Food Chemistry 175 (2015) 1–9

Contents lists available at ScienceDirect

Food Chemistry

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

Rapid Communication

Authentication of beef versus horse meat using 60 MHz ^{1}H NMR spectroscopy

W. Jakes [a,1], A. Gerdova [b], M. Defernez [a], A.D. Watson [a], C. McCallum [a,2], E. Limer [a,3], I.J. Colquhoun [a], D.C. Williamson [b], E.K. Kemsley [a,*]

[a] Analytical Sciences Unit, Institute of Food Research, Norwich Research Park, Norwich NR4 7UA, UK
[b] Oxford Instruments Industrial Analysis, Tubney Woods, Abingdon, Oxfordshire OX13 5QX, UK

ARTICLE INFO

Article history:
Received 4 June 2014
Received in revised form 24 October 2014
Accepted 19 November 2014
Available online 26 November 2014

Keywords:
Speciation
Beef
Horse
Meat
60 MHz ^{1}H NMR
Bench-top NMR
Triglyceride
Authenticity
Chemometrics
Non-targeted screening

ABSTRACT

This work reports a candidate screening protocol to distinguish beef from horse meat based upon comparison of triglyceride signatures obtained by 60 MHz ^{1}H NMR spectroscopy. Using a simple chloroform-based extraction, we obtained classic low-field triglyceride spectra from typically a 10 min acquisition time. Peak integration was sufficient to differentiate samples of fresh beef (76 extractions) and horse (62 extractions) using Naïve Bayes classification. Principal component analysis gave a two-dimensional "authentic" beef region ($p = 0.001$) against which further spectra could be compared. This model was challenged using a subset of 23 freeze–thawed training samples. The outcomes indicated that storing samples by freezing does not adversely affect the analysis. Of a further collection of extractions from previously unseen samples, 90/91 beef spectra were classified as authentic, and 16/16 horse spectra as non-authentic. We conclude that 60 MHz ^{1}H NMR represents a feasible high-throughput approach for screening raw meat.

© 2014 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

1. Introduction

In January 2013, the Food Safety Authority of Ireland announced the discovery of horse meat in a number of beef burgers, heralding a pan-European meat authenticity crisis. In the UK, an urgent investigation by the Food Standards Agency (FSA) found several beef products that contained horsemeat, resulting in large-scale removal of products from supermarket shelves (Food Standards Agency, 2013). Several retailers and suppliers were embroiled in the crisis, as more and more beef products were found to contain undeclared horse meat.

There was no suggestion that horse meat is a health hazard *per se*. However, the presence of horse meat in a food chain where none should be present implies failures in mechanisms designed to guarantee food provenance and safety. Such failings open the door to health issues, since meat unfit for human consumption might be able to enter the food supply chain. Also, an incidence such as this constitutes a fraud – the consumer is paying for one thing but being sold a cheaper substitute.

As with all types of authenticity, policing and prevention depends, in part, on reliable means of testing either for product purity or for the presence of an adulterant. There are several ways of detecting horse meat as an adulterant in beef. The original results from Ireland relied on DNA, and in the UK the FSA has accumulated results from tens of thousands of DNA-based tests for horse in beef products. DNA testing has the potential advantage that it is species specific, but it is relatively slow and expensive. Species determination of meat via mitochondrial DNA is relatively straightforward given a target species, although the same methods will not give reliable 'weight for weight' (w/w) quantitation results for meat adulteration. DNA-based methods, particularly the issue of quantitation, are reviewed in Ballin, Vogensen, and Karlsson (2009).

Other methods target proteins. Of these the best known is ELISA, an immunological technique able to give species detection

* Corresponding author. Tel.: +44 (0)1603 255014.
E-mail address: kate.kemsley@ifr.ac.uk (E.K. Kemsley).
[1] Present address: School of Chemistry, University of East Anglia, Norwich Research Park, NR4 7TJ, UK.
[2] Present address: Worcester College, University of Oxford, Oxford OX1 2HB, UK.
[3] Present address: Oriel College, University of Oxford, Oxford OX1 4EW, UK.

http://dx.doi.org/10.1016/j.foodchem.2014.11.110
0308-8146/© 2014 The Authors. Published by Elsevier Ltd.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/3.0/).

and which, like DNA-based testing, is readily available commercially. A range of analytical methods including HPLC, GC and mass spectrometry have been employed to examine protein and various other properties of meat (Ballin, 2010; von Bargen, Dojahn, Waidelich, Humpf, & Brockmeyer, 2013).

In this work we focus on the triglyceride content of meat. The idea of exploiting triglyceride content as a marker for horse meat is not new: Paschke (1938) introduced a chemical method for the detection of horse meat in mixtures with beef, mutton or pork based on the relatively high level of linolenic acid, C18:3, in horse fat. Since then, numerous authors have reported the triglyceride composition of horse meat, including some that make comparisons with other meats (Chernukha, 2011; He, Ishikawa, & Hidari, 2005; Lisitsyn, Chernukha, & Ivankin, 2013; Lorenzo et al., 2014). Relative to beef, in addition to higher levels of linolenic acid, horse meat is higher in polyunsaturated fatty acid (PUFA), but lower in saturated fatty acid (SFA) and monounsaturated fatty acid (MUFA). For example, for C18:3, He et al. quote 1.47% of total detected fatty acid (longissimus dorsi muscle) for horse versus 0.15% for beef (Holstein steer). The factor of ~10 difference is indicative of linolenic acid's potential as a horse versus beef marker (He et al., 2005). He et al. also quote SFA (horse = 34.37 versus beef = 42.83%, total detected fatty acid), MUFA (horse = 50.43 versus beef = 52.80%) and PUFA (horse = 15.20 versus beef = 4.37%) for particular groups of animals with specified diet. Note that data from different authors shows considerable scattering: for example, the intramuscular fat level of C18:3 ω-3 (α-linolenic) fatty acid level in Galician foals, a horse relic from Ice Age times, has been quoted at 23.87% of total fatty acid content (Lorenzo, Fucinos, Purrinos, & Franco, 2010).

High-resolution, low-field (1.4 T, 60 MHz) bench-top spectrometers are a relatively recent development in NMR technology, which we have previously found to be effective for the analysis of another class of triglyceride-rich samples, vegetable and nut oils (Parker et al., 2014). High-field NMR, on the other hand, is well-established for the study of edible oils (Fang, Goh, Tay, Lau, & Li, 2013; Guillen & Ruiz, 2001; Johnson & Shoolery, 1962; Longobardi et al., 2012). Several authors have quantified the triglyceride mix of edible oils, and in some cases animal fats, based on the integration of spectrum peak areas (Barison et al., 2010; Guillen & Ruiz, 2003a, 2003b; Knothe & Kenar, 2004; Sedman, Gao, Garcia-Gonzalez, Ehsan, & van de Voort, 2010; Shiao & Shiao, 1989). Siciliano et al. used peak area integration to study pork fatty acid composition of two salami products during ripening, though such meat-specific applications are rare in the literature (Siciliano et al., 2013). Peak-area based quantitation has also been used in a low-field environment in a medical context. For example, Szczepaniak et al. used a 1.5 T whole-body NMR scanner to measure intracellular triglyceride stores *in vivo* (Szczepaniak et al., 1999). The key point underpinning the peak area approach is that the area of a spectrum peak is proportional to the number of protons associated with that peak.

