Analytical Methods

ROYAL SOCIETY
OF **CHEMISTRY**

PAPER
Royston Goodacre *et al.*
Rapid, high-throughput, and quantitative determination of orange juice
adulteration by Fourier-transform infrared spectroscopy

175 YEARS

PAPER

CrossMark
click for updates

Rapid, high-throughput, and quantitative determination of orange juice adulteration by Fourier-transform infrared spectroscopy†

David I. Ellis,[a] Joanne Ellis,[a] Howbeer Muhamadali,[a] Yun Xu,[a] Andrew B. Horn[b] and Royston Goodacre*[a]

Orange juice is a hugely popular, widely consumed, and high price commodity typically traded in a concentrate form making it highly susceptible to adulteration. It has been consistently shown to be one of the leading food categories of reported cases of food fraud. One of the many forms of adulteration is dilution which can then be disguised with sugar solutions, or juices from other fruits or vegetables, which mimic the natural fruit sugars in this juice. Here, we demonstrate Fourier transform infrared (FT-IR) spectroscopy as a rapid, high-throughput and quantitative method for the determination of orange juice adulteration. Initial experiments involved the simple adulteration of pure orange juice with 0.5–20.0% water disguised with glucose, fructose or sucrose individually. This was followed by more complex mixtures of these three sugars at appropriate concentrations found in freshly prepared orange juice established using GC-MS; a total of 41 samples were prepared and all experiments undertaken in triplicate. Principal components-discriminant function analysis (PC-DFA) was undertaken on raw spectral data followed by partial least squares regression (PLSR) for quantification of the level of adulteration. Results from these chemometric analyses showed that infrared spectra contained information allowing for the discrimination and quantification between the three naturally occurring sugars in orange juice to disguise adulteration *via* dilution. Furthermore, it was clearly demonstrated that FT-IR in combination with PLSR is able to predict the levels of adulteration with excellent accuracy; the typical error on these predictions for test samples was 1.7%. We believe that the further development of these and other rapid methods could have an important role to play in the area of food authenticity and integrity, and food analysis in general.

1. Introduction

It is important for consumers, retailers and food regulatory bodies that food and beverage products (such as fruit juices) are of a consistently high quality, authentic, and have not been subjected to adulteration by any lower-grade material either by accident or for economic gain.[1] This practice, more recently termed food fraud or food crime[2] (as well as economically motivated adulteration (EMA)[3]), can be said to be an emerging issue which has gained much more prominence since high profile incidents such as the melamine scandal in China[4] and the so-called horsegate scandal in the UK and Europe,[5] in addition to many others which have occurred frequently over the last few decades.[6–9] Orange juice is a hugely popular, widely

consumed, and high price commodity typically traded in a concentrate form that is highly susceptible to adulteration[10] *via* substitution with (*e.g.*) lower grade sugars, fruits, or juice from another geographical region,[11] as well as reconstituted juice being labelled/sold as freshly squeezed juice. Indeed, due to the relatively high cost of this large volume product (estimated global forecasts for production of 65 °Brix in 2015/16 is 1.8 million metric tons [12]), orange juice is consistently named and been shown to be one of the leading food categories of reported cases of food fraud, with an incidence of 4%.[13] As well as being brazenly dishonest and unscrupulous, these forms of adulteration can have health implications, such as partial substitution of orange juice with the cheaper grapefruit juice resulting in negative and serious pharmacological interactions.[14,15] As grapefruit juice contains naringin, a major flavonoid glycoside (which is metabolized to the flavonone naringenin in humans) which affects the clinical modulation of drug transport, altering their bioavailability and hence effectiveness with potentially serious consequences.[6]

In general, the vast majority of orange juice sold has been reconstituted in a relatively uncomplicated process, involving extraction of juice from fruit, then enzyme inactivation *via* heat

[a]*School of Chemistry, Manchester Interdisciplinary Biocentre, University of Manchester, 131 Princess Street, Manchester, M1 7ND, UK. E-mail: roy.goodacre@manchester.ac.uk*

[b]*School of Chemistry, University of Manchester, Oxford Road, Manchester, M13 9PL, UK*

treatment to prevent loss of quality. Juice is then heated in high vacuum evaporators to remove water, with the remaining concentrate consisting of sugars and other solids. The volume of these juice soluble solids within fresh juice is approximately 12%, whereas it is 65% in concentrate. The concentrate is then frozen until required or sold on as frozen orange juice concentrate (FOJC).

