Analytical Methods

Orange Juice

?

GRAPEFRUIT

ROYAL SOCIETY
OF CHEMISTRY

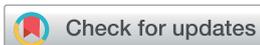**PAPER**
Royston Goodacre *et al.*
Rapid detection and quantification of the adulteration of
orange juice with grapefruit juice using handheld Raman
spectroscopy and multivariate analysis

Indexed in
Medline!

Check for updates

Rapid detection and quantification of the adulteration of orange juice with grapefruit juice using handheld Raman spectroscopy and multivariate analysis

Mehrvash Varnasseri,† Yun Xu ⟨iD⟩ † and Royston Goodacre ⟨iD⟩ *

Detecting food adulteration has always been an important task for food safety, especially when grapefruit is the adulterant as components in the juice have undesired interactions with many medicines. In this study we employed a handheld Raman device to detect adulteration of orange juices with grapefruit juices. Fresh fruits of orange and grapefruit were purchased from five different sources and fruit juices were made using a handheld juicer. The extracted juices were then mixed in a way that concentrations of grapefruit juices varied from 0% to 100% in 5% increments. In order to study the impact of the different sources of the fruits, three different sets of mixtures were prepared based on their spectral similarity and dissimilarity. Raman spectra were collected using a handheld instrument with an excitation laser at 785 nm and data analysed using principal component analysis (PCA), principal component-discriminant function analysis (PC-DFA) and partial least squares regression (PLS-R). PLS-R models were trained and validated on: (i) the full data set from the three different mixture sets, and (ii) each set of the three mixtures separately. The results showed that a good calibration model was obtained using full data which had a coefficient of determination (Q^2) of 0.81 and a root mean square error of prediction (RMSEP) of 12.5%. Such results were improved when the PLS-R model was trained and validated on the three separate mixture combinations, where the Q^2 varied from 0.85 to 0.89 and RMSEP varied from 9.9% to 11.6%. Finally, we adopted a two step approach in which a partial least squares for discriminant analysis (PLS-DA) was trained first to classify the three sample sources and then three different PLS-R models were subsequently trained on samples from the same source. This resulted in a Q^2 of 0.83 and RMSEP of 12.0%. In conclusion, we have demonstrated that Raman spectroscopy can be used as a portable and rapid analytical tool for detecting adulteration of grapefruit juice added to orange juice.

Introduction

In modern society consuming juices, particularly fruit-based products, is increasing. According to the Association of the Industry of Juices and Nectars of the European Union (AIJN), global consumption of fruit juice and nectars equated to 36.2 billion litres in 2017. There is also a dramatic rise in the EU's requirement for freshly squeezed juices due to the huge demand for premium juices and within the EU fruit juice and nectar consumption was 9.2 billion litres in the same year.[1,2] Such high demands inevitably invite economically motivated adulteration of fruit juice. Orange juice is one of the most targeted food commodities for adulteration and one such adulteration method is to blend it with grapefruit juice.[3–5] Such

adulteration gives the culprits unfair advantage in the market, damage consumers' experience and more importantly it may have serious health implications.[6,7] This is due to the fact that grapefruit juice has been reported to interact with numerous drugs and inadvertent consumption of grapefruit juice by some consumers can result in negative and potentially serious pharmacological interactions.[8–10] Indeed, many patient information leaflets specifically warn about avoiding grapefruit consumption. This highlights the need for a cheap, rapid and portable analytical tool for early detection of grapefruit adulteration to orange juice.

Classical analytical platforms such as HPLC-DAD-MS/MS,[11] UPLC-QToF MS, FT-IR and NMR spectroscopy had been reported for fruit juice product authentications.[12–16] A few molecular DNA based methods have also reported in the literature for the same purpose.[17–21] However, all these methods involve lengthy sample preparations and bulky instruments which make them not suitable for rapid on-site detection. By contrast, Raman spectroscopy, especially using portable

Department of Biochemistry and Systems Biology, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, L69 7ZB, UK. E-mail: Roy.Goodacre@liverpool.ac.uk

† These authors contributed equally to the work.

handheld Raman platforms, has many attractive features such as being non-destructive, with little to no sample preparation required, resulting in chemical information rich spectra and is insensitive to water which means that the juice can be measured directly. All of these features have made Raman spectroscopy particularly suitable for such applications.[22,23]

In this study we employed a CBEx handheld Raman spectrometer from Snowy Range (Laramie, Wy, U.S.A.) to measure freshly squeezed orange juices adulterated with different concentrations of grapefruit juices. The collected Raman spectra were then subjected to multivariate analysis including principal component analysis (PCA) and partial least squares regression (PLS-R) to detect and quantify the levels of adulterations of grapefruit juices added orange juices. Considering the fact that the same fruit from different origins (*e.g.*, site of purchase) may also have subtle differences in their chemical composition (*e.g.*, due to storage conditions and length of storage) and this may affect their Raman spectra. To assess such effect, orange and grapefruit were purchased from five different sources and their Raman spectra were analysed and compared.

