Food Standards Agency
food.gov.uk

Review of capability of methods for the verification of country of origin for food and feed

March 2024

Authors: Helen H. Grundy, Hez J. Hird, M. Rosario Romero, Katharina Heinrich, Mark Harrison, Adrian J. Charlton, Emma L. Bradley (Fera Science Ltd.)

DOI: 10.46756/sci.fsa.ple720

Contents

List of figures

List of tables

Acknowledgements

This report has been produced by Fera Science Ltd. after exercise of all reasonable care and skill but is provided without liability in its application and use. The research was commissioned and funded by FSA and Defra. The views expressed reflect the research findings and the authors interpretation; they do not necessarily reflect Government policy or opinions.

Glossary

ADHB	Agriculture and Horticulture Development Board
AES	Atomic emission spectrometer
AI	Artificial intelligence
ANN	Artificial Neural Network
AOAC	Association of Analytical Chemists
Approx.	Approximate
ATR	Attenuated total reflection IR spectroscopy
BMPA	British Meat Processors Association
BP-NN	Back propagation artificial neural networks
CD	Corona Discharge
CEN	European Committee for Standardisation
CODEX	CODEX ALIMENTARIUS – international food standards, guidelines and codes of practice
CO2-eq-IRMS	Carbon dioxide-Equilibration-Isotope Ratio Mass Spectrometry
DA	Diode Array (detector)
DD	Data driven
DNA	Deoxyribonucleic acid
$(D/H)_I$, $(D/H)_{II}$, $\delta^{18}O$ and $\delta^{13}C$	Criteria used for authenticating for example wine or fruit juices: Deuterium/Hydrogen ratio at carbon position 1 and 2 of ethanol, δ-value of stable oxygen isotopes ($^{18}O/^{16}O$) and carbon isotopes ($^{13}C/^{12}C$) expressed in per mille.

EA-IRMS	Elemental analyser coupled to an isotope ratio mass spectrometer
eAmbrosia	European Union (EU) Geographical Indications register
EEM	Excitation-emission matrix
E-nose	Electronic nose
ESI	Electrospray ionization
EU	European Union
EVOO	Extra virgin olive oil
FAMEs	Fatty acid methyl esters
FI	Flow Infusion
FIRMS	Forensic Isotope Ratio Mass Spectrometry (network)
FTIR	Fourier transform infra-red
DAD	Diode array detection
GBS	Genotyping-by-sequencing
GC-C/Py - IRMS	Gas chromatograph coupled via combustion furnace (in pyrolysis mode) to isotope ratio mass spectrometer
GC	Gas chromatography
GI	Geographical indication
GO	Geographical origin
HIS	Hyperspectral imaging
HS	Head space
HPLC	High performance liquid chromatography

HRMS	High resolution mass spectrometry
ICP	Inductively coupled plasma
IMS	Ion mobility spectrometry
IRMS	Isotope ratio mass spectrometry
IST	Intelligent sensory technology
LC	Liquid chromatography
LDA	Linear discriminant analysis
MALDI	Matrix assisted laser desorption ionization time-of-flight mass spectrometry
MIR	Mid infra-red
MP	Microwave plasma
MS	Mass spectrometry
NASIR	National Association of Security and Investigative Regulators
NHE/NTP	Number of honeydew elements/number of total pollen
NIR	Near infra-red
NIST	National Institute of Standards and Technology
NGS	Next generation sequencing
NMR	Nuclear magnetic resonance
OES	Optical emission spectrometry
OIV	International Organization of Vine and Wine
OTU	Operational taxonomic unit

PARAFAC	Parallel factor analysis
PC1/2	Principal components; PC1 representing the most variation in the data and PC2 representing the second most variation in the data, when performing Principal Component Analysis (PCA)
PCR	Polymerase chain reaction
PCR-RFLP	Polymerase chain reaction coupled with restriction fragment length polymorphism
PDO	Protected Designation of Origin
PGI	Protected Geographical Indication
PLS-DA	Partial least squares discriminant analysis
PTR	Proton transfer
RASFF	Rapid Alert System for Food and Feed
REIMS	Rapid Evaporative Ionisation Mass Spectrometry
RFID	Radio Frequency Identification
RSD	Relative standard deviation
SIMCA	Soft independent modelling of class analogy
SIRA	Stable isotope ratio analysis
SNIF-NMR®	Site-specific natural isotope fractionation nuclear magnetic resonance
SNP	Single nucleotide polymorphism
SPME	Solid phase microextraction

TEEM	Transmission and fluorescence excitation-emission matrix
TOF	Time of flight
TRACE	Tracing Food Commodities in Europe - Framework 6 EU project
TSB	Technical Strategy Board
TSG	Traditional Speciality - UK protected
UK	United Kingdom
VE-tongue	Voltametric electronic tongue
VOO	Virgin olive oils
‰	Units of per mil (per thousand units)

Executive Summary

Background

This project, funded jointly by FSA and Defra, involved an up-to-date review of the availability, capability and limitations of methods of analysis for the verification of the country of origin of food and feed. The project also engaged with stakeholders to understand their views relating to the geographical origin of food and feed. Following this review of the current status of origin testing, recommendations are made to build upon previous country of origin research to support the food and feed industry going forward.

Country of origin is defined as the country where food or feed is entirely grown, produced, or manufactured, or, if produced in more than one country, where it last underwent a substantial change. In the UK, EU-assimilated legislation states that indication of the country of origin is a mandatory labelling requirement for food and feed, including products such as meat, vegetables, eggs, honey and wine.

The country of origin claim plays an important role for consumers who tend to relate certain country of origin labelling to superior quality or brand identity. Patriotism (or ethnocentrism) can also play a role in consumer food choice. In Europe, there are 3500 products with a specific geographical origin and their production methods are officially protected (Protected Designation of Origin = PDO; Protected Geographical Indication = PGI; Geographical Indication (for spirit drinks) = GI). These goods often carry a premium price. In addition to customer preference and sale price, country of origin claims are important to businesses when they seek to (i) monitor food miles (carbon footprint), (ii) ensure sustainable sourcing of, for example soy and palm oil (including new Regulation (EU) 2023/1115 on deforestation-free products), (iii) avoid trading of goods which are subject to sanctions, (iv) reassure consumers over concerns of farming and animal welfare standards, (v) avoid foods which are linked to exploitation of farm workers, enforced, or child labour.

‘Verification’ of geographical origin involves testing against a database to confirm that the data for a sample are consistent with those for that geographical location as claimed on a product label. Verification therefore does not involve testing a sample from an unknown location to unequivocally identify its provenance, as such methods are not available or are extremely limited in scope.

Aims

The main aims of this project were two-fold. Firstly, to conduct an up-to-date review of the literature on analytical techniques used for the verification of the country of origin of food and feed commodities. The results of the literature review are included in this report. Each section of the review concentrates on a type of analytical technique and draws together the reports on commodities which have been analysed using that technique. A critique of the techniques is given together with recommendations on their capability and limitations. An outlook section for each technique provides insight for future development potential together with a summary of the most mature methods which demonstrate capability to verify origin for various food commodities.

Secondly, the project engaged with stakeholders. This part of the project ran in parallel to the literature review. The majority of stakeholders consulted are involved in food supply and enforcement with an aim to understand the benefits and challenges in origin verification to the industry. An additional stakeholder was included representing a successful global origin verification organisation in order to inform as a model example of managing geographical origin testing. Information from each of these stakeholders is detailed in this report.

As a smaller contribution to the project, data were also accessed to provide information regarding the number of global official notifications which have been logged for incidents relating to geographic origin declarations in food and feed over the last decade. These data demonstrate the scale of issues relating to geographical origin mislabelling (incorrect origin detailed on the label) in food and feed according to commodity and according to the nation notifying on the issue. Finally, information regarding costs to set up laboratories to verify origin claims, and also charges for commercial testing services, have been included to inform regarding future support in this area which would involve official control laboratories.

Findings and limitations regarding the technologies

Methods to verify origin are wide-ranging with most data usually resulting from the testing of stable isotope ratios combined with trace element profiles, in conjunction with chemometrics. No single methods are usually unequivocal for determining origin, but data can be used to verify if the origin testing outputs are consistent with the origin declared on a food label or in the paper trail. Databases of, for example, Stable Isotope Ratio Analysis (SIRA) and trace element data have been developed for certain commodities, the geographical scope of which varies greatly, often from small areas in a single country to a small number of areas within several countries. The quality and up-to-date status (to account for annual and seasonal variation) of these databases is critical to the relevancy of testing outputs. In addition to stable isotope ratio and trace element profiling, depending on the product, other technologies can often add value to origin data, including metabolomics, genomics, various lower-cost spectroscopy methodologies, blockchain and proteomics. Certain technologies will be more applicable than others, depending on the commodity.

A list of the most promising technologies for the main commodities for which research has been conducted is provided below:

- **Cereals:** Trace element and SIRA analysis
- **Cocoa:** Near infra-red (NIR) spectroscopy and sensory techniques with AI.
- **Coffee:** SIRA analysis in combination with trace element analysis.
- **Fish and shellfish:** While fish and shellfish represent the third most notified commodity for misdescription of origin, there is a particular challenge in assessing geographical origin of these commodities. Trace element profiles vary greatly due to a combination of natural and anthropogenic activities, climate change, the inherent fluidity of the marine and fresh water environment and depending on harvest time. It is therefore likely that the combination of several technologies will provide the most informative models overall, including trace element analysis, NIR and Rapid Evaporative Ionisation Mass Spectrometry (REIMS) study of lipid markers. Due to the wide range of varying factors, it will be imperative that databases are constantly updated to account for these variations.

- **Fruit juice:** Trace element analysis.
- **Honey:** Pollen analysis using light microscopy, study of volatile compounds, sugars, organic acids and amino acids have been used to differentiate between floral type of honey. Also SIRA and trace element analysis, metabolomic and genomic approaches in combination with blockchain.
- **Meat:** SIRA, trace element analysis, with fatty acid profiling, along with radio frequency identification (RFID) to monitor livestock movement.
- **Olive oil and other edible oils:** SIRA combined with NMR and profiling of phenolic compounds, fatty acid profile (including fatty acid methyl esters (FAMEs)), sterols, triacylglycerol (TAGs), volatile compounds and colour. The potential of Fourier transform infra-red (FTIR) spectroscopy ATR should also be considered.
- **Rice:** SIRA with trace element analysis.
- **Tea:** Trace element profiling.
- **Whisky:** Gas chromatography (GC), liquid chromatography (LC), spectroscopy and trace element analysis.
- **Wine:** SIRA and site-specific natural **isotope fractionation** nuclear magnetic resonance (SNIF-NMR) methodologies, with possible support from trace element analysis.

The commodities for which the largest numbers of incidents have been notified over the last decade relating to geographical origin are honey and wine, followed by fish, meat, olive oil and food/dietary supplements.

While blockchain could aid paperchain traceability in a weight-of-evidence approach and can be searched in seconds, a common opinion is that implementation costs of these systems are high. The second common concern is that blockchain requires honest participation in terms of, for example, correctly declaring the commodity, the reputability of the handling party and where they are located. At present, there is no means of validating that information logged in a blockchain system is genuine.

Stakeholder engagement

There were two parts to the stakeholder engagement. First, a stakeholder representing a successful working model of origin testing was interviewed to inform this project. World Forest ID, primarily a timber origin authentication organisation but also working increasingly in food origin verification for Forest Risk Commodities, was included in the stakeholder engagement to understand how this organisation successfully conducts origin testing.

Key aspects to the success of this model include:

- Curation of the data by a single organisation.
- Building databases which span global geographies.
- Ensuring the authenticity of samples in the databases.
- Regularly expanding databases to account for seasonal and annual variation, for example relating to climate change and other natural and anthropogenic variation.
- Training to provide standardised methods to prepare the data.
- Testing by harmonised methods to provide confidence in the accuracy and reproducibility of the data.
- Inter-laboratory data comparisons to monitor reproducibility.
- Sharing of data so that databases they are publicly accessible and transparent.

