMR

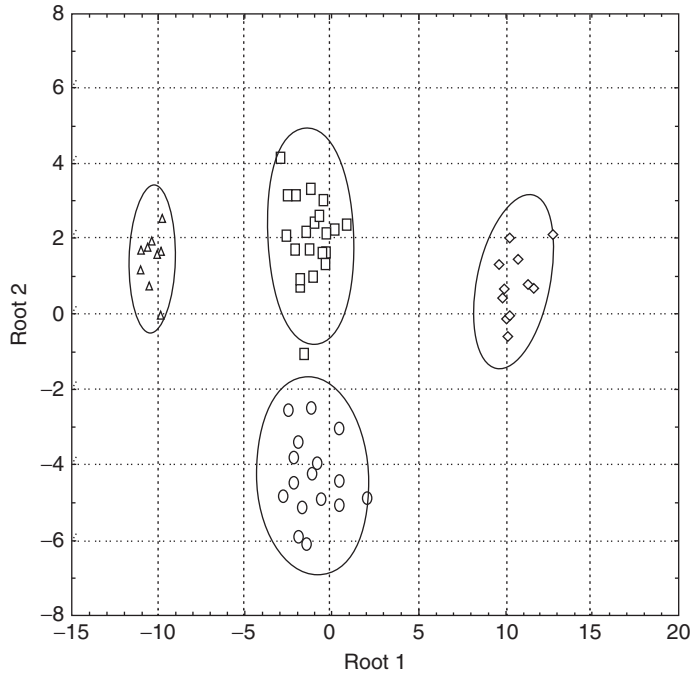



**FIGURE 4.27** LDA score plot of 60 Sicilian extra virgin olive oils based on 36 $^{13}C$ peaks. Samples labeled with the same symbol are from the same cultivar: dot, Nocellara; box, Biancolilla; diamond, Cerasuola; and triangle, Tonda Iblea. (From Mannina *et al.*, 2003.)

data were successfully proposed as methodology to detect this fraud. The $^{1}H$ data on major and minor components combined with $^{13}C$ NMR data were able to characterize vegetable oils according to the acyl positional distribution in the glycerol moiety, and were crucial in detecting hazelnut oil addition at percentages higher than 8%. In this chapter, genuine hazelnut oils, mixtures at 17–20% with olive oils, a blend of olive and hazelnut oils, genuine olive oils, and mixtures with 2–6% of hazelnut oil were analyzed and divided into three clusters. All of the blends were correctly classified with the exception of two samples adulterated at 8%.

$^{1}H$ and $^{13}C$ nuclei are not the only nuclei that can be informative in the study of olive oil, and in particular in fraud detection. Interestingly, a $^{31}P$ NMR study used for detection of adulteration obtained with addition of lampante and refined oils to EVOO was presented by Fragaki *et al.* (2005). In this study, the authors addressed the adulteration process by monitoring the phosphorylated fatty acids obtained by chemical derivatization and in particular (a) the 1,2- and 1,3-diacylglycerols, (b) total diacylglycerols, (c) the ratio between 1,2 diacylglycerols and total

diacylglycerols, (d) total free sterols, and (e) free acidity. By using one-way ANOVA, five selected variables were considered from the $^{31}$P NMR spectra. Fifty-nine samples of olive oils of different degrees and varieties from Greece were then classified by DA. Furthermore, artificial mixtures were detected with an adulteration degree as low as 5% for refined and lampante oils and test samples of real mixtures were also checked.

## F. Fish and meat

If several nuclei could be observed in high-resolution NMR techniques to monitor similarities or differences in both chemical shifts or integrals, other parameters can be monitored by using LF $^{1}$H NMR. In this case, relaxation parameters are usually measured as intrinsic discriminating values. As pointed out in several studies, T2 relaxation decay has a multi-exponential decay in both muscles and fish tissues. This suggests the presence of different ''pools'' in tissues and water distribution was assumed to be present in three distinct compartments, namely (a) ''bound water,'' (b) ''entrapped water,'' and (c) ''free water.'' In those three pools, water acts with different relaxation times because it can be bound to proteins, involved in the conversion of muscle to meat and entrapped by weak surface forces, showing relaxation values in the range of 1–10, 10–100, and 100–400 ms, respectively.

The LF was applied by Jepsen *et al.* (1999) for determination of the oil and water content of salmon and for determining the water holding capacity (WHC) and water content in cod. NMR relaxation data were analyzed by PCA and PLS models, validated against reference quality parameters, and biexponential fitting of data was adopted. Fresh and frozen cod was used, while only fresh salmon was investigated with both T2 and T1 determination. These latter data seemed not suitable for calibration in both the PCA and PLS models. The results demonstrated nevertheless the capability of LF combined to multivariate analysis as rapid, noninvasive analytical technique for the fish industry, giving rise to precise WHC determinations, and for water and oil determinations, especially when compared to other previous spectroscopic determinations (NIR). In a recent paper (Aursand *et al.*, 2008), changes in water distribution in lean (Atlantic cod) and fatty (Atlantic salmon) fishes during salting with different brines were addressed by means of T2 relaxation time measured by LF NMR spectroscopy. NMR data were processed with PCA, continuous distribution analysis, and biexponential fitting. The authors observed an increase in water mobility when salting in 15% brines was used and in general, this increase occurred more rapidly in lean than in fatty fishes, most likely due to the different fat content, which is known to act as a diffusion inhibitor. Interestingly, the comparison of relaxation times with physicochemical data showed a good correlation.

In another paper (Brown *et al.*, 2000), the close correlation of water and technological parameters in pork meat was investigated. WHC is a primary parameter also for meat quality. The sensory characteristics for storage and processing are strongly influenced by the bound water. ANOVA analysis of T2 relaxation times measured at LF indicated the inadequacy of a discrete model to explain all of the features observed, while a continuous model is more appropriate. In fact, the discrete model revealed the presence of three relaxation components, indicating three types of water, spanning T2 values from 20 up to 260 ms, while the continuous model indicated a large number of kinds of water that cluster into two broad populations.

## G. Dairy products

### 1. Milk

Milk quality plays a key role not only for its wide consumption but also because milk is involved in a large series of derived products, such as yogurt, butter, cheese, and ice cream. In addition, some of these derived products are becoming PDO and PGI products, thus implying very high quality levels for these foods. In this respect, for example, an important determination could be to assess the different composition between cow and buffalo milk, the latter used for the PDO ''buffalo mozzarella.'' The lipid fraction of both cow and buffalo milk collected in different areas of Apulia region, in the south of Italy, was extensively analyzed by the use of $^{13}C$, $^{1}H$, COSY, HMBC, and HMQC experiments that allowed the complete assignment of the TAG fraction (Brescia *et al.*, 2004). The FA integrated proton resonances were evaluated with PCA and HCA, while a further DA was performed revealing the possibility of distinguishing between the two milk species.

Recently, dairy products were reviewed in terms of analytical methods and chemometric tools (Karoui and De Baerdemaeker, 2007) for assessing both quality and authenticity. In this review, spectroscopic techniques such as IR (mid and far), front face fluorescence spectroscopy (FFFS), and NMR combined with statistical analysis tools were summarized and discussed in terms of advantages against traditional techniques in accordance with the food type. Transversal relaxation rates were evaluated in investigating the structural state of milk powder and its reconstitution process in water. Davenel *et al.* (2002) showed that the poor tendency to reconstruct the native cow phosphocasein micelles, concentrated by tangential microfiltration and powdered by spray-drying could be significantly improved by the addition of whey proteins, polydextroses, or NaCl before spray-drying of the powder, without significantly affecting their micellar structure. Furthermore, modification of the casein structure upon addition of citrate or phosphate solutions to the retentate

was underlined by a decrease in relaxation rate, while addition of the Ca salts strongly disturbed the micellar organization and led to the formation of insoluble structures during spray-drying. Recently, the fermentation of soymilk by lactic acid bacteria was investigated from the metabolite point of view (Yang *et al.*, 2009). Fermented soymilk enriched the aglycone levels and changed the isoflavone profile. The content of phenolics and flavonoids was thus evaluated by means of NMR data and PCA. Free radical scavenging activities were also quantified. The authors found a higher free radical scavenging activity after 12 h of fermentation, confirmed by the higher content of phenolics and flavonoids determined by using the metabolomic approach.