This process leaves this fruit juice product vulnerable and offers several opportunities to those wishing to tamper and adulterate orange juice for economic gain, by increasing product yields for example. A simple dilution with water is one option but this is easily detected *via* a refractometer measurement of the final dilution ratio (°Brix). However, sugars,[16] or juices from other fruits or vegetables, can be added to the FOJC which mimic natural fruit sugars in orange juice and thus conceal the addition of water. Whilst some of the least expensive sugars to use, including corn and cane sweeteners for example, are easy to detect *via* stable isotope ratio analysis (SIRA),[17] due to the different pathways used to fix CO_2, other approaches have also been attempted. These include invert sugars, such partially invert sucrose, or beet sugars which have been widely used as they closely mimic the isotope ratio of the natural orange juice sugars, though these can be detected *via* chromatographic and/or MS methods.[18] Other forms of adulteration include the addition of amino acids to make the protein profile of the diluted or mislabelled juice appear normal,[19] as well as the addition of citric acid and trace minerals to adjust the acid and chemical profiles of the orange juice and make them appear within expected ranges.

Whilst a number of analytical methods have been used to detect adulteration of reconstituted orange juice, many of these are time-consuming, relatively expensive, could not be considered high-throughput and unlikely to be made portable in the near future (*e.g.*, SIRA or chromatographic methods).[10,18] Here we propose and demonstrate Fourier-transform infrared (FT-IR) spectroscopy[20–22] as a rapid, non-destructive, novel, inexpensive, and high-throughput system for the quantitative detection of sugar adulterants in orange juice.

2. Experimental

Sample preparation – 10 South African variety Valencia oranges were purchased from a retail outlet on the day of each experiment, then manually squeezed using a hand held juicer. The resultant juice was then sieved to remove any solids and placed into separate 10 mL centrifuge tubes and spun at 3080 *g* for 5 min at 4 °C in a centrifuge (Jouan CR3.22). These were transferred into 2 mL tubes and centrifuged for 3 min at 15 871 *g* using an Eppendorf Microcentrifuge 5424R with an FA-45-24-11 rotor (Eppendorf Ltd, Cambridge, UK). The juice (supernatant) was then stored at −80 °C until required.

Prior to spectral analysis, the experiments were divided into two main parts:

(i) Analysis of adulteration by each of the 3 naturally occurring sugars in orange juice (sucrose, glucose & fructose). Stock solutions (11.7% (w/v); based on expected levels from popular shop bought labels (data not shown)) of sucrose, glucose, and fructose were prepared by dissolving 10.5 g of each of the three sugars

respectively in 90 mL of water. Each of the stock solutions where then used to adulterate the orange juice in 0.5% increments so that a range from 0–20% (v/v) was achieved. Samples were then vortexed for 20 s and 10 μL of each sample spotted onto a 96 well silicon sample plate and oven dried at 50 °C for 30 min in order to fix the sample and remove as much water as possible. However, due to the nature of orange juice containing high levels of various sugars, all dried spots retained traces of water which is also apparent in Fig. S1† (broad peak centred at 3370 cm^{-1}).

(ii) Adulteration of mixed sugars in orange juice. GC-MS was used to ascertain the sugar levels in the freshly prepared orange juice; see ESI† for details. A stock solution was then prepared of a mixture of sucrose, glucose, and fructose to mimic the typical proportions of these sugars in orange juice. Glucose (2.45 g), sucrose (5.90 g), and fructose (1.90 g) were added to 100 mL water to give an overall sugar solution of 10.25% (w/v). This solution was used to adulterate orange juice in 0.5% increments ranging from 0–20%. A total of 41 samples were prepared. This whole process was repeated twice further to give three sample sets.

3. Chemometric methods

All multivariate analyses were conducted in Matlab R2014b (The MathWorks, Natick, MA, USA) and we used principal components-discriminant function analysis (PC-DFA) and partial least squares regression (PLSR) as described by Gromski *et al.*[23]

PC-DFA was performed directly on the raw (unprocessed) infrared spectra. Briefly, PCA was first used to project these spectra into new principal component (PC) axes based on maximum variance in these infrared data. These new PCs along with *a priori* information about the sample technical replicates (class structure) were used by the DFA algorithm. In DFA the algorithm uses projection to separate samples based on this class structure and, since the information used was based on the technical replicates, this does not bias the analysis towards either the type of sugar or the level of adulterant in these analyses. The resultant PC-DFA scores are plotted to show the separations achieved and the corresponding PC-DFA loadings plots provide information on which spectral features were used to construct the scores plots.

