LWT - Food Science and Technology 60 (2015) 1017–1024

Contents lists available at ScienceDirect

LWT - Food Science and Technology

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

Traditional balsamic vinegar and balsamic vinegar of Modena analyzed by nuclear magnetic resonance spectroscopy coupled with multivariate data analysis

Giulia Papotti [a], Davide Bertelli [a, *], Riccardo Graziosi [a], Annalisa Maietti [b], Paola Tedeschi [b], Andrea Marchetti [c], Maria Plessi [a]

[a] Dipartimento di Scienze della Vita, Università di Modena e Reggio Emilia, via Campi 183, 41125 Modena, Italy
[b] Dipartimento di Scienze Chimiche e Farmaceutiche, Università di Ferrara, via Fossato di Mortara 17/19, 44100 Ferrara, Italy
[c] Dipartimento di Scienze Chimiche e Geologiche, Università di Modena e Reggio Emilia, via Campi 183, 41125 Modena, Italy

ARTICLE INFO

Article history:
Received 12 February 2013
Received in revised form
26 November 2013
Accepted 15 October 2014
Available online 23 October 2014

Keywords:
Balsamic vinegar of Modena
Traditional balsamic vinegar of Modena
^{1}H NMR spectroscopy
PCA
GDA

ABSTRACT

Balsamic vinegar of Modena (BVM) and traditional balsamic vinegar of Modena (TBVM) are highly appreciated typical Italian products. The quality control and authentication assurance of both these balsamic vinegars are very important topics. In the recent years, the interest to develop new and standardized analytical procedures, able to further enhance the quality and commercial value of these typical and unique products and to preserve them from possible sophistications and adulterations, is increased. In this work, 76 samples of both BVM and TBVM were analyzed by ^{1}H NMR spectroscopy coupled with multivariate data analysis. The spectral data were analyzed by principal component analysis (PCA), general discriminant analysis (GDA) and classification tree analysis (CTA). The best and very promising model was obtained by a GDA which shows 98.6% of total variance explained by the first canonical function and a predictive capacity of 98.4% with a good separation between clusters. The signals of 5-HMF, α-glucopyranose, malic acid, succinic and acetic acids and the signal at 3.3 ppm were found to be the most statistically significant variables.

1. Introduction

The use of high resolution NMR in food authenticity is a subject of great interest and in recent years a widespread diffusion of this technique has been observed. In particular, the usability of this technique as a fingerprint analysis tool coupled with multivariate data analysis was widely discussed (Bertelli, Papotti, Bortolotti, Marcazzan, & Plessi, 2012; Bertelli et al., 2010; Koda, Furihata, Wei, Miyakawa, & Tanokura, 2012; Krishnan, Kruger, & Ratcliffe, 2005; Lopez-Rituerto et al., 2012; Mannina et al., 2012).

Balsamic vinegar of Modena (BVM) and traditional balsamic vinegar of Modena (TBVM) are considered typical and well known Italian products, highly appreciated all over the world. Although BVM and TBVM have some common characteristics, they are different products. TBVM is a Protected Designation of Origin (PDO) product (Reg. CE n. 813/2000 April 17, 2000), owing to its typical production procedure and the well-defined area of origin. Nowadays, local vinegar houses, often founded by small family-run business, produce the TVBM, according to the ancient methods of production, whose origins are to be found in the Modenese traditions. Long fermentation and aging procedures, which require expertise and caution in respect of the maturation state of the product, contribute to develop the unique and unmistakable characteristics that we recognize today in this very valuable product. BVM has recently obtained the registration with Protected Geographical Indication (PGI) status, granted by the European Union (Reg. CE n. 583/2009 July 3rd, 2009). They are both obtained from the alcoholic and acetic fermentation of cooked and concentrated grape musts, and this is the main characteristic that distinguishes balsamic vinegars from other vinegars, which are generally produced from alcoholic solution. TBVM and BVM mainly differ in the aging process and the production procedures. TBVM is aged in characteristic wooden barrels and may be found on the market in two different products according to the aging process: old (>12 and <25 years) and extra old (>25 years). During the production, the inoculation of colonies of acetic bacteria is allowed, while the use of any extra additive is forbidden. BVM is a cheaper product, with maturation in wooden barrels from two months up to 3 years, and it

* Corresponding author. Tel.: +39 059 2055761; fax: +39 059 2055131.
E-mail address: davide.bertelli@unimore.it (D. Bertelli).

http://dx.doi.org/10.1016/j.lwt.2014.10.042
0023-6438/

is allowed to add vinegar obtained by wine acetification (10% v/v minimum) and caramel (2% v/v maximum) for color correction (Decreto Ministeriale, 3 Dicembre 1965). Normally, the quality control and authentication assurance of both these balsamic vinegars was performed by means of sensorial analysis and by very simple chemical–physical property determinations, like total acidity, density and dry matter. Here arises the interest to develop new and standardized analytical procedures, able to further enhance the quality and commercial value of these typical and unique products and to preserve them from possible sophistications and adulterations. These new approaches, coupled with chemometric analysis, may provide helpful classification models for authentication and commercial quality characterization.

Several analytical studies focused their attention on the balsamic vinegars characterization and, some of them aim to investigate the aging process and the related changes. (Antonelli, Chinnici, & Masino, 2004; Chiavaro, Caligiani, & Palla, 1998; Cocchi, Lambertini, Mancini, Marchetti, & Ulrici, 2002; Cocchi et al., 2006; Del Signore, Stancher, & Calabrese, 2000; Gullo, Caggia, De Vero & Giudici, 2006; Plessi, Bertelli, & Miglietta, 2006; Plessi, Monzani, & Coppini, 1989; Theobald, Muller, & Anklam, 1998). However, they are often time consuming approaches, not compatible with routine analyses. Among many advantages that the ^{1}H NMR spectroscopy offers, it may simultaneously determine the different metabolites of vinegar in few minutes, as required for food authenticity and quality control. To our knowledge, only few studies have been carried out on Traditional Balsamic Vinegar of Modena (TBVM) and Balsamic Vinegar of Modena (BVM), regarding their quality evaluation and valorization, using NMR (Cirlini, Caligiani, & Palla, 2009; Consonni, Cagliani, Benevelli, et al., 2008; Consonni, Cagliani, Rinaldini, & Incerti, 2008; Consonni & Cagliani, 2007). Here a characterization of both BVM and TBVM using high-resolution ^{1}H NMR spectroscopy coupled with multivariate statistical data analysis is presented. Besides, we aim to build a discriminant model able to characterize TBVM according to the aging process, which is actually the most required information for quality assessment, and, nowadays, no objective analytical techniques have been officially defined. The particular climatic characteristics, the soil and the grape varieties typically grown in Modena strongly contribute to make the TBVM a unique and unmistakable product. These characteristics and the particular production procedures used by the local producers, which follow the Modenese traditions of cooking musts and drawing and topping up procedures among the wooden casks, make it is difficult to obtain statistical models which are representative of the intrinsic variability and peculiarities of TBVM. In this work, we applied supervised pattern recognition procedures, never used before, to a very significant number of samples, that is non-obvoius, also considering the high price of the samples.