## 2. Cheese

Different types of cheeses have been subjected to NMR investigations with different aims. Grana-type cheeses are considerable commercially important products and have been characterized by both solid and solution state (HR-MAS and HR-NMR) NMR spectroscopies. These traditional cheeses are essentially Grana Padano and Parmigiano Reggiano; both are PDO products and are highly appreciated all around the world. Parmigiano Reggiano cheese is a particularly rich food in terms of both organoleptic and nutritional aspects. It experiences different ripening stages, starting from 14, through 24 and up to 30 or more months.

The ripening process has been largely studied with different techniques, focused mainly on the evolution of proteins, but also by measuring the water-soluble composition by NMR. In a recent study, the ripening of Parmigiano Reggiano was investigated by monitoring the different water-soluble components of cheese samples after 14, 24, and 30 months of ripening (Consonni and Cagliani, 2008b). The OPLS-DA performed on $^{1}$H NMR data led to a clear differentiation of samples corresponding to their aging (Fig. 4.28). In particular, the authors observed an increase of threonine and a decrease of leucine during the ripening. Their data were in agreement with previous results of the Resmini group (Resmini *et al.*, 1988) and confirmed also by the Shintu group (Shintu and Caldarelli, 2005; Shintu *et al.*, 2004) on solid samples. As a matter of fact, the catabolism of the branched amino acids (leucine, isoleucine, and valine) degrades them into volatile compounds, adding important flavor attributes to the older cheese. The NMR amino acid's profile of the Grana Padano cheese ripening was investigated (De Angelis Curtis *et al.*, 2000) in an aqueous extract, and showed almost the same composition as Parmigiano Reggiano cheese, with the exception of citrulline, which was not present in Grana Padano cheese (Fig. 4.29). An increase of serine, alanine, methionine, and phenylalanine was observed as a function of the ripening stage, together with a decrease of glutamate, leucine, and valine.

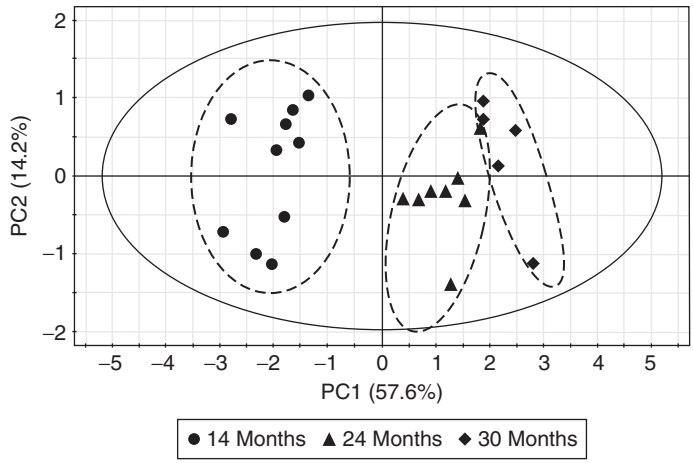



**FIGURE 4.28** O-PLS score plot performed by considering 23 Italian samples of different ripening stages of Parmigiano Reggiano. (From Consonni and Cagliani, 2008b.)

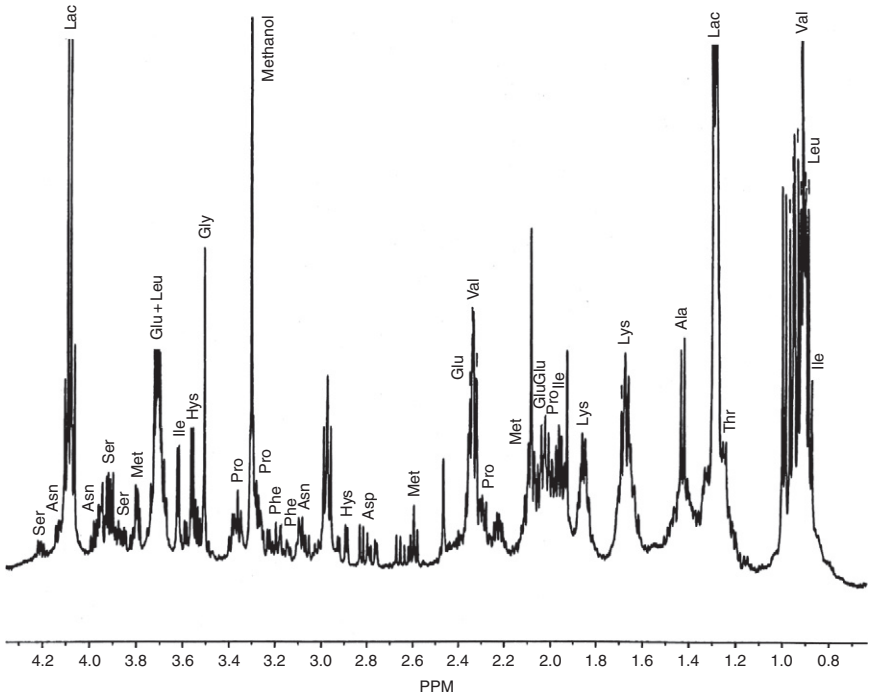



**FIGURE 4.29** Aliphatic region of $^{1}$H NMR spectrum of Grana Padano cheese in aqueous extract at 12 months of ripening recorded at AM 500 spectrometer. (From De Angelis Curtis *et al.*, 2000.)

The chemical composition of the Italian soft cheese ''robiola'' has been investigated during its degradation process over time, by means of HR and MAS spectroscopies (Lamanna *et al.*, 2008). In this study, statistically significant differences in the metabolite content were detected when a storage time of longer than 1 day was used. In particular, quantitative determination of 15 NMR signals corresponding to relative metabolites was evaluated during the degradation process either ''in air'' or ''in package'' conditions, and those compounds corresponding to amino acids (leucine, alanine, valine, isoleucine, creatine, choline, and phosphoryl choline), organic acids (GABA, malic, acetic, and lactic), glucose, and three unknown compounds. The authors observed a decrease in valine, alanine, GABA, malic acid, and the unknown compounds during the ''in air'' degradation time, while only isoleucine, valine, and GABA decreased during the ''in the package'' degradation time. An increase in all other metabolites was observed during both degradation processes.

Another popular PDO Italian cheese, ''Asiago,'' has been investigated to distinguish the products of alpine farms from lowland and mountain industrialized factory products. In this work (Schievano *et al.*, 2008), chloroform extract samples, obtained from different farms, were analyzed by using $^{1}H$ and $^{13}C$ NMR spectroscopy. By carrying out a PCA protocol on the integrated signals, a differentiation among the different cheeses was achieved mainly on the basis of the composition of fatty acids.