These studies demonstrate that ^{1}H NMR is a useful tool for both triglyceride quantitation and sample classification. In the present work, we combine these threads to develop low-field ^{1}H NMR as an authentication tool based on the triglyceride content of meats from different species (patent pending). Specifically, we propose that NMR can provide a compositional profiling approach to verify beef authenticity against a known potential adulterant, horsemeat. Bearing in mind the aims, constraints and limitations of high-throughput screening, a simple chloroform-only extraction was used and spectra acquired with a high-resolution, low-field bench-top spectrometer. Spectral information relevant to the characterisation of beef versus horse meat is extracted and modelled. We report here on the success and robustness of this approach.

2. Materials and methods

2.1. Samples

Fresh meat samples were purchased from a variety of outlets (supermarkets and butchers) in England, France and Belgium. Additional frozen samples were obtained via commercial importers. The stated meat origin was UK or Ireland (meat bought in England), France or Belgium (bought there) and South America or France (commercial importers). The samples included a variety of cuts as well as mince. Meat that had been further processed (e.g. sausages) was generally avoided, as it would be impossible to confirm the species of such samples through visual inspection.

Three collections of triglyceride extracts were prepared, as summarized below. Further details on the source, nature, storage and replication of the samples are given in Table 1. The sample preparation procedure is described in Section 2.2.

2.1.1. 'Training Set' samples

Researchers at Oxford Instruments ('Lab 1') purchased 9 beef and 4 horse samples, from which 46 and 20 extracts were prepared for NMR analysis, respectively. Researchers at the Institute of Food Research ('Lab 2') purchased 10 beef and 15 horse samples, from which 30 and 42 extracts were prepared, respectively. Since only small quantities of meat are required for each extraction, the remainders of each of Lab 2's samples were stored at −40 °C.

2.1.2. 'Test Set 1' – a freeze–thawed subset of the Training Set samples

Approximately 6 weeks after acquiring the Training Set data, Lab 2 recovered and thawed 8 beef and 15 horse samples from the −40 °C store and prepared single extractions from each.

2.1.3. 'Test Set 2' – new samples

Lab 1 purchased a further 27 beef samples, from which 79 extracts were prepared for NMR analysis. Lab 2 purchased 4 beef and 6 horse samples, from which 12 and 16 extractions were prepared, respectively. The total numbers of beef and horse extracts prepared across both labs were 91 and 16, respectively. The role of these test samples was to challenge the authenticity model created from the Training Set samples.

In addition to extracts from meat samples, Lab 2 prepared a small collection of samples from three laboratory-grade triglycerides (Sigma–Aldrich): glyceryl tristearate (C18:0), glyceryl trioleate (C18:1) and glyceryl trilinolenate (C18:3). A stock mixture was prepared containing 15% w/w C18:0 and 85% w/w C18:1. This was used to make four triglyceride mixtures containing 0%, 10%, 20% and 30% w/w of C18:3, respectively. These were diluted with approximately 80% by volume of chloroform before NMR analysis.

2.2. Meat sample preparation

Both Lab 1 and Lab 2 used similar, simple preparation and extraction procedures, with the aim of establishing a protocol appropriate for a low-cost, high-throughput screening scenario. No attempt was made to determine the extraction efficiency, since the objective was to obtain representative compositional profiles suitable for speciation, rather than absolute quantitation. The extraction agent was deuterated chloroform (Lab 1) or chloroform (Lab 2), which is well-suited for the extraction of neutral lipids such as triglycerides.

The preparation for the Training Set samples at Lab 1 was as follows: a small amount of meat was cut into pieces (~1 cm^3) and homogenised in a food processor (Kenwood mini-chopper) for 30 s. Next, 1.5 ml of deuterated chloroform (Sigma–Aldrich) was added to 3–6 g homogenised meat (depending on fattiness; the

Table 1
Horse and beef samples' description and extraction numbers.

Approximate batch preparation and spectral acquisition date	Lab[a]	Species	F versus FT[b]	Supplier	Cut of meat[c]	No. samples	No. extractions	No. spectra
Training Set samples								
August 2013	1	Beef	F	UK supermarket	Mince	3	5	15
October 2013	1	Beef	F	UK supermarket	Mince	2	5	10
		Beef	F	UK supermarket	Steak	3	5	15
		Beef	F	UK supermarket	Steak	1	6	6
		Horse	F	French butcher	Diced	1	5	5
		Horse	F	French butcher	Sausage	1	5	5
		Horse	F	French supermarket	Steak	2	5	10
December 2013	2	Beef	F	UK supermarket	Mince	3	3	9
		Beef	F	UK supermarket	Steak	7	3	21
		Horse	F	French butcher	Diced	7	3	21
		Horse	F	French butcher	Roasting joint	1	2	2
		Horse	F	French butcher	Steak	5	3	15
		Horse	F	French butcher	Steak	2	2	4
Test Set 1, freeze–thawed Training Set samples								
January 2014	2	Beef	FT	UK supermarket	Mince	2	1	2
		Beef	FT	UK supermarket	Steak	6	1	6
		Horse	FT	French butcher	Diced	7	1	7
		Horse	FT	French butcher	Roasting joint	1	1	1
		Horse	FT	French butcher	Steak	7	1	7
Test Set 2, new samples								
January 2014	1	Beef	F	UK supermarket	Mince	1	1[d]	1
		Beef	F	UK supermarket	Mince	2	2[d]	4
		Beef	F	UK supermarket	Mince	1	3[d]	3
		Beef	F	UK supermarket	Mince	1	11[d]	11
		Beef	F	UK supermarket	Steak	6	1[d]	6
		Beef	F	UK supermarket	Steak	5	2[d]	10
		Beef	F	UK supermarket	Steak	5	3[d]	15
		Beef	F	UK supermarket	Steak	2	4[d]	8
		Beef	F	UK supermarket	Steak	3	5[d]	15
		Beef	F	UK supermarket	Steak	1	6[d]	6
January 2014	2	Horse	FT	French butcher	Steak	1	1	1
March 2014	2	Beef	F	UK supermarket	Mince	1	3	3
		Beef	F	UK supermarket	Steak	3	3	9
		Horse	F	Belgium butcher	Diced	1	3	3
		Horse	F	Belgium butcher	Steak	2	3	6
		Horse	FT	Commercial importer	Mince	1	3	3
		Horse	FT	Commercial importer	Steak	1	3	3

[a] Lab 1 = Oxford Instruments, Lab 2 = Institute of Food Research.
[b] F = fresh meat, FT = meat supplied frozen then thawed, or supplied fresh then frozen in-house and subsequently thawed.
[c] Diced meat tended to be of a visibly higher fat content than steak.
[d] Samples for which the NMR analysis entailed variable numbers of scans and relaxation delay (RD) times.

lowest quantities were used for visibly fatty samples) and the mixture vortexed for 10 min before being refrigerated for 1 h at 4 °C. The solvent extract was then recovered by pipette, filtered through paper tissue and placed in a 5 mm disposable NMR tube (Sigma–Aldrich). All samples were stored at 4 °C until NMR data were collected. Replicate extractions were obtained by homogenising a representative cut of meat, and then preparing separate extractions from discrete subsamples. The order in which extracts were presented to the spectrometer was randomised within each batch.