Experimental methods and materials

Grapefruit and orange fruits were purchased from five different sources, denoted as A, B, C, D and E. These sources are local shops popular to local residents who wish to purchase fruits and vegetables. As such the storage conditions and length of storage are unknown. The three natural sugars that are found in orange juices and grapefruit juice – fructose, glucose and sucrose – were acquired from Thermo Fisher Scientific, UK.

Grapefruit and oranges were manually squeezed using a handheld juicer. The extracted juices were then filtered to remove solid debris and centrifuged at $3100g$ for 5 min. The supernatants were further centrifuged at $15\,871g$ for 3 min. All centrifuge steps were kept at 4 °C and the processed juice (supernatants) were labelled and stored in at -80 °C until further analysis. Mixtures of two types of juices were prepared by mixing orange juice with grapefruit juice in glass vials and the concentration of grapefruit juice varied from 0% to 100% v/v in 5% increments which resulted in 21 different concentrations of grapefruit juices within the mixtures. More details are provided in the results section.

A CBEx handheld Raman spectrometer from Snowy Range (Laramie, WY, U.S.) was employed to measure Raman spectra of juices. The spectrometer operates using a 70 mW (on sample) 785 nm laser with detection on a 2048 element CCD array, resulting in a 12–14 cm^{-1} spectral resolution. All spectra were collected in the spectral range 400–2300 cm^{-1} with an acquisition time of 2 s.

Prior to multivariate statistical analysis, all spectra were normalized using standard normal variate (SNV) normalization so that each spectrum has a mean of 0 and a standard deviation of 1. Principal component analysis (PCA) and principal component-discriminant function analysis (PC-DFA) were used to visualize the pattern of samples. Partial least squares regression (PLS-R) was then employed to quantify the level of adulterations of grapefruit juice added to orange juice. The data

were split into training and test set based on concentrations of grapefruit juice added to the mixture. The samples with 0%, 10%, 20%, …, 100% grapefruit juice were used as training sets and the samples with 5%, 15%, 25%, …, 95% grapefruit juice were used as test sets. The number of latent variables (LVs) of PLS models was optimised by performing a *k*-fold ($k = 11$ which is the number of different concentrations in the training set) cross-validation on the *training set only*. The optimal number of LVs was set as the one which resulted in minimal root mean squares errors of cross-validation (RMSECV). Once this process was completed the trained PLS-R models were then applied to the *test set* and the results were reported as coefficient of determination Q^2 and root mean squares errors of prediction (RMSEP), as given in eqn (1) and eqn (2), of the test set.

$$Q^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{1}$$

$$\mathrm{RMSECV/RMSEP} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \tag{2}$$

where y_i is the known concentration of sample i; $\hat{y}_i$ is the predicted concentration of sample i; $\bar{y}$ is the mean of all known concentrations and n is the number of samples use for cross-validation or that in the test set.

Based on the PLS-R model, a limit of detection was also estimated using net analyte signal approach as described by Olivieri *et al.*[24]

All data analysis works were carried out in MATLAB 2020a (the MathWorks, MA, U.S.) environment and relevant in-house MATLAB functions are available freely on our Github repository at https://github.com/biospec.

Results

It appeared that most meaningful Raman signal of juices were within the range of 480 to 1200 cm^{-1} and thus all the multivariate analysis were carried out within this region. The Raman spectra of pure orange and grapefruit juices are shown in Fig. 1, along with those of the three pure sugars that are predominantly found in these juices: *viz.* fructose, glucose and sucrose. The tentative band assignments are provided in Table 1 with the abundance of the bands relative to the three pure sugars. It can be seen that most major bands of these three sugars are clearly visible in juice spectra with visible differences in the relative ratios of these bands between two types of juice. This indicates that the relative ratios of these sugars are different in these two types of juices which is consistent with the finding of a previous study.[25] The study had analysed 286 orange juice and 38 grapefruit juice samples collected in Florida, California, New Jersey and Texas. The results showed that the relative ratio of glucose, fructose and sucrose in orange juice, on average, was 1 : 1 : 2 while that of grapefruit juice on average was 1 : 1 : 1.