## H. Honey

The importance of the nutritional and therapeutic effects of honey increased its cost and consumption in recent years. As a consequence of the increased market demand, the adulteration of honey, especially in botanical origin, increased as well. Honey is essentially a complex sugar mixture of variable composition based on different brands; minor components are amino acids and aromatic compounds differently distributed according to the different varieties. The floral origin of honey is actually determined by melissopalynological analysis and organoleptic characteristics. NMR was initially introduced to this food with the aim of quality characterization. The first NMR study applied $^{13}C$ NMR spectroscopy for detection of minor components and tutin, a poison found in New Zeeland honey, due to the presence of bees on the tutu plants (Blunt *et al.*, 1979). In 1995, another $^{13}C$ NMR study (Mazzoni *et al.*, 1995) presented a quantitative determination of the sugar content applied on artificial sugar mixtures and on different varieties of Corsican honey. In this chapter, the authors focused their attention on the opportunity to record good carbon spectra and quantitative measurements. Several sugars (up to eight carbohydrates) and their isoforms were detected in the authentic water/honey solution, including mono-, di-, and trisaccharides. In particular, in addition

to fructose and glucose, maltulose, turanose, maltose, isomaltose, nigerose, melezitose, and isomaltotriose as well as the fructose/glucose ratio (important parameter to evaluate the crystallization process) were measured. Other studies were presented, mainly focused on NMR isotopic methods for authenticity determination (Cotte *et al.*, 2007; Giraudon *et al.*, 2000; Lindner *et al.*, 1996). The use of selective TOCSY NMR experiment was suggested for quantification of the minor components of honey, like amino acids (Sandusky and Raftery, 2005). In particular, this technique was largely used for characterization of overlapped signals and to access spin system identification by using specific coupling constants. By applying this tricky technique in a quantitative way, the authors found a reasonable correlation between floral origin and amino acids content: alanine, proline, threonine, phenylalanine, and tyrosine together with ethanol. A more recent paper investigated the floral classification of Italian honeys by the use of HMBC (heteronuclear multiple bond correlation spectroscopy) experiments combined with chemometric analysis (Lolli *et al.*, 2008). This interesting application of bidimensional heteronuclear NMR experiments was unfortunately intrinsically affected by low signal-to-noise ratio, notwithstanding use of reverse acquisition mode and organic solvent instead of water. Nevertheless, hydroxylated carbons were detected. The authors demonstrated the capability to resolve different floral origins by these NMR experiments by combining GDA and PCA. Finally, in the already mentioned paper of Consonni and Cagliani (2008a), a clear discrimination between acacia and polyfloral honeys was obtained by PCA performed on $^{1}$H NMR data (Fig. 4.30).

## I. Coffee

Coffee is a complex mixture whose flavor is often obtained by blending different coffee beans to achieve the desired organoleptic properties. The chemical composition of the major constituents of ''expresso'' coffee was investigated in recent years (Bosco *et al.*, 1999). Three different varieties of coffee were analyzed in both water and organic solvent solutions, resulting in quite complex $^{1}$H spectra, particularly in the aromatic region, with signals of caffeic, ferulic, quinic, chlorogenic acids present. More recently, quality control and authenticity of instant coffee was proposed by using $^{1}$H NMR and multivariate statistical analysis. In this approach, PCA and LDA were employed to identify different processes on the basis of 5-(hydroxymethyl)-2-furaldehyde content (Charlton *et al.*, 2002).

## J. Vanilla

Vanillin is one of the widely used flavoring compounds adopted in the food industry. Naturally, it is obtained from *Vanilla planifolia*, largely cultivated in the tropical area. Vanilla can be synthesized by different

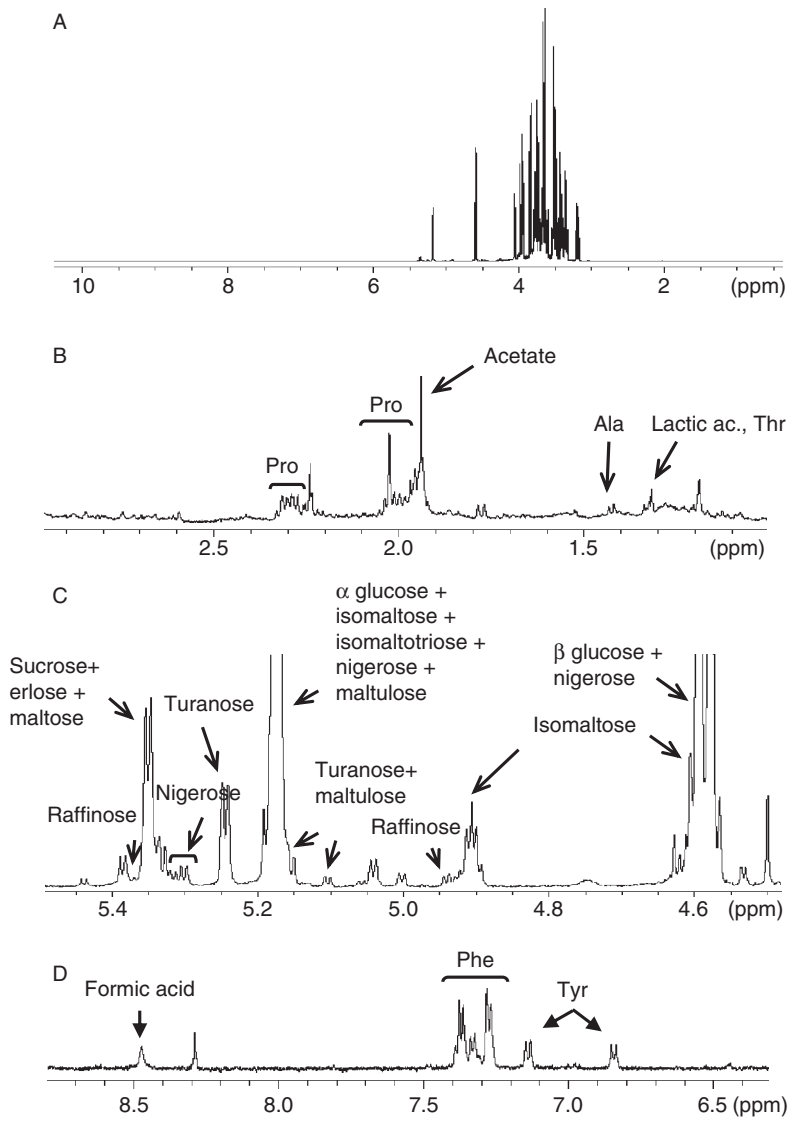



**FIGURE 4.30** (A) Complete $^{1}$H NMR spectrum (500 MHz) of polyfloral honey sample dissolved in water. Expansions of the aliphatic (B), anomeric (C), and aromatic (D) regions showing principal spin system assignments.

chemical transformations of other phenylpropanoids, like ferulic acid or lignin, or from guaiacol, or by biotechnology processes. Obviously, the natural product has the highest price and for this reason, fraudulent substitution with product of synthetic origin is often encountered.

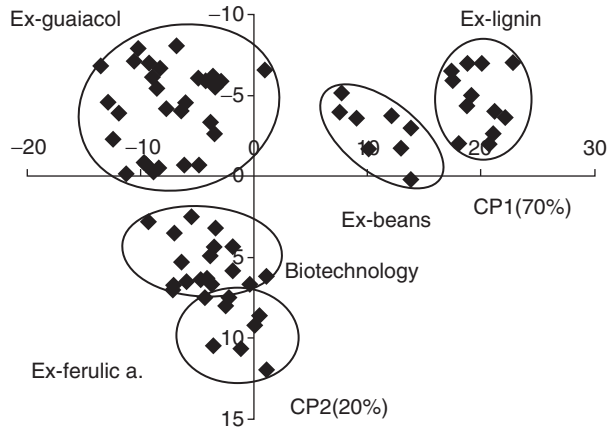



**FIGURE 4.31** Score plot of PCA performed on partial reduced molar fractions fi/Fi of sites 1 and 5–8 of vanillin calculated from $^{13}C$ NMR spectra (125 MHz) acquired with D = 21 s. The 21 samples are represented in the plane of the two main axes, and the relative weights are indicated in parentheses. (From Tenailleau *et al.*, 2004.)