For the Test Set 2 samples, Lab 1's procedure was modified slightly. In particular, the amount of sample mixed with deuterated chloroform was not weighed, and the mixture was not refrigerated after vortexing.

Lab 2's preparation for all meat samples was the same as that used by Lab 1 for the Training Set samples, with the following variations. Approximately 10 g of meat was homogenised. For each extraction, non-deuterated chloroform (analytical grade, Sigma–Aldrich) was added to a 5 ± 0.05 g subsample of the homogenised meat. The extract was filtered through compacted, non-absorbent cotton wool (Fisher Scientific).

2.3. 60 MHz ^{1}H spectra

60 MHz ^{1}H NMR spectra were acquired on Pulsar low-field spectrometers (Oxford Instruments, Tubney Woods, Abingdon, Oxford, UK) running SpinFlow software (v1, Oxford Instruments). Both Lab 1 and Lab 2 had their own instrument. The sample temperature was 37 °C, and the 90° pulse length was ~7.2 μs as determined by the machine's internal calibration cycle. No resolution enhancement methods were applied to the spectral data.

At Lab 1, a variable number of FIDs were collected, with the aim of achieving a target signal-to-noise ratio. This strategy was inspired by the relatively poor signal-to-noise ratio of the horse extract spectra, which is in turn due to the low fat content of horse

meat. For the Training Set, the relaxation delay (RD) was set to 30 s but for the Test Set 2 samples, Lab 1 varied the RD from 2 to 30 s, the time range arising from balancing the need to reach relaxation equilibrium against the drive for a short total acquisition time.

In contrast, at Lab 2, the same acquisition parameters were used throughout. Sixteen FIDs were collected from each extraction with a fixed RD of 30 s, resulting in a standard acquisition time of ~10 min per extract. Lab 1 performed more shimming and pulse calibration runs than Lab 2. The different approaches reflect the emphasis in Lab 2 on standardisation and cost minimisation, in contrast with Lab 1's emphasis on spectral quality.

In all cases, the FIDs were Fourier-transformed, co-added and phase-corrected using SpinFlow and MNova (Mestrelab Research, Santiago de Compostela, Spain) software packages to present a single frequency-domain spectrum from each extract. Lab 1 also used MNova to manually improve the phase correction whereas Lab 2 did not, opting instead for a less subjective, automated approach. All spectra were initially referenced to chloroform at 7.26 ppm.

2.4. High-field ^{1}H spectra

For the purpose of comparison, a high-field 600 MHz ^{1}H NMR spectrum was collected at Lab 2 from an extract of horse (randomly chosen from Test Set 1), using a Bruker Avance III HD spectrometer running TopSpin 3.2 software and equipped with a 5 mm TCI cryoprobe. The original sample was dried down and the lost chloroform replaced with deuterated chloroform. The probe temperature was regulated at 27 °C. The spectrum was referenced to chloroform at 7.26 ppm.

2.5. Data analysis

All data visualisation and processing of the frequency-domain spectra was carried out in Matlab (The Mathworks, Cambridge, UK).

Before any quantitative analysis, spectra were re-aligned on the frequency scale by sideways shifting using the glyceride peak maximum as the reference point (Parker et al., 2014). The area of the group of glyceride resonances was used to normalise the intensity of each spectrum. To develop the authentication models, selected regions corresponding to the olefinic, glyceride, bis-allylic and terminal CH_3 resonances were extracted from each spectrum to form a dataset of reduced size. Each region was baseline-corrected separately. For the olefinic and glyceride peaks, baselines were calculated using polynomial fitting. For the bis-allylic and terminal CH_3 resonances, which are not well isolated, baselines were fitted using a Lorentzian function to account for contributions from the wings of neighbouring resonances. The integrated olefinic and bis-allylic peak areas were used in a Naïve Bayes classification model. The olefinic, bis-allylic and terminal CH_3 regions were concatenated and used as input in a principal component analysis (PCA).

3. Results and discussion

Visual assessment indicated that the meat samples varied quite considerably in their fat content. This affected the concentration of triglycerides present in the NMR tube, manifesting as large variations (up to an order of magnitude) in the intensity of the triglyceride signals and hence signal-to-noise across the collection of raw spectra. The Lab 1 protocol mitigated this effect somewhat, by collecting and co-adding FIDs until a nominal minimum signal-to-noise was achieved, although in some instances this entailed total acquisition times of several hours. At Lab 2, in contrast, only 16 FIDs were co-added throughout, so very low-fat samples in particular exhibit comparatively poor signal-to-noise. However, in Lab 2 the spectral acquisition time was kept to ~10 min for all samples.

The data normalisation step scaled the raw responses in each spectrum so that they could be readily examined on a single set of axes. Furthermore, through division by the glyceride peak areas, the responses were mapped onto a meaningful "per-glyceride" vertical scale. This means that the concentrations of chemical species present in different samples can be directly compared by examining the normalised spectra plotted on a common set of axes.

An exemplary collection of spectra (Training Set, Lab 2 data) is shown in Fig. 1. For clarity, the groups of spectra from the two meat species are vertically offset with respect to one another. In broad terms, these are typical 60 ^{1}H MHz spectra of triglycerides that contain a range of long-chain fatty acids with differing amounts of unsaturation. Some of the key spectral regions are indicated, based on the assignment given for 60 MHz ^{1}H NMR of triglycerides by Parker et al. (2014). It can be seen that there is more variation amongst the spectra from horse samples compared with those from beef and, furthermore, that some of the former are considerably noisier and thus are distinguished more easily in the overlaid spectra of Fig. 1. This is likely a consequence of the generally lower fat content of horse compared to beef.

The regions outlined by dotted rectangles can be attributed to distinct chemical species. The peaks centred at ~4.2 ppm ("glyceride") arise from ^{1}H nuclei attached to carbon at positions 1 and 3 on the glycerol backbone. This is a useful group of peaks, as its integrated area provides a direct measure of the glyceride concentration in the sample, hence its use as an internal reference in our pre-processing procedure.

The set of peaks at ~5.2 ppm ("olefinic") were largely from the ^{1}H nuclei attached to carbons involved in a double bond. This signal is thus related to the total number of unsaturated bonds in a triglyceride, regardless of whether these are located within mono-unsaturated or poly-unsaturated chains. The olefinic region contains a ^{13}C satellite peak at ~5.5 ppm attributable to the use of non-deuterated chloroform by Lab 2.

The very small signals at ~2.7 ppm ("bis-allylic") arose from bis-allylic protons from the $-CH_2-$ groups located between pairs of double bonds and thus provides a measure of the number of poly-unsaturated fatty acid chains present in the sample. Note that these are visible only in the spectra from horse.

Finally, the region around 0.9 ppm ("terminal methyl, CH_3") arises from the protons attached to the terminal carbon of each fatty acid chain. For a triglyceride there will be contributions from each of the three terminal CH_3 groups per single glycerol backbone.

Fig. 1 suggests that there are systematic differences between the spectra from the two species, but this becomes much more apparent when selected parts of the spectrum are viewed on a magnified scale. Fig. 2 shows the olefinic, glyceride, bis-allylic and terminal CH_3 regions, each on an appropriate vertical scale, from the entire collection of Training Set spectra, presented separately for each species and Lab.