The PCA scores plot of the 10 pure juices is shown in Fig. 2(a). Separations of two types of juices and separations between
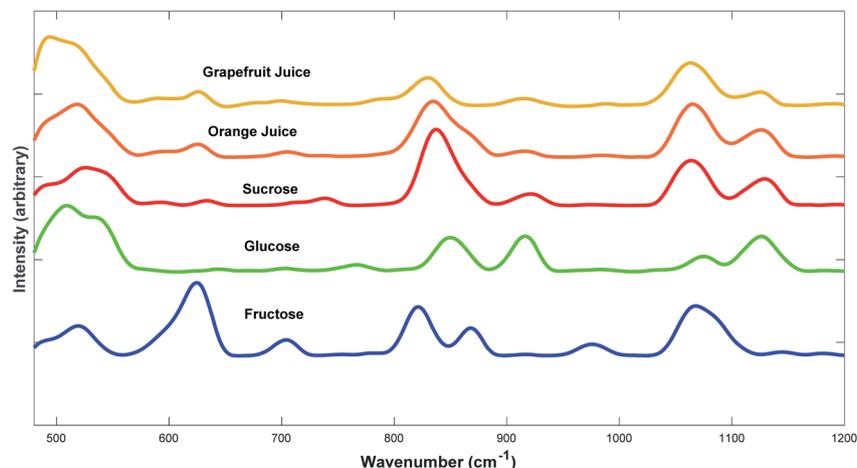
Fig. 1 Typical Raman spectra of grapefruit and orange juices (both from source A, showed spectra were the average of five repeats) in comparison with the three dominant sugars found in these fruit juices: sucrose, glucose and fructose. The y-axis is offset for visual clarification.

Table 1 Vibrational modes assignment

Raman band (cm^{-1})	Source of vibration[a]	Assignment
509–524	GG/F/S	Skeletal vibration
593–600	S	Skeletal vibration
625–631	FF/s	Ring deformation
705	F	Skeletal vibration
822	FF	C–OH stretch
837–850	SS/G	Unknown
868	F	C–O–C cyclic alkyl ethers
916–921	f/G/s	CH, COH bend
976–988	F/s/g	Ring "breathing"
1065–1074	FF/SS/G	C–O–C cyclic alkyl ethers
1127–1129	SS/GG	C–OH deformation

[a] Sources: G for glucose, F for fructose and S for sucrose. Strength of vibration is denoted as follows: x weak; X strong; XX dominant.

difference sources within each type of juice can both be observed. It is also interesting to see that the distances between two types of juices in the PCA space varied significantly between the different sources of the fruits. The differences between orange from source A and grapefruit juice from source B appeared to be the most different in fruits purchased from source A (highlighted in red) while such difference is much smaller in orange purchased from source C and grapefruit purchased from source E (highlighted in blue). This suggested that fruits from different sources could have subtle differences in their Raman spectra and that this could affect the capability of detecting adulteration levels of grapefruit juices added to orange juices. The loadings plot of this PCA model is shown in Fig. 2(b). The most significant bands are located between 800–900 cm^{-1} which corresponding to the major sugar bands as shown in Fig. 1, suggesting subtle sugar contents differences also contributes to the separations between fruits from different sources. With this factor in mind, three sets of mixtures of orange and grapefruit juices were prepared: (1) orange from source A mixed with grapefruit from source B, denoted as A1–B2; (2) orange from source D and grapefruit from source C, denoted as D1–C2 and (3) orange from source C and grapefruit source E, denoted as C1–E2. These represent three scenarios where orange-grapefruit juice mixtures in PCA space were most different (A1–B2), moderately different (D1–C2) and least different (C1–E2). For each mixture set, a series of orange and grapefruit juice mixtures were prepared with the concentration of grapefruit juice varying from 0% to 100% v/v in 5% increments, as described in the experimental section. The PCA scores plot of one set of mixture, D1–C2, is given in Fig. 3(a) and it is clear that there was a gradient corresponding to the amount of adulterant within the mixture along PC 1 axis. The PC 1 loadings plot is shown in Fig. 3(b) and again the most significant bands are located between 800–900 cm^{-1}, pointing to sugar differences are the main factor in differentiating these two types of fruits. The PCA scores and loadings plots of the other two mixtures also showed similar trends (data not shown). A more interesting pattern was observed in PC-DFA scores plot where all three mixtures were analysed together (Fig. 4). It appears that the first and therefore most significant discriminant function was responsible for the amount of grapefruit juice added to orange juice while the source differences were observed on the 2nd and 3rd discriminant functions. This suggest that although the origin of fruit might have some impact in quantifying the grapefruit juice added to orange juice, such impact was not significant enough to make such quantification impossible. We also note in Fig. 4 that the three pure orange juices (A1, C1, D1) cluster more closely together compared to the three pure grapefruit juices (B2, C2, E2).