2. Materials and methods

2.1. Materials and sample preparation

A total of 76 samples of both TBVM and BVM have been analyzed. Among them, 23 were extra old (>25 years of aging) TBVM, 17 were old (>12 and <25 years of aging) TBVM and 36 were BVM of unknown aging (Table 1). All the TBVM and several BVM were provided by local vinegar houses, while the other BVM were purchased on the market. All TBVMs and BVMs are labeled as PDO and PGI products respectively. Samples were prepared by dissolving 0.1 g exactly weighed of vinegar in 500 μL of dimethyl sulphoxide-d_6 (DMSO-d_6) (Sigma–Aldrich, Milan, Italy), and transferred into the Wilmad NMR tube (5 mm, Ultra-Imperial grade, 7 in. L, 526-PP, Sigma–Aldrich, Milan, Italy). Twenty μL of tetramethylsilane (TMS) was added as reference compound. Standard compounds for metabolite assignments were from Sigma–Aldrich (Milan, Italy).

2.2. Physical and chemical determinations

Undiluted samples were used for °Brix measures, which were carried out with refractometer. Total acidity (g/100 mL of acetic acid) was determined by the titration with sodium hydroxide 0.5 M using the method reported in the Resolution OIV-OENO 52-2000. *R ratio*, which is the rate between °Brix and Total Acidity, and indicates the balance among sweet and sour tones, correlating in this way the density with the acidity (Gullo & De Vero, 2004, pp. 93–107; Satrioni, 2010), was also calculated for TBVM samples. This parameter is often used as a tool to correctly conduct vinegar houses and, in this work, to identify possible outliers among the samples, before performing NMR analysis. All the determinations were performed in triplicate.

2.3. NMR spectroscopy

To characterize samples ^{1}H NMR, ^{13}C NMR, two-dimensional ^{1}H–^{13}C heteronuclear multiple-bond correlation (HMBC) and

Table 1
Values of °Brix and *R ratio* in Traditional Balsamic Vinegar of Modena (TBVM) and °Brix in Balsamic Vinegar of Modena (BVM) samples ($n = 3$).

TBVM samples	Type[a]	°Brix	*R ratio*[b]	BVM samples	Brix
1	Old	65 ± 0.2	10.50 ± 0.2	41	32 ± 0.1
2	Old	67 ± 0.2	12.91 ± 0.4	42	22 ± 0.2
3	Old	61.5 ± 0.3	9.70 ± 0.1	43	15.2 ± 50.1
4	Old	69.5 ± 0.1	11.94 ± 0.2	44	26.5 ± 0.2
5	Old	68.2 ± 0.1	9.50 ± 0.5	45	24 ± 0.1
6	Extra old	71 ± 0.2	10.88 ± 0.4	46	19 ± 0.1
7	Extra old	71.75 ± 0.1	11.00 ± 0.3	47	27 ± 0.2
8	Extra old	71 ± 0.3	11.14 ± 0.5	48	23.5 ± 0.1
9	Extra old	73 ± 0.2	11.09 ± 0.6	49	15.3 ± 0.1
10	Extra old	70.5 ± 0.1	12.56 ± 0.4	50	20 ± 0.4
11	Old	68.5 ± 0.2	11.01 ± 0.2	51	19.5 ± 0.2
12	Extra old	73 ± 0.4	9.92 ± 0.1	52	19 ± 0.1
13	Extra old	74 ± 0.1	12.07 ± 0.1	53	29.5 ± 0.2
14	Old	70 ± 0.1	11.59 ± 0.7	54	38.5 ± 0.1
15	Extra old	72 ± 0.1	9.28 ± 0.2	55	32 ± 0.2
16	Extra old	71 ± 0.1	11.75 ± 0.2	56	36.5 ± 0.1
17	Old	60.5 ± 0.2	8.75 ± 0.6	57	44.5 ± 0.2
18	Old	65.5 ± 0.2	9.86 ± 0.2	58	40 ± 0.1
19	Old	68 ± 0.2	10.33 ± 0.2	59	28.5 ± 0.2
20	Extra old	72 ± 0.1	12.16 ± 0.1	60	39 ± 0.2
21	Extra old	72.5 ± 0.1	10.63 ± 0.1	61	38.5 ± 0.1
22	Extra old	70 ± 0.1	10.08 ± 0.1	62	53.5 ± 0.1
23	Extra old	73 ± 0.2	10.08 ± 0.	63	38.5 ± 0.1
24	Extra old	72.5 ± 0.1	10.54 ± 0.2	64	42.5 ± 0.1
25	Extra old	71.8 ± 0.2	11.21 ± 0.3	65	36 ± 0.1
26	Old	65 ± 0.1	9.88 ± 0.2	66	18 ± 0.1
27	Old	63 ± 0.1	9.36 ± 0.1	67	36 ± 0.2
28	Old	64 ± 0.1	10.24 ± 0.2	68	31 ± 0.1
29	Old	63.5 ± 0.2	9.92 ± 0.3	69	29 ± 0.1
30	Old	65 ± 0.1	10.25 ± 0.4	70	30 ± 0.2
31	Extra old	71.5 ± 0.2	10.87 ± 0.5	71	38.5 ± 0.2
32	Extra old	71 ± 0.1	9.89 ± 0.1	72	53 ± 0.3
33	Extra old	73.5 ± 0.1	11.38 ± 0.2	73	19 ± 0.1
34	Extra old	72 ± 0.2	11.41 ± 0.2	74	38 ± 0.2
35	Extra old	71 ± 0.3	10.64 ± 0.1	75	25 ± 0.1
36	Extra old	72.5 ± 0.2	11.49 ± 0.2	76	30 ± 0.1
37	Extra old	70.5 ± 0.3	10.72 ± 0.1		
38	Extra old	71 ± 0.1	8.75 ± 0.3		
39	Old	60.8 ± 0.2	12.90 ± 0.2		
40	Old	62 ± 0.1	9.92 ± 0.1		

[a] Age is indicated only for known aging process samples. Extra old >25 years; old >12 and <25 years. For BVM the aging is unknown, however it is < 3 years.
[b] °Brix/Total acidity (g/100 ml).