PLSR was programmed as initially described by Martens and Næs.[24] In PLSR we used different combinations of training and test sets to allow us to assess the reproducibility of the model and robustness of the analyses. For this we used bootstrap analysis ($n = 1000$). In each iteration on average 63.2% of the data were used for training the PLSR model and leave one out calibration was used within each training set to decide on the optimum number of PLS factors (latent variables) to use (the maximum possible was 25). After this process the test set (comprising on average 36.8% of the data) were used to test the calibrated model. This process was repeated a total of 1000 times and statistics were calculated on the test set predictions.

4. Results and discussion

The initial experiment was designed to adulterate pure orange juice with 0.5–20.0% water disguised with glucose, fructose or
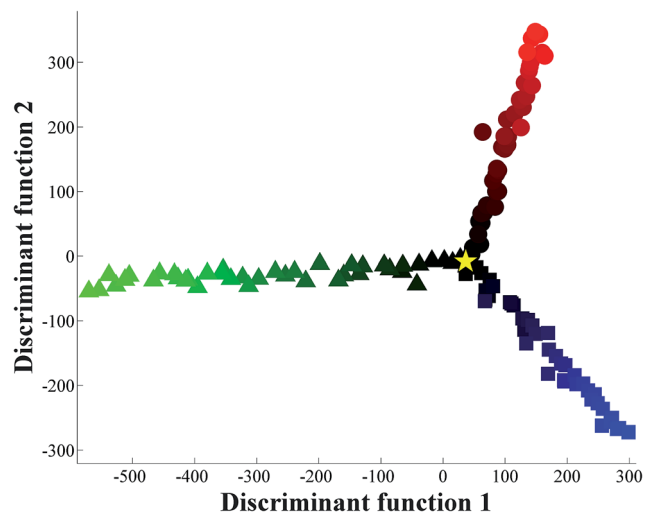
Fig. 1 PC-DFA scores plot of the FT-IR spectral data, showing the relationship between these juice samples when one of three single sugar solutions (11.7% (w/v) are added to pure orange juice (signified by a yellow star). In this analysis PCs 1–7 were used and these account for 94.0% of the total explained variance (TEV); the averages of 4 technical replicates are plotted. The colours shading represent the level of adulteration (from 0–20%) where: fructose – black to red circles (red is highest addition of fructose); glucose – black to green triangles (green is highest addition of glucose); sucrose – black to blue squares (blue is highest addition of sucrose).

constant. When the spectra (Fig. S1†) are perceived visually (that is by eye and not by machine vision[25]), two things become readily apparent. The first is that this is a highly reproducible technique (Fig. S1† shows all 363 raw FT-IR spectra), and the second is that without any chemometric processing it would be impossible to quantify the level of additional water spiked into the pure orange juice. Clearly there are some subtle visible differences in the spectra between 1200 and 900 cm^{-1}, which is what one would expect as this area of the mid-infrared region is predominantly associated with polysaccharides.[6,26]

The next stage was to use chemometrics to investigate if there was any inherent spectral structure in these data that would allow one to discriminate between the additions of the different sugar solutions, as well as to assess the level of excess water used to dilute the pure orange juice. PC-DFA was used to generate a scores plot which shows the relationship between these juice samples (Fig. 1). In this plot freshly squeezed pure orange juice (yellow star in the figure) is located at the origin. There are then three trajectories radiating from the origin point of pure juice which can be seen as three 'spokes'. The first highlighted by a green gradation shows the incremental addition of water with glucose and this trajectory is evident in the negative part of the first discriminant function (DF1). The second highlighted in red shows the addition of fructose and this trajectory occurs in the positive DF2 direction. The blue gradation in the negative part of DF2 indicates the addition of the disaccharide sucrose which is a combination of the monosaccharides glucose and fructose. It is clear from this chemometric analysis that these infrared spectra do indeed contain sufficient information to allow one to discriminate, and to quantify, between the three sugars naturally found in orange juice that have been used individually to disguise the addition