Due to normalisation, the glyceride peak areas are the same (equal to unity) in all spectra. Fig. 2 reveals that the peaks from Lab 1 are slightly sharper than those from Lab 2. This is probably attributable to known technical improvements in Lab 1's spectrometer relative to the instrument used in Lab 2, and also a more comprehensive strategy of magnet shimming and pulse calibration by Lab 1.

It can be seen that horse spectra consistently exhibit larger olefinic and much larger bis-allylic peaks than beef, indicating a higher unsaturated fat content in the horse samples. This is in agreement with reports in the literature relating to distinct fatty acid compositions of different species (Dobranic, Njari, Miokovic, Fleck, & Kadivc, 2009; He et al., 2005; Lisitsyn et al., 2013; Tonial

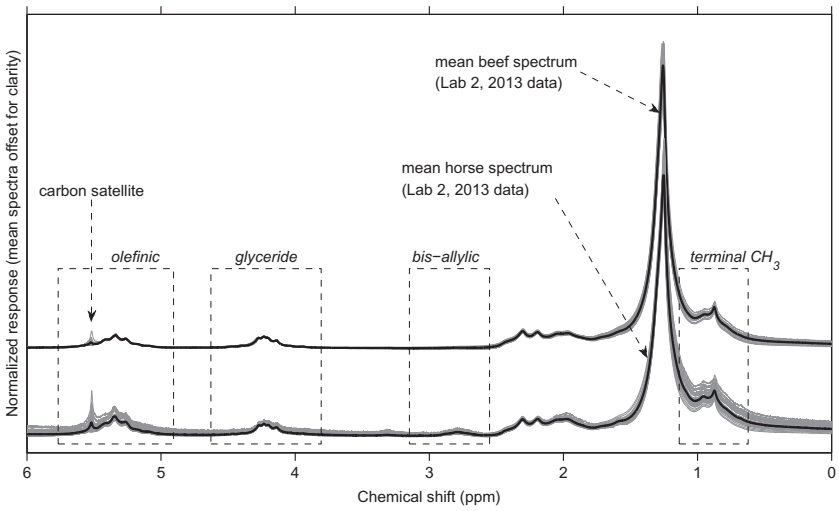

Fig. 1. ^{1}H 60 MHz NMR spectra (Lab 2 training data) for beef (upper trace) and horse (lower trace), displaced vertically for clarity and normalised to the glyceride peak area. Rectangles highlight four regions of interest.

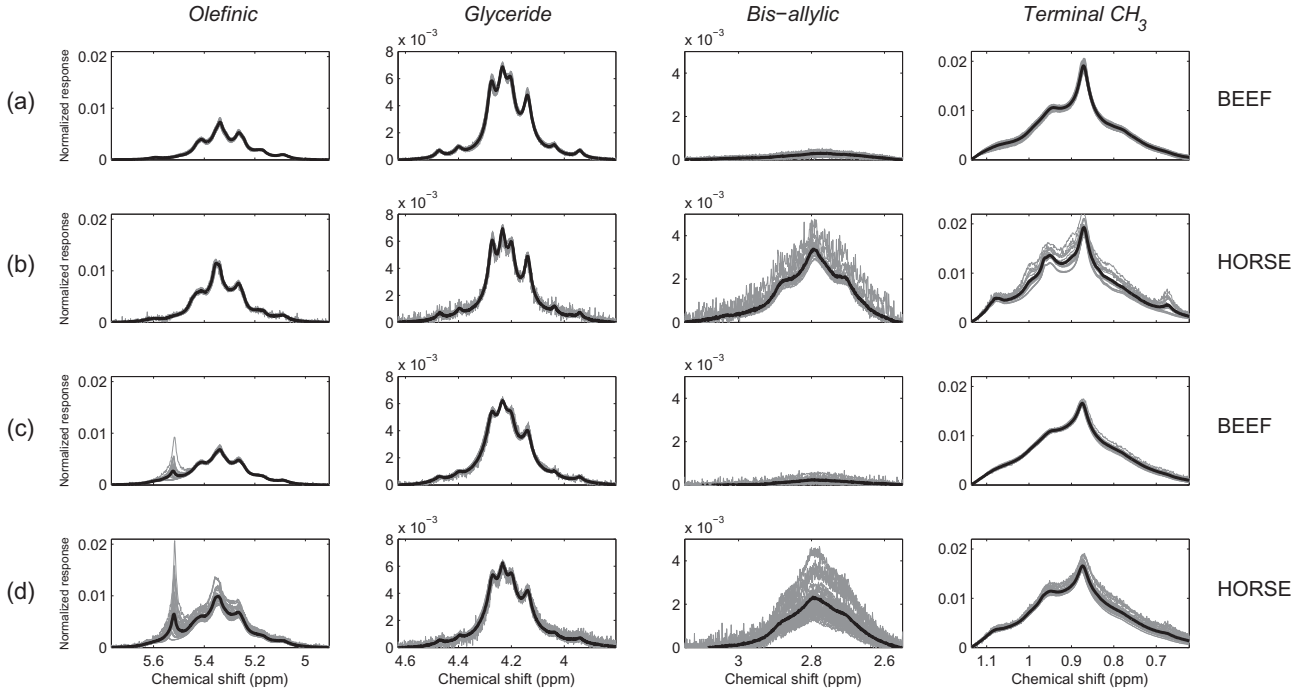

Fig. 2. The four spectral regions of interest for the entire training dataset; Lab 1 (a) beef and (b) horse and Lab 2 (c) beef and (d) horse.

et al., 2009) and suggests that simple integrated peak areas may be used to distinguish species in a quantitative manner. Naïve Bayes classification was applied to the integrated olefinic and bis-allylic peak areas only, calculated from the Training Set data. 100% correct classifications were obtained for both the beef and horse groups. Furthermore, the method employed crossover validation: Lab 1 data were used to predict Lab 2, and vice versa. Not only is this a promising outcome in terms of efficacy of the methodology, it also implies that the difference between Labs (extraction procedure, researcher and spectrometer) is not adversely affecting the ability to distinguish species.

There are additional differences in the two species' data in the terminal CH_3 region, highlighted by the larger number of peaks visible in the horse spectra, especially for Lab 1 data. Fig. 3(a) shows the mean of the Training Set beef and horse spectra from Lab 1. To aid in annotation, these were compared with a high-field 600 MHz ^{1}H NMR spectrum of a single randomly chosen horse sample from Lab 2 (Fig. 3(b); peaks annotated based on Vinaixa et al. (2010)), and with spectra from the series of triglyceride mixtures prepared at Lab 2 (Fig. 3(c)).