The problem of fraud detection was approached in the past by using isotopic determination. The potential of IRMS to distinguish natural from guaiacol- and lignin-derived vanillin was evaluated. In this particular case, due the economic advantage of using synthetic vanillin, the fraud became very sophisticated providing specific isotopically enriched products. Then this problem was overcome by online pyrolysis, but only partially. Finally, the $^{13}C/^{12}C$ natural ratio method was proposed, showing also differences between origins, but this latter method was quite critical in terms of feasibility. Recently, the determination of site-specific $^{13}C$ abundance and PCA was proposed to estimate the vanillin origin as a rapid screen and for more refined discrimination (Tenailleau *et al.*, 2004). In this chapter, the partial reduced molar fractions of 21 vanillin samples from different origin (natural, synthetic, and biotechnology) were calculated from $^{13}C$ decoupled spectra. Only five out of eight $^{13}C$ calculated fractions were projected in the PCA plane, leading to a clear origin differentiation (Fig. 4.31). When all of the reduced molar fractions were used, again samples were separated according to their origin and the analysis of PCs indicated sites 1 and 8 as the most discriminating.

## IV. CONCLUSIONS

In this chapter, NMR and chemometrics were reviewed as applied to food quality and geographical origin determination. Different food types were deeply characterized. Alternative methods for fraud detection were also

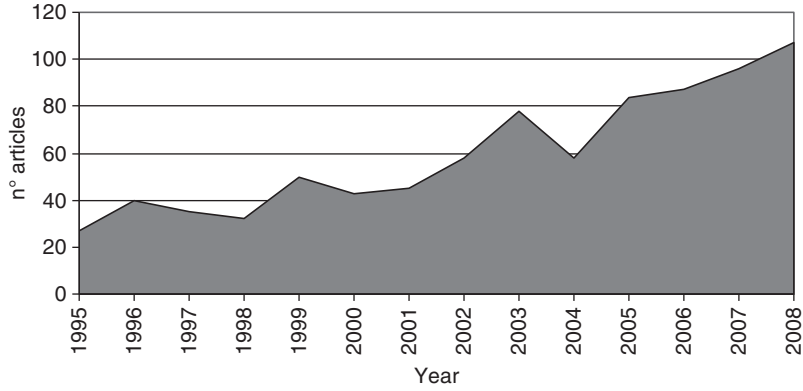



**FIGURE 4.32** Trends in published articles dealing with NMR and food from 1995 to 2008. (ISI web source, 2009.)

proposed as well as geographical origin assessments. Wine, olive oils, beer, and juices are included, being among the most attractive foods from the economic point of view, while a growing interest is only emerging for the other foods. Moreover, the NMR methods appeared mostly focused on $^{1}H$ high-resolution experiments, performed in solution, with the addition of very few bidimensional versions of correlated spectroscopy experiments with the aim of resonance assignment. Few solid-state experiments and very limited examples of LF techniques were found; notwithstanding, these preliminary investigations were quite interesting and promising. Among all possible multivariate methods, PCA is strongly preferred, most likely being an unsupervised technique used with explorative aims. Progress was quite evident over the last 15 years concerning the use of NMR spectroscopy (ISIWEB source), and it is representative of the increasing noteworthiness of this technique in food science (Fig. 4.32). As time went by, NMR coupled to chemometrics established a leading role in food research, well documented by all of the papers in the literature from the groups of different research leaders. We believe that this approach needs to be properly taken into account by controlling agencies and accredited laboratories, in order to support or even substitute some of the old fashioned analytical protocols now used in food authenticity determination.

## REFERENCES

Amiot, M. T., Fleuriet, A., and Macheix, J. T. (1986). Importance and evolution of phenolic compounds in olive during growth and maturation. *J. Agric. Food Chem.* **34**, 823–826.

Andrew, E. R., Bradbury, A., and Eades, R. G. (1958). Nuclear magnetic resonance spectra from a crystal rotated at high speed. *Nature* **182**, 1659–1660.

Aursand, I. G., Gallart-Jornet, L., Erikson, U., Axelson, D. E., and Rustad, T. (2008). Water distribution in brine salted cod (*Gadus morhua*) and salmon (*Salmo salar*): A low-field $^{1}$H NMR study. *J. Agric. Food Chem.* **56**, 6252–6260.

Aursand, M., Standal, I. B., Praël, A., McEvoy, L., Irvine, J., and Axelson, E. (2009). $^{13}$C NMR pattern recognition techniques for the classification of Atlantic salmon (*Salmo salar* L.) according to their wild, farmed, and geographical origin. *J. Agric. Food Chem.* **57**, 3444–3451.

Avenoza, A., Busto, J. H., Canal, N., and Peregrina, J. M. (2006). Time course of the evolution of malic and lactic acids in the alcoholic and malolactic fermentation of grape must by quantitative $^{1}$H NMR (qHNMR) spectroscopy. *J. Agric. Food Chem.* **54**, 4715–4720.

Bertocchi, F. and Paci, M. (2008). Applications of high-resolution solid-state NMR spectroscopy in food science. *J. Agric. Food Chem.* **56**, 9317–9327.

Blunt, J. W., Munro, M. H. G., and Swallow, W. H. (1979). C-13 NMR analysis of tutin and related substances: Application to the identification of minor components of toxic honey. *Aust. J. Chem.* **32**, 1339–1343.

Bogdanov, S. (1997). Nature and origin of the antibacterial substances in honey. *LWT Food Sci. Technol.* **30**, 748–753.

Bogdanov, S. and Martin, P. (2002). Honey authenticity. *Mitt. Geb. Lebens-mittelunters. Hyg.* **93**, 232–252.

Bosco, M., Toffanin, R., De Paolo, D., Zatti, L., and Segre, A. (1999). High-resolution $^{1}$H NMR investigation of coffee. *J. Sci. Food Agric.* **79**, 869–878.

Brescia, M. A., Caldarola, V., De Giglio, A., Benedetti, D., Fanizzi, F. P., and Sacco, A. (2002a). Characterization of the geographical origin of Italian red wines based on traditional and nuclear magnetic resonance spectrometric determinations. *Anal. Chim. Acta* **458**, 177–186.

Brescia, M. A., Di Martino, G., Fares, C., Di Fonzo, N., Platani, C., Ghelli, S., Reniero, F., and Sacco, A. (2002b). Characterization of Italian durum wheat semolina by means of chemical analytical and spectroscopic determinations. *Cereal Chem.* **79**, 238–242.

Brescia, M. A., Alviti, G., Liuzzi, V., and Sacco, A. (2003a). Chemometric classification of olive cultivars based on compositional data of oils. *JAOCS* **80**, 945–950.

Brescia, M. A., Košir, I. J., Caldarola, V., Kidrič, J., and Sacco, A. (2003b). Chemometric classification of Apulian and Slovenian wines using $^{1}$H NMR and ICP-OES together with HPICE data. *J. Agric. Food Chem.* **51**, 21–26.

Brescia, M. A., Mazzilli, V., Sgaramella, A., Ghelli, S., Fanizzi, F. P., and Sacco, R. (2004). $^{1}$H NMR characterization of milk lipids: A comparison between cow and buffalo milk. *JAOCS* **81**, 431–436.

Brescia, M. A., Monfreda, M., Buccolieri, A., and Carrino, C. (2005). Characterization of the geographical origin of buffalo milk and mozzarella cheese by means of analytical and spectroscopic determinations. *Food Chem.* **89**, 139–147.