sucrose individually. Three sugar solutions were prepared at 11.7% (w/v) as this was identified from inspection of packaging at a local supermarket to be the typical level of sugar found in freshly squeezed orange juice. These sugar solutions were then added to freshly squeezed orange juice to keep the overall °Brix
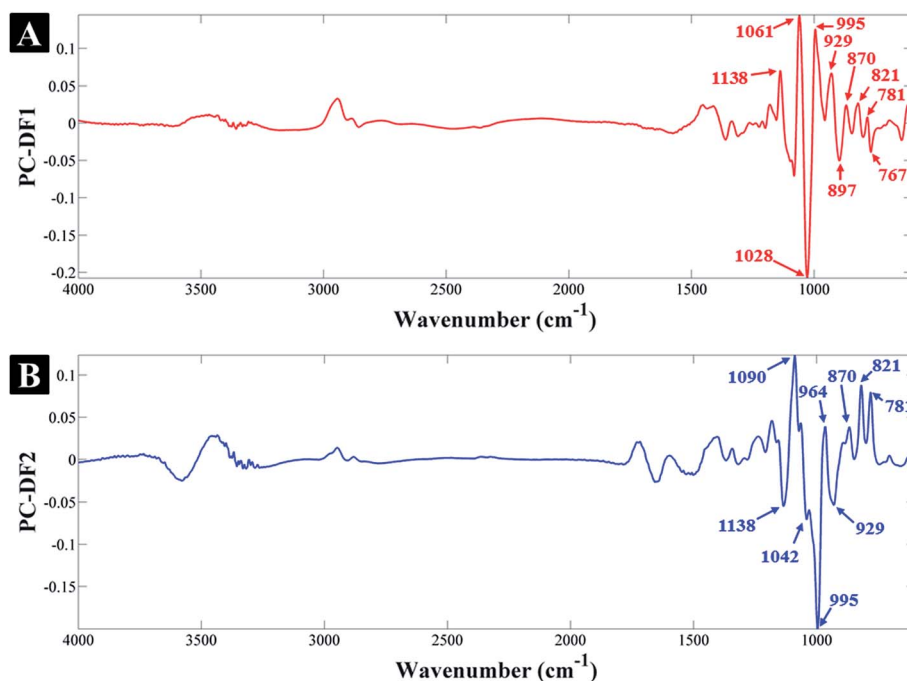

Fig. 2 PC-DFA loadings plots for the (A) first and (B) second discriminant functions.

of up to 20% water. The PC-DFA analyses relies on the projection of the original infrared data onto this PC-DFA scores plot (Fig. 1) and the features that are used to achieve this model can be seen in the associated PC-DFA loadings plots which are shown in Fig. 2.

According to the PC-DF1 loadings (Fig. 2A), the most significant vibrational bands contributing to the clustering of glucose-containing samples (on the negative side of PC-DF1) and its separation from fructose and sucrose samples, include: 767, 897 and 1028 cm^{-1} corresponding to different vibrational modes specific to glucose, which are also present in the FT-IR spectra of the standard glucose solution (in water) in Fig. S2.† The list of all the significant vibrational bands and their corresponding assignments are provided in Table 1. Similarly, the most significant peaks associated with samples clustering on the positive side of PC-DF1, include: 781, 821, 870, 929, 995, 1049 and 1138 cm^{-1} which are specific to fructose and sucrose-containing samples, but are not present in the glucose samples (Fig. S2†). The above vibration assignments and their significance to specific sugars also applies to the PC-DF2 loadings where on the positive side of PC-DF2 the peaks at 781, 821, 870 and 964 cm^{-1} are associated with fructose samples, and on the negative side the peaks at 929, 995, 1042 and 1138 cm^{-1} are associated with sucrose samples, which is again in agreement with their reference spectra (Fig. S2†).

As it was relatively easy to discriminate the addition of the different sugars individually the next stage was to make the adulteration more difficult to detect, by adding water disguised by the sugars in combination and in the correct ratio, whilst keeping the overall °Brix constant. GC-MS was used to determine the sugar content (see ESI†) accurately in our freshly squeezed orange juice and we therefore prepared a sugar solution containing 2.45 g glucose, 5.90 g sucrose and 1.90 g fructose per 100 mL. This solution was used to generate three series of adulterated pure orange juice with final concentration ratios ranging from 0.5–20.0% (in 0.5% increments). We again

performed PC-DFA and it is clear from the scores plot in Fig. 3 that whilst there is a gradient starting from the bottom left of the plot to the top right that is associated with the increasing addition of water, the three experimental repeats are not congruent. We therefore decided to use more powerful chemometrics for quantitative predictions.