The horse spectrum in Fig. 3(a) is qualitatively very similar to the spectra of mixtures with a C18:3 constituent (Fig. 3(c)), consistent with the presence of an appreciable C18:3 component in the extracts from horse meat. Comparison with the high-field spectrum in Fig. 3(b) helps interpretation. Linolenic acid C18:3 ω-3 (α-linolenic acid) contains a double bond close to the terminal CH_3 that is known to cause a shift to higher ppm values (from 0.87 to 0.97, high-field NMR values) (Alonso-Salces, Holland, & Guillou, 2011). We found peaks at both 0.87 and 0.97 ppm in the high-field horse meat spectrum (Fig. 3(b)) and in the low-field spectra of both horse and C18:3 containing mixtures (Fig. 3(a) and (c)). Note that the outer peaks of the two triplets in panel (b) derive from a coupling constant value in Hz that is independent of field strength, which is why they appear at different values in

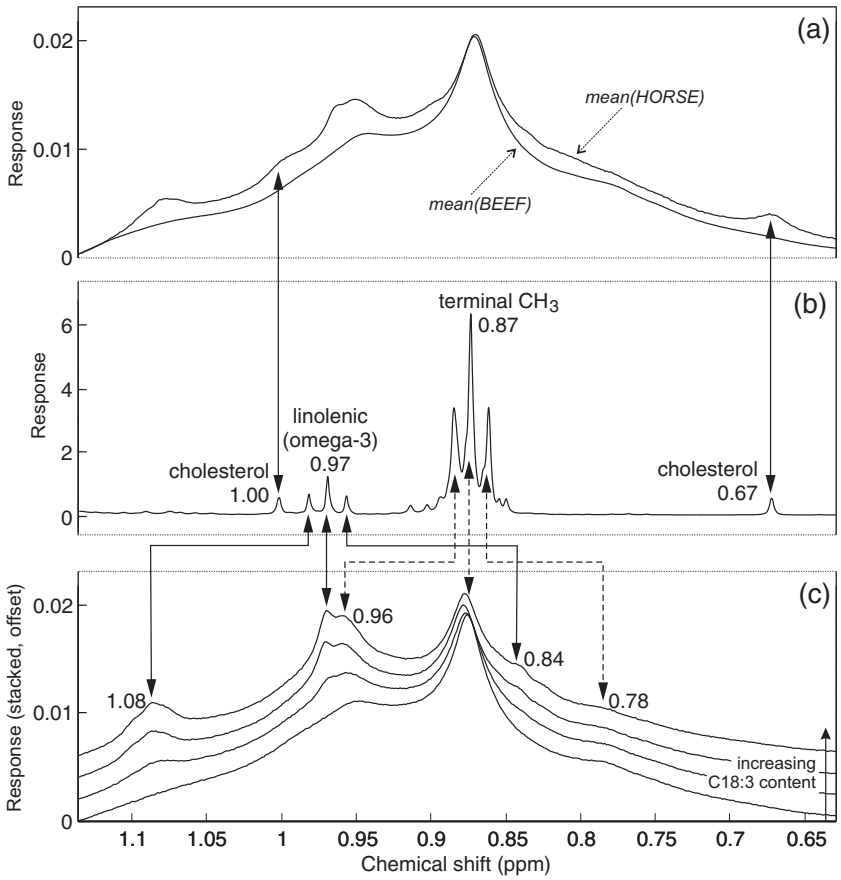

Fig. 3. Terminal CH_3 region for (a) horse and beef 60 MHz mean spectra, compared to (b) horse 600 MHz spectrum and (c) triglyceride mixtures 60 MHz spectra. Numbers indicate the chemical shifts of various peaks and arrows indicate peaks identified across panels (note that the outer peaks of triplets appear at different ppm values for 600 and 60 MHz data).

600 MHz (b) and 60 MHz (c) spectra. This also results in the third peak of the α-linolenic acid triplet appearing at 0.84 ppm in the 60 MHz spectra and being obscured by a terminal CH_3 peak at 0.78 ppm.

In contrast, the beef spectrum more closely resembles that of the C18:0 + C18:1 mixture. This is consistent with beef having essentially no C18:3 content. Therefore, linolenic acid, previously identified as a marker for horse meat versus beef, has an NMR signature in the form of a shifted terminal CH_3 peak combined with a bis-allylic peak. Note however that in the C18:3 ω-6 (γ-linolenic acid) isomer, the relevant double bond is further away from the CH_3 terminal so does not give rise to the same shift. Therefore, for C18:3 ω-6 (γ-linolenic acid) the CH_3 peak is at 0.866 ppm, indistinguishable from those for saturated, oleic and linoleic acids. In other words, the NMR shifted-CH_3 marker is not related to total linolenic acid, but specifically to the α-linolenic acid content.

The high-field data also helps to identify two peaks visible in the mean horse spectra, but absent in the beef extracts and triglyceride mixtures. These are at 0.67 and 1.00 ppm, and are due to cholesterol (Vinaixa et al., 2010). Such cholesterol peaks appear in some, but not all, of the individual horse spectra and are most apparent in those extracts with the lowest overall triglyceride concentration. This is a consequence of the inflating effect of normalising by the glyceride peak area. This effect is most pronounced for the weakest samples, which are consistently the extracts from horse, as the leaner of the two meat species.

PCA was applied to datasets of normalised intensities obtained by concatenating the olefinic (NB: truncated at 5.39 ppm to exclude the carbon satellite region), bis-allylic and terminal CH_3 regions of Fig. 2, treating each Lab's Training data separately. The first two PC scores are plotted against one another in Fig. 4(a) and (b), with symbols coded according to species. In both cases, the first dimension contains most of the relevant information relating to the difference between the two species. Furthermore, regions of the loading corresponding to the olefinic and bis-allylic peaks are positively associated with horse samples (Fig. 4(c) and (d)); this is as expected, given the performance of the Naïve Bayes classification using just these integrated peak areas reported above. The loadings in the terminal CH_3 region show considerable detail, including peaks at 1.08 ppm, 0.96 ppm and 0.84 ppm that tally with those in Fig. 3 and are associated with increasing C18:3 content, and peaks at 1.00 ppm and 0.67 ppm linked to cholesterol.

For comparison, Fig. 4(c) and (d) also include second traces showing the covariance of each dataset with the group membership data; projections onto this vector have scores with maximally separated group means (Kemsley, 1996). The similarity between these covariance vectors and the first PC loadings confirm that the greatest source of variation in both datasets arises from the difference between the two species.

From these results, we concluded that any effects due to differences between the Labs (arising from extraction procedures, researchers, instrumentation, etc.) were insignificant compared with the variance due to species. Thus the Training Set data from both Labs were combined and used to develop a single authentication model. PCA was applied to this pooled dataset. The scores on the first two axes are shown in Fig. 5(a). Plotting the horse data from each Lab with different symbols confirms that there is no systematic difference between labs to be seen (note there is too much