In light of results of PCA and PC-DFA, PLS-R modelling was carried out in three different ways. First, PLS-R models were trained and tested on the whole data set, regardless of the origins of the mixtures. This represents the expected results if the source of fruits were ignored, which would be more real-world as it may be preferred to have a single model that fits all these data. Then three separate PLS-R models were trained and tested on each set of mixtures separately and this represents the scenario when *a priori* knowledge about the source of
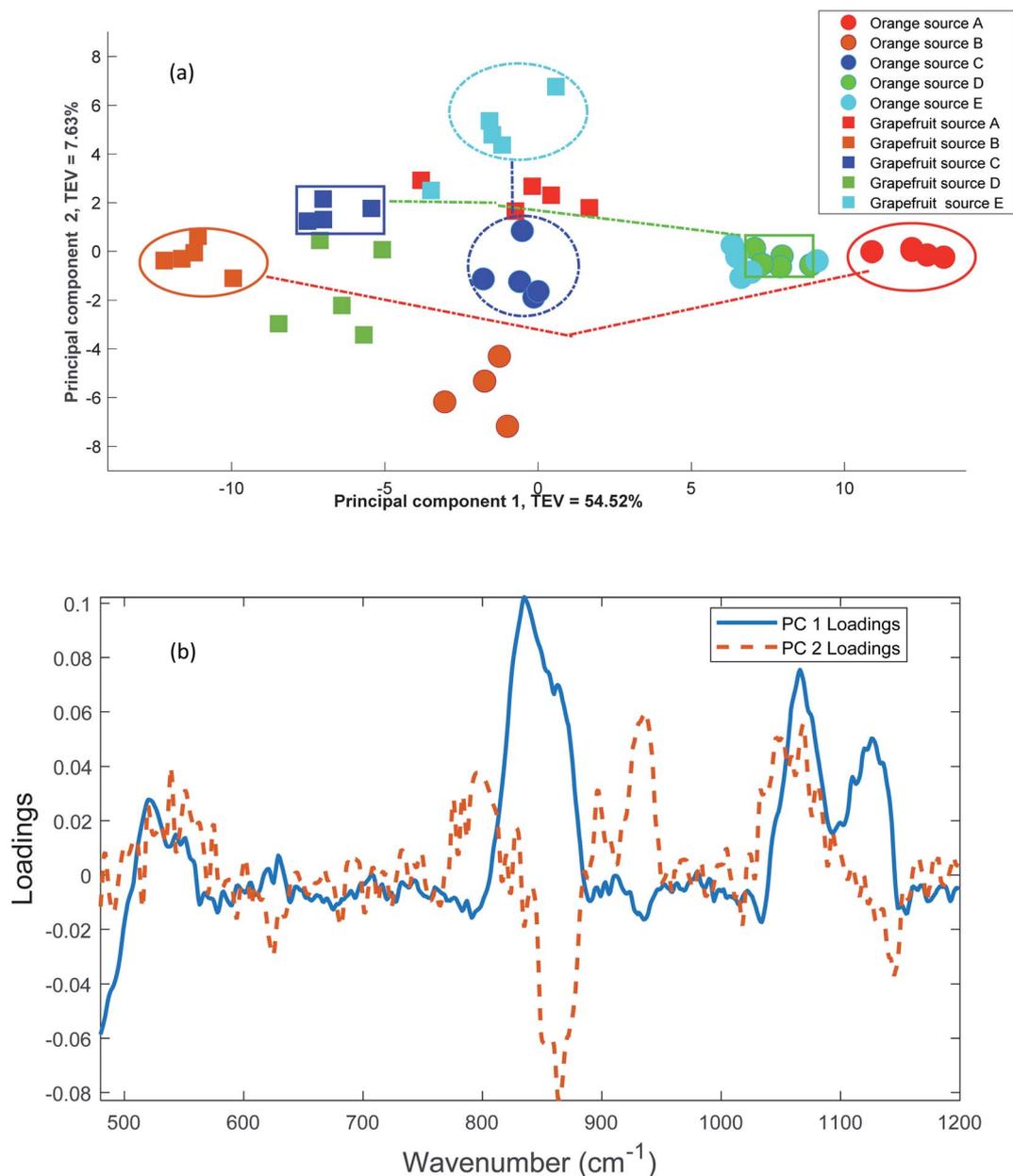
Fig. 2 PCA plots of orange and grapefruit juices. (a) Scores plot (TEV = total explained variance). Dotted lines highlight the 3 mixtures that were then prepared from the different sources of the orange and grapefruit juices; (b) loadings plot.