PLSR is a popular multivariate regression approach that is employed for quantitative calibration.[24] We therefore used PLSR on the whole orange juice data set and used bootstrapping for validation purposes. In bootstrapping resampling of the spectral data with replacement[6,23] is used to generate two data sets randomly: one is termed the training set and is used to calibrate the PLSR model; the second is used as a test set which is employed to test the predictive ability of the model. This process is repeated to generate different combinations of training and test sets. In our study, we used 1000 bootstraps for PLSR and the results for the 1000 test sets only (not the calibration/training sets) are shown in Fig. 4. It is clear from this figure that FT-IR in combination with PLSR is able to predict the adulteration of the pure orange juice with water showing excellent accuracy with a typical error for predicting test samples of 1.7%. According to the European Union directive 657/2002/EC, the error should be smaller than 5%[27] (see Fig. S3 and S4† for the full statistics on these models). We do of course note that 16.5% and 17.0% are poorly under predicted and the error is large. We have re-inspected these data and suspect that ironically enough considering a study of this nature, that this could be a simple labelling error and visibly demonstrates these deviations from the norm.
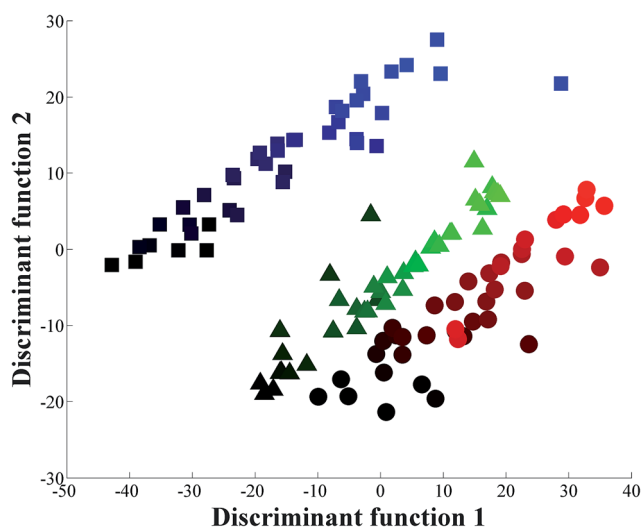
Table 1 List of the significant vibrations contributing to the PC-DFA clustering patterns, identified by the loadings plots, with corresponding assignments to mid-infrared spectra of sugars

Wavenumber (cm^{-1})	Assignment[a]	Sugars
767	C–H and C–O deformation	Glucose
781	C–H and C–O deformation	Fructose
821	C–H deformation	Fructose
870	C–H deformation, C–C stretching	Fructose
897	C–H deformation	Glucose
929	C–C stretching	Sucrose
964	C–O stretching	Fructose
995	C–O stretching	Sucrose
1028	C–O stretching	Glucose
1042	C–C stretching	Sucrose
1061	C–C and C–O stretching	Sucrose, Fructose
1138	C–O stretching	Sucrose

[a] Ref. 28 and 38–40.

Fig. 3 PC–DFA scores plot of the FT-IR spectral data, showing the relationship between the three experiments where the mixture of the three sugars is added to pure orange juice, from 0–20 in 0.5% increments. In this analysis PCs 1–20 were used and these account for 99.7% TEV; the averages of 4 technical replicates are plotted. The colours represent the three different repeats for the three experimental repeats: 1 – black to red circles (red is highest addition of mixture in repeat 1); 2 – black to green triangles (green is highest addition of mixture in repeat 2); 3 – black to blue squares (blue is highest addition of mixture in repeat 3).
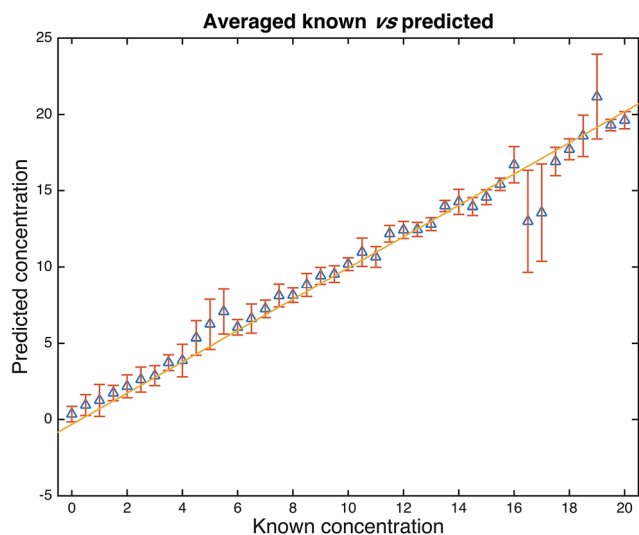
Fig. 4 PLSR predictions from the 1000 bootstrap resamplings. Data are from the test set only and the points are the means with standard deviation error bars shown. The statistics for these models are as follows: average number of latent variables (PCs) used for PLS = 16.259 (min–max: 7–25), the average RMSECV (in the training set) was 1.6851, average RMSEP (test set) was 1.7391. The correlations of the predictions with the known solutions were $R^2 = 0.9142$ (training set) and $Q^2 = 0.9081$ (test set). See Fig. S2 and S3† for the full distributions.