$^{1}H-^{13}C$ heteronuclear single quantum coherence (HSQC) spectra were acquired with a Bruker FT-NMR Avance 400 spectrometer (Bruker Biospin GmbH Rheinstetten, Karlsruhe, Germany) operating at 400.13 MHz for ^{1}H. All of the experiments were performed at 300 K and nonspinning. ^{1}H NMR data were acquired using the Bruker spin–echo sequence "cpmg1d" (Carr–Purcell–Meiboom–Gill, Bruker Library). This sequence allows to suppress all broad signals, including the water signal, which may be removed without direct suppression, and enhances narrow resonances. Acquisition parameters were as follows: time domain (number of data points), 32K; dummy scans, 4; acquisition time, 3.4210 s; delay time, 3.0 s; number of scans, 64; spectral width, 4789.27 Hz; fidres, 0.1461. Total acquisition time was 7 min and 46 s. The assignments have been carried out on the basis of the ^{13}C NMR.

The acquisition parameters of the ^{13}C NMR experiments were as follows: number of scans, 8K; dummy scans, 4; time domain (number 142 of data points), 32K; spectral width, 22075.055 Hz; acquisition time, 0.7422 s; delay time, 1.5 s; fidres, 0.6737 Hz. Total acquisition time was 5 h, 14 min, and 59 s. The acquisition parameters of the HMBC experiments were as follows: number of scans, 32; dummy scans, 16; time domain, 3K in the acquisition or direct HMBC dimension F2 (^{1}H) and 100 in indirect HMBC dimension F1 (^{13}C); spectral width, 5592.841 Hz in F2 (^{1}H) and 20124.465 Hz in F1 (^{13}C); digital resolution, 1.8206 Hz in F2 (^{1}H) and 201.245 Hz in F1 (^{13}C); acquisition time, 0.2747 s; delay time, 0.5 s; HMBC delay time, 62.5 ms. Total acquisition time was 82 min and 11 s. The acquisition parameters of the HSQC experiments were as follows: number of scans, 4; dummy scans, 12; time domain, 1K in the acquisition or direct HSQC dimension F2 (^{1}H) and 256 in indirect HSQC dimension F1 (^{13}C); spectral width, 5995.204 in F2 (^{1}H) and 19118.721 F1 (^{13}C); digital resolution, 5.855 Hz in F2 (^{1}H) and 74.682 Hz in F1 (^{13}C); acquisition time, 0.0854 s; delay time, 1.5 s. Total acquisition time was 27 min and 50 s. The chemical shifts were reported as δ_H (ppm) relative to TMS.

2.4. Spectral preprocessing

The classification models were obtained using the ^{1}H NMR spectra as intensity. The complexity of the spectra, generated by applying the ^{1}H NMR technique, makes necessary the use of preprocessing and chemometric methods. Each spectrum generated a 16K data points corresponding to time domain, that is the number of points acquired and digitalized by the instrument along the spectral width, and then converted into a frequency domain spectrum by Fourier transform; these files were collected in a data set consisting of 16K variables and 76 samples. All ^{1}H NMR spectra were phased and calibrated using the TMS signal by the XWinNMR software package (Bruker Biospin GmbH Rheinstetten). To reduce the inhomogeneous proton NMR chemical shift, all spectra were aligned using the toolbox Icoshift 1.0 for MATLAB (Mathworks Inc., Natick, MA, USA) (Savorani, Tomasi, & Engelsen, 2009). Finally, the spectra were baseline corrected by PLS_Toolbox version 5.2.2 for use with MATLAB (eigenvector Research Inc., Wenatchee, WA, USA). All the spectral regions devoid of signals and the residual solvent (DMSO-d_6) signal (region from 2.45 to 2.55 ppm) were not considered. The resulting data set refers to the complete spectral region (12362 variables). Two other data sets have been prepared, the first one referred to the low-frequency spectral region between 0.65 and 2.70 ppm (2586 variables), which principally contains the signals of acidic and aliphatic compounds, and the second one, which contains the signals of the mid-frequency region, between 2.70 and 5.50 ppm (3499 variables).

2.5. Statistical analysis

Before the spectral analyses, all data were normalized, mean-centered, and scaled by the pareto-scaling method (Winning et al., 2009). To achieve a reliable classification, unsupervised and supervised pattern recognition procedures were applied to the data sets. Principal component analysis (PCA) was performed to verify the intrinsic variation in the data sets. Factor analysis (FA) (Burt, 1950) and general discriminant analysis (GDA) (McLachlan, 1992) were used to classify the vinegars according to their NMR fingerprint. To perform GDA, a reduction in variables with respect to complete data sets was necessary. For the complete spectral region data set the number of variables was reduced considering only the signals which presented a factorial weight during FA > |0.8|, and the resulting data set with 1453 variables and 76 samples was obtained. For the other two data sets, a less severe variables reduction was applied, by simply reducing in spectral resolution. The number of variables in fact was halved, compared to the original data sets, thus ensuring that each peak maintains its shape. After the construction of the models, to evaluate the classification performance, the leave-one out method was used as a validation procedure (Henrion & Henrion, 1994). The most significant signals resulted from GDA were integrated, using the software Amix 3.7.10 (Bruker Biospin GMBH, Rheinstetten, Germany), and used for classification tree analysis (CTA). The aim was to build a discriminant model able to fit adequately the aging process and the type of sample, in order to find a separation between the extra old TBVM and the old TBVM and identify the signals that allow such separation.

CTA is used to predict membership of cases or objects in the classes of a categorical dependent variable from their measurements on one or more predictor variables, and, in the past years, it has been successfully used in different areas of healthcare (Harper & Shahani, 2002; Harper & Winslett, 2004; Ridley et al., 1998), and food science (Bertelli, Plessi, Sabatini, Lolli, & Grillenzoni, 2007; Cirlini, Caligiani, Palla, & Palla, 2010). Three different building tree methods were applied to datasets: (a) Discriminantbased Univariate Splits (DUS); (b) Discriminant-based Linear Combination Splits (DLCS); (c) Classification & Regression Tree-style Exhaustive Search for Univariate splits (C&RT). For all of these three building methods, the FACT-style direct stopping was used as stopping rule (Loh & Vanichestakul, 1988). To estimate the prediction capacity of the models, a one-third cross-validation method was applied.