Other working examples of food origin testing include SIRA and trace element databases funded by the consortium of producers of Grana Padano (PDO) and Parmigiano Reggiano (PDO) cheeses to monitor the origin of retail samples, with successful inter-laboratory validation.

The main outputs of the engagement with stakeholders from elsewhere in the food and feed industry were as follows:

- Stakeholders in the food and feed industry are aware that there is no ‘silver bullet’ for origin testing.

- In terms of testing activities, food safety testing is the main focus of the industry followed by species identification. Origin testing is a lower priority.
- Stakeholders are aware that stable isotope ratio analysis and trace element analysis tend to be the most applicable technologies for assessing the origin of the majority of commodities and that these tests can be used to interrogate the accuracy of food and feed traceability documents.
- Stakeholders highlighted the lack of accredited services and proficiency testing as a barrier to their origin testing.
- Stakeholders emphasised that databases can be out of date since they have often not been added to for some time which is necessary to account for recent variations and that the data are therefore less than ideal.
- It was highlighted that many databases are not publicly accessible, accessed only by testing service providers, and stakeholders reported that the quality, composition, applicability to different matrices and current relevance of the databases is not always transparent.
- Stakeholders stated that, while many of the larger members of the food supply chain are aware of the importance of accurate origin labelling and perform testing to challenge label information, more support is needed for smaller sized organisations. Education at these organisations may enable more issues to be picked up earlier in the supply chain and ultimately reduce risk of more widespread fraud.

Conclusions

- There is no unequivocal single technique to verify the claimed country of origin of food and feed.
- Testing methods tend to inform as to whether data are in line with labelling or certification claims but rarely categorically assign origin.
- The maturity of the testing for origin varies greatly across commodities and geographical locations.

- The validity of studies undertaken to date is highly dependent on the number and authenticity of the samples used as a reference collection.
- Combined SIRA and trace element analyses are the most widely used. Depending on the commodity, other technologies such as genomics, metabolomics and spectroscopy can add value.
- Multivariate data generally improve classification rates and robustness.
- Livestock can move geographically during their lifetime. From an analytical perspective, it is therefore challenging to identify meat origin, and both geography and feeding regimes must be considered.
- Validity of studies undertaken to date is highly dependent on the number, nature and geographical range of the samples used in a reference database. In many cases the reference databases are currently inadequate or outdated.
- Curation of databases, and standardisation of data capture, and methods of analysis, is key and is lacking.
- The lack of continued expansion of previously established databases renders datasets out-of-date due to lack of account of seasonal and annual variation. Before the databases can be used in the future, old data must be tested against contemporary samples to examine temporal trends or to establish whether data have remained stable in the intervening period. While such multivariate tests can be completed relatively easily, this adds delay if an issue arises in the supply chain, during which time samples cannot be tested in the probable time requirement for fresh food or to investigate the early stages of an issue before the issue becomes widespread. Ideally, databases would be maintained so that they are fit-for-use as and when needed.
- Databases should be regularly challenged with new/additional samples to ensure continued relevance.
- Obtaining authentic matrix-matched reference materials to support testing is a critical problem which requires funding to achieve best practice under the principle of identical treatment.

- Database sharing is poor due to investment and IP concerns. This can limit progress in verification of origin.
- Standardisation activities are under development by AOAC, CEN, CODEX, however these activities take time for completion.

Other barriers to geographical origin verification relate to:

- Costs – the main mature methods used require mass spectrometric analysis which requires high-cost instruments and skilled analysts. Investment in building and maintaining databases along with a program of regular food testing would mean that a higher number of samples is tested and the methods would become more cost effective. Investment is also required to continue to develop and assess lower cost methods such as spectroscopy.
- Demand for proficiency testing is lacking for geographical origin determination and this is impeding the trustworthiness of the data. Transferability of data from one lab to another must be demonstrated by proficiency trials.
- Few methods to verify origin are accredited and the accreditation process can take twelve months. Means to fast-track methods for accreditation are needed when a method is specifically developed to address and investigate a known issue in the supply chain.
- There is a challenge to present multivariate data in court to support origin fraud. Only limited methods have been considered ready to be used in court.
- Due to the nature of the methods, should a product be a composite product or be adulterated and contain a commodity from a mixture of origins, this is problematic for origin verification and the achievable limits of testing for adulterated samples must be understood.

Future direction

Recommended future direction is detailed below:

All stakeholders highlighted challenges with a lack of quality data or up-to-date data in the various databases which have been prepared over the years in support of food provenance testing. In the future, investment could be made to support the building of robust datasets, necessitating the harmonised collection and analysis of a high number of authentic samples from a large (e.g. global) geographical range for commodities which are vulnerable to origin fraud.

These datasets should be curated by a single organisation to facilitate harmonisation. Datasets should be regularly expanded with samples to account for seasonal and annual variation. The single organisation should be funded so as to allow continued expansion and curation of the datasets over time so that a database does not lose relevance and is fit-for-use as soon as required for any supply chain issue which may arise. The suitability of the datasets should be challenged each year by the testing of additional samples (i.e. not the samples already included in the prediction models) and all testing facilities involved should take part in proficiency trials and achieve acceptable data, demonstrated through statistical measurements, for example z-scores.

Facilitate the fast-tracking of methods for accreditation when a method is needed to address and investigate a known issue in the supply chain.

More matrix-matched certified reference materials should be available which requires investment.

Demand for proficiency testing in origin analyses should be encouraged and as a result of this, certified testing schemes, which are currently lacking, would be initiated.

There is much existing data from other surveillance testing exercises which are not yet/rarely used in origin verification and which could have value for origin determination. This could be explored. Investment could be made to initiate the collation of these existing data and metadata and investigate their utility for origin verification purposes. Prediction models could be built to determine the relevance of these data for addressing origin issues. Such existing data could include environmental and geological data, pesticide residue data, dioxin and polychlorinated biphenyl levels (particularly relevant to foods including salmon), viral screen data and fungi species presence information.

There is scope for the emergence of other methods, either for initial, fast screening or more detailed analysis. While hand-held spectroscopy technologies have been used in small proof-of-principle studies, multivariate analysis may be considered as a natural next step in the quest to verify origin, fusing data from various technologies.

Opportunities for public, private and inter-nation funding to improve methods in geographical provenance could be sought to protect our supply chain.

Government funding and private and EU funding opportunities in origin determination could be considered to expand collaborations between those working on building datasets for a commodity (or range of commodities) across different geographical locations and, in turn, to expand the geographical range of datasets. Specific high value commodities that generate large tax incomes could be prioritised for engagement since these industries could more readily support the testing costs. An example of such includes Scotch whisky.

1. Background to the project

The constantly evolving network of global food/feed trade has led to an unprecedented variety of foods available to consumers. At the same time consumers are increasingly interested in the provenance of food, especially in view of various food scandals, from contaminated eggs, substitution of beef with horse meat in ready meals, mis-labelled beverages and more recently 'Operation Hawk'[1] (pork from South America and Europe falsely labelled as British) and FSA National Food Crime Unit investigations into the geographical provenance of corned beef.

The country of origin claim plays an important part for businesses when they seek to (i) monitor food miles (carbon footprint), (ii) ensure sustainable sourcing of for example soy, palm oil and (iii) avoid trading of goods which are subject to sanctions. In addition to the above, consumers tend to relate certain country of origin labelling to superior quality or brand identity (for example Parma ham, New Zealand Mānuka honey and Modena Balsamic vinegar) or safety (linked to expected high animal welfare standards and control procedures in place), so that products are marketed at higher prices than similar products of other or unknown provenance. High volume products such as fruit and grain are also vulnerable to origin fraud due to the large amounts of related value. Fragmented supply chains, often due to geopolitical situations can also result in the supply chain becoming more vulnerable to origin fraud (for example Chinese raspberries being described as Chilean. Unscrupulous producers might buy cheaper raw materials from other regions /countries and illegally sell the product with different origin labels for financial gain. Since leaving the EU, the UK government is considering the introduction of labels on UK-grown food to protect farmers and consumers over concerns of welfare standards in some food imports.

Country of origin is defined as the country where food or feed is entirely grown, produced, or manufactured, or, if produced in more than one country, where it last underwent a substantial change. In the UK EU-assimilated legislation states that

---

[1] National Food Crime Unit Annual Update (food.gov)

indication of the country of origin is a mandatory labelling requirement for food and feed, including products such as meat, vegetables, eggs, honey and wine[2],[3] (Appendix 1).

In Europe, there are 3500 products[4] with a specific geographical origin and their production methods are officially protected (Protected Designation of Origin = PDO; Protected Geographical Indications = PGI) – EU 664/2014[5]. These are registered in a public database that was launched in 2019: eAmbrosia[6] - the European Union (EU) Geographical Indications register. Since January 2021, following the EU exit, the UK GI (geographical indication) schemes and logos[7] have been established:

Designated origin - UK protected (PDO), Geographic origin - UK protected (PGI) and Traditional Speciality - UK protected (TSG). The principles and benefits of the GI scheme remain the same and continue to provide intellectual property protection.

'Verification' of geographical origin involves testing against a database to confirm that the data for a sample are consistent with those for that geographical location. Verification therefore does not involve testing a sample from an unknown location and unequivocally identifying its provenance.

In partnership with Defra, the FSA conducted a study in 2014[8] which used stable isotope ratio analysis (SIRA) and paper-based approaches to assess whether foods

---

[2] Regulation (EU) 2017/625 of the European Parliament and of the Council of 15 March 2017 (eur-lex.europa.eu)

[3] Regulation (EU) No 1169/2011 of the European Parliament and of the Council on the provision of food information to consumers (legislation.gov.uk)

[4] Protecting local food and drinks: 3 500 geographical indications registered (agriculture.ec.europe.eu)

[5] Commission Delegated Regulation (EU) No 664/2014 (eur-lex.europa.eu)

[6] eAmbrosia - the EU geographical indications register (ec.europa.eu)

[7] Guidance: Protected geographical food and drink names: UK GI schemes (gov.uk)

[8] Country of origin of foods study published | food.gov.uk (nationalarchives.gov.uk)

claiming to originate from the UK and Ireland were as claimed. It examined 96 food samples of which 78 samples were shown to be consistent with the origin claimed, while 18 were identified for follow-up investigation and later confirmed as consistent using traceability documentation. The study demonstrated that SIRA had potential to verify country of origin claims when used in combination with traceability and other evidence, but had some limitations, including a need for comprehensive databases to support the method[9]. This triggered the supporting project (FA0155) during the EU FoodIntegrity Project[10] (2014-2018, Collaborative Project: 613688, Seventh Framework Programme, KBBE.2013.2.4-01: Assuring quality and authenticity in the food chain) which aimed to address database issues raised in this study to maximise their sharing, harmonisation and use to verify food origin, so contributing to assuring food integrity.

Defra has also commissioned several research projects to look at country of origin verification methods for a variety of foods between 2012 and 2019[11,12,13,14,15,16]. Techniques investigated included isotopic analysis, deoxyribonucleic acid (DNA) speciation, single nucleotide polymorphism (SNP) genotyping, metagenomics, and

---

[9] Study: No Misleading Claims of Food's Country Origins (foodingredientsfirst.com)

[10] Ensuring the Integrity of the European food chain (cordis.europa.eu)

[11] Geographic traceability tools for commercial fish and fish products - FA0118 (sciencesearch.defra.gov.uk)

[12] Development of Analytical Methods to Verify Labelling Claims Relating to Egg Production - FA0130 (sciencesearch.defra.gov.uk)

[13] Metagenomics for determination of origin part 2 (full study) - FA0141 (sciencesearch.defra.gov.uk)

[14] The development of isotopic and fingerprinting techniques to verify the production origin and geographical origin of food and feed - FA0159 (sciencesearch.defra.gov.uk)

[15] Development of Metagenomic Methods for Determination of Origin - phase 2 - FA0160 (sciencesearch.defra.gov.uk)

[16] British Beef Origin Project 2 – improvement of the British beef Isotope Landscape Map (Isoscape) - FA0152 (sciencesearch.defra.gov.uk)

commodities investigated included fish/fish products, egg, oysters, chicken, cheese (Blue Stilton) and beef. In general, processed products for example sausages available containing multiple ingredients present additional challenges and determining where ‘substantial changes’ may have occurred during production can be complicated. For these techniques robust reference databases and reference materials can be limited.