Brescia, M. A., Sacco, D., Sgaramella, A., Pasqualone, A., Simeone, R., Peri, G., and Sacco, A. (2007). Characterization of different typical Italian breads by means of traditional, spectroscopic and image analyses. *Food Chem.* **104**, 429–438.

Brown, R. J. S., Capozzi, F., Cavani, C., Cremonini, M. A., Petracci, M., and Placucci, G. (2000). Relationship between $^{1}$H NMR relaxation data and some technological parameters of meat: A chemometric approach. *J. Magn. Reson.* **147**, 89–94.

Bylesjö, M., Rantalainen, M., Cloarac, O., Nicholson, J. K., Holmes, E., and Trygg, J. (2006). OPLS discriminant analysis: Combining the strengths of PLS-DA and SIMCA classification. *J. Chemom.* **20**, 341–351.

Caligiani, A., Acquotti, D., Palla, G., and Bocchi, V. (2007). Identification and quantification of the main organic components of vinegars by high resolution $^{1}$H NMR spectroscopy. *Anal. Chim. Acta* **585**, 110–119.

Charlton, A. J., Farrington, W. H. H., and Brereton, P. (2002). Application of $^{1}$H NMR and multivariate statistics for screening complex mixtures: Quality control and authenticity of instant coffee. *J. Agric. Food Chem.* **50**, 3098–3103.

Cirlini, M., Caligiani, A., and Palla, G. (2009). Formation of glucose and fructose acetates during maturation and ageing of balsamic vinegar. *Food Chem.* **112**, 51–56.

Clark, S., Barnett, N. W., Adams, M., Cook, I. B., Dyson, G. A., and Johnston, G. (2006). Monitoring a commercial fermentation with proton nuclear magnetic resonance spectroscopy with the aid of chemometrics. *Anal. Chim. Acta* **563**, 338–345.

Clement, A., Dorais, M., and Vernon, M. (2008). Multivariate approach to the measurement of tomato maturity and gustatory attributes and their rapid assessment by Vis-NIR spectroscopy. *J. Agric. Food Chem.* **56**, 1538–1544.

Codex Alimentarius Commission (2002). Codex standard 12 Revised Codex Standard for honey. *Stan. Methods* **11**, 1–8.

Comon, P. (1994). Independent component analysis: A new concept? *Sig. Proc.* **36**, 287–314.

Consonni, R. and Cagliani, L. R. (2007). NMR relaxation data for quality characterization of Balsamic vinegar of Modena. *Talanta* **73**, 332–339.

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

Consonni, R. and Cagliani, L. R. (2008b). Ripening and geographical characterization of Parmigiano Reggiano cheese by $^{1}$H NMR spectroscopy. *Talanta* **76**, 200–205.

Consonni, R. and Gatti, A. (2004). $^{1}$H NMR studies on Italian Balsamic and Traditional Balsamic vinegars. *J. Agric. Food Chem.* **52**, 3446–3450.

Consonni, R., Cagliani, L. R., Benevelli, F., Spraul, M., Humpfer, E., and Stocchero, M. (2008c). NMR and chemometric methods: A powerful combination for characterization of Balsamic and Traditional Balsamic vinegar of Modena. *Anal. Chim. Acta* **611**, 31–40.

Consonni, R., Cagliani, L. R., Rinaldini, S., and Incerti, A. (2008d). Analytical method for authentication of Traditional Balsamic vinegar of Modena. *Talanta* **75**, 765–769.

Consonni, R., Cagliani, L. R., Stocchero, M., and Porretta, S. (2009). Triple concentrate tomato paste: Discrimination between Italian and Chinese products. *J. Agric. Food Chem.* **57**, 4506–4513.

Cordella, C., Moussa, I., Martel, A. C., Sbirrazzuoli, N., and Lizzani-Cuvelier, L. (2002). Recent developments in food characterization and adulteration detection: Technique-oriented perspectives. *J. Agric. Food Chem.* **50**, 1751–1764.

Cotte, J. F., Casabianca, H., Lhéritier, J., Perrucchietti, C., Sanglar, C., Waton, H., and Grenier-Laustalot, M. F. (2007). Study and validity of $^{13}$C stable carbon isotope ratio analysis by mass spectrometry and $^{2}$H site-specific natural isotopic fractionation by nuclear magnetic resonance isotopic measurements to characterize and control the authenticity of honey. *Anal. Chim. Acta* **582**, 125–136.

Council Directive 2001/110/EC of 20 December 2001 relating to honey (2002). *Off. J. Eur. Commun.* **L10**, 47–52.

Cuny, M., Vigneau, E., Le Gall, G., Colquhoun, I., Lees, M., and Rutledge, D. N. (2008). Fruit juice authentication by $^{1}$H NMR spectroscopy in combination with different chemometrics tools. *Anal. Bioanal. Chem.* **390**, 419–427.

Davenel, A., Schuck, P., Mariette, F., and Brule, G. (2002). NMR relaxometry as non-invasive tool to characterize milk powders. *Lait* **82**, 465–473.

Davis, R. A., Charlton, A. J., Oehlschlager, S., and Wilson, J. C. (2006). Novel feature selection method for genetic programming using metabolomic H-1 NMR data. *Chemom. Intell. Lab.* **81**, 50–59.

Day, M. P., Zhang, B., and Martin, G. J. (1995). Determination of the geographical origin of wine using joint analysis of elemental and isotopic composition. II-Differentiation of the principal production zones in France for the 1990 vintage. *J. Sci. Food Agric.* **67**, 113–123.

De Angelis Curtis, S., Curini, R., Delfini, M., Brosio, E., D'Ascenzo, F., and Bocca, B. (2000). Amino acid profile in the ripening of Grana Padano cheese: A NMR study. *Food Chem.* **71**, 495–502.

D'Imperio, M., Mannina, L., Capitani, D., Bidet, O., Rossi, E., Bucarelli, F. M., Quaglia, G. B., and Segre, A. (2007). NMR and statistical study of olive oils from Lazio: A geographical, ecological and agronomic characterization. *Food Chem.* **105**, 1256–1267.

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

Du, Y. Y., Bai, G. Y., Zhang, X., and Liu, M. L. (2007). Classification of wines based on combination of $^{1}$H NMR spectroscopy and principal component analysis. *Chin. J. Chem.* **25**, 930–936.

Duarte, I., Barros, A., Belton, P. S., Righelato, R., Spraul, M., Humpfer, E., and Gil, A. M. (2002). High-resolution nuclear magnetic resonance spectroscopy and multivariate statistical analysis for the characterization of beer. *J. Agric. Food Chem.* **50**, 2475–2481.

Duarte, I. F., Barros, A., Almeida, C., Spraul, M., and Gil, A. M. (2004). Multivariate analysis of NMR and FTIR data as a potential tool for the quality control of beer. *J. Agric. Food Chem.* **52**, 1031–1038.

Eads, T. M. and Bryant, R. G. (1986). High-resolution proton NMR spectroscopy of milk, orange juice and apple juice with efficient suppression of water peak. *J. Agric. Food Chem.* **34**, 834–837.

European Communities (2000). Regulation 1760/2000. *Off. J. Eur. Comm. L* **204**(43), 1–10.

Falcone, P. M., Baiano, A., Conte, A., Mancini, L., Tromba, G., Zanini, F., and Del Nobile, M. A. (2006). Imaging techniques for the study of food microstructure. A review. *Adv. Food Nutr. Res.* **51**, 205–263.

Food and Agricultural Organization of the United Nations (2007). The State of World Fisheries and Aquaculture. FAO, Rome, Italy.