the fruit is available and one had trained models specifically to each source. Lastly, a two-step hierarchical model was adopted in which first a classification model of PLS-DA was trained on the training set to classify the source of fruits, *i.e.* A1–B2, D1–C2 or C1–E2. Then three PLS-R models were trained on each set of mixtures separately using corresponding subsets of samples in the training set. In the test phase, each sample in the test set was first subjected to the PLS-DA models to predict its source of origin and based on the prediction of the source, this test sample was then assigned to a corresponding PLS-R model to predict the concentration of grapefruit adulteration. This means that if a source was incorrectly predicted in the first

classification stage then the wrong PLS-R model would be used for quantification.

A predicted *versus* known plot of PLS-R model predictions on all data (*i.e.*, source of origin ignored) is given in Fig. 5. A good agreement between known and predicted concentrations were observed. Such plot for all other models showed similar pattern and thus omitted for brevity. The number of latent variables (LVs), RMSECV, Q^2, RMSEP and estimated LoDs of all PLS-R models are summarised in Table 2. When PLS-R modelling was performed on all data without considering the source of fruit, the averaged Q^2 was 0.8171, RMSEP was 12.47% with an estimated LoD of 11.70%. The prediction accuracy improved when PLS-R model was trained and tested on the data from the
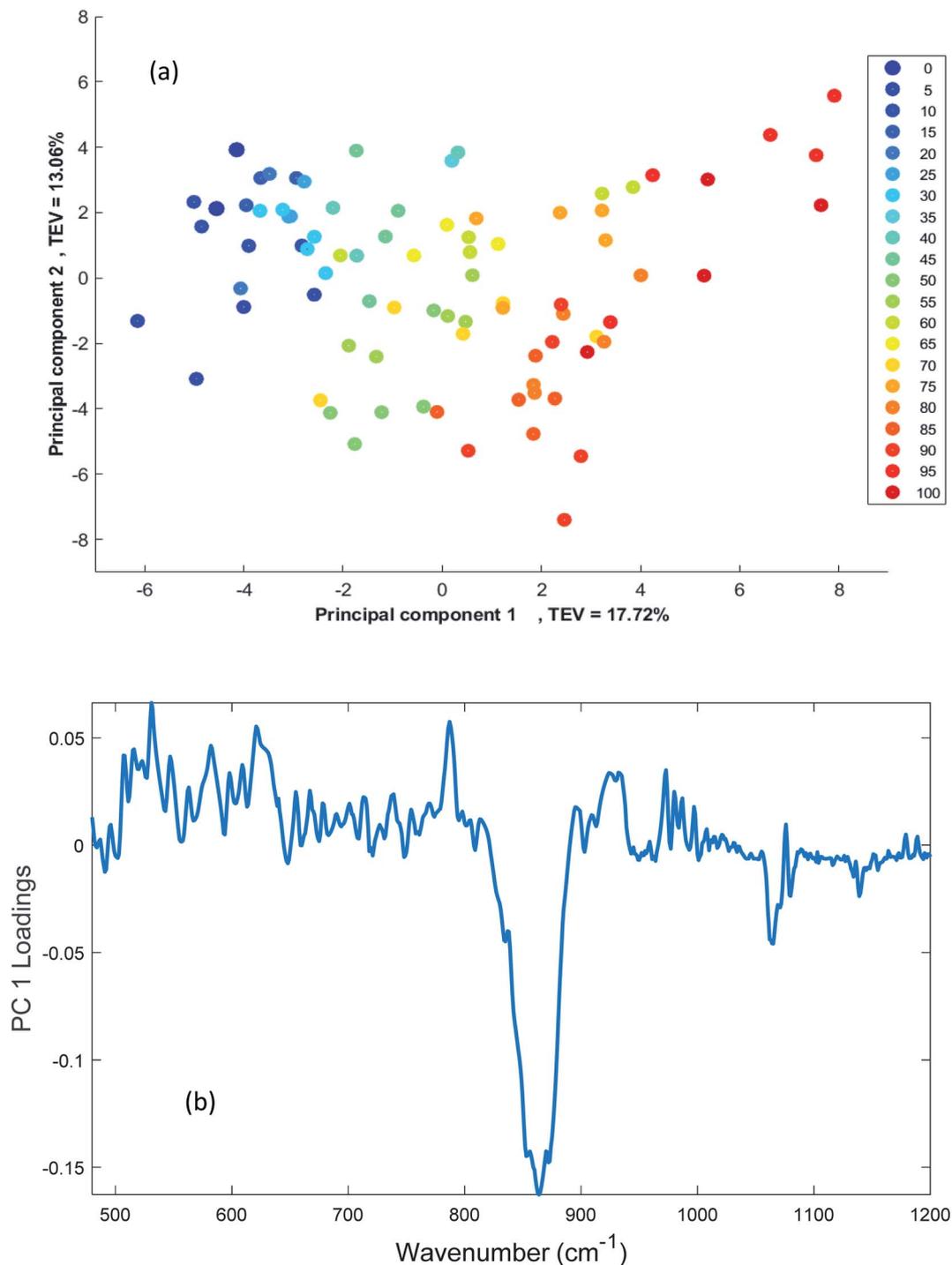
Fig. 3 PCA plots of the fruit juice combination of D1–C2. (a) Scores plot, the colours represent the level of grapefruit added. And the details of this are provided within the figure. (b) PC 1 loadings plot.