All calculations were performed using the PLS_Toolbox version 5.2.2 for Matlab, Statistica 6.1 (StatSoft® Italia, Vigonza, Italy) and SPSS 13.0 (SPSS Inc., Chicago, IL, USA).

3. Results and discussion

One of the main differences that BVM and TBVM show at macroscopic level is the sugar concentration, which directly influences the density (°Brix). As well-known (Masino, Chinnici, Franchini, Ulrici, & Antonelli, 2005), this parameter shows an increase with aging, and this is confirmed also by the density values measured in our samples, reported in Table 1. Another important parameter, which correlates density with total acidity, is *R ratio* (Gullo & De Vero, 2004, pp. 93–107; Satrioni, 2010). *R ratio* is used only for traditional balsamic vinegars (TBVM), and, for vinegars produced according to set rules, the optimal value of *R ratio* is between 7 and 10 (Gullo & De Vero, 2004, pp. 93–107). The 76 samples show *R* values close to this range, between 8.75 and 12.91 (Table 1). Although our results are slightly higher, all the TBVM samples were considered valid and used in the following analyses, since similar *R* values were obtained also by other authors (Consonni, Cagliani, & Rinaldini, et al., 2008).

A ^{1}H NMR spectrum recorded at 400 MHz of a TBVM sample, with expansions of aliphatic/alcoholic, and sugar regions, is shown in Fig. 1.

As can be observed in Fig. 1, broad signals were suppressed, including the water signal. Unfortunately also the large signals of the hydroxyl groups are suppressed; anyway the spectra are very rich in information. This sequence thus provides a sensitive means of investigating the composition of the balsamic vinegar samples object of this study. The signals were assigned on the basis of additional 2D NMR experiments, and by recording NMR spectra of pure compounds. The assignments were confirmed by comparing our results with literature data (Cirlini et al., 2009; Consonni, Cagliani, Rinaldini, et al., 2008; Consonni, Cagliani, Benevelli, et al., 2008; Consonni & Cagliani, 2007). The principal metabolites assignments are summarized in Table 2. The simple comparison of the calibrated and normalized spectra (Fig. 2) confirmed that BVM, TBVM old and TBVM extra old show different spectral characteristics, certainly due to the different production techniques, therefore to the consequent different composition. In the case of TBVMs, also the aging process may probably influence the spectral aspect of these samples. The main differences in the spectral intensities distribution are related to the signals of acetic acid, ethanol and acetoin. In TBVM, these compounds seem to be less concentrated in the majority of more aged samples as reported also by other authors. Besides, 5-HMF, fructose and glucose, 6-acetyl glucose and 2,3-butanediol were subjected to larger increase. These compounds have been described in the literature as the most significant ones for monitoring the aging process in TBVs (Caligiani, Acquotti, Palla, & Bocchi, 2007; Consonni, Cagliani, Benevelli, et al., 2008; Consonni, Cagliani, Rinaldini, et al., 2008; Cirlini et al., 2009; Masino et al., 2005; Theobald et al., 1998). When a comparison between balsamic and traditional balsamic vinegars is performed, it is necessary to specify that decreases and increases in metabolite concentrations are not closely related to the aging process, therefore, it is more cautious, and more correct, to affirm that these

Table 2
Metabolites and ^{1}H chemical shifts identified.[a]

Compound	Group	δ (ppm)	Multiplicity[b]	J (Hz)
Acetic acid	$C2H_3$	1.90	s	–
Acetoin	$C4H_3$	1.13	d	7.0
	$C1H_3$	2.08	s	–
2,3-Butanediol	$C1H_3$	0.93	d	5.9
	$C4H_3$	0.98	d	5.9
Ethanol	$C2H_3$	1.04	t	7.2
Formic acid	HCOOH	8.16	s	–
β-Fructofuranose	C1H	3.44	m	–
	C6H	3.40	m	–
β-Fructopyranose	C1H	3.32; 3.48	dd	11.2; 5.4
	C3H C4H	3.62	m	–
	C5H	3.68	m	–
	C6H	3.82	d	12.0
Glucose acetate signals[c]	CH_3CO	1.93–2.01	m	–
α–Glucopyranose	αC1H	4.89	d	3.5
	C4H	3.09	m	–
	C6H	3.62	m	–
β-Glucopyranose	βC1H	4.26	d	7.9
	C2H	2.87	t	8.4
5-HMF	C1H	9.56	s	–
	C3H	7.52	d	3.5
	C4H	6.63	d	3.3
	$C6H_2$	4.47	s	–
Lactic acid	$C3H_3$	1.19	d	6.7
Malic acid	C2H	2.38	m	–
	C2'H	2.59	m	–
	C3H	4.07	m	–
Succinic acid	$C2H_2$ $C3H_2$	2.40	s	–
Tartaric acid	C2H C3H	4.31	s	–
Valine	CγH3	0.87	d	6.8

[a] Assignments were from heteronuclear single quantum coherence (HSQC) and heteronuclear multiple-bond correlation (HMBC) experiments. The chemical shifts were expressed as relative values to those of tetramethylsilane (TMS) at 0 ppm.
[b] Peak multiplicities: s, singlet; d, doublet; t, triplet; dd, doublet of doublets; m, multiplet.
[c] Esters of glucose (6-acetylglucose) in the two anomeric forms. See Cirlini et al. for more details.