5.   Concluding remarks

Orange juice is a hugely popular, widely consumed and high-volume product, and could be said to be relatively expensive when directly compared to many other popular fruit juices. This alone makes it a potential target for economically motivated fraud. With this vulnerability to fraudulent tampering/stretching further exacerbated by the fact that the vast majority of this product is usually reconstituted from orange concentrate. Earlier studies using attenuated total reflectance (ATR) stated that this method showed promise as a rapid screening technique for a range of sugar adulterants in apple juice.[28] Here, we deliver on this promise and show the rapid, quantitative and high-throughput screening potential[29] of this technique for orange juice stretching with the PLSR model providing results in accordance with international validation guidelines.[27] Though for the first time, in the much higher volume and more popular orange juice, a product which has repeatedly been shown to be one of the leading food categories of reported cases of food fraud.[13] Whilst many other techniques have been applied to this area,[30–34] we concur with others in forwarding the potential of vibrational spectroscopy[35] and sincerely believe that the further development of these, and other rapid methods[25,36] could have an important role to play in the area of food authenticity and integrity,[37] and food analysis in general. With the versatility of these methods being such that their placement could be within food processing or production facilities, regulatory laboratories, and due to the inherent portability of these techniques out in the field, or indeed anywhere within food supply chains.[2]

6.   Acknowledgements

References

1 D. I. Ellis, D. Broadhurst, S. J. Clarke and R. Goodacre, *Analyst*, 2005, **130**, 1648–1654.

2 D. I. Ellis, H. Muhamadali, S. A. Haughey, C. T. Elliott and R. Goodacre, *Anal. Methods*, 2015, **7**, 9401–9414.

3 J. C. Moore, J. Spink and M. Lipp, *J. Food Sci.*, 2012, **77**, R118–R126.

4 J. S. Chen, *Chin. Med. J.*, 2009, **122**, 243–244.

5 C. Elliott, *Elliott Review into the Integrity and Assurance of Food Supply Networks – Final Report – A National Food Crime Prevention Framework*, 2014.

6 D. I. Ellis, V. L. Brewster, W. B. Dunn, J. W. Allwood, A. P. Golovanov and R. Goodacre, *Chem. Soc. Rev.*, 2012, **41**, 5706–5727.

7 Europol, Largest-ever seizures of fake food and drink in Interpol-Europol operation, https://www.europol.europa.eu/content/largest-ever-seizures-fake-food-and-drink-interpol-europol-operation, accessed 1 April 2016.

8 D. K. Trivedi, K. A. Hollywood, N. J. W. Rattray, H. Ward, D. K. Trivedi, J. Greenwood, D. I. Ellis and R. Goodacre, *Analyst*, 2016, **141**, 2155–2164.

9 C. E. Handford, K. Campbell and C. T. Elliott, *Compr. Rev. Food Sci. Food Saf.*, 2016, **15**, 130–142.

10 Z. Jandric, D. Roberts, M. N. Rathor, A. Abrahim, M. Islam and A. Cannavan, *Food Chem.*, 2014, **148**, 7–17.

11 N. Cicero, C. Corsaro, A. Salvo, S. Vasi, S. V. Giofre, V. Ferrantelli, V. Di Stefano, D. Mallamace and G. Dugo, *Nat. Prod. Res.*, 2015, **29**, 1894–1902.