same source. Q^2 increased to 0.8467, 0.8693 and 0.8891 for three sets of A1–B2, D1–C2 and C1–E2, respectively, and their RMSEP reduced to 11.55%, 10.18% and 9.90%, respectively. The LoDs of these three models were also improved to 8.10%, 7.67% and 6.51% respectively. Such differences highlighted the adverse effect of difference source of origins of fruits on quantifying the level of adulteration of grapefruit juice in orange juice. The

prediction accuracy of PLS-DA model in predicting source of origins was 87.0% and a detailed confusion matrix is given in Table 3. The following prediction in concentrations of grapefruit juice resulted in a Q^2 of 0.8322 and a RMSEP of 11.95%. These were worse than the results of single source PLS-R model albeit better than those of the PLS-R model which completely ignored source of origin. This demonstrated that by introducing
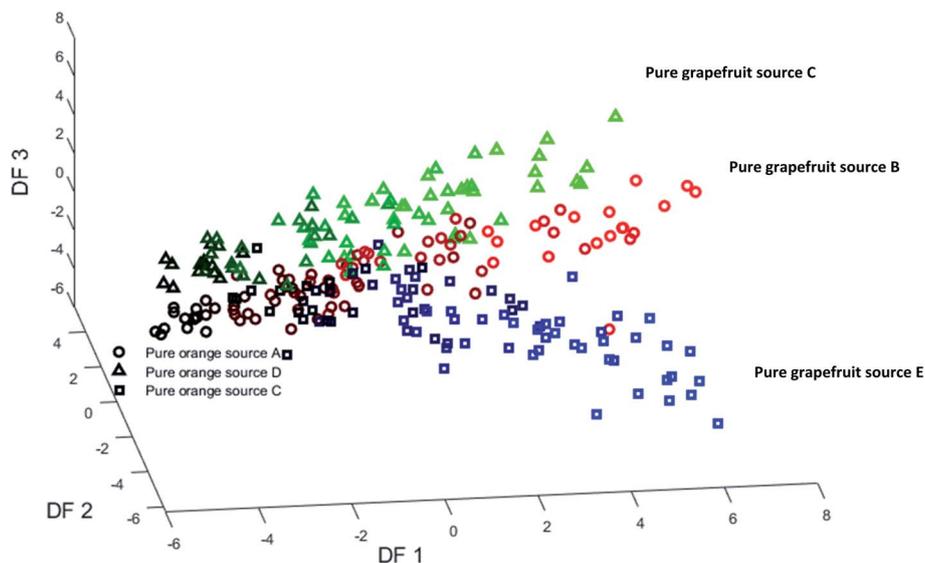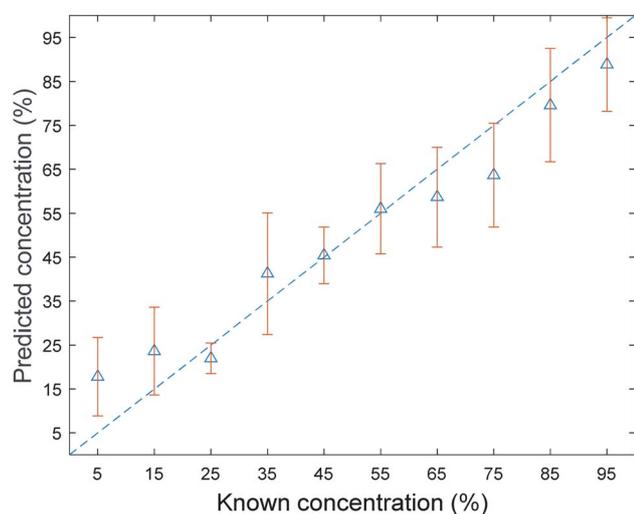
Fig. 4 PC-DFA scores plot.