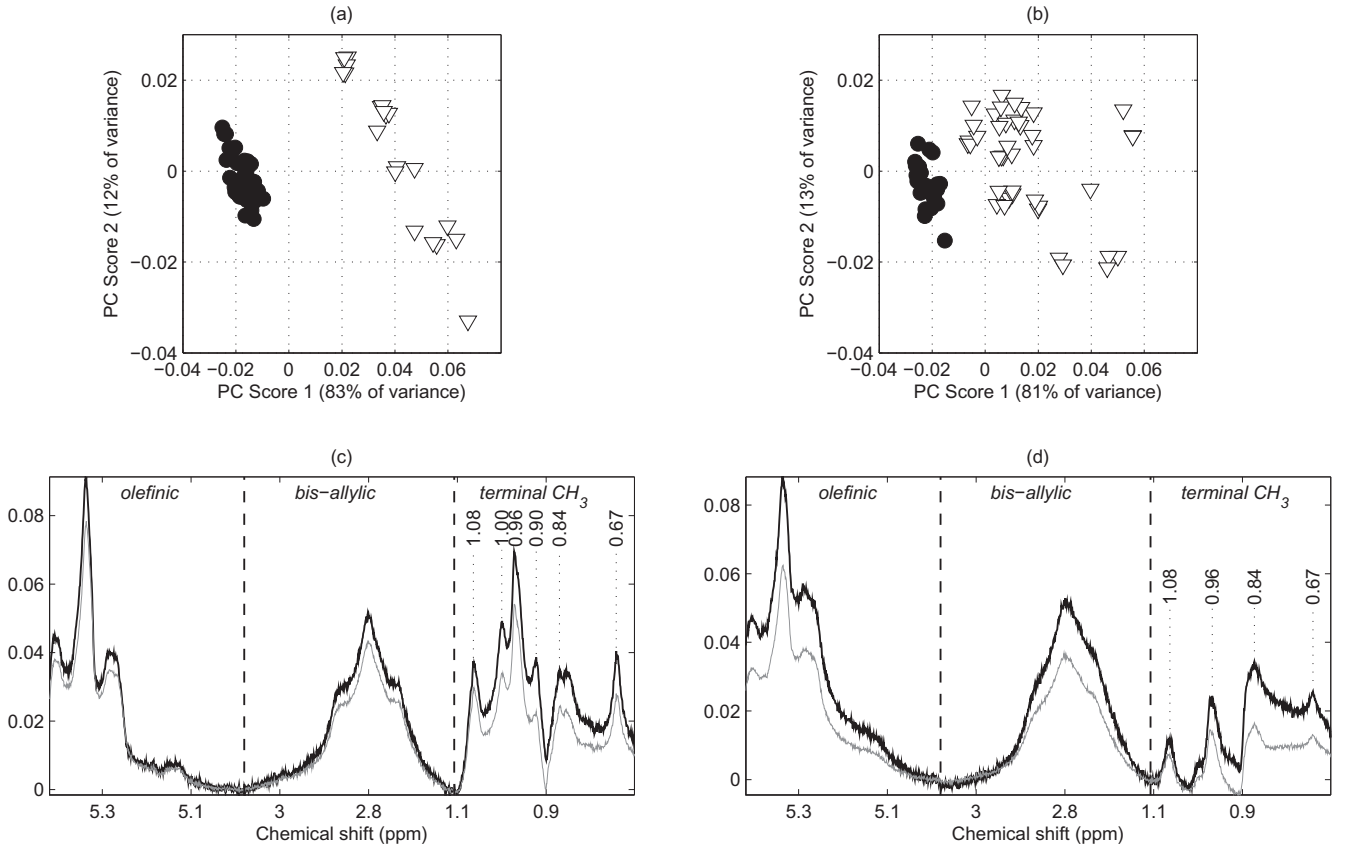

Fig. 4. First versus second principal component scores plots for (a) Lab 1 Training Set data, and (b) Lab 2 Training Set data (black disks = beef, open triangles = horse). (c) and (d) Corresponding loadings plots (black trace), together with the covariance of each dataset with the group membership (grey trace) and peaks picked from the loadings in the CH_3 region.

overlap of points to illustrate this clearly for the beef samples). The loading vectors (data not shown) are highly similar to those from the Training Set data treated separately, as might be expected.

Note again that ~95% of the information content is contained in the first two PC dimensions, thus the scores can be used to represent the beef and horse groups in a compact way. The relative spreads of the two groups indicates much greater variability of horse compared with beef samples. This is also evident when plotting the normalised, integrated areas of the olefinic versus the bis-allylic peaks (data not shown). We do not believe this is attributable to experimental or data processing issues (see discussion of Fig. 5(d) below); rather, the lower variability shown by the triglyceride content of our beef samples is likely due to the similar diets, gender, geographic origin and age of slaughter of the cattle. In contrast, our horse samples come from several different countries with potentially greater variation in farming practices and, in turn, fatty acid composition (Lorenzo et al., 2010; Lorenzo et al., 2014).

Whilst successful outcomes were obtained in the Naïve Bayes analyses reported above, the underlying assumption of equal group variances is potentially open to challenge given the higher variance of the horse data relative to beef. An alternative to the two-group classification approach is to focus on the 'authentic' group only, here beef, and consider anything else as 'non-authentic'. In this study, horse is used as an exemplary non-authentic material, because it has been a key undeclared ingredient in recent incidences of fraud. The non-authentic group could of course encompass any meats that are not pure beef.

Conceptually the approach is as follows: for any given spectrum, the null hypothesis H_0 is that it belongs to the authentic group; H_0 is then tested at the desired significance level by calculating some statistic and comparing it with a critical value. Working in the PC coordinate system, we can equate this to a boundary drawn around the authentic group, derived from the covariance matrix of the authentic samples and expressed as a line of constant Mahalanobis D^2 from the group centre. Using just the first two PC dimensions, since these contain ~95% of the original information content, the boundary is represented by an ellipse, shown in Fig. 5(a) for the $p = 0.001$ critical value, corresponding to $D^2 = 13.82$ (an assumption in this approach is that the D^2 values come from a χ^2 distribution with two degrees of freedom, and this was confirmed by a probability plot (not shown) of D^2 versus χ^2).

Note the choice of significance level is arbitrary and can be chosen to meet the needs of the application under consideration. Using $p = 0.001$, the chance of rejecting an authentic beef sample (i.e. incorrectly rejecting H_0, a Type I error) is 0.1%. It can be seen from Fig. 5(a) that none of the beef samples fall outside this boundary – since only 76 samples are included here, this is consistent with the significance level.

It is harder to estimate the chance of incorrectly accepting a non-authentic (substituted or adulterated) sample as authentic beef (i.e. of incorrectly accepting H_0, a Type II error). This is the case for all problems of this nature, since the non-authentic population is open-ended. The pragmatic solution is simply to state the error rate obtained from the samples belonging to specific types of non-authentic samples.

We investigated the fitness of our model by confronting it with sets of unseen data (Test Sets 1 and 2, see Table 1). These data were pre-processed and reduced as described above, and then rotated into PC space using the parameters (centering and loading vectors) obtained from the combined Training Set data.