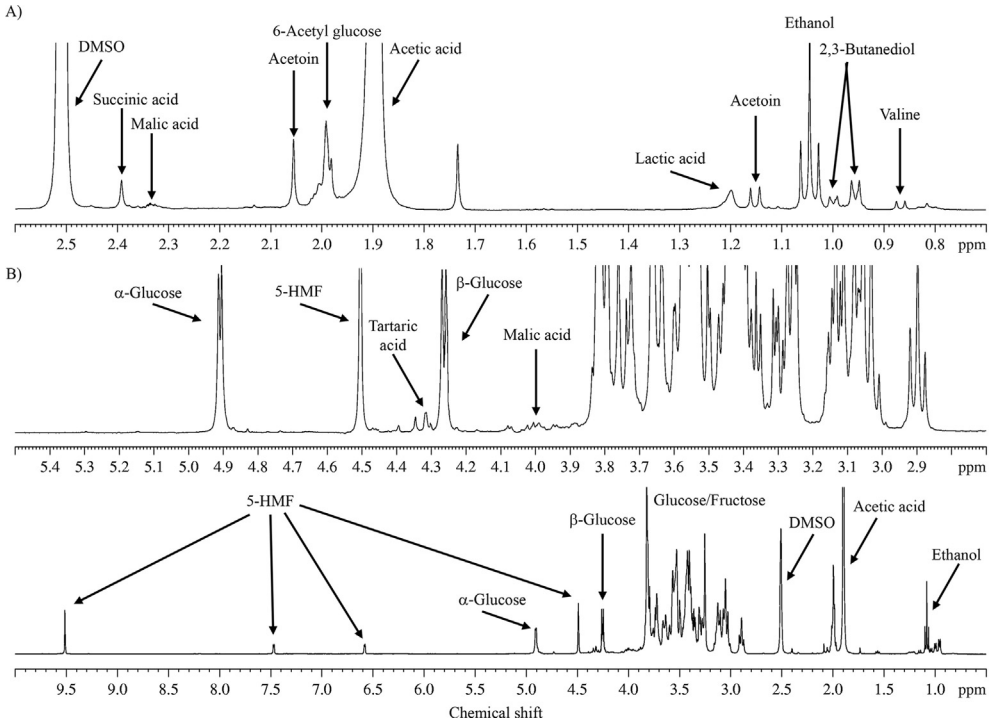

Fig. 1. Typical ^{1}H NMR spectra recorded at 400 MHz of a TBVM sample; expansions of (A) aliphatic/alcoholic, and (B) sugar regions with metabolite assignments (see Table 2).

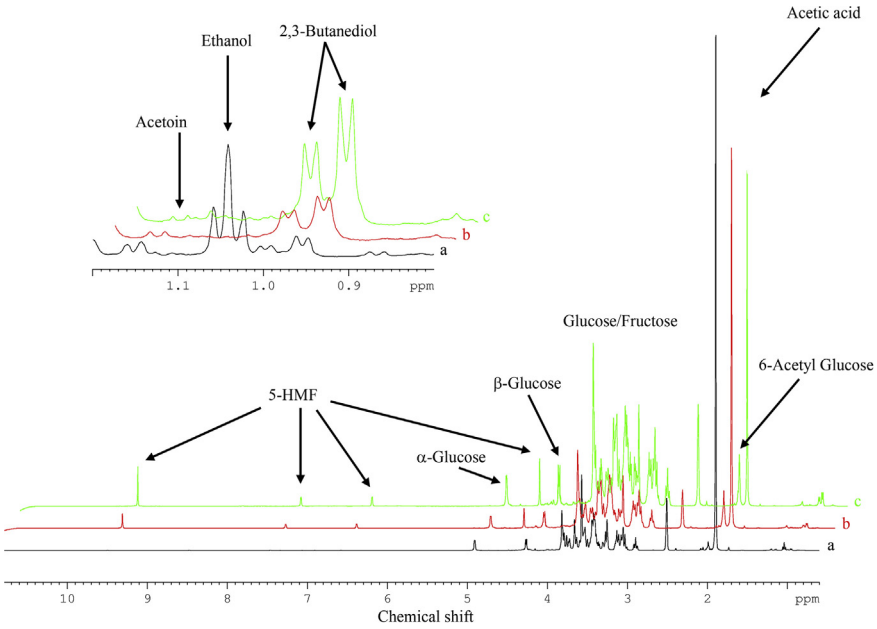

Fig. 2. Comparison of the ^{1}H NMR spectra calibrated and normalized using the TMS signal of BVM (a), TBVM old (b) and TBVM extra old (c).

differences are probably more related to the intrinsic diversity of products.

The PCA was performed on the ^{1}H NMR complete spectral region to check possible sample grouping (Fig. 3). This model resulted in 8 PCs explaining 96.91% of the total variance, and it was able to discriminate only the BVMs from TBVMs, without any differentiation between old and extra old TBVMs. The same results were obtained applying PCA on the low-frequency and mid-frequency regions. This result demonstrated the considerable complexity of the system; therefore, in order to give a clearer interpretation of PCA model, and to verify whether the application of a supervised multivariate statistical analysis was able to classify the samples, GDA was applied. The analysis was performed on complete and also on expanded spectral region data sets, since a less severe variables reduction was necessary in the aliphatic/alcoholic and sugar regions. The best results were obtained using the forward stepwise procedure. The GDA model obtained from the complete spectra was able to group the samples in three evident clusters, corresponding to the type of sample (Fig. 4). The first two canonical discriminant function (DF) explains 98.6% of the total variance, and the results of the leave-one out cross-validation show a predictive capacity of 98.4%. The first two DFs are particularly correlated with the signals of 5-HMF (C1H, C3H, C4H), α-glucopyranose (αC1H), malic acid (C3H), with a signal in the glucose and fructose region at 3.3 ppm, succinic and acetic acids. These results were confirmed by the models obtained by analyzing the expanded spectral regions, and were both able to group the vinegars in three evident clusters (score plots not shown). The first DFs of aliphatic/alcoholic and sugar regions data sets explain 98.3% and 89.7% of the total variance respectively, and the results of the leave-one out cross-validation show a predictive capacity of 90.5% and 96.8% respectively. The first two DFs of each model are particularly correlated with the signals of succinic acid, malic acid (C2H), 6-acetyl glucose and the

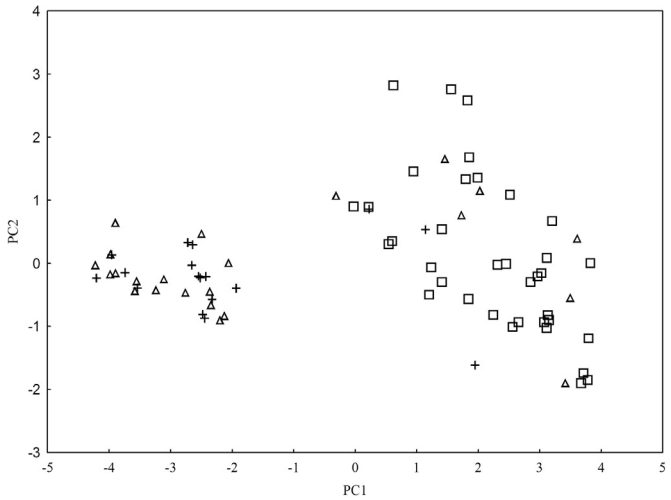

Fig. 3. PCA score plot performed on complete spectral region data set: TBVM extra old (Δ), TBVM old (+), and BVM (□).