Fig. 5 PLS-R predicted plotted against known concentrations of grapefruit juice adulteration for PLS-R models trained on the full data from all three combinations of sources. Error bars indicates standard deviations in predictions between all samples with the same level of adulteration.

Table 2 Results of PLS-R models[a]

	No. of LV	Q^2	RMSECV (%)	RMSEP (%)	LoD (%)
All data	3	0.8171	13.20	12.47	11.70
C1–E2	2	0.8891	10.13	9.90	6.51
D1–C2	2	0.8693	10.44	10.18	7.67
A1–B2	2	0.8467	11.23	11.55	8.10
PLS-DA/PLS-R	2/3	0.8322	—	11.95	—

[a] RMSECV and LOD are omitted for two-models results as it used the same single source PLS-R models.

Table 3 Confusion matrix of PLS-DA for predicting the three different fruit juice combinations

	C1–E2 (predicted)	D1–C2 (predicted)	A1–B2 (predicted)
C1–E2 (known)	**86.7%**	6.7%	6.6%
D1–C2 (known)	7.9%	**89.5%**	2.6%
A1–B2 (known)	10%	5.0%	**85.0%**

a classification model the effect of sources was partially mitigated, yet because the classification model was not perfect, the final accuracy was still worse than models on single source.

Discussion and conclusions

Food adulteration has always been a serious threat to consumers' rights and wellbeing. Especially for cases like grapefruit which due to the chemical naringin within the juice, has undesired interactions with multiple medicines.[26–29] Therefore, it is important to have a rapid and portable detection tool to be able to detect such adulteration. Raman spectroscopy had rarely been used as an analytical platform for juice analysis in the past. This is mainly due to pigments in juices have strong fluorescent signals and could easily overwhelm Raman signals.[30,31] However, in this study we demonstrated that a portable Raman spectrometer utilising a NIR excitation laser and collect a relatively narrow range of frequencies, fluorescent interferences can be avoided. The measured spectra combined with chemometrics models can be used to detect and quantify adulteration of grapefruit juice added to orange juice. The prediction errors varied from 9.9 to 12.5% and estimated limit of detection varied from 6.5 to 11.7% which are relatively high compared to other quantification applications using more

sophisticated instrument such as LC-MS. Considering low level adulteration may have little to no economical gain, it is reasonable to expect that deliberate economically motivated adulteration of foodstuff will occur in significant quantities; therefore, a relatively high LoD is acceptable. As, handheld Raman spectrometers can be deployed remotely on site and thus there is no need to transport the juice to a dedicated laboratory for testing. Moreover, as the analysis measures the sample directly and sophisticated sample preparation is not needed which makes it an ideal solution for on-site detection. In conclusion, handheld Raman devices like the one used in this study are ideal for on-site policing within the food supply chain, and such approaches are exciting 'capable guardians' for food security.[32] In addition, we also demonstrated that the fruits' origin has an effect in quantification of the level of adulterations, and this may be due to transportation and storage conditions, length of storage, and/or seasonal effects (note that as these fruits were bought from a local supplier this information was not available to us). A dedicated model for a specific source can achieve better results when applied to samples from the same source. Therefore, in real world applications, one may want to take such variation into account, and it will be beneficial if one can document source of origin and train dedicated models for each of the sources.

Author contributions

Conceptualization, R. G.; methodology, R. G. and M. V.; software: Y. X., R. G.; validation, R. G., Y. X.; formal analysis: M. V.; investigation: M. V., Y. X.; resources, M. V., R. G.; data curation: M. V., Y. X.; writing – original draft preparation, Y. X., M. V.; writing – review and editing: Y. X., M. V., R. G.; supervision, R. G.; project administration, R. G.; funding acquisition, R. G.; All authors have read and agreed to the published version of the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 M. A. Faria, A. Magalhães, M. E. Nunes and M. B. P. P. Oliveira, *Food Control*, 2013, **33**, 136–141.