Fig. 5(b) shows the scores for Test Set 1 samples (see Table 1). Recall these were all originally fresh beef and horse samples used

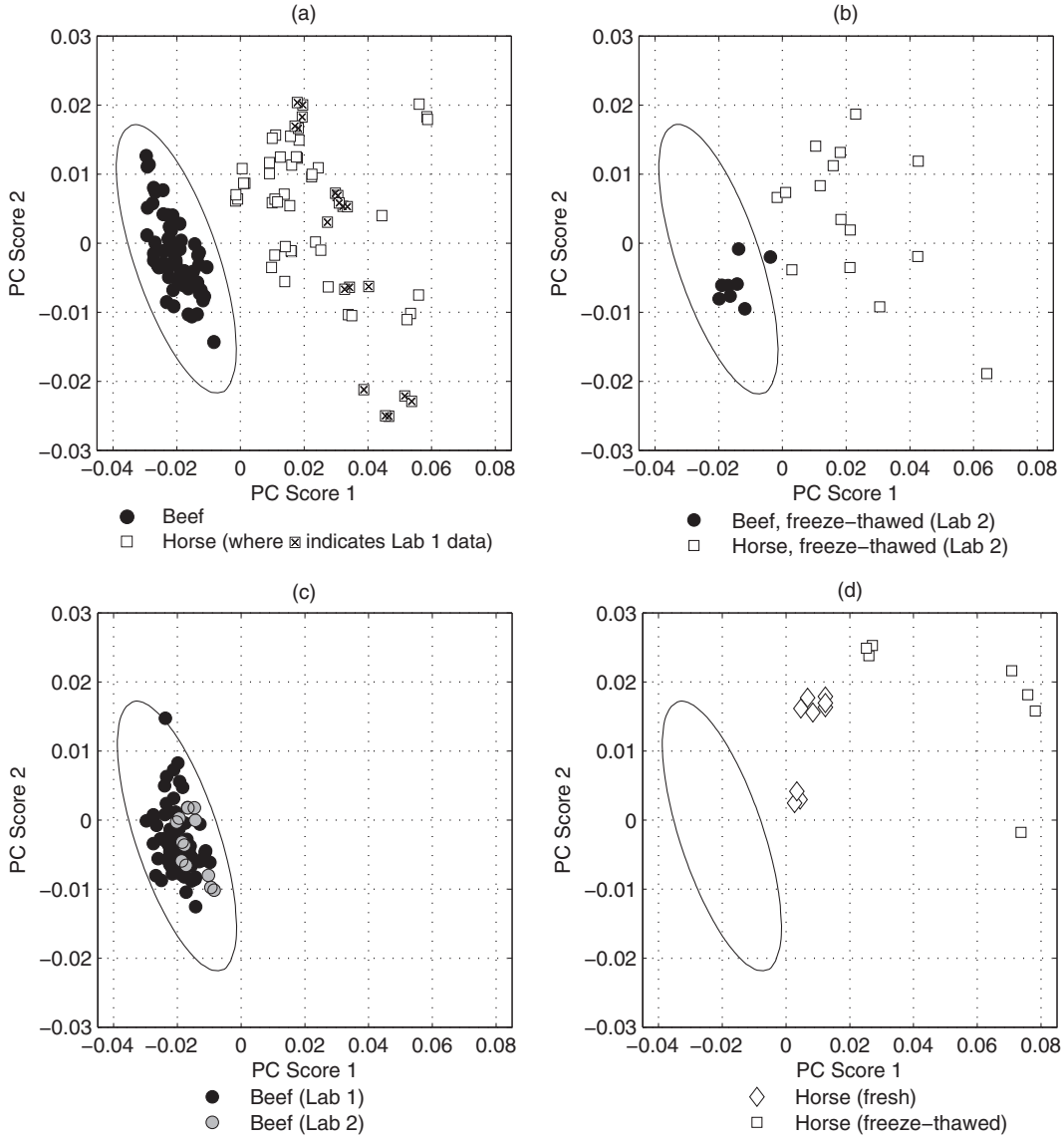

Fig. 5. First versus second principal component plots of: (a) the entire Training Set data (black disks = beef, squares = horse (containing 'x' for Lab 1, open for Lab 2)); (b) Test Set 1, (c) Test Set 2, beef, (d) Test Set 2, horse. On all plots, an ellipse is shown indicating the line of constant Mahalanobis distance (D^2 = 13.82) from the beef group centre.

in the Lab 2 Training Set, but were then frozen, stored and thawed to become Test Set 1. A single beef data point lies just outside the ellipse. This represents a Type I error, the rejection of an authentic sample. No horse data points appear inside the ellipse, meaning that there are no Type II errors. From this we conclude that freeze–thawing samples does not impact on the model's capacity to identify samples as authentic beef or 'non-authentic'.

Fig. 5(c) and (d) shows the outcomes for Test Set 2 samples (see Table 1), for beef and horse, respectively. Panel (c) shows combined data from both labs from a collection of new, independent beef samples, all analysed as fresh samples. From a total of 91 beef data points, just one lies outside the boundary, constituting a single Type I error. Therefore, of the new extracts presented to the model, all but one are correctly classified as 'authentic'. Panel (d) shows the outcome of challenging the method with new, independent horse samples; this includes both fresh and freeze–thawed meats (6 independent samples corresponding to 16 extracts in total). All are correctly classified as non-authentic, that is, there are no type II errors. We note in passing that the 5 clusters each containing 3 points in close juxtaposition in Fig. 5(d) correspond to 5 independent samples, where each sample had been used to produce 3 replicate extractions. This gives an impression of the technical repeatability of the methodology, and implies that the variance shown by the dataset as a whole is due mainly to variation across meat samples and not to experimental sampling, extraction or data processing issues.

4. Conclusions

In this work we have demonstrated that 60 MHz ^{1}H NMR is able to differentiate between beef and horse meat by exploiting the differences in their triglyceride compositions. A simple, cheap and fast chloroform-based extraction protocol was shown to yield classic low-field NMR triglyceride spectra, with no more than a 10 min spectral acquisition time required for all but the leanest samples. Three signals (bis-allylic, olefinic and the terminal CH_3 peak) were particularly useful in characterising differences between horse and beef meat. Using these three signals, training samples were used to model the 'authentic' (beef) group. Applying the model to 107 extracts prepared from new, completely independent samples resulted in all but one being correctly authenticated.

A primary goal in the development of the methodology has been to ensure that it is readily transferable into an industrial setting, and this has influenced certain aspects of the experimental designs. First, sample preparations and spectral acquisitions were performed at two laboratories using slightly different protocols and instrumentation, with one of the labs focusing on minimising the time and cost of the analysis, an important objective for any potential high-throughput screening system. It was found that the variation in outcomes dues to these differences was insignificant relative to the observed dissimilarity between the two species. Second, we showed that freeze–thawing meat samples did not undermine the analysis, an important point to establish since the supply chain involves both chilled and frozen meat.

We envisage that our approach will be suitable as a screening technique early in the food supply chain, before cuts or chunks of raw beef are processed into mince or other preparations. A candidate point for detecting adulteration is in large (up to ~4000 kg) frozen blocks of meat trimmings. Such blocks could be core-sampled (in the same way as for currently used ELISA or DNA testing) and discrete fragments of tissue analysed using the NMR-based approach to determine whether they are authentic or not. Further, the level of confidence in the authenticity of the entire block could be established through standard statistical sampling strategies.

Although not investigated in the work presented here, the methodology could in principle be extended to quantifying beef–horse mixtures. However, differences in the overall fat content of the two species presents a considerable challenge. Since horse meat is generally leaner than beef, the extract composition is likely to be dominated by the triglycerides originating from the beef component. However, it is probable that horse meat used as an adulterant would comprise relatively fatty cuts rather than lean steak, so there could be value in simulating such scenarios in future work.

For a technique to be useful as a high throughput screening tool, in addition to being fast and inexpensive, it must be simple to use. Framing our analysis as a classic single-group authenticity problem, we have implemented software that simply reports the results on a test sample as either 'authentic' or 'non-authentic', without any analysis or interpretation on the part of the operator. In a hypothetical universe containing just beef and horse, we have established that 60 MHz ^{1}H NMR can report this outcome with virtually complete accuracy.