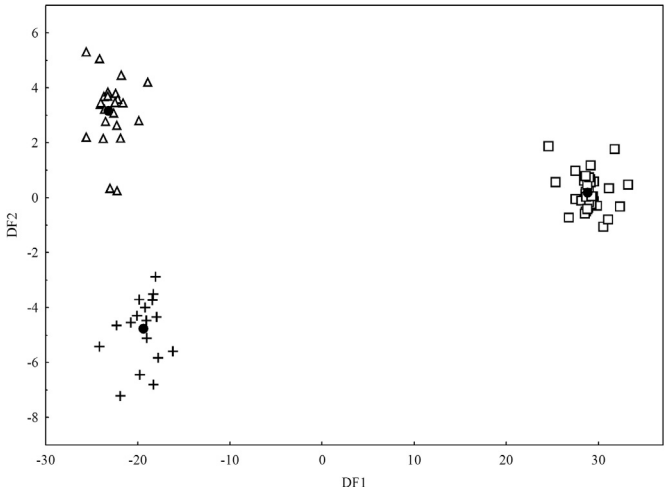

Fig. 4. Vinegar classification using GDA on complete spectral region data set, showing the separation of samples in three evident clusters: TBVM extra old (Δ), TBVM old (+), BVM (□) and group centroids (●).

Table 3
Calibration and validation results using classification tree analysis with building tree methods on the data set obtained by integrating the most significant signals resulted by general discriminant analysis.

Building tree method	All samples calibration ($n = 76$)		One-third validation ($n = 50 + 26$)				
			Calibration set		Validation set		
	Terminal nodes	Correct classification (%)	Terminal nodes	Correct classification (%)	Correct classification n (%)		
					1[a] $n = 4$	2[a] $n = 8$	3[a] $n = 14$	Total $n = 26$
DUS[b]	23	100	12	100	3 (75)	5 (62.5)	14 (100)	22 (84.6)
DLCS[c]	5	100	4	100	0 (0)	5 (62.5)	14 (100)	19 (73.1)
C&RT[d]	12	100	9	100	4 (100)	4 (50)	14 (100)	22 (84.6)

[a] 1 = Traditional Balsamic Vinegar of Modena (old) 2 = Traditional Balsamic Vinegar of Modena (extra old) 3 = Balsamic Vinegar of Modena.
[b] Discriminant-based Univariate Splits.
[c] Discriminant-based Linear Combination Splits.
[d] Classification & Regression Tree-style Exhaustive Search for Univariate splits.

unknown signal at 1.74 ppm for the aliphatic/alcoholic region data set, α-glucopyranose (αC1H), 5-HMF ($C6H_2$), tartaric acid, β-glucopyranose (βC1H), β-fructopyranose (C6H), 2 signals at 3.3 ppm and 3.03 ppm referred to the glucose and fructose region for the sugar region data set. The signals, corresponding to the most significant compounds in GDA, were integrated and used to build a new CTA discriminant model, in order to evaluate their effects on the balsamic vinegars discrimination. The obtained results are summarized in Table 3. In general, when the results of CTA were judged, the number of terminal nodes must be considered and a tree with a small number of terminal nodes must be preferred if the same capacity of classification is reached by different approaches. Considering the classification results and the number of terminal nodes, C&RT seems to be a good compromise between classification capacity and the complexity of the tree. In Fig. 5 the best tree obtained by C&RT method is reported. As evident, the model is able to discriminate between the extra old TBVM and the old TBVM and to identify the signals that allow such separation. The most interesting finding resulted from CTA is that in this model the discrimination between old TBVM and extra old TBVM is essentially due to the signals of acetic acid, 5-HMF (C4H), malic acid (C3H), β-glucopyranose (βC1H), already identify as significant ones also in the literature (Caligiani et al., 2007; Consonni, & Gatti, 2004; Consonni, Cagliani, Benevelli, et al., 2008; Consonni, Cagliani, Rinaldini, et al., 2008; Theobald et al., 1998), and to the unknown compound at 1.74 ppm.

4. Conclusion

The NMR spectroscopy, coupled with multivariate analysis, has demonstrated to be a powerful tool in BVM and TBVM characterization and quality control. The application of NMR in food characterization and control is a very powerful tool due to several good reasons. This technique is suitable to analyze samples without any manipulation, so that, in most cases, extraction and purification are unnecessary, and it is not a destructive analysis. Owing to these advantages, this technique has undergone a great development in food science, mainly concerned with the qualitative interpretation of the NMR spectra. This is particularly true for liquid food, such as wine (Papotti et al., 2013), fruit juice (Clausen, Pedersen, Bertram, & Kidmose, 2011), beverages (Lachenmeier et al., 2005; Maes, Monakhova, Kuballa, Reusch, & Lachenmeier, 2012) and vinegar

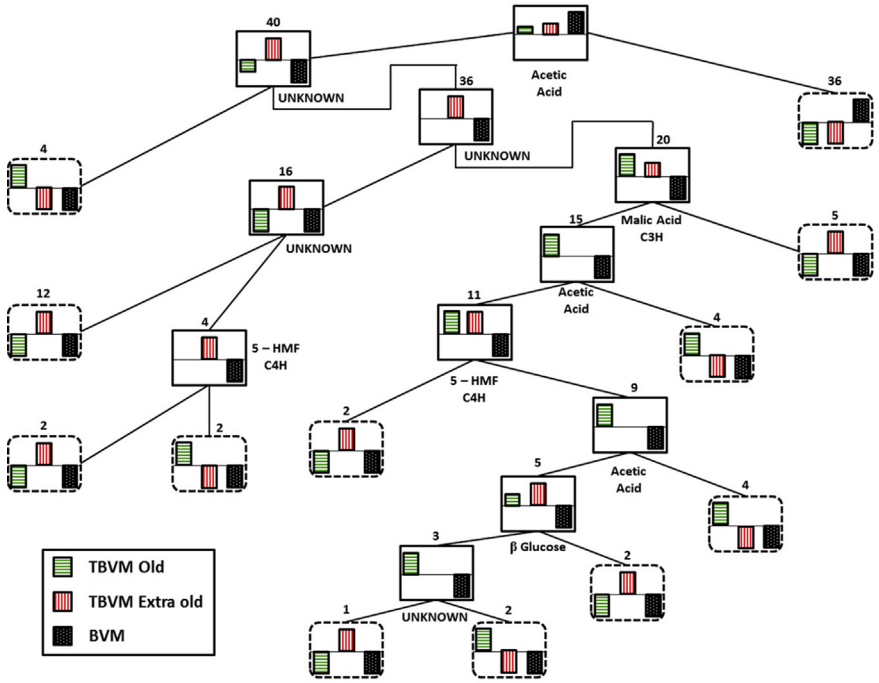

Fig. 5. Vinegar classification using CTA (C&RT-style method) on the data set obtained by integrating the most significant signals resulted by GDA.