2 *AIJN Liquid Fruit Market Report*, http://www.aijn.org/publications/facts-and-figures/aijn-market-reports, accessed January 2022.

3 D. R. Petrus, P. J. Fellers and H. E. Anderson, *J. Food Sci.*, 1984, **49**, 1438–1443.

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

5 R. M. Horowitz, in *The Orange, its Biochemistry and Physiology*, ed. W. B. Sinclai, University of California, Riverside, 1961, pp. 334–372.

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

7 M. Chen, S. Y. Zhou, E. Fabriaga, P. H. Zhang and Q. Zhou, *J. Food Drug Anal.*, 2018, **26**, S61–S71.

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

9 Z. Jandrić, M. Islam, D. K. Singh and A. Cannavan, *Food Control*, 2017, **72**, 181–188.

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

11 B. Abad-García, L. A. Berrueta, S. Garmón-Lobato, A. Urkaregi, B. Gallo and F. Vicente, *J. Agric. Food Chem.*, 2012, **60**, 3635–3644.

12 J. M. Obón, M. C. Díaz-García and M. R. Castellar, *J. Food Compos. Anal.*, 2011, **24**, 760–771.

13 D. Giuffrida, P. Dugo, A. Salvo, M. Saitta and G. Dugo, *Fruits*, 2010, **65**, 1–8.

14 S. Ehling and S. Cole, *J. Agric. Food Chem.*, 2011, **59**, 2229–2234.

15 E. Muntean, *Chromatographia*, 2010, **71**, 69–74.

16 M. Twohig, D. Krueger, A. Gledhill and J. Yang, *Agro Food Ind. Hi-Tech*, 2011, **22**, 23–26.

17 M. Spraul, B. Schütz, P. Rinke, S. Koswig, E. Humpfer, H. Schäfer, M. Mörtter, F. Fang, U. C. Marx and A. Minoja, *Nutrients*, 2009, **1**, 148–155.

18 N. S. Vedeanu, D. A. Magdas, L. Bolojan and G. Damian, *Chem. Pap.*, 2011, **66**, 612–616.

19 C. Kurz, M. Leitenberger, R. Carle and A. Schieber, *Food Chem.*, 2010, **119**, 806–812.

20 N. Ogrinc, K. Bat, I. J. Kosir, T. Golob and R. Kokkinofta, *J. Agric. Food Chem.*, 2009, **57**, 6764–6769.

21 L. Palmieri, E. Bozza and L. Giongo, *Nutrients*, 2009, **1**, 316–328.

22 D. I. Ellis, H. Muhamadali, M. Chisanga and R. Goodacre, in *Encyclopedia of Food Chemistry*, ed. L. Melton, F. Shahidi and P. Varelis, Elsevier, Amsterdam, 2019, vol. 1, pp. 364–370.

23 M. Varnasseri, H. Muhamadali, Y. Xu, P. I. C. Richardson, N. Byrd, D. I. Ellis, P. Matousek and R. Goodacre, *Appl. Sci.*, 2021, **11**, 8347–8360.

24 A. C. Olivieri, H. C. Goicoechea and F. A. Iñón, *Chemom. Intell. Lab. Syst.*, 2004, **73**, 189–197.

25 J. Zhang and M. A. Ritenour, *Proc. Fla. State Hortic. Soc.*, 2016, **129**, 178–180.

26 B. Ameer and R. A. Weintraub, *Clin. Pharmacokinet.*, 1997, **33**, 103–121.

27 U. Fuhr and A. L. Kummert, *Clin. Pharmacol. Ther.*, 1995, **58**, 365–373.

28 M. A. Alam, N. Subhan, M. M. Rahman, S. J. Uddin, H. M. Reza and S. D. Sarker, *Adv. Nutr.*, 2014, **5**, 404–417.

29 J. Renugadevi and S. M. Prabu, *Toxicology*, 2009, **256**, 128–134.

30 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.

31 E. Arendse, H. Nieuwoudt, L. S. Magwaza, J. F. I. Nturambirwe, O. A. Fawole and U. L. Opara, *Food Bioprocess Technol.*, 2021, **14**, 1–22.

32 D. I. Ellis, H. Muhamadali, D. P. Allen, C. T. Elliott and R. Goodacre, *Curr. Opin. Food Sci.*, 2016, **10**, 7–15.