Fragaki, G., Spyros, A., Siragakis, G., Salivaras, E., and Dais, P. (2005). Detection of extra virgin olive oil adulteration with lampante olive oil and refined olive oil using nuclear magnetic resonance spectroscopy and multivariate statistical analysis. *J. Agric. Food Chem.* **53**, 2810–2816.

Frank, I. E. and Friedman, J. H. (1989). Classification: Oldtimers and newcomers. *J. Chemom.* **3**, 463–475.

Friedman, J. H. (1989). Regularized discriminant analysis. *J. Am. Stat. Assoc.* **84**, 165–175.

Fumagalli, E., Baldoni, E., Abbruscato, P., Genga, A., Lamanna, R., and Consonni, R. (2009). NMR techniques coupled with multivariate statistical analysis: Tools to analyze *Oryza sativa* metabolic content under stress conditions. *J. Agron. Crop Sci.* **195**, 77–88.

García-Gonzáles, D. L., Mannina, L., D'Imperio, M., Segre, A. L., and Aparico, R. (2004). Using $^{1}$H and $^{13}$C NMR techniques and artificial neural networks to detect the adulteration of olive oil with hazelnut oil. *Eur. Food Res. Technol.* **219**, 545–548.

Geladi, P. and Kowalski, B. R. (1986). Partial least-squares regression: A tutorial. *Anal. Chim. Acta* **185**, 1–17.

Giraudon, S., Danzart, M., and Merle, M. H. (2000). Deuterium nuclear magnetic resonance spectroscopy and stable carbon isotope ratio analysis/mass spectrometry of certain monofloral honeys. *J. AOAC Int.* **83**, 1401–1409.

Guillou, C., Koziet, J., Rossmann, A., and Martin, G. J. (1999). Determination of the $^{13}$C content of organic acids and sugars in fruit juices: An inter-comparison study. *Anal. Chim. Acta* **388**, 137–143.

Hernandez, C. V. and Rutledge, D. N. (1994). Characterization of cocoa masses: Low resolution pulse NMR study of the effect of geographical origin and roasting on fluidification. *Food Chem.* **49**, 83–93.

Higdon, J. V. and Frei, B. (2003). Tea catechins and polyphenols: Health effects, metabolism, and antioxidant functions. *Crit. Rev. Food Sci. Nutr.* **43**, 89–143.

Hirschfeld, T. (1980). The hyphenated methods. *Anal. Chem.* **52**, 297A–312A.

Hopfield, J. J. (1982). Neural networks and physical system with emergent collective computational abilities. *Proc. Natl. Acad. Sci. USA* **79**, 2554–2558.

Humpfer, E., Schäefer, H., Vervoort, J., Hofmann, G., Mörtter, M., Keller, S., Spraul, M., Lachenmeier, D., Rinke, P., and Duarte, I. (2008). Identification of an apple juice compound directly in the $^{1}$H-spectrum of the mixture with hyphenated techniques and push button NMR: Example fruit juice screening. *Planta Med.* **74**, 900.

Jackson, J. E. (1991). A User's Guide to Principal Components. John Wiley & Sons, New York.

Jepsen, S. M., Pedersen, H. T., and Engelsen, S. B. (1999). Application of chemometrics to low-field $^{1}$H NMR relaxation data of intact fish flesh. *J. Sci. Food Agric.* **79**, 1793–1802.

Karoui, R. and De Baerdemaeker, J. (2007). A review of the analytical methods coupled with chemometrics tools for the determination of the quality and identity of dairy products. *Food Chem.* **102**, 621–640.

Kirwan, G. M., Clark, S., Barnett, N. W., Niere, J. O., and Adams, M. J. (2008). Generalized 2D-correlation NMR analysis of a wine fermentation. *Anal. Chim. Acta* **629**, 128–135.

Košir, I. and Kidrič, J. (2001). Identification of amino acids in wines by one- and two-dimensional nuclear magnetic resonance spectroscopy. *J. Agric. Food Chem.* **49**, 50–56.

Košir, I. J. and Kidrič, J. (2002). Use of modern nuclear magnetic resonance spectroscopy in wine analysis: Determination of minor compounds. *Anal. Chim. Acta* **458**, 77–84.

Lachenmeier, D. W., Frank, W., Humpfer, E., Schäfer, H., Keller, S., Mörtter, M., and Spraul, M. (2005). Quality control of beer using high resolution nuclear magnetic spectroscopy and multivariate analysis. *Eur. Food Res. Technol.* **220**, 215–221.

Lamanna, R., Piscioneri, I., Romanelli, V., and Sharma, N. (2008). A preliminary study of soft cheese degradation in different packaging conditions by $^{1}$H NMR. *Magn. Reson. Chem.* **46**, 828–831.

Lauterbur, P. C. (1973). Image formation by induced local interactions: Examples employing nuclear magnetic resonance. *Nature* **242**, 190–191.

Le Gall, G., Puaud, M., and Colquhoun, I. J. (2001). Discrimination between orange juice and pulp wash by $^{1}$H nuclear magnetic resonance spectroscopy: Identification of marker compounds. *J. Agric. Food Chem.* **49**, 580–588.

Le Gall, G., Colquhoun, I. J., and Defernez, M. (2004). Metabolite profiling using $^{1}$H NMR spectroscopy for quality assessment of green tea, *Camellia sinensis* (L.). *J. Agric. Food Chem.* **52**, 692–700.

Lindner, P., Bermann, E., and Gamarnik, B. (1996). Characterization of honey by deuterium NMR. *J. Agric. Food Chem.* **44**, 139–140.

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

López-Rituerto, E., Cabredo, S., López, M., Avenoza, A., Busto, J. H., and Peregrina, J. M. (2009). A thorough study on the use of quantitative $^{1}$H NMR in Rioja red wine fermentation processes. *J. Agric. Food Chem.* **57**, 2112–2118.

Lowe, I. J. (1959). Free induction decay of rotating solids. *Phys. Rev. Lett.* **2**, 285–287.

Luykx, D. M. A. M. and van Ruth, S. M. (2008). An overview of analytical methods for determining the geographical origin of food products. *Food Chem.* **107**, 897–911.

Mannina, L., Patumi, M., Proietti, N., and Segre, A. L. (2001a). PDO (Protected Designation of Origin): Geographical characterization of Tuscan extra virgin olive oils using high-field H-1 NMR spectroscopy. *It. J. Food Sci.* **13**, 53–63.

Mannina, L., Patumi, M., Proietti, N., Bassi, D., and Segre, A. L. (2001b). Geographical characterization of Italian extra virgin olive oils using high-field $^{1}$H NMR spectroscopy. *J. Agric. Food Chem.* **49**, 2687–2696.

Mannina, L., Dugo, G., Salvo, F., Cicero, L., Ansanelli, G., Calcagni, C., and Segre, A. (2003). Study of the cultivar- composition relationship in Sicilian olive oils by GC, NMR, and statistical methods. *J. Agric. Food Chem.* **51**, 120–127.

Mannina, L., D'Imperio, M., Lava, R., Schievano, E., and Mammi, S. (2005). Caratterizzazione NMR e analisi statistica di oli di oliva DOP veneti. *La rivista italiana delle sostanza grasse* **LXXXII**, 59–63 March/April.

Marini, F. (2009). Artificial networks in foodstuff analyses: Trends and perspectives. A review. *Anal. Chim. Acta* **635**, 121–131.

Martin, M. L. and Martin, G. J. (1991). A site-specific and multi-element isotopic approach to origin inference of sugars in food and beverages. *Mikrochim. Acta* **2**, 81–91.

Martin, G. J., Danho, D., and Guillou, C. (1990). NMR and MS stable isotope studies of fruit juice adulteration. *Abstr. Paper Am. Chem. Soc.* **200**, 154, AGFG.