12 USDA, *Citrus: World Markets and Trade*, U. S. D. o. Agriculture and F. A. Service, United States of America, 2016.

13 R. Johnson, *Food Fraud and "Economically Motivated Adulteration" of Food and Food Ingredients, Report R43358*, Congressional Research Service, Washington, D.C., 2014.

14 D. B. Kell and R. Goodacre, *Drug Discovery Today*, 2014, **19**, 171–182.

15 D. G. Bailey, *Br. J. Clin. Pharmacol.*, 2010, **70**, 645–655.

16 J. Zidkova and J. Chmelik, *J. Mass Spectrom.*, 2001, **36**, 417–421.

17 S. Rummel, S. Hoelzl, P. Horn, A. Rossmann and C. Schlicht, *Food Chem.*, 2010, **118**, 890–900.

18 L. Vaclavik, A. Schreiber, O. Lacina, T. Cajka and J. Hajslova, *Metabolomics*, 2012, **8**, 793–803.

19 J. L. Gomez-Ariza, M. J. Villegas-Portero and V. Bernal-Daza, *Anal. Chim. Acta*, 2005, **540**, 221–230.

20 D. Cozzolino, *Anal. Methods*, 2015, **7**, 9390–9400.

21 D. I. Ellis, D. Broadhurst and R. Goodacre, *Anal. Chim. Acta*, 2004, **514**, 193–201.

22 H. Muhamadali, D. Weaver, A. Subaihi, N. AlMasoud, D. K. Trivedi, D. I. Ellis, D. Linton and R. Goodacre, *Analyst*, 2016, **141**, 111–122.

23 P. S. Gromski, H. Muhamadali, D. I. Ellis, Y. Xu, E. Correa, M. L. Turner and R. Goodacre, *Anal. Chim. Acta*, 2015, **879**, 10–23.

24 H. Martens and T. Næs, *Multivariate Calibration*, John Wiley & Sons, Chichester, UK, 1991.

25 X. Zou, H. Xiawei and M. Povey, *Analyst*, 2016, **141**, 1587–1610.

26 D. I. Ellis, G. G. Harrigan and R. Goodacre, *Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis*, 2003, pp. 111–124.

27 European Commission, *Off. J.*, 2002, 8–36.

28 J. F. D. Kelly and G. Downey, *J. Agric. Food Chem.*, 2005, **53**, 3281–3286.

29 C. Black, S. A. Haughey, O. P. Chevallier, P. Galvin-King and C. T. Elliott, *Food Chem.*, 2016, **210**, 551–557.

30 M. Aldeguer, M. Lopez-Andreo, J. A. Gabaldon and A. Puyet, *Food Chem.*, 2014, **145**, 1086–1091.

31 F. Ammari, L. Redjdal and D. N. Rutledge, *Food Chem.*, 2015, **168**, 211–217.

32 Y. J. Hu, C. Zhu, G. Q. Chen, Y. Zhang, F. B. Kong, R. Li, Z. W. Zhu, X. Wang and S. M. Gao, *Spectrosc. Spectral Anal.*, 2014, **34**, 2143–2147.

33 M. Navarro-Pascual-Ahuir, M. J. Lerma-Garcia, E. F. Simo-Alfonso and J. M. Herrero-Martinez, *J. Agric. Food Chem.*, 2015, **63**, 2639–2646.

34 M. A. Pardo, *Food Chem.*, 2015, **172**, 377–384.

35 G. Downey, in *New Analytical Approaches for Verifying the Origin of Food*, ed. P. Brereton, Woodhead Publ Ltd, Cambridge, 2013, pp. 94–116.

36 P. Brereton, in *New Analytical Approaches for Verifying the Origin of Food*, ed. P. Brereton, Woodhead Publ Ltd, Cambridge, 2013, pp. 3–11.

37 D. I. Ellis and R. Goodacre, *Anal. Methods*, 2016, **8**, 3281–3283.

38 B. Brizuela, L. C. Bichara, E. Romano, A. Yurquina, S. Locatelli and S. A. Brandán, *Carbohydr. Res.*, 2012, **361**, 212–218.

39 M. Grube, M. Bekers, D. Upite and E. Kaminska, *Spectroscopy*, 2002, **16**, 289–296.

40 M. Ibrahim, M. Alaam, H. El-Haes, A. F. Jalbout and A. d. Leon, *Ecletica Quim.*, 2006, **31**, 15–21.