(Boffo, Tavares, Ferreira, & Ferreira, 2009; Caligiani et al., 2007; Consonni & Gatti, 2004; Dell'Oro, Ciambotti, & Tsolakis, 2012; Thomas & Jamin, 2009; Van-Diep et al., 2011), which may be directly analyzed, after the addition of a deuterated solvent and an internal standard. Other advantages of NMR are the relatively easy and rapid acquisition of data (few minutes are required to acquire a simple ^{1}H NMR spectrum), the remarkable selectivity and identification of unknown compounds at a molecular level with high reproducibility and repeatability, the ability to furnish structural and quantitative information on a wide range of chemical species in a single NMR experiment. Besides, it is considered a useful fingerprint method for authentication analysis.

Statistical analysis showed that the signals of 5-HMF, α and β-glucopyranose, malic, succinic, tartaric and acetic acids, 6-acetyl glucose, a signal in the glucose and fructose region at 3.3 ppm were the most statistically significant variables. All these compounds are described in the literature as relevant ones for discriminating the balsamic vinegars and for monitoring the aging process. Theobald et al. (1998) established that the highest concentration of 5-HMF was found in TBVM samples. The changes in sugars and in acetic acid contents during the aging process of TBVMs are a well-known finding (Caligiani et al., 2007; Consonni, Cagliani, Benevelli, et al., 2008; Consonni, Cagliani, Rinaldini, et al., 2008). In 2004, as a part of NMR studies on Italian balsamic and traditional balsamic vinegars, Consonni and Gatti (2004) confirmed that, among several compounds, the highest content of malic acid was found for the older BVM and TBVM, while much lower values were found for the younger BVs. Cocchi, Lambertini, Mancini, Marchetti, and Ulrici (2002) proved that succinic acid increases in the young vinegars and decreases in the old ones. The hypothesis of formation of glucose acetates during maturation and aging of balsamic vinegar was verified by Cirlini et al. in 2009. The results of that work show an increase of glucose acetate during aging of traditional balsamic vinegars due to the progressive reaction between sugars and acetic acid. Moreover, the glucose acetate formation is strictly related to the initial glucose amount.

All these findings provide interesting additional knowledge on TBVMs, usable to reinforce and safeguard the wealth of Modenese traditions, represented by these unique products.

Acknowledgments

The authors express their gratitude towards the “Assessore all'Agricoltura e alla Qualità del Territorio rurale” of the province of Modena for the financial efforts provided for our research. We also express our thanks to the staff of C.I.G.S. (Centro Interdipartimentale Grandi Strumenti, Modena, Italy) for assistance during the experimental work and the Fondazione Cassa di Risparmio di Modena, for financial support given towards the purchase of the Bruker Avance 400 spectrometer.

References

Antonelli, A., Chinnici, F., & Masino, F. (2004). Heat-induced chemical modification of grape must as related to its concentration during the production of traditional balsamic vinegar: a preliminary approach. *Food Chemistry, 88*, 63–68.

Bertelli, D., Lolli, M., Papotti, G., Bortolotti, L., Serra, G., & Plessi, M. (2010). Detection of honey adulteration by sugar syrups using one-dimensional and two-dimensional high-resolution nuclear magnetic resonance. *Journal of Agricultural and Food Chemistry, 58*, 8495–8501.

Bertelli, D., Papotti, G., Bortolotti, L., Marcazzan, G. L., & Plessi, M. (2012). ^{1}H-NMR simultaneous identification of health-relevant compounds in propolis extracts. *Phytochemical Analysis, 23*, 260–266.

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

Boffo, E. F., Tavares, L. A., Ferreira, M. M. C., & Ferreira, A. G. (2009). Classification of Brazilian vinegars according to their ^{1}H NMR spectra by pattern recognition analysis. *LWT–Food Science and Technology, 42*, 1455–1460.

Burt, C. (1950). The factorial analysis of qualitative data. *British Journal of Psychology, 3*, 166–185.

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

Chiavaro, E., Caligiani, A., & Palla, G. (1998). Chiral indicators of ageing in balsamic vinegars of Modena. *Italian Journal of Food Science, 4*, 329–337.

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

Cirlini, M., Caligiani, A., Palla, L., & Palla, G. (2010). HS-SPME/GC-MS and chemometrics for the classification of balsamic vinegars of Modena of different maturation and aging. *Food Chemistry, 124*, 1678–1683.

Clausen, M. R., Pedersen, B. H., Bertram, H. C., & Kidmose, U. (2011). Quality of sour cherry juice of different clones and cultivars (*Prunus cerasus L.*) determined by a combined sensory and NMR spectroscopic approach source. *Journal of Agricultural and Food Chemistry, 59*, 12124–12130.

Cocchi, M., Durante, C., Grandi, M., Lambertini, P., Manzini, D., & Marchetti, A. (2006). Simultaneous determination of sugars and organic acids in aged vinegars and chemometric data analysis. *Talanta, 69*, 1166–1175.

Cocchi, M., Lambertini, P., Mancini, D., Marchetti, A., & Ulrici, A. (2002). Determination of carboxylic acids in vinegars and in aceto balsamico tradizionale di Modena by HPLC and GC methods. *Journal of Agricultural and Food Chemistry, 50*, 5255–5261.

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

Consonni, R., Cagliani, L. R., Benevelli, F., Spraul, M., Humpfer, E., & Stocchero, M. (2008). NMR and chemometric methods: a powerful combination for characterization of balsamic and traditional balsamic vinegar of Modena. *Analytica Chimica Acta, 611*, 31–40.

Consonni, R., Cagliani, L. R., Rinaldini, S., & Incerti, A. (2008). Analytical method for authentication of traditional balsamic vinegar of Modena. *Talanta, 75*, 765–769.