Martin, G. J., Mazure, M., Jouitteau, C., Martin, Y. L., Aguile, L., and Allain, P. (1999). Characterization of the geographic origin of Bordeaux wines by a combined use of isotopic and trace element measurements. *Am. J. Enol. Vitic.* **50**, 409–417.

Mazzoni, V., Bradesi, P., Tomi, F., and Casanova, J. (1995). Direct qualitative and quantitative analysis of carbohydrate mixtures using $^{13}C$ NMR spectroscopy: Application to honeys. *Magn. Reson. Chem.* **35**, S81–S90.

McLachland, G. J. (1992). Discriminant Analysis and Statistical Pattern Recognition. Wiley, New York.

Miller, J. C. and Miller, J. N. (1993). Statistics for Analytical Chemistry. 3rd ed. Ellis Horwood, PTR Prentice-Hall, New York.

Molan, P. C. (1998). The limitation of the methods of identifying the floral source of honeys. *Bee World* **79**, 59–68.

Montedoro, G. F. and Garofolo, L. (1984). The qualitative characteristics of virgin olive oils. The influence of variables such as variety, environment, preservation, extraction, conditioning of the finished product. *J. Agric. Food Chem.* **61**, 157–168.

Montedoro, G., Servili, M., Baldioli, M., Selvaggini, R., Miniati, E., and Macchioni, A. (1993). Simple and hydrolyzable compounds in virgin olive oil. 3. Spectroscopic characterizations of the secoiridoid derivatives. *J. Agric. Food Chem.* **41**, 2228–2234.

Ogrinc, N., Košir, I. J., Kocjančic, M., and Kidrič, J. (2001). Determination of authenticity, regional origin, and vintage of Slovenian wines using a combination of IRMS and SNIF-NMR analyses. *J. Agric. Food Chem.* **49**, 1432–1440.

Oliveira, C. M., Ferreira, A. C. S., Guedes de Pinho, P., and Silva, A. M. S. (2008). New qualitative approach in the characterization of antioxidants in white wines by antioxidant free radical scavenging and NMR techniques. *J. Agric. Food Chem.* **56**, 10326–10331.

Pereira, G. E., Gaudillere, J. P., Van Leeuwen, C., Hilbert, G., Lavialle, O., Maucourt, M., Deborde, C., Moing, A., and Rolin, D. (2005). $^{1}H$ NMR and chemometrics to characterize mature grape berries in four wine-growing areas in Bordeaux, France. *J. Agric. Food Chem.* **53**, 6382–6389.

Pereira, G. E., Gaudillere, J. P., Pieri, P., Hilbert, G., Maucourt, M., Deborde, C., Moing, A., and Rolin, D. (2006). Microclimate influence on mineral and metabolic profiles of grape berries. *J. Agric. Food Chem.* **54**, 6765–6775.

Perez, E., Rodriguez-Malaver, A. J., and Vit, P. (2006). Antioxidant capacity of Venezuelan honey in wistar rat homogenates. *J. Med. Food* **9**, 510–516.

Petrakis, P., Touris, I., Liouni, M., Zervou, M., Kyrikou, I., Kokkinofta, R., Theocharis, C. R., and Mavromoustakos, T. M. (2005). Authenticity of the traditional Cypriot spirit ''Zivania'' on the basis of $^{1}H$ NMR spectroscopy diagnostic parameters and statistical analysis. *J. Agric. Food Chem.* **53**, 5293–5303.

Petrakis, P. N., Agiomyrgianaki, A., Christophoridou, S., Spyros, A., and Dais, P. (2008). Geographical characterization of Greek virgin olive oil (Cv. Koroneiki) using $^{1}H$ and $^{31}P$ NMR fingerprinting with canonical discriminant analysis and classification binary trees. *J. Agric. Food Chem.* **56**, 3200–3207.

Povlsen, V. T., Rinnan, Å., Van Den Berg, F., Andersen, H. J., and Thybo, A. K. (2003). Direct decomposition of NMR relaxation profiles and prediction of sensory attributes of potato samples. *Lebensm. Wiss. U. Technol.* **36**, 423–432.

Remaud, G. S., Martin, Y. L., Martin, G. G., Naulet, N., and Martin, G. J. (1997). Authentication of mustard oils by combined stable isotope analysis (SNIF-NMR and IRMS). *J. Agric. Food Chem.* **45**, 1844–1848.

Resmini, P., Pellegrino, L., Hogemboom, J., and Bertuccioli, M. (1988). Gli amminoacidi liberi nel formaggio Parmigiano Reggiano stagionato in ricerca triennale sulla composizione e su alcune peculiari caratteristiche del formaggio Parmigiano Reggiano. In ''Consorzio del Formaggio Parmigiano Reggiano'', Reggio Emilia, Italy 41-57.

Rezzi, S., Axelson, D. E., Héberger, K., Reniero, F., Mariani, C., and Guillou, C. (2005). Classification of olive oils using high throughput flow $^{1}$H NMR fingerprinting with principal component analysis, linear discriminant analysis and probabilistic neural networks. *Anal. Chim. Acta* **552**, 13–24.

Rezzi, S., Giani, I., Héberger, K., Axelson, D. E., Moretti, V. M., Reniero, F., and Guillou, C. (2007). Classification of gilthead sea bream (*Spaurus aurata*) from $^{1}$H NMR lipid profiling combined with principal component and linear discriminant analysis. *J. Agric. Food Chem.* **55**, 9963–9968.

Romesburg, H. C. (1984). Cluster Analysis for Researchers. Robert E. Krieger Publishing Co., Malabar, FL.

Sacchi, R., Addeo, F., and Paolillo, L. (1997). $^{1}$H and $^{13}$C NMR of virgin olive oil. An overview. *Magn. Reson. Chem.* **35**, S133–S145.

Sacchi, R., Mannina, L., Fiordiponti, P., Barone, P., Paolillo, L., Patumi, M., and Segre, A. (1998). Characterization of Italian extra virgin olive oils using $^{1}$H-NMR spectroscopy. *J. Agric. Food Chem.* **46**, 3947–3951.

Sacco, A., Neri Bolsi, I., Massini, R., Spraul, M., Humpfer, E., and Ghelli, S. (1998). Preliminary investigation on the characterization of durum wheat flours coming from some areas of south Italy by means of $^{1}$H high resolution magic angle spinning nuclear magnetic resonance. *J. Agric. Food Chem.* **46**, 4242–4249.

Sacco, A., Brescia, M. A., Liuzzi, V., Reniero, F., Guillou, C., Ghelli, S., and van der Meer, P. (2000). Characterization of Italian olive oils based on analytical and nuclear magnetic resonance determinations. *JAOCS* **77**, 619–625.

Sacco, D., Brescia, M. A., Buccolieri, A., and Caputi Jambrenghi, A. (2005). Geographical origin and breed discrimination of Apulian lamb meat samples by means of analytical and spectroscopic determinations. *Meat Sci.* **71**, 542–548.

Sacco, D., Brescia, M. A., Sgaramella, A., Casiello, G., Buccolieri, A., Ogrinc, N., and Sacco, A. (2009). Discrimination between Southern Italy and foreign milk samples using spectroscopic and analytical data. *Food Chem.* **114**, 1559–1563.

Sandusky, P. and Raftery, D. (2005). Use of selective TOCSY NMR experiments for quantifying minor components in complex mixtures: Application to the metabolomics of amino acids in honey. *Anal. Chem.* **77**, 2455–2463.

Schievano, E., Arosio, I., Lava, R., Simionato, V., Mammi, S., and Consonni, R. (2006). Olio di oliva DOP del lago di Garda: uno studio NMR e analisi statistica multivariata. *La rivista italiana delle sostanza grasse* **LXXXIII**, 14–17.