In the EU, new regulations were brought about following the horsemeat scandal, including Regulation (EU) 2017/625 that focuses on enhancing the food control systems of the EU and its member states. This regulation not only includes food safety, but also food authenticity (Kulling *et al.*, 2019). The European Commission for the first time recommended setting up so-called European Union reference centers for the authenticity and integrity of the agri-food chain, with responsibilities to: (1) provide specialised knowledge in relation to food authenticity and integrity and to the methods for detecting fraud; (2) provide specific analyses designed to identify the segments of the agri-food chain that are more vulnerable; (3) establish and maintain collections or databases of authenticated reference materials; (4) disseminate research findings and technical innovations. One of these centres was created in Germany, NRZ-Authent, as part of the Max Rubner Institut (MRI), the Federal Bureau of Statistics and public prosecutor’s offices. NRZ-Authent comprises infrastructure modules dedicated to IT, databases, accreditation and National Reference Laboratory functions, and also, thematic modules, one of which focuses on geographical origin verification. Standardisation and validation of methods is a key objective of this centre, as is the creation of a database of analytical methods available to food control authorities, including information about relevant control authorities, academic and private laboratories and companies. This way, they aim to have an up-to-date map of capabilities in Germany that will facilitate the identification of research gaps and decision making regarding further developments. Having left the EU, the UK must plan strategies to verify the geographical origin of food.

This project aims to conduct an up-to-date review of the availability of methods of analysis suitable for the verification of the country of origin of food and feed, and to provide recommendations on their capability and limitations, and applicability to the delivery of official controls, taking into account the current regulatory position and the fact that many public analysts do not have in-house access to many of the more sophisticated techniques such as IRMS. The project will provide practical recommendations in future direction to progress the UK's capability for origin verification. This will be delivered through review of the literature and engagement with stakeholders.

2. Literature review

2.1 Introduction

A wide range of analytical techniques have been used to support the determination of geographical origin of food and feed. In 2008, an overview of these analytical methods was published by Luykx and van Ruth (Luykx & van Ruth, 2008) that included mass spectrometry, spectroscopy, separation methods and other techniques. This article highlighted that a combination of methods measuring different types of analytes seemed to be the best approach to establish the geographical origin of food, with chemometric analysis of the 'fused' data supporting the interpretation of such a multifactorial approach. Since then, many more articles have been published on the topic, and the main technologies are still at the forefront, but many others have been applied to the determination of geographical origin of food and feed, and important advances have taken place on computational tools to support the analytical technologies.

A review of published international literature has been conducted in this project to identify current and emerging methods of analysis for the verification of the country of origin of food and feed, covering the past 10 years. This includes peer-reviewed literature and grey literature (the latter being searched by Google and including, for example, information from the media, institute proceedings and trade journals).

This review provides a critical assessment of techniques for country of origin verification of food and feed, including SIRA, trace element analysis, DNA methods, SNP genotyping, proteomics and metabolomics. Emerging methods to verify geographical origin and availability of quality reference databases, reference materials and proficiency testing schemes are also addressed.

2.2 Methodology

A list of relevant keywords, text phrases and the date range to be used in the literature searching elements of the project was agreed with FSA and Defra (Table 1). These keywords and text phrases were used to search Web of Science. Additional web searches to identify literature were also completed.

Table 1: Search terms

Keywords separated by 'OR'	-	Keywords separated by 'OR'	-	Keywords separated by 'OR'
analy* method* detection	AND	geographic* origin* provenance traceability geographical indication geographical authentication country of origin	AND	beef cattle pork lamb poultry chicken turkey meat products ham bacon sausage burger shellfish cod haddock

Keywords separated by 'OR'	-	Keywords separated by 'OR'	-	Keywords separated by 'OR'
				salmon caviar chicken eggs vegetable oil olive oil sunflower oil oilseed rape oil palm oil soy* oil wine whisky whiskey whey ginger fruit juice vanilla oregano tomato turmeric

Keywords separated by 'OR'	-	Keywords separated by 'OR'	-	Keywords separated by 'OR'
				soya saffron za'atar garlic grain* wheat barley rice coffee cocoa pepper honey tofu

A framework to grade each article was agreed by the project team prior to filtering. Each article was assigned a grade relating to relevance and the robustness of the data. An initial 5 star rating was applied (star rating 3 to 5 for large authentic database/established technique or standard/review respectively), articles which were deemed non-relevant (star rating 1, for example not fitting the country of origin topic) or of non-robust data (star rating 2, for example method only validated in house with fewer than 50 samples) were filtered out. As a quality check, a second team member reviewed the list of articles which had been retained and filtered out.

Since the number of articles selected through the rating system was too high, a second round of selection was applied, giving more focus to recent articles and reviews and relevant papers therein.

3. Review of techniques

The content of the review has been structured around technology categories, and within these, the information has been split by commodity where possible.

3.1 Stable Isotope Ratio Analysis (SIRA)

Traditionally, SIRA is the method of choice to determine geographical origin of food and feed, using an elemental analyser coupled to an isotope ratio mass spectrometer (EA-IRMS) for bulk analysis (for example defatted meat protein, olive oil, wine, honey), since the isotope signals from the bio-elements (H, C, N, O, S) present in local feed and water transfer to the animal and plant tissue, and are in turn linked to their local environment. There are three scenarios in which SIRA can be applied to verify geographical origin, (i) upfront investment to generate and curate a database for a specific food (for example EU wine database (Christoph *et al.*, 2015), (ii) keeping a library of reference samples to compare to (for example Agriculture and Horticulture Development Board (ADHB) pork and (iii) simple batch to batch comparison on an *ad hoc* basis.

A more niche application concerns compound-specific analysis (for example vanillin, other flavours or organic acids) which is performed by a gas chromatograph coupled to IRMS (GC-IRMS). More recently, a system of liquid chromatography coupled to IRMS (LC-IRMS) has evolved from questions of food adulteration (for example individual sugars of honey) for geographical characterisation of Italian grape musts (Perini & Bontempo, 2022); (Perini *et al.*, 2023).

Of the manuscripts identified as using SIRA to determine geographical origin only 15 manuscripts related to methods for which a database was available (n>200 per isotope), reflecting that the majority of approaches published have not yet been tested on a sufficiently large number of samples to prove their potential to successfully determine geographical origin for food commodities. A table summarising details of the databases covered by these 15 articles is presented below (Table 2), and further details are shown in Appendix 2. Methods tended to show improved classification rates and robustness when SIRA was used in combination with other analytical techniques to verify geographical origin, namely trace elements, fatty acids, nuclear magnetic resonance (NMR) and DNA or with environmental data of isoscapes (maps representing the natural variation in stable isotope, or indeed trace element, composition) for δ^2H and $\delta^{18}O$ of precipitation and /or ground water. Analysis by EA-IRMS is more prominent than GC-IRMS and LC-IRMS.

Table 2: Summary of databases with large numbers of samples found in the literature

Food	**Country of origin**	**Number of samples**	**Analytical technique**	**Reference**
Honey	Italy (botanical varieties)	265	EA-IRMS and SNIF-NMR	(Bontempo *et al.*, 2017)
Honey	Turkey (pine)	373	EA-IRMS	(Uçurum *et al.*, 2023)
Beef	13 regions in 8 European countries	224	T/C and EA-IRMS	(Bontempo *et al.*, 2023)
Wheat	Regions within Hesse (Germany)	132	EA-IRMS	(Gatzert *et al.*, 2021)
Olive Oil (EVOO)	Greece	210	EA and Py-IRMS	(Karalis *et al.*, 2020)
Olive Oil (VOO)	Italy (9 regions)	387	EA and Py-IRMS	(Portarena *et al.*, 2014)
Rice	China (17 provinces)	900	EA and Py-IRMS	(C. L. Li *et al.*, 2022)
Rice	China (17 provinces)	794 + 132	EA and Py-IRMS	(Sheng *et al.*, 2022)
Pork	Korea (non-Korea)	252 (94)	EA-IRMS	(Park *et al.*, 2018)
Wine	Austria (10 sites in 4 regions)	349	EA-IRMS and CO2-eq-IRMS	(Philipp *et al.*, 2018)
Wine	Venetian (7 regions)	2019	MC-ICP-MS	(Rapa *et al.*, 2023)
Wine	Global 8 countries)	600 (red, imported)	GC-C-IRMS and CO_2-eq-IRMS	(Wu *et al.*, 2019)
Wine	France (4 regions)	240 (red)	GC-C-IRMS and CO_2-eq-IRMS	(Wu *et al.*, 2021)
Chicken	Global (17 countries)	384 (incl 5% turkey)	EA-IRMS	(Rees *et al.*, 2016)

Food	Country of origin	Number of samples	Analytical technique	Reference
Scallops	China (7 coastal regions)	575	EA-IRMS	(Zhang *et al.*, 2019)

Table 3: Databases held for verification of origin of commodities. (Data taken from Food Authenticity Network Authenticity databases library)

Commodity	Geographical scope	Company	Technology	Status of database	Funding status to maintain database
Cereals (various)	UK and various EU countries	Fera Science	SIRA	Proprietary	No database funding received
Cheese	UK and various EU countries	Fera Science	SIRA	Proprietary	No database funding received
Chicken	UK and various EU countries	Fera Science	SIRA	Proprietary	No database funding received
Chicken	UK and various EU countries	Food Forensics	Mass spectrometry	Proprietary	Maintained

Commodity	**Geographical scope**	**Company**	**Technology**	**Status of database**	**Funding status to maintain database**
Cotton	Various, including China, Uzbekistan, Turkmenistan, Egypt, US, Brazil, and India	Food Forensics	Mass spectrometry	Proprietary	Maintained
Honey	UK and various EU countries	Fera Science	SIRA	Proprietary	No database funding received
Lamb	UK and various EU countries	Fera Science	SIRA	Proprietary	No database funding received
Milk	UK and various EU countries	Fera Science	SIRA	Proprietary	No database funding received
Olive oil	UK and various EU countries	Fera Science	SIRA	Proprietary	No database funding received

Commodity	Geographical scope	Company	Technology	Status of database	Funding status to maintain database
Pork	UK and various EU countries	Fera Science	SIRA	Proprietary	No database funding received

Various food fraud databases are listed by the European Commission[17]. Those which particularly hold data to address food and feed geographical origin testing include EU Wine DataBank, Oleum DataBank (olive oil), Honey DataBank, Horse meat database.

A summary of articles selected for the top commodities (highest percentage of publications within the time frame of the literature review, namely: wine, meat, rice, olive oil and coffee) found in the searches is presented in the sections below.