Consonni, R., & Gatti, A. (2004). ^{1}H NMR studies on italian balsamic and traditional balsamic vinegars. *Journal of Agricultural and Food Chemistry, 52*, 3446–3450.

Del Signore, A., Stancher, B., & Calabrese, M. (2000). Characterisation of balsamic vinegars by amino acid content using a multivariate statistical approach. *Italian Journal of Food Science, 12*, 317–332.

Dell'Oro, V., Ciambotti, A., & Tsolakis, C. (2012). Quality and safety of Italian wine vinegar. *Italian Journal of Food Science, 24*, 182–187.

Gullo, C., Caggia, M., De Vero, L., & Giudici, P. (2006). Characterization of acetic acid bacteria from traditional balsamic vinegar. *International Journal of Food Microbiology, 106*, 209–212.

Gullo, M., & De Vero, L. (2004). *Ricerche finalizzate alla tutela della tipicità dell'Aceto Balsamico Tradizionale di Reggio Emilia*. Reggio Emilia (Italia): Mucchi Editore.

Harper, P. R., & Shahani, A. K. (2002). Modelling for the planning and management of bed capacities in hospitals. *Journal of the Operational Research Society, 53*, 11–18.

Harper, P. R., & Winslett, D. J. (2004). Classification trees: a possible method for maternity risk grouping. *European Journal of Operational Research, 169*, 146–156.

Henrion, R., & Henrion, G. (1994). Uberwachte klassifikation. In *MultiVariate Datenanaly* (pp. 71–73). Berlin (Germany): Springer-Verlag.

Koda, M., Furihata, K., Wei, F., Miyakawa, T., & Tanokura, M. (2012). NMR-based metabolic profiling of rice wines by F-2-selective total correlation spectra. *Journal of Agricultural and Food Chemistry, 60*, 4818–4825.

Krishnan, P., Kruger, N. J., & Ratcliffe, R. G. (2005). Metabolite fingerprinting and profiling in plants using NMR. *Journal of Experimental Botany, 56*, 255–265.

Lachenmeier, D. W., Frank, W., Humpfer, E., Schäfer, H., Keller, S., Mörtter, M., et al. (2005). Quality control of beer using high-resolution nuclear magnetic resonance spectroscopy and multivariate analysis. *European Food Research and Technology, 220*, 215–221.

Loh, W. Y., & Vanichestakul, N. (1988). Tree-structured classification via generalized discriminant analysis (with discussion). *Journal of the American Statistical Association, 83*, 715–728.

Lopez-Rituerto, E., Savorani, F., Avenoza, A., Busto, J. H., Peregrina, J. M., & Balling Engelsen, S. (2012). Investigations of La Rioja terroir for wine production using H-1 NMR metabolomics. *Journal of Agricultural and Food Chemistry, 60*, 3452–3461.

Maes, P., Monakhova, Y. B., Kuballa, V., Reusch, H., & Lachenmeier, D. W. (2012). Qualitative and quantitative control of carbonated cola beverages using ^{1}H NMR spectroscopy. *Journal of Agricultural and Food Chemistry, 60*, 2778–2784.

Mannina, L., D'Imperio, M., Gobbino, M., D'Amico, I., Casini, A., & Sobolev, A. P. (2012). Nuclear magnetic resonance study of flavoured olive oils. *Flavour and Fragrance Journal, 27*, 250.

Masino, F., Chinnici, F., Franchini, G. C., Ulrici, A., & Antonelli, A. (2005). A study of the relationships among acidity, sugar and furanic compound concentrations in set of casks for Aceto Balsamico Tradizionale of Reggio Emilia by multivariate techniques. *Food Chemistry, 92*, 673–679.

McLachlan, G. J. (1992). *Discriminant analysis and statistical pattern recognition*. New York: John Wiley & Sons, Inc.

Papotti, G., Bertelli, D., Graziosi, R., Silvestri, M., Bertacchini, L., Durante, C., et al. (2013). Application of one- and two-dimensional NMR spectroscopy for the characterization of protected designation of origin Lambrusco wines of

Modena. *Journal of Agricultural and Food Chemistry, 61*, 1741–1746. http://dx.doi.org/10.1021/jf302728b.
Plessi, M., Bertelli, D., & Miglietta, F. (2006). Extraction and identification by GC-MS of phenolic acids in traditional balsamic vinegar from Modena. *Journal of Food Composition and Analysis, 19*, 49–54.
Plessi, M., Monzani, A., & Coppini, D. (1989). Quantitative determination of acids and derivates in balsamic and other vinegars. *Sciences des Aliments, 9*, 179–183.
Ridley, S., Jones, S., Shahani, A., Brampton, W., Nielsen, M., & Rowan, K. (1998). Classification trees for iso-resource grouping in intensive care. *Anaesthesia, 53*, 833–840.
Satrioni, F. (2010). *Notiziario Della Consorteria Dell'aceto Balsamico Tradizionale*. http://www.museodelbalsamicotradizionale.org/ita/pages/consorteria/giornale-settembre10.pdf (17th October, 2012).
Savorani, F., Tomasi, G., & Engelsen, S. B. (2009). Icoshift: a versatile tool for the rapid alignment of 1D NMR spectra. *Journal of Magnetic Resonance, 202*, 190–202.
Theobald, A., Muller, A., & Anklam, E. (1998). Determination of 5-hydroxymethylfurfural in vinegar samples by HPLC. *Journal of Agricultural and Food Chemistry, 46*, 1850–1854.
Thomas, F., & Jamin, E. (2009). ^{2}H NMR and ^{13}C-IRMS analyses of acetic acid from vinegar, ^{18}O-IRMS analysis of water in vinegar: International collaborative study report. *Analytica Chimica Acta, 649*, 98–105.
Van-Diep, L., Zheng, X.-W., Ma, K., Chen, J.-Y., Han, B.-Z., & Nout, M. J. R. (2011). Characterization of Fen-Daqu through multivariate statistical analysis of ^{1}H NMR spectroscopic data. *Journal of the Institute of Brewing, 117*, 516–522.
Winning, H., Roldan-Marín, E., Dragsted, L. O., Viereck, N., Poulsen, M., Sanchez-Moreno, C., et al. (2009). An exploratory NMR nutri-metabonomic investigation reveals dimethyl sulfone as a dietary biomarker for onion intake. *Analyst, 134*, 2344–2351.