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

Segre, A. L. and Mannina, L. (1997). $^{1}$H NMR study of edible oils. *Recent Res. Dev. Oil Chem.* **1**, 297–308.

Sequi, P., Dell'Abate, M. T., and Valentini, M. (2007). Identification of cherry tomatoes growth and origin by means of magnetic resonance imaging. *J. Sci. Food Agric.* **87**, 127–132.

Shaw, A. D., Di Camillo, A., Vlahov, G., Jones, A., Bianchi, G., Rowland, J., and Kell, D. B. (1997). Discrimination of the variety and the region of extra virgin olive oils using $^{13}C$ NMR and multivariate calibration with variable reduction. *Anal. Chim. Acta* **348**, 357–374.

Shintu, L. and Caldarelli, S. (2005). High-resolution MAS NMR and chemometrics: Characterization of the ripening of Parmigiano Reggiano cheese. *J. Agric. Food Chem.* **53**, 4026–4031.

Shintu, L. and Caldarelli, S. (2006). Toward the determination of the geographical origin of Emmental(er) cheese via high resolution MAS NMR: A preliminary investigation. *J. Agric. Food Chem.* **54**, 4148–4154.

Shintu, L., Ziarelli, F., and Caldarelli, S. (2004). Is high-resolution magic angle spinning NMR a practical speciation tool for cheese samples? Parmigiano Reggiano as a case study. *Magn. Reson. Chem.* **42**, 396–401.

Shintu, L., Caldarelli, S., and Franke, B. M. (2007). Pre-selection of potential molecular markers for the geographic origin of dried beef by HR-MAS NMR spectroscopy. *Meat Sci.* **76**, 700–707.

Solinas, M. (1987). HRGC analysis of phenolic components in virgin olive oil in relation to the ripening and the variety of olives. *Magn. Reson. Chem.* **64**, 255–262.

Son, H. S., Kim, K. M., Van Den Berg, F., Hwang, G. S., Park, W. M., Lee, C. H., and Hong, Y. S. (2008). $^{1}H$ Nuclear magnetic resonance-based metabolomic characterization of wines by grape varieties and production areas. *J. Agric. Food Chem.* **56**, 8007–8019.

Son, H. S., Hwang, G. S., Kim, K. M., Ahn, H. J., Park, W. M., Van Den Berg, F., Hong, Y. S., and Lee, C. H. (2009a). Metabolomic studies on geographical grapes and their wines using $^{1}H$ NMR analysis coupled with multivariate statistics. *J. Agric. Food Chem.* **57**, 1481–1490.

Son, H. S., Hwang, G. S., Kim, K. M., Kim, E. Y., Van Den Berg, F., Park, W. M., Lee, C. H., and Hong, Y. S. (2009b). $^{1}H$ NMR-based metabolomic approach for understanding the fermentation behaviors of wine yeast strains. *Anal. Chem.* **81**, 1137–1145.

Son, H. S., Hwang, G. S., Park, W. M., Hong, Y. S., and Lee, C. H. (2009c). Metabolomic characterization of malolactic fermentation and fermentative behaviors of wine yeast in grape wine. *J. Agric. Food Chem.* **57**, 4801–4809.

Standal, I. B., Praël, A., McEvoy, L., Axelson, D. E., and Aursand, M. (2008). Discrimination of cod liver oil according to wild/farmed and geographical origins by CG and $^{13}C$ NMR. *J. Am. Oil Chem. Soc.* **85**, 105–112.

Tarachiwin, L., Masako, O., and Fukusaki, E. (2008). Quality evaluation and prediction of *Citrullus lanatus* by $^{1}H$ NMR-based metabolomics and multivariate analysis. *J. Agric. Food Chem.* **56**, 5827–5835.

Tenailleau, E. J., Lancelin, P., Robins, R. J., and Akoka, S. (2004). Authentication of the origin of vanillin using quantitative natural abundance $^{13}C$ NMR. *J. Agric. Food Chem.* **52**, 7782–7787.

Thybo, A. K., Andersen, H. J., Karlsson, A. H., Dønstrup, S., and Stødkilde-Jørgensen, H. (2003). Low-field NMR relaxation and NMR-imaging as tools in differentiation between potato sample and determination of dry matter content in potatoes. *Lebensm. Wiss. U. Technol.* **36**, 315–322.

Trygg, J. and Wold, S. (2002). Orthogonal projections to latent structures (OPLS). *J. Chemom.* **16**, 119–128.

Viggiani, L. and Castiglione Morelli, M. A. (2008). Characterization of wines by nuclear magnetic resonance: A work study on wines from the Basilicata region in Italy. *J. Agric. Food Chem.* **56**, 8273–8279.

Vlahov, G., Shaw, A. D., and Kell, D. B. (1999). Use of $^{13}C$ nuclear magnetic resonance distortionless enhancement by polarization transfer pulse sequence and multivariate analysis to discriminate olive oil cultivars. *JAOCS* **76**, 1223–1231.

Vlahov, G., Schiavone, C., and Simone, N. (2001). Quantitative $^{13}C$ NMR method using the DEPT pulse sequence for the determination of the geographical origin (DOP) of olive oils. *Magn. Reson. Chem.* **39**, 689–695.

Vlahov, G., Del Re, P., and Simone, N. (2003). Determination of geographical origin of olive oils using $^{13}C$ nuclear magnetic resonance spectroscopy I—Classification of olive oils of the Puglia region with denomination of protected origin. *J. Agric. Food Chem.* **51**, 5612–5615.

Vogels, J. T. W. E., Terwel, L., Tas, A. C., Van Den Berg, F., Dukel, F., and Van Der Greef, J. (1996). Detection of adulteration in orange juices by a new screening method using proton NMR spectroscopy in combination with pattern recognition techniques. *J. Agric. Food Chem.* **44**, 175–180.

Wang, Y., Tang, H., Nicholson, J. K., Hylands, P. J., Sampson, J., Whitcombe, I., Stewart, C. G., Caiger, S., Oru, I., and Holmes, E. (2004). Metabolomic strategy for the classification and quality control of phytomedicine: A case study of chamomile flower (*Matricaria recutita* L.). *Planta Med.* **70**, 250–255.

Watson, D. G., Peyfoon, E., Zheng, L., Lu, D., Seidel, V., Johnston, B., Parkinson, J. A., and Fearnley, J. (2006). Application of principal components analysis to $^{1}H$-NMR data obtained from propolis samples of different geographical origin. *Phytochem. Anal.* **17**, 323–331.

Winning, H., Viereck, N., Wollenweber, B., Larsen, F. H., Jacobsen, S., Søndergaard, I., and Engelsen, S. B. (2009). Exploring abiotic stress on asynchronous protein metabolism in single kernels of wheat studied by NMR spectroscopy and chemometrics. *J. Exp. Bot.* **60**, 291–300.

Wishart, D. S. (2008). Metabolomics: Application to food science and nutrition research. *Trends Food Sci. Tech.* **19**, 482–493.

Wold, S. (1976). Pattern recognition by means of disjoint principal component models. *Patt. Recognit.* **8**, 127–139.

Wold, S., Ruhe, A., Wold, H., and Dunn, W. I. (1984). The collinearity problem in linear regression. The partial least squares approach to generalized inverses. *SIAM J. Sci. Stat. Comput.* **5**, 735–743.

Yang, S. O., Kim, S. H., Lee, J. H., Kim, Y. S., Yun, S. S., and Choi, H. K. (2009). Classification of fermented soymilk during fermentation by $^{1}H$ NMR coupled with principal component analysis and elucidation of free-radical scavenging activities. *Biosci. Biotechnol. Biochem.* **73**, 1184–1188.