3.1.1 Wine

In the EU, each member state analyses the isotopic composition of a certain number of wine samples per vintage year as a commitment to the EU wine data bank which was established in 1991 for the control of the wine sector. The accepted analytical methods for the determination of the required isotope ratios ($(D/H)_I$, $(D/H)_{II}$, R-value, $\delta^{18}O$ and $\delta^{13}C$) are IRMS and SNIF-NMR. On the basis of the isotope data of the vintage 1997 to 2004, average values for Vienna, Lower Austria, Styria and Burgenland could be calculated (Philipp *et al.*, 2018) and some trends identified. However, prior knowledge of the vintage year remains advantageous in order to verify wine region of origin. The EU Wine Databank is unique in that it was initially developed to address general wine authenticity issues (detecting adulterations with sugar or water) using SIRA and SNIF-NMR. However, the databank was extended to address geographical origin issues. The

---

[17] Food Fraud Databases (knowledge4policy.ec.europa.eu)

vintage and region of production is usually explicitly claimed on the label and so an appropriate sub-section of a larger database can be used for comparison and verification. The $(D/H)_I$ and $(D/H)_{II}$ ratios of the methyl group in the ethanol of a wine was later exploited for origin verification purposes. The $(D/H)_I$ ratio is linked to botanical origin (C3 plant), while the $(D/H)_{II}$ ratio of the methylene group gives indication of the fermentation water and therefore the geographical origin of a wine.

The integrity of 50 European wines sold in China was checked, based on anthocyanin composition, stable isotopes and glycerol (Müller *et al.*, 2021). Glycerol is a by-product of fermentation that provides a perception of sweetness and improves 'mouth feel'. It is sometimes added to wine, as an adulterant, for these reasons. As stated by these authors, in general, wine that has been diluted with tap water shows decreased $\delta^{18}O$ values due to discrimination against the heavier isotopes in the plants during evapotranspiration. There is a consensus that $\delta^{18}O$ in wine water in southern Europe typically does not fall below -2 ‰. In colder and generally more humid central Europe, wines usually do not show $\delta^{18}O$ values below -5 ‰. The authors reported that European-labelled wines from China ranged between $\delta^{18}O$ -6.1 ‰ and 8.7 ‰. On this basis, it was concluded that five samples were clearly mislabelled as French wines as they showed $\delta^{18}O$ values below -2 ‰, two of them even below -5 ‰.

Another research group, Wu *et al.*, combined SIRA and trace element profiling in conjunction with carbon and oxygen isoscapes, applying partial least squares discriminant analysis (PLS-DA) and using an Artificial Neural Network (ANN) to verify the origin of 240 French red wines from four regions (Bordeaux, Burgundy, Languedoc-Roussillon and Rhone) with verification accuracy of 98.2% (Wu *et al.*, 2021) and 600 red wine from 8 countries (France, Italy, Spain, USA, Chile, South Africa, Australia and China) with overall discrimination accuracy of 83.9% (Wu *et al.*, 2019).The wines had been sourced from Shenzhen Customs and also via reliable wine importers prior to comparison with Chinese wines.

3.1.2 Meat

There are many reports where stable isotope data has been used to support the development of models for the determination of geographical origin of beef, pork and lamb. For example, Park *et al.* verified the origin of 252 defatted Korean (Korean Institute for Animal Products) and non-Korean (n= 94, Korean Meat Import Association) pork samples, by combining trace element content and their isotopes (27 isotopes of 19 elements) with stable isotope ratios of carbon and nitrogen, and resulting in a discrimination rate of 100% for their multivariate chemometric linear discriminant analysis (LDA) approach (Park *et al.*, 2018). Particularly, the isotope ratios of the heavy elements: $^{88}Sr/^{86}Sr$, $^{88}Sr/^{84}Sr$, $^{208}Pb/^{207}Pb$,$^{208}Pb/^{206}Pb$, $^{114}Cd/^{111}Cd$, $^{114}Cd/^{110}Cd$, $^{112}Cd/^{111}Cd$, and $^{112}Cd/^{110}Cd$, for Korean pork samples have shown that they are greater compared to the non-Korean ones (Chilean, Mexican, Spanish, Canadian, German, US American).

Pianezze *et al.* (2021) critically reviewed publications, which applied SIRA to tracing the geographical origin of lamb and confirmed that isotope ratios of carbon mainly discriminate difference in animal diets, while the isotopic ratios of deuterium and oxygen tend to discriminate geographical origin. Models described in their review could reach total accuracy of 90% or more. The authors stated that these results could potentially be improved even further by considering additional variables such as fatty acids and trace mineral elements. They recommended to create a lamb-based reference material to enable the comparability of isotope ratio data which is specific to PGI and PDO lamb products across multiple laboratories (Pianezze *et al.*, 2021).

Most recently, Bontempo and Perini published the evaluation by principal component analysis (PCA) of the hydrogen ($^{2}H/^{1}H$), carbon ($^{13}C/^{12}C$), nitrogen($^{15}N/^{14}N$), and sulphur ($^{34}S/^{32}S$) isotope ratio data from 227 authentic and defatted beef samples (Framework 6 EU project 'TRACE, Tracing Food Commodities in Europe', Grant Number 6942), which were collected from 13 sites in eight countries. Whereas country of origin could not be determined in all cases, coastal and inland regions, areas of different latitudes and breeding systems were successfully discriminated. The PC1/2 plot showed 77.5 % data variation and revealed a clustering into three groups a) coastal areas at lower latitudes, using C_4-based diet (Barcelona, Chalkidiki, Sicily and Tuscany), b) coastal areas at higher latitudes (Cornwall, Orkney and Bohernagore) and c) inland areas at lower latitudes, using a mixed diet (Limousin, Frankonia, Allgäu, Gäuboden, Mühlviertel and Trentino). The results presented by the authors are encouraging and they recommend enlarging the isotope databanks for each country, as this would help to verify the authenticity claims of beef samples and not only protect, but also promote PDO and PGI products, for example Orkney beef (Bontempo *et al.*, 2023).

3.1.3 Rice

Rice is an important staple food in China and its authenticity is closely associated with nutrition and safety (Z. M. Li *et al.*, 2022). In 2017, 900 Japonica and Indica rice samples from 17 Chinese provinces were collected to establish a comprehensive database for stable isotopes (H,C,N,O) and trace elements (K, Mg, Ca, Na, Fe, Zn, Mn, Cu, Ni, Cr and Mo), with the goal of origin discrimination of four regions, Middle-Lower Yangtze Plain, Northeast, Southwest and Southeast. Qualitative models were built using back propagation artificial neural networks (BP-NN). All the samples were randomly divided into a training set (70%), a validation set (15%) and a testing (blind) set (15%). Classification of Japonica and Indica rice from different regions was achieved with a high accuracy (97.2% and 77.9%, respectively), but annual climatic variations and environmental factors influencing these fingerprints need further exploration (C. L. Li *et al.*, 2022).

Sheng *et al.* (2022) used a similar set of 794 rice samples from 2017 to build isoscape models with focus on environmental similarity which allowed the prediction of the geospatial distribution of δ^{13}C, δ^{2}H and δ^{18}O values for Chinese rice in 2018 and to validate the results with 132 actual samples from 2018. The climate data was derived by interpolating gauged daily temperatures, precipitation, relative humidity and sunshine duration from 825 Chinese Meteorological Administration stations[18] in 2017. For the predictive model rice samples were randomly divided into 555 training samples (70 %) and 239 validation samples (30 %) and these dataset divisions were repeated 10 times. The training samples were used to develop an isoscape model, and then the validation samples were used to evaluate the prediction model. The mean correlation coefficients (r) for observed and predicted δ^{13}C, δ^{2}H and δ^{18}O values were 0.82, 0.81 and 0.63 respectively, meaning that the four main production regions in China: Northeast China, middle to lower Yangtze River plain, Southwest China and Southeast China could be distinguished (Sheng *et al.*, 2022).

---

[18] Chinese Meteorological Administration stations (data.cma.cn)

3.1.4 Olive oil

Olive oil and in particular extra virgin olive oil are highly priced and linked to fraudulent activities by blending with lesser seed/nut oils or mislabelling of country of origin, for example trading Spanish as Italian olive oil. The latter holds the highest price on the current olive oil market[19].

There are numerous examples in the literature of application of SIRA for geographical verification of olive oils. In a study by Jiménez-Morillo and co-workers, virgin olive oils (VOO, n= 138, 2016 and 2017), prepared within 24h from harvest of olives, from three different Mediterranean countries (Portugal, France and Turkey) were analysed. The countries could be discriminated on the basis of multivariate statistical analysis of geoclimatic and isotopic data ($\delta^{13}C$, $\delta^{2}H$ and $\delta^{18}O$) based on bulk analysis by EA-IRMS (Jiménez-Morillo *et al.*, 2020).

During the harvest period 2015-2016 of Koroneiki olives in Greece, a total of 210 extra virgin olive oil (EVOO) from 6 different regions were analysed for isotopic values of ^{18}O (bulk oil) and ^{13}C (biophenolic extracts and bulk oil) (Karalis *et al.*, 2020). For comparison, data also included harvests from 2005/2006/2010 which allowed to monitor climatic changes over time. A relationship between $\delta^{18}O$ Water (springs used for irrigation) and $\delta^{18}O$ EVOO was also demonstrated. Regarding the $\delta^{13}C$ values, results for bulk EVOO and biophenolic extracts were very similar and can be interchangeably used for differentiating Ionian Island olive oils from olive oils originating from Crete and Chalkidiki, when applied in conjunction with $\delta^{18}O$ EVOO.

The addition of fatty acid profiles (concentration and $\delta^{13}C_{FA}$) to the stable isotope ratio of the bulk EVOO (H, C and O) has been recommended for improving the differentiation of vegetable oils from the Southern and Northern hemisphere (Spangenberg, 2016).

---

[19] Prices continue to rise throughout the Mediterranean, affecting Olive Oil trade and consumption trends globally (certifiedorigins.com)

Paolini and collaborators describe their in-house method validation of FAMEs analysis of olive oil triglycerides by GC-C/Py -IRMS for $\delta^{13}C$ and $\delta^{2}H$ values as a new tool for improving the capability to differentiate bulk EVOOs geographically. They generated repeatability data (n=10) for an EVOO and FAME standards (methyl linoleate, methyl oleate, methyl palmitate and methyl stearate) and reproducibility data for five olive oils, whilst using reference materials icosanoic acid methylesters USGS70 and USGS71 for normalising the raw data. They achieved suitable measurement uncertainty of ± 0.3‰ and ± 3‰ for the $\delta^{13}C$ and $\delta^{2}H$, respectively. The authors also established that the contribution of the methylation group (established by EA-IRMS and SNIF-NMR) to the δ-values was negligible (Paolini *et al.*, 2017).

3.1.5 Coffee

Coffee is highly vulnerable to food fraud, and beyond the traditional analytical methods, an effective and practical toolbox for coffee geographical origin determination is needed (Sim *et al.*, 2023a).

For a novel approach, Sim and co-workers performed near infrared spectral imaging of green arabica coffee beans (6 samples per region) from 10 global regions in Africa (Ethiopia, Kenya) and South America (Columbia, Peru) to demonstrate that compositions of 5 stable isotope ratios (H, O, C, N, S) and 33 trace elements (l, B, Ba, Ca, Cd, Ce, Co, Cr, Cs, Cu, Dy, Er, Eu, Fe, Gd, K, La, Li, Mg, Mn, Mo, Na, Nd, Ni, Pb, Pr, Rb, Se, Sm, Sn, Sr, Y, and Zn) could be predicted via modelling (PLS-Toolbox & SOLO 9.0-Eigenvector and R-Studio version 4.2.0). The R^2 ranged from 0.69 and 0.93, with five elements (Mn, Mo, Rb, B, La) being moderately to well predicted whereas three isotope ratios ($\delta^{13}C$, $\delta^{18}O$, $\delta^{2}H$) were well predicted by NIR. Three elements were found as particular markers for country of origin (set of 60 samples), with chromium and lead only being found in African, and selenium in South American samples (Sim *et al.*, 2023b).

A recent article described a novel approach whereby stable isotope and trace element analyses were combined with non-linear machine-learning data analysis to improve coffee origin classification and marker selection (Sim *et al.*, 2023a). Specialty green coffee beans (2021 -2022) were sourced from three continents (Africa, Central America, South America), eight countries, and 22 regions alongside metadata (altitude, mean annual temperature, annual rainfall). By using ensemble decision trees, random forest and extreme gradient boost, accuracies for continental (0.94) and Central American (0.89) regional models showed improved data compared to partial least squares discriminant analysis (PLS-DA) on its own.

The oxygen isotope ratio of α-cellulose extracted from roasted coffee beans (n = 49 from 21 different countries) has been described as a useful indicator to determine region-of-origin, as it integrates source water and climate signals (Driscoll *et al.*, 2020). Cellulose is unaffected by the roasting process and thus the $\delta^{18}O$ values of cellulose can be used at point of sale as no measurable isotopic variability arises from roasting-related changes in the chemical composition of the coffee. The authors analysed the α-cellulose to calculate the $\delta^{18}O$ value of coffee bean water based on the known relationship between the $\delta^{18}O$ value of cellulose and that of the water at the site of synthesis. Hence, the ^{18}O enrichment of coffee bean water was modelled as a function of local relative humidity, temperature and source water $\delta^{18}O$ value. This function was incorporated into a mechanistic model of cellulose $\delta^{18}O$ values to predict the $\delta^{18}O$ values of coffee bean cellulose across coffee-producing regions globally (isoscape). The $\delta^{18}O$ values of the α-cellulose ranged from 22 to 42 ‰ and the modelled values fell within +/- 2.3 ‰ of the measured values. This excellent agreement of practice and theory will allow for the author's model to be used to limit the scope of possible origins prior to the measurement of additional isotopic or trace element parameters.

3.1.6 Honey

Honey is a high value product and assessing its authenticity, geographical origin and floral source is of high interest. Baroni *et al.* (2015) combined data for 33 trace elements (including K/Rb and Ca/Sr ratios) and carbon (bulk and protein) and strontium isotopes of Argentinean honeys (n= 79, harvested across 2007 to 2009 by bee keepers) for three different regions (Buenos Aires, Córdoba, and Entre Ríos) and demonstrated the link of honey to soil and water by applying different chemometric tools (generalised procrustes analysis – 91.5% and canonical correlations – $r^2 = 0.99$) (Baroni *et al.*, 2015).

Kalashnikova and Simonova (2022) studied hydrogen, oxygen and carbon stable isotope ratios of unadulterated honey samples (n=91) taken from Russian regions with different climatic characteristics. The values of isotope compositions varied from -29.5 to -24.2 ‰ for carbon, from -116.6 to -34. ‰ for hydrogen, and from 12.7 to 25.7‰ for oxygen. They found that the average δ^2H and δ^{18}O values in honey correlated with the average δ-values of atmospheric precipitation in the regions of the honey origins. The isotopic composition of carbon was also affected by climate; three zones of "isotopic landscape" for regions of Russia were identified: Siberian honey samples had the lowest δ^{13}C, δ^2H and δ^{18}O values; honeys from the European part of Russia had intermediate δ-values, whilst honeys from the Black Sea region had the highest δ-values (Kalashnikova & Simonova, 2022).

During 2015 to 2017, 373 Turkish pine honey (honeydew honey) samples, produced under controlled conditions without sugar feeding, were collected from 47 sites across 3 different regions (mostly from the Aegean coast, with some from the Marmara and Mediterranean shore). $\delta^{13}C_{protein}$, $\delta^{13}C_{honey}$, C4% sugar, sugars (fructose, glucose, sucrose, and maltose) and a multitude of physicochemical properties - electrical conductivity, moisture, ash, free acidity, colour CIELAB attributes, optical rotation $[\alpha]^{20}$, proline and diastases activities - were determined alongside melissopalynological analyses [number of honeydew elements /number of total pollen (NHE/NTP) ratios (> 3 to be classed as honeydew honey)]. The results showed that all physicochemical parameters exhibited similar values to those of pine and honeydew honeys from other countries, except for C4% sugar (with a mean value of 6.40 ± 5.00 %). By applying PCA, the origin of the pine honey samples could be assigned to three different groups (South Anatolia, North Agean and high-altitude locations) (Uçurum *et al.*, 2023).

3.1.7 Cheese

A working example of using SIRA (and trace element analysis) to protect food against mislabelling relates to PDO cheese such as Parmigiano Reggiano and Grana Padano. Camin *et al.*, 2015 reported an international collaborative study based on blind duplicates of seven hard cheeses was performed according to the IUPAC protocol and ISO Standards 5725/2004 and 13528/2005. The H, C, N and S stable isotope ratios of defatted cheese determined using Isotope Ratio Mass Spectrometry (IRMS), alongside trace element analysis, were measured in 13 different laboratories. The average standard deviations of repeatability (sr) and reproducibility (sR) were 0.1 and 0.2 ‰ for δ^{13}C values, 0.1 and 0.3 ‰ for δ^{15}N values, 2 and 3 ‰ for δ^{2}H values, and 0.4 and 0.6 ‰ for δ^{34}S values, thus comparable with results of official methods and the literature for other food matrices.

3.1.8 Government Research Projects using SIRA

In addition to the above research, UK government-funded projects based on SIRA have been undertaken to address issues in the determination of geographical origin, as discussed below.

3.1.8.1 Projects FA0205 and FS515009 British Beef Origin Projects – British beef Isotope Landscape Map (Isoscape)

Given the mandatory indication of origin for beef and beef products and also concerns following the 2013 horse meat crisis, research was funded to develop methods to independently verify the declared origin of meat on Country of Origin Labels.

An existing Isotope landscape or food origin map (Isoscape) had been created for British beef in the form of a geographical origin decision making web tool with previous funding from the FSA and Defra Seedcorn development funding (ref. Q01123, later transferred to Defra jurisdiction, ref. FA0205). Project (FS515009)[20], conducted by Fera Science, used the methodology established in FA0205 (n= 292 samples from 2007-2010) to improve the robustness of the decision-making tool, which was initially limited by a lack of authentic beef stable isotope data from East Anglia, the Midlands and the South East of England.

The results from this project were used to successfully augment the beef isotope database in 2014 with 200 samples taken from East Anglia, the Midlands and the South East of England in spring and autumn (FA0152). From analysis of blind test samples to determine their geographical origin, it was possible to identify the production region to within an average distance of 202 km (standard deviation = 94). To develop the database into a tool that could be used by enforcers to check geographical origin of a suspect sample, a statistical metric was developed using chi-squared values to determine if a sample was consistent with a specific production area. In this case the

---

[20] FS515009 (British Beef Origin Project 2) (foodstandards.gov.scot)

model was able to correctly determine that samples were consistent with their production origin in 95% of the cases.

The database was further augmented in 2014-2015 with beef data collected from Scotland (n=250) and Northern Ireland (n=49) funded by Food Standards Scotland (FSS, Ref. FS515009)[20] and conducted by Fera Science. Those isotopes that were shown to have utility in the geographical discrimination of beef were hydrogen, carbon, nitrogen, and sulfur. The isotope ratios that were most likely to be affected by seasonality were carbon (relating to feed source, particularly if grass-based feed is enriched with maize or corn during winter months) and deuterium (relating to climate and changes in temperature). The effects of seasonality were found to be minimal for all samples (from England, Scotland and Northern Ireland). The final database of 791 samples was challenged with 15 blind samples (retail samples randomly selected by FSS, country of origin was revealed to analysts post analysis). The procedure identified the 2-letter postcode regions where the samples could have originated from. The method correctly confirmed all samples were consistent with their actual postcode. It was noted though, that the isotopic profiles of several samples were also consistent with other regions of the United Kingdom. Consequently, the procedure was updated to provide options for classification, using different confidence settings. The procedure was also tested with stable isotope profiles from a limited number of these blind samples of non-UK origin, namely from Ireland, Germany, Brazil, Australia, and New Zealand. Using this procedure, Brazilian, Australian and New Zealand samples were clearly distinguished from Scotch beef. A German beef sample could be differentiated from certain parts of Scotland but not from all of them. An Irish beef sample could not be differentiated from Scotch beef. The pasture and/or fodder eaten by the cattle affect the carbon isotope signature of the meat. Plants following a C3 photosynthetic pathway have a carbon isotope signature of around -27 to -24 ‰, whereas C4 plant material possess a more positive carbon isotope signature of -16 to -9 ‰ (BBOP-FA0205). Brazilian, Australian and New Zealand samples come from cattle that are almost exclusively fed on C4 plants (maize), and therefore their carbon isotope signatures were different to those from Scottish cattle which are mostly fed on C3 plants(grass).

The stable isotope ratio analysis data set for Scotch beef was then updated and incorporated into a procedure (UK based web-tool), which was able to (i) confirm the origin of Scotch beef and (ii) detect potential food fraud in geographical mislabelling of Scotch beef, in form of a screening method. Considering that the majority of global production of beef is from the Americas (45% of Global production, 2013, FAOSTAT) this tool would offer protection from the mislabelling of beef as Scotch. Since the last data entry was in 2015, the database would need to be updated and challenged by testing more samples as this would identify any possible variations due to changes in feeding regimes or potential impacts of climate change on the isotopic signatures. The webtool would also need to be updated to support these activities. This scenario exemplifies current weaknesses in origin testing in that investment is made to prepare prediction models for the duration of a project but funding is not available to update, expand and maintain the database in a usable form.

3.1.8.2 Project FA0206 - Assessing the origin of wine using existing compositional information

The primary focus of the EU funded project entitled “Establishing a WINE Data Bank for analytical parameters for wines from Third Countries” (WINE-DB project, G6RD-CT-2001-00646-WINE-DB) was the discrimination of wine samples with respect to their geographical origin using only a few chemical parameters. In a section of the project funded by Defra, Fera Science investigated the possibility of discriminating the wines in the data bank according to their harvesting seasons and grape varieties. Several chemometric methods were selected and evaluated for this purpose. These were discriminant partial least squares, classification and regression trees, uninformative variable elimination discriminant partial least squares and neuro-fuzzy systems. With classification and regression trees, it was possible to identify a few chemical parameters including isotopic ratios (for example $\delta^{18}O$), biogenic amines and rare earth elements that discriminate between vintages and some grape varieties for wines produced in a particular country such as Czech Republic, Hungary, Romania or South Africa. These parameters were used in evaluating the authenticity of wines. It must be noted that this work provided proof-of-principle data only due to the small dataset and was not challenged with samples outside of the databank. This study has highlighted the complexity of defining the composition of wine by the grape variety used and by vintage. Whilst discrimination of different vintages and varieties is possible, it is clear that the variability in wine composition between vintages/varieties is not systematic, requiring different measurements to be made to define a fingerprint for each vintage and variety. This is most probably due to, amongst other factors, fluctuations in climate, variability in the manufacturing process and the sourcing of raw materials. Whilst clearly a consensus definition of a particular wine can be reached in terms of its chemical composition, this study has indicated that this is likely to be specific to the wine that is being analysed unless a wide range of measurements are made. Therefore, false labelling claims will only be identified by the continued maintenance of regulatory databases and/or the archiving of authentic materials, along with the current practice of expert interpretation of the data.

3.1.8.3 Project FA0159 - The development of isotopic and fingerprinting techniques to verify the production origin and geographical origin of food and feed (eggs, poultry and pork)

The aim of this project (2015-2019) was to build on the pilot studies undertaken (FA0130, Development of analytical methods to verify labelling claims relating to egg production) to identify markers, develop stable isotope ratio methods and build reference databases to verify age (and thereby freshness), production origin and country of origin labelling for eggs and poultry. The availability of these analytical tools would support enforcers in responding to anecdotal evidence of mislabelling of the origin of these products.

The geographical origin of feed used by poultry farms may vary, depending on cereal prices within the UK and elsewhere. Therefore, the stable isotope signatures for hydrogen, carbon, nitrogen and sulphur (HCNS) of a farm may fluctuate significantly with the change in geographical origin of the feed, which might provide a challenge for traditional stable isotope databases of solid food products (for example chicken). These are based on the data of the freeze-dried and defatted extract = protein (HCNS isotope ratios) alone, using EA-IRMS. To overcome this, this project by Fera Science took the novel approach to use CO2-eq-IRMS, which is commonly applied to aqueous samples (for example wine, juice). We measured $^{18}O/^{16}O$ isotope ratios of liquid albumen, without any sample treatment, alongside the water on the egg farms. In this way generated $\delta^{18}O$ data allowed us to develop an egg/water model, which can detect substitution of eggs from one source with a secondary source based on a statistically robust minimum sample size, for example 12 chickens, assuming a sufficient $\delta^{18}O$ difference between sources (for example 1.78 ‰). The advantage of this methodology is that a database of samples is not required to determine that substitution has occurred. This egg/water model can be applied to any suspect egg samples as long as drinking water from the egg production site can be sourced for confirmatory analysis. As a final stage to this work, the robustness of the model requires future investigation.

Fera Science also developed a statistical chicken (broiler) model, which uses ^{2}H isotope ratios from farm water and HCNS isotope ratios from feed and chicken. Without the requirement of in-depth databases, substitution/product extension in the UK chickens can easily be detected. This is possible, when other feeding regimes (high maize content affecting ^{13}C signature) compared to the UK were followed, and/or chickens were reared in warmer climates (affecting ^{2}H signature). Our model looked at a range of substitution scenarios (from 5 to 50 %) and the sampling (6 to 300 chickens) required for detecting any substitution. It therefore allows to select a statistically robust minimum sample size for the detection of alien chickens, for example 12 chickens if 30% substitution occurs for a given isotopic difference in the population of 13‰ (^{2}H) or 1.01‰ (^{13}C) and this would confirm that ^{2}H and ^{13}C isotope ratios are most suited for detecting fraudulent activities. We recommend to conduct a small retail survey to put the robustness of the chicken model to the test. Access to feed and drinking water from chicken farm and subsequent analysis will validate the approach.

An outcome of the project highlighted that it has become clearly apparent that there is a need for governmental bodies to provide funding in order to generate and curate relevant isotopic databases, ideally as open access. The EU wine databank is a prime example of an EU membership database, where national reference laboratories receive governmental funding to contribute data. One cannot see the data gathered by other countries but can request verification by the member states of measured data, which allows detection of fraudulent wines. Therefore, government support may be needed to make progress on database sharing issues given the significant challenges that remain. The project highlighted that AHDB holds the largest and most extensive database relating to pork origin which can only be accessed via the British Meat Processors Association pork scheme. Database surveys conducted during this project highlighted that, beside this Intellectual Property issue of commercial databases, the lack of funding and the comparability of the data (standard reference materials and authenticity of samples) present hurdles for creation and sharing of data at the national, European and global levels.

3.1.8.4 Project FA0179 Review of analytical methods, Horizon Scanning and Capabilities for Food Authenticity (Food composition, labelling and Standards)

This project (2018-2022), funded by Defra and conducted by Fera Science, reviewed (i) the current landscape of tools that can be used to verify food authenticity, (ii) what approaches are used in practice within the food industry and (iii) an associated gap analysis where tools are not implemented. In addition, a review of technologies that are implemented in orthogonal sectors was undertaken to identify where already deployed analytical technologies may be sought, if required, in case of an emerging issue. A specific review into UK based commodities with protected food name status was also undertaken to determine whether the current EU system provides sufficient protection of these high value commodities and what aspects could be further developed as the UK sets up its own protected food name scheme after leaving the EU.

The review discovered areas of future considerations for ‘product specifications with geographical indication’ for the UK. In the current format the product specification describes how a product is linked to a given geography (geographical indication). EU 1151/0212 defines what information should be in the product specification, including a description of the product in question and how to maintain the record of a particular specification. One new aspect could be the inclusion of analytical measurements. This would enable screening of products in a laboratory environment with subsequent follow-up by inspectors, where specifications were not met. It was noted that expert knowledge is required to identify appropriate approaches to determine origin depending on the PGO/PDI product. Therefore, it is not possible to generally recommend a single approach; instead approaches should be considered on a case-by case-basis. Some technologies are more suitable to geographical origin verification for example stable isotope ratios, but one needs to acknowledge that isotopic signatures, representing geographical origin, are influenced by various contributing factors, for example having fed animals with feed of different geographical origin than the rearing location. Given the complexities and expert knowledge required around different commodities, it was suggested to have continued discussions with Defra and the UK GI scheme teams to ensure all aspects are adequately controlled; such as (i) achievable verification means, (ii) practicalities of auditing products (both within and external to the UK) and the (iii) desire to promote products outside of the UK for increased exports.

As part of this project, a number of UK product specifications under the EU protected food name schemes (PGO, PGI, TSG) were examined. It was noted that these specifications only included a product description, but not details on possible markers of authenticity or testing methods to verify a produce. Delegated control bodies will have access to authentic samples (produce controls) to verify that a product meets the requirements of its specification. These products command premium prices and therefore present an increased incentive for adulteration. The verification of such products is a complex task and requires knowledge of the aspects that distinguish them from conventional, non-premium equivalents. As the UK has now established its own protected food name scheme, consideration should be given for including recommended analytical tests for product authentication. Given the complexity of the challenge, this scheme may also benefit from external stakeholder engagement and involvement of industry and research providers in this field.

Stakeholder engagement during the project demonstrated that geographical origin is already an important characteristic for consumers when considering food labelling. One of the conclusions of the report was that a dedicated department or even independent body could drive best practices as a central focal point for UK protected designation food produce.

3.1.8.5 EU Framework 6 Project TRACE: Tracing (the Origin of) Food Commodities in Europe (2005 – 2009)

Defra participated in the European Research Project “TRACE”, coordinated by Central Science Laboratory (CSL, now Fera Science Ltd), with 52 participating countries. During the TRACE project, working group 1 (consisting of experts in stable isotope, trace element and strontium ratio analysis) performed analysis of light stable isotopes (H, O, C, N, S), strontium ratios and trace element profiling to monitor the uptake of these profiles from the soil, water, feed (cereal/wheat), olive oil and meat (lamb/beef). The sampling (n=20 region and year) was conducted in 11 European regions in two consecutive years.

The work at CSL was supplemented by similar analyses at a further 15 scientific institutes across Europe to provide analysis of 12,200 food, soil and water samples (over 647,000 element concentration and isotopic data points generated). The data were used to exploit current annual geological and groundwater maps to generate a European Food Origin Map. An extensive multi-element SIRA database of European honeys was also established. There is a large database of unreleased data generated by SIRA and TE for beef, lamb, chicken, wheat, honey and others. Most of this has been published as summaries and predates the review window for this report.

Concerning lamb, the results permitted differentiation of lamb meat, from most production regions, achieving a correct classification rate of 78% (Camin *et al.*, 2007).

The geological origin of the olive oils was characterised based on the content of fourteen elements (Mg, K, Ca, V, Mn, Zn, Rb, Sr, Cs, La, Ce, Sm, Eu, U). By combining the three isotopic ratios (H, C, and O) with the fourteen elements and applying a multivariate discriminant analysis, a good discrimination between olive oils from 8 European sites was achieved, with 95% of the samples correctly classified into the production site (Camin *et al.*, 2010).

Regarding cereals, more than 500 cereal samples collected over 2 years from 17 sampling sites across Europe and representing an extensive range of geographical and environmental characteristics were analysed. For the first time, the potential usefulness of combining Sr, C, N and O isotopic signatures, alone or with key element concentrations (Na, K, Ca, Cu and Rb, progressively identified out of 31 sets of results), was investigated through multiple step multivariate statistics.

From the classification categories compared for cereals (north/south; proximity to the Atlantic Ocean/to the Mediterranean Sea/to else; bed rock geologies) the first two were the most efficient (particularly with the ten variables selected together). In some instances, element concentrations made a greater impact than the isotopic tracers. Validation of models included external prediction tests on 20% of the data randomly selected. This allowed to study the robustness of these multivariate data treatments and to determine measurement uncertainties of the results. With the models tested it was possible to individualise 15 of the 17 sampling sites and therefore this approach demonstrated much potential (Goitom Asfaha *et al.*, 2011).

It was investigated whether honeys produced in regions with different climatic and geological characteristics could be discriminated on the basis of the isotopic data. The H, C, N and S stable isotope ratios of 516 authentic honeys from 20 European regions were analysed. The mean hydrogen isotopic ratios of the honey protein were found to be significantly correlated with the mean hydrogen isotopic ratios of precipitation and groundwater in the production regions. Carbon isotopic ratios were influenced by climate. The sulphur stable isotope composition was clearly influenced by geographical location due to sea spray and surface geology of the production regions. The results show that the ratios of these stable isotopes can be applied to verify the origin of honey. Carbon and sulphur were identified as providing the maximum discrimination between honey samples (canonical discrimination analysis). For seven regions (Allgaeu-Germany, Lakonia-Greece, Sicily-Italy, Cornwall-England, Algarve-Portugal, Orkneys-Scotland and Poland), the percentage of correct classified samples is greater than 70%. It was concluded that the methodology in its current state can be used to provide reliable origin information. The authors highlighted that the discriminatory success of the method may be further improved in the future with consideration of combining the data with that of trace elements, polymerase chain reaction (PCR) analysis and pollen analysis (Schellenberg *et al.*, 2010).

3.1.8.6 Technical Strategy Board Project: Authentication Scheme for novel and Premium British Food and Drink (Authentick)

In 2011 the Technical Strategy Board (TSB) funded the ‘Authentication Scheme for novel and Premium British Food and Drink (Authentick) under the Nutrition for Life programme.

The very ambitious aim was to develop isoscape models for Scottish Whisky, British meat (beef, pork, lamb and chicken), British eggs and British honey, based on stable isotope data (H, O, C, N, and S). Data sets for statistical analyses (isoscape) were supplemented with results for samples from parallel funded Defra projects (FA0152 and FA0129). Sufficient data (n = 100) was only available for meat and egg commodities to allow for the isoscapes to be created and tested by close loop validation. The following success rates were achieved to correctly assign samples at post code/government office resolutions: beef 25%, lamb 5 %, pork 6 % and egg 30%. The error rate in geographical origin assignment at both postcode and government office resolution was unacceptable and is likely to be attributed to factors such as poor sample distribution (not enough sampling of production areas), and/or inclusion in the isoscape and insufficient number of samples. The data are now held in a private database by a private company.

3.1.8.7 EU Framework 7 Project FOODINTEGRITY: Ensuring the integrity of the European food chain (2014 – 2018)

FoodIntegrity was a 5-year, interdisciplinary project funded under Framework 7 and led by Fera Science, that aimed to assure the integrity of our food. The project comprised 60 participants from EU Member States, China and Argentina. The main impact of FoodIntegrity was that all stakeholders are now better informed, share best practice, have better networks, better tools, improved methods and systems for addressing food fraud and for improving consumer confidence in the food they eat. A large database was generated which encompasses numerous methods for geographical origin testing. This database is held by the European Commission Joint Research Centre (JRC) and is no longer online so was unavailable to consult for the duration of this project. The individual methods were instead captured by the extensive literature review, incorporating the range of technologies discussed in this report.

3.1.9 Limitations and gaps in SIRA analysis

As for all analytical techniques, SIRA does not always provide unequivocal information about the geographical source of a food sample. As concluded in project FS515009[21], the statistical comparison of a profile to a database of samples with a known origin can provide a very high confidence that a sample has been mislabelled (for example >95%); in other cases mislabelling may be indicated, but with lower confidence (for example 80%). Hence, users of this approach (and similar approaches) need to make their own judgement as to whether the data are inconsistent with the declared geographical origin (indicating incorrectly declared origin), often backed up with other

---

[21] FS515009 (British Beef Origin Project 2) (foodstandards.gov.scot)

forms of evidence such as chain of custody evidence about the appropriate response to results, which show (with a particular confidence) that a sample may have been mislabelled.

Generating and curating stable isotope databases requires a significant investment, hence the majority of such databases are not open access. For this reason, published articles report only the mean or median values for a collection of samples and not individual data points. This means that statistical evaluation of the data by others cannot be easily performed. An exception is the precipitation database of the university of Utah, which enables access to isoscapes for ^{18}O and ^{2}H. SIRA employs reference materials (RMs) to determine stable isotope ratios of a given sample. In an ideal scenario, these should (a) match the matrix of the sample in question and (b) be available in pairs of the same matrix to stretch the entire δ-scale of the samples to be analysed. The first steps have been made with a combined effort of multiple institutes to produce reference materials produced by the United States Geological Survey, reference numbers USGS82 to 91, which relate to (i) two honeys from Canada and tropical Vietnam, (ii) two flours from C3 (rice) and C4 (millet) plants, (iii) four vegetable oils from C3 (olive, peanut) and C4 (corn) plants, and (iv) two collagen powders from marine fish and terrestrial mammal origins (Schimmelmann *et al.*, 2020). Zhao and co-workers produced two defatted beef reference materials, CAAS-1901 and CAAS-1802 (Zhao *et al.*, 2019), with substantially different δ^{13}C values (due to difference in dietary intake) which were measured and confirmed by nine international laboratories. However, δ^{15}N values fell within the analytical error of the method (0.2 ‰) and therefore, these materials on their own are not suitable for nitrogen data scale correction. Three laboratories (Chartrand *et al.*, 2022) assigned δ^{13}C values for two vanillin certified reference materials, VANA-1 (-31.3 +/-0.06 ‰) and VANB-1 (-25.85 +/-0.05 ‰) and recommended to establish a third vanillin reference material close to -15 ‰ to cover the entire carbon delta scale.

It is important that robust databases are created, capturing natural variability, for example by including samples gathered across seasons and years and including pure and processed foods.

It is clear from the literature review that there is a lack of examples of food industry funding research in country of origin verification. Should industry become more involved, sampling would capture the natural variability of pure and processed foods. Also, there are few cited examples of implementing databases into working systems to monitor supply chains. There would be great benefit in this form of implementation, especially at the points of greatest vulnerability. Ideally, an infrastructure would be in place for vulnerable food commodities, with databases being prepared and maintained, in order to quickly test a given commodity to identify issues and address these issues before they become a problem.

There are areas where databases have been developed and these can easily be translated to similar products. As a hypothetical example, for corned beef there is a quick ^{13}C test to detect South American beef in UK corned beef. Generally, South American beef cattle are fed on maize (C4 plant with δ^{13}C range -14 to -12 ‰) and UK beef cattle are fed almost entirely on grass (C3 plant with δ13C range -30 to -23 ‰) with very little supplementation of maize. This should be reflected in the isotopic signature of the meat. The proportion of C3 versus C4 photosynthetic plant material in the cattle's diet determines whether beef has an isotope signature of around -27 to -24 ‰ or of around -16 to -9 ‰, respectively, if cattle are almost exclusively fed on C4 plants (BBOP-FA0205). Hence, there should be no need to generate a specific database for corned beef to answer the question of whether UK corned beef has in fact been substituted for South American beef. Without requiring a database for corned beef, one could analyse a suspect corned beef sample and compare it directly to the Fera UK beef database, developed during FSA project FS515009.

An additional limitation is the lack of proficiency testing schemes for geographical origin verification. To our knowledge there are several proficiency testing schemes (for example organised by Eurofins since 1999, International Atomic Energy Agency (IAEA); FIRMS network facilitated by Axio/LGC) in place which allow stable isotope laboratories to demonstrate the accuracy and precision of measurements or ability to detect adulteration in honey (determining the C4 sugar addition, when comparing the carbon isotope ratios of honey with honey protein – AOAC 998.12 method), but none of these focus on geographical origin. Indeed, to our knowledge and from accessing the global proficiency testing providers database EPTIS, there are no proficiency testing schemes for any technology which address geographical origin. Schemes tend to be undertaken when there is sufficient participant interest when approached by PT providers. This lack of schemes may therefore highlight a very large gap in capability and/or engagement.

3.1.10 Conclusions for SIRA analysis

Overall, SIRA is still the method of choice for verification of geographical origin of food and feed, despite being costly to run compared to handheld spectroscopy devices; as long as established databases, built on stable isotope data (for example EU wine databank, AHDB pork, Italian cheeses and intellectual property (IP) from commercial laboratories), fulfil their purpose. Ensuring that databases capture natural variation is key and the data should be updated regularly to account for natural variation. Once databases are established, investment must be made to implement them into working systems to monitor supply chains, especially at points identified as being of greatest vulnerability. In most cases SIRA is combined with trace elements and/or strontium ratio analysis, followed by application of chemometrics. In future, the collaboration of analysts and statisticians will be paramount for selecting the most appropriate approach for fusing data from different analytical techniques and origins (sample or environmental factors). On the one hand, stable isotope analysts might not always possess an in-depth knowledge of statistical modelling or be aware of the required size of data sets to perform meaningful statistics and, on the other hand, statisticians might lack the understanding of the analytical data and their significance to verify country of origin.

Regarding the maintenance of databases and updating the samples on a regular basis, having a break in the collection of samples for a particular commodity does not necessarily invalidate using historical data for origin determinations. Instead, ‘old data’ should be considered and may be tested against contemporary data to examine temporal trends, or indeed to establish if data has remained stable in the intervening period. Such multivariate tests can be completed relatively easily to classify old data and vice-versa assuming the methods of analysis e.g. sample preparation and fractions analysed have remained the same. However, this would incur a delay should a method be required to address an issue which arises in the supply chain. Therefore there are benefits in maintaining databases on a regular basis.

3.1.11 Outlook for SIRA analysis

We envisage that in the first instance, non-destructive screening techniques (for example NIR spectroscopy, multispectral imaging) alongside paper trail or artificial intelligence (AI) will be used to verify geographical origin, followed by SIRA in conjunction with trace elements and/or strontium ratios including chemometrics to confirm results if necessary. This scenario allows the confirmatory methods which are more costly to run, only to be used when absolutely needed; whilst continuing to deter people from mislabelling activities through the latest screening methods.

3.2 Trace element analysis

Profiling of foodstuffs using inductively coupled plasma (ICP) – mass spectrometry (MS), ICP – optical emission spectrometry (OES) and microwave plasma (MP) – atomic emission spectrometers (AES) to determine the trace element concentrations has been used to determine / support the verification of the geographical origin of a range of food and beverages. The trace elements present within the soil depend on a number of factors including the type of rock the soil originated from, the pH, moisture content, clay content, topographic features of an area, climate, time and human activity. Elements such as rubidium, strontium and calcium which are associated with geology (rock formation, bottom of soil layers) provide valuable sources of data which are not influenced by soil geochemistry. Other elements are more linked to soil chemistry, occurring in the top layer of soil.

The elemental profile of plants is linked to that of the soil in which the plants are grown which in turn impact the elemental profile of the grazing animals. In addition to the geographical origin, the trace element composition can also be influenced by the genotype, environment and their interactions. As these factors impact the concentrations of the trace elements, it is not always possible to use this data in isolation to confirm the origin of a food or beverage. Examples of commodities for which working databases include trace elements are cotton and PDO cheeses (Grana Padano and Parmigiano Reggiano).

Katerinaopolou and co-workers published a “systematic literature review of geographical origin authentication by elemental analytical techniques”, focused on papers published between 2015 and 2019. Of the 155 papers selected as part of this review 28 described the contribution of ICP-MS as a tool to support the identification of the geographic origin. In many cases the trace element data was combined with other techniques, in particular IRMS, to demonstrate the origin. The review referenced examples of origin determination for potatoes, lettuce, peppers, tomato, onions, rice, flour, cereals, oranges, herbs, olive oil, scallops, eel, sea bass, clams, honey, lamb, milk and cheese using trace element and isotope ratio data (Katerinopoulou *et al.*, 2020).

Applying the search criteria previously described, 47 papers were selected for review. Additional information was derived from grey literature such as instrument manufacturers application notes and other papers of interest as referenced in the selected manuscripts. The foods and beverages in which trace element analysis has been applied for origin identification in these publications are discussed.

3.2.1 Cereals

Ten varieties of wheat were cultivated in three regions of China (Hebei Province, Henan Province and Shaanxi Province). The samples (270 samples, 90 from each region harvested across three years) were randomly selected and were analysed for thirteen isotopes: ^{24}Mg, ^{27}Al, ^{44}Ca, ^{55}Mn, ^{56}Fe, ^{63}Cu, ^{66}Zn, ^{75}As, ^{88}Sr, ^{95}Mo, ^{111}Cd, ^{137}Ba and ^{208}Pb by high resolution ICP-MS. Differences in element concentrations were observed for all except Cu and Pb. Multiway analysis of variance demonstrated that geographic origin was the most important source of variation for Mn, Sr, Mo and Cd, genotype for Ba and the harvest year for the remainder of the elements tested. The concentrations of the elements in the soil correlated with those in the wheat for Mg, Ca, Mn, Sr and Cd. Conversely, the concentrations of As and Mo in the soil and wheat were negatively correlated. There was no correlation for the other elements measured. The authors (Liu *et al.*, 2017) concluded that four elements (Mn, Sr, Mo and Cd) were “closely related to geographic origin” and that these elements could be used to “reliably and effectively trace wheat geographical origin”.

Zhang *et al.* (2021a) analysed eight elements in highland Tibetan highland barley and soil from five regions across Tibet. Tibetan highland barley is claimed to be rich in nutrients and have unique health care functions (Yu *et al.*, 2016) leading to the research to establish if origin can be determined. As well as trace element analysis the authors determined crude starch, crude protein and crude fibre. 126 Highland barley and associated soil samples were collected from 5 Tibetan cities and analysed by ICP-

OES. Multivariate statistical analysis was carried out (principal component analysis, partial least squares discriminant analysis and orthogonal partial least squares discriminant analysis) which the authors state can “clearly classify samples from different regions”. There was a positive correlation between the zinc and iron content in the barley and the soil and a negative correlation for potassium, manganese and phosphorus (Zhang *et al.*, 2021a).

3.2.2 Cocoa

Almost 10 years ago, the International Cocoa Organization (2014) reported that cocoa beans are a top ranked commodity with their price linked to geographical origin and therefore, the ability to identify the geographic origin of cocoa beans is required to prevent fraudulent activity.

Bertoldi *et al.* (2016) demonstrated the application of multi element analysis to determine the geographical origin of cacao beans and cocoa products. They analysed 61 cacao bean samples from 23 countries across Africa, Asia and Central and South America using ICP-MS. 29 elements (Ag, As, Ba, Be, Bi, Ca, Cd, Co, Cr, Cs, Cu, Fe, Ga, Hg, K, Li, Mg, Mn, Na, Ni, P, Rb, Se, Sr, Th, Tl, U, Y and Zn) were included in the model derived. Where sample numbers were high enough the statistics applied reclassified all of the cocoa beans to the macro-area of origin, i.e. Africa, Asia and Central and South America (Bertoldi *et al.*, 2016).

Chocolate, made from seeds of the cacao tree, can be contaminated with cadmium and lead by polluted soil. Lead can also contaminate the cocoa beans after harvest, potentially from dust and soil during drying of the cacao beans. There have recently been reports of high levels of these elements in dark chocolate (which contains higher levels of cacao solids). Work has separately been carried out to investigate the link between the metal concentrations in chocolate to the country of origin and in so doing determining the source of the highest levels of lead and cadmium in cacao beans. Following the analysis of 139 single origin chocolates for Ca, K, Mg, Na, P, S and the trace elements Al, As, B, Ba, Be, Cd, Co, Cr, Cs, Cu, Fe, Ga, In, Li, Mn, Mo, Ni, Sb, Se, Sn, Sr, Pb, Ti, Tl, U, V, Zn and Z, researchers (Vanderschueren *et al.*, 2019) developed a decision tree that allowed the differentiation of chocolate from Africa, Asia Pacific, Central America and South America following Classification and Regression Tree (CART) analysis. The concentrations of all elements in the chocolate, bar Cr and V, correlated with the cacao content indicating the cacao as the source of the elements. 16 of the samples tested exceeded the European limit for cadmium for chocolate with 50% or higher cacao content. It was primarily the high Cd content in the South American samples that differentiated their origin with the addition of Mo content allowing separation from Central American samples. Cd concentrations in African and Asian Pacific samples were lower. Ba, Sr and Zn content could be used to further differentiate the samples. It was not possible to classify the Asia Pacific samples which the authors proposed to be variable soil composition and climate in the region (which included samples from India, Indonesia, Papua New Guinea, Samoa, Vanuata and Vietnam). The overall misclassification rate (all origins) was 23% based on the concentrations of the Cd, Mo, Ba, Sr and Zn (Vanderschueren *et al.*, 2019).

3.2.3 Coffee beans

There have been numerous studies to determine the geographical origin of coffee beans mainly applying isotopic methods to determine the ratios of $^{11}B/^{10}B$, $^{13}C/^{12}C$, $^{15}N/^{14}N$, $^{18}O/^{16}O$, $^{34}S/^{32}S$, and $^{87}Sr/^{86}Sr$ [(Rodrigues *et al.*, 2011a), (Rodrigues *et al.*, 2011b), (Rodrigues *et al.*, 2009), (Serra *et al.*, 2005), (Wieser *et al.*, 2001)].

Work by Rodrigues and collaborators showed that for coffee bean samples from the Hawi'ian islands, the combination of S, O, C, N, and Sr isotope analyses with multielement analysis allowed differentiation of the different Hawai'ian coffee-producing regions. The results indicated relationships between environmental variables and the green coffee bean isotopic composition. The authors state that additional work is needed to clarify the mechanisms underlying many of these relationships, however, the results suggested that the isotopic composition of coffees from different regions may, to some degree, be predictable. If so, this would support the use of stable isotopes as a tool for the verification of coffee origin. In addition, the coffee plant seeds' isotopes may contribute to tracing environmental impacts occurring in Hawaii, in particular if related with volcanic activity, distance to the ocean, and altitude (Rodrigues *et al.*, 2011a). However, others have reported that the isotope ratio data for Sr alone did not differentiate the samples sufficiently to determine the country of origin (Liu *et al.*, 2014). Instead, they analysed 21 Arabica beans for B, Rb, Sr, Ba, Fe, Mn and Zn and the $^{11}B/^{10}B$ and $^{87}Sr/^{86}Sr$ isotopes. Samples, with assured origin, were obtained from Africa, America and Asia. The authors reported that the concentrations of Rb, Sr and Ba could be used to classify the origin, but it was the isotope ratio data for Sr and B that provided the more sensitive information linked to the origin.

3.2.4 Fish and seafood

Fish and shellfish represent the third most notified commodity for misdescription of origin. There has been a lot of research carried out on the authenticity of fish following reports of species substitution, however with geographical location attracting higher market prices, methodology for geographical origin determination is needed in addition to the DNA testing used for determination of fish speciation. Elemental concentrations vary greatly in the marine environment and are controlled by a combination of natural and anthropogenic activity (Rainbow, 2017).

Han *et al.* (2021) determined the concentrations 14 elements (Al, Ca, Co, Cr, Cu, Fe, Ga, K, Mg, Mn, Na, Ni, Sr, and Zn) in salmonids with the aim to determine the geographical origin of these species. Factors affecting the concentrations in the samples (n = 96) included harvest time (Fe, K, Na, Mn and Zn), fish size (Al, Fe, Ga, Mn and Zn) and whether the samples were freshwater-cultured or seawater-cultured (Ga, K, Mg, Na, Ni and Sr) as well as the geographical location (Al, Ca, Co, Cr, Fe, Ga, K, Mg, Mn, Na, Ni, Sr, and Zn) . Samples (freshwater and seawater) were collected from various regions in China and Chile. Analysis was carried out by ICP-AES. It has previously been reported that the elemental composition of fish is impacted by climatic variation (Mendil *et al.*, 2010) and (Fallah *et al.*, 2011). In this study the concentrations of Mn, Na and Zn were influenced by harvest time and climate, however, Fe and K were also impacted when harvested at different times but cultured in a controlled system. Thus, the differences observed could not be fully attributed to differences in the climatic conditions. Models developed following LDA of the data from 12 elements (Ca, Co, Cr, Fe, Ga, K, Mg, Mn, Na, Ni, Sr, and Zn) were able to predict geographic origin with an accuracy of >90%. The authors acknowledge that the study “had some limitations” based on limited sampling points, individual samples, and influencing factors and proposed that additional work is required to refine the models to increase the success rate for the identification of origin (Han *et al.*, 2021).

More recently, the use of machine learning techniques has been applied to the determination of origin. Using data from elemental analysis Bai and co-workers were able to determine the origin of crayfish discriminating between three sites in China. 10 elements (V, Fe, Al, Ga, Co, Zn, Cs, Rb, Ba, and Sr) were found to be characteristic elements of red swamp crayfish in different regions of China (Bai *et al.*, 2022).

REIMS and ICP-MS have been applied to classify salmon production methods and origin (Hong *et al.*, 2023). Samples (n = 522) produced by two production methods (farmed and wild caught) were collected from four regions (Alaska, Norway, Iceland and Scotland). Using trace element data coupled with lipid marker data and applying chemometric modelling (PCA) with machine learning all test samples (n = 17) were correctly classified for both production method and origin. 20 elements were used to develop the model (Li, B, Al, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, As, Se, Rb, Sr, Nb, Mo, Cd, Cs, and Ta). Without the fusion of data (with the lipid marker data) the accuracy of the origin determination was lower (66%). Within this article the authors also summarised the findings of other research groups using ICP-MS to determine origin including the combination of ICP-MS and NIR spectroscopy that was able to differentiate between Chilean-farmed and Norway-origin salmon (Fu *et al.*, 2021).

A similar approach has also been applied to scallops (Morrison *et al.*, 2019) but in this example trace element data alone was able to discriminate between the harvesting sites (3 locations on the west coast of Ireland). Ba, B, Cr, Ob, Mn, Mo and Se were the elements included in the model. The authors highlighted the importance of “periodical verification of reference chemical signatures” and that any “reference library must therefore be continuously updated to enable successful classification over time”.

3.2.5 Fruit juice

The concentrations of 25 elements (K, Na, P, Ca, Mg, Al, B, Ba, Be, Co, Cr, Cu, Fe, Li, Mn, Mo, Ni, Sb, Se, Si, Sn, Ti, Tl, V, Zn) in 36 prickly pear juice samples and the soil in which they were grown (3 regions in Greece) were determined by ICP-OES (Karabagias, 2019). There was correlation between the trace element concentrations in the soil and the fruit and the geographical origin could be predicted once the trace element data (seven elements were used in the model) was combined with the data derived from the analysis of the volatile substances (21 tentatively identified volatile substances were included in the model). The resulting model classified the prickly pear juice to the correct geographical origin with a success rate of >85%.

Multivariate analysis of trace element concentrations measured 482 in Australian and Brazilian orange juice samples showed a clear differentiation between them (Simpkins *et al.*, 2000). Both juices and orange peel samples were analysed. Regional differences were observed in Australia linked to differences in element concentrations in the soil as well as between Australian juices and Brazilian juice concentrate samples. Ru, Ba and B contribute the main differences between Australian and Brazilian samples. Peel samples could similarly be differentiated between the two countries.

3.2.6 Garlic

Analytical screening techniques have been used to determine the composition of Chinese garlic. 34 trace elements, 68 volatile compounds and 854 metabolites were detected using a suite of analytical techniques and used to develop a chemometric model (Mi *et al.*, 2021). A large number of parameters were required to differentiate the Chinese garlic and it was not possible, in this study, to achieve this using the trace element data alone The authors also reviewed other studies [(Camargo *et al.*, 2010), (Ahn *et al.*, 2019), (D'Archivio *et al.*, 2019), (Liu *et al.*, 2018), (Vadalà *et al.*, 2016)] that have been carried out to determine the geographical origin (across a number of regions) of garlic and from these studies summarised that trace element profiles could, to some extent, be applied to determine the origin.

Ahn and collaborators (Ahn *et al.*, 2019) investigated the difference between domestic garlic from South Korea and imported garlic from China for pH, moisture content, total flavonoid content, and all trace minerals except for manganese and magnesium. They used logistic regression analysis to determine the geographical origin (South Korea or China) of garlic after selecting the appropriate independent variables. As a result, the calculated logistic regression equation from the analysis of copper, iron, phosphorus, zinc, and sucrose could be used to determine whether the geographical origin of garlic was South Korea or China. In contrast, Camargo and colleagues (Camargo *et al.*, 2010) determined the mineral content of ten garlic cultivars from Argentina and attempted to establish a relationship between the mineral contents of the garlic cultivars and their geographic origins through the use of neutronic activation analysis and PCA. Whilst they were able to demonstrate that cultivars cultivated under identical agri-environmental conditions could be categorised into groups by using just four metals as variables they were unable to demonstrate that the use of mineral profiles constituted an adequate tool for determining the geographic origin of garlic.

Similarly, in another publication (Vadalà *et al.*, 2016) the results of the analysis of garlic samples from Sicily, Tunisia and Spain by ICP-MS were compared. Twelve samples were analysed. Despite the small sample set the authors described the differences in the metal content of the samples linked to the agri-environmental conditions in which they were grown. The main finding of the study was the higher levels of Se in the Nubia Red Garlic samples from Sicily which the authors stated “is also useful to demonstrate that Nubia Red Garlic shows important health qualities and could be used as an anticarcinogenic agent”.

3.2.7 Honey

Four papers were selected that describe the determination of the geographical origin of honey based on the elemental composition. A further article of interest was identified in the preparation of this review. Honey is another high value commodity for which fraud has been reported. Mānuka honey is a premium priced product due to its purported antibacterial properties and health benefits attributed to it. Therefore, numerous research groups have been working to define methodologies to confirm the origin of a given honey whether it be Mānuka honey from New Zealand or another source.

Grainger and collaborators carried out a study to see if New Zealand honey could be differentiated from honey samples produced in the rest of the world based on the element profile. 352 honey samples from 34 different countries were sourced from across the globe (Grainger *et al.*, 2023). All samples were derived from nectar and a combination of mono- and multi-floral botanical origins. The samples were analysed by ICP-MS for B, Na, Mg, Al, K, Ca, Cr, Mn, Fe, Co, Ni, Cu, Zn, Ga, Rb, Sr, Cd, Cs, Ba, Hg, Tl, ^{206}Pb, ^{207}Pb, ^{208}Pb. Using a decision tree approach with five terminal nodes, the New Zealand samples could be distinguished from the rest of the world with 92% accuracy. Therefore, the results show that the approach may be a promising tool to determine whether a honey sample originated in New Zealand. However, additional samples will need to be analysed and added to the database to better capture the global variation.

Differences in element concentrations in honey samples prepared in different years is also needed to improve the robustness of the model.

Not all studies have been able to demonstrate the applicability of elemental analysis for origin determination when considering regions within a given country. Researchers in Uruguay (Berriel *et al.*, 2019) analysed 25 samples of honey collected from apiaries throughout the country. Although the concentrations of the elements differed in the honey samples, the statistical analysis carried out could not differentiate between regions. This was attributed to the variability