Standard DNA-based methods require separate tests for each adulterant a product is being screened for. In contrast, our framework lends itself to development such that a single NMR-based test could potentially detect a whole host of non-authentic samples: horse, beef-horse mixtures, or other animal species entirely. Estimating the expected Type II error rates for different types of non-authentic samples would naturally require further targeted studies; however, preliminary work (data not shown) has indicated that a comparable Type II error rate is likely to be obtained for pork.

Acknowledgements

The authors acknowledge the support of Innovate UK (formerly the Technology Strategy Board; Project Number 101250) and the Biotechnology and Biological Sciences Research Council (Grant Number BBS/E/F/00042674).

References

Alonso-Salces, R. M., Holland, M. V., & Guillou, C. (2011). H-1-NMR fingerprinting to evaluate the stability of olive oil. *Food Control, 22*(12), 2041–2046.

Ballin, N. Z. (2010). Authentication of meat and meat products. *Meat Science, 86*(3), 577–587.

Ballin, N. Z., Vogensen, F. K., & Karlsson, A. H. (2009). Species determination – Can we detect and quantify meat adulteration? *Meat Science, 83*(2), 165–174.

Barison, A., da Silva, C. W. P., Campos, F. R., Simonelli, F., Lenz, C. A., & Ferreira, A. G. (2010). A simple methodology for the determination of fatty acid composition in edible oils through H-1 NMR spectroscopy. *Magnetic Resonance in Chemistry, 48*(8), 642–650.

Chernukha, I. (2011). Comparative study of meat composition from various animal species. *Tehnologija Mesa, 52*(1), 167–171.

Dobranic, V., Njari, B., Miokovic, B., Fleck, Z. C., & Kadivc, M. (2009). Chemical composition of horse meat. *Meso, 11*(1). 32–36, 62–97.

Fang, G. H., Goh, J. Y., Tay, M., Lau, H. F., & Li, S. F. Y. (2013). Characterization of oils and fats by H-1 NMR and GC/MS fingerprinting: Classification, prediction and detection of adulteration. *Food Chemistry, 138*(2–3), 1461–1469.

Food Standards Agency. (2013). Report of the investigation by the Food Standards Agency into incidents of adulteration of comminuted beef products with horse meat and DNA.

Guillen, M. D., & Ruiz, A. (2001). High resolution H-1 nuclear magnetic resonance in the study of edible oils and fats. *Trends in Food Science & Technology, 12*(9), 328–338.

Guillen, M. D., & Ruiz, A. (2003a). H-1 nuclear magnetic resonance as a fast tool for determining the composition of acyl chains in acylglycerol mixtures. *European Journal of Lipid Science and Technology, 105*(9), 502–507.

Guillen, M. D., & Ruiz, A. (2003b). Rapid simultaneous determination by proton NMR of unsaturation and composition of acyl groups in vegetable oils. *European Journal of Lipid Science and Technology, 105*(11), 688–696.

He, M. L., Ishikawa, S., & Hidari, H. (2005). Fatty acid profiles of various muscles and adipose tissues from fattening horses in comparison with beef cattle and pigs. *Asian-Australasian Journal of Animal Sciences, 18*(11), 1655–1661.

Johnson, L. F., & Shoolery, J. N. (1962). Determination of unsaturation and average molecular weight of natural fats by nuclear magnetic resonance. *Analytical Chemistry, 34*(9), 1136–1139.

Kemsley, E. K. (1996). Discriminant analysis of high-dimensional data: A comparison of principal components analysis and partial least squares data reduction methods. *Chemometrics and Intelligent Laboratory Systems, 33*(1), 47–61.

Knothe, G., & Kenar, J. A. (2004). Determination of the fatty acid profile by H-1-NMR spectroscopy. *European Journal of Lipid Science and Technology, 106*(2), 88–96.

Lisitsyn, A. B., Chernukha, I. M., & Ivankin, A. N. (2013). Comparative study of fatty acid composition of meat material from various animal species. *Scientific Journal of Animal Science, 2*(5), 124–131.

Longobardi, F., Ventrella, A., Napoli, C., Humpfer, E., Schutz, B., Schafer, H., et al. (2012). Classification of olive oils according to geographical origin by using H-1 NMR fingerprinting combined with multivariate analysis. *Food Chemistry, 130*(1), 177–183.

Lorenzo, J. M., Fucinos, C., Purrinos, L., & Franco, D. (2010). Intramuscular fatty acid composition of "Galician Mountain" foals breed. Effect of sex, slaughtered age and livestock production system. *Meat Science, 86*(3), 825–831.

Lorenzo, J. M., Victoria Sarries, M., Tateo, A., Polidori, P., Franco, D., et al. (2014). Carcass characteristics, meat quality and nutritional value of horsemeat: A review. *Meat Science, 96*(4), 1478–1488.

Parker, T., Limer, E., Watson, A. D., Defernez, M., Williamson, D., & Kemsley, E. K. (2014). 60 MHz ^{1}H NMR spectroscopy for the analysis of edible oils. *TrAC, Trends in Analytical Chemistry, 57*, 147–158.

Paschke, B. (1938). Nachweis von Pferdefett in Mischungen mit Schwin-, Rinder- und Hammelfett. *Zeitschrift fur Untersuchung der Lebensmittel, 76*(5), 476–478.

Sedman, J., Gao, L., Garcia-Gonzalez, D., Ehsan, S., & van de Voort, F. R. (2010). Determining nutritional labeling data for fats and oils by H-1 NMR. *European Journal of Lipid Science and Technology, 112*(4), 439–451.

Shiao, T. Y., & Shiao, M. S. (1989). Determination of fatty acid compositions of triacylglycerols by high resolution NMR-spectroscopy. *Botanical Bulletin of Academia Sinica, 30*(3), 191–199.

Siciliano, C., Belsito, E., De Marco, R., Di Gioia, M. L., Leggio, A., & Liguori, A. (2013). Quantitative determination of fatty acid chain composition in pork meat products by high resolution H-1 NMR spectroscopy. *Food Chemistry, 136*(2), 546–554.

Szczepaniak, L. S., Babcock, E. E., Schick, F., Dobbins, R. L., Garg, A., Burns, D. K., et al. (1999). Measurement of intracellular triglyceride stores by H-1 spectroscopy: Validation in vivo. *American Journal of Physiology-Endocrinology and Metabolism, 276*(5), E977–E989.

Tonial, I. B., Aguiar, A. C., Oliveira, C. C., Bonnafe, E. G., Visentainer, J. V., & de Souza, N. E. (2009). Fatty acid and cholesterol content, chemical composition and sensory evaluation of horsemeat. *South African Journal of Animal Science, 39*(4), 328–332.

Vinaixa, M., Rodriguez, M. A., Rull, A., Beltran, R., Blade, C., Brezmes, J., et al. (2010). Metabolomic assessment of the effect of dietary cholesterol in the progressive development of fatty liver disease. *Journal of Proteome Research, 9*(5), 2527–2538.

von Bargen, C., Dojahn, J., Waidelich, D., Humpf, H. U., & Brockmeyer, J. (2013). New sensitive high-performance liquid chromatography tandem mass spectrometry method for the detection of horse and pork in halal beef. *Journal of Agricultural and Food Chemistry, 61*(49), 11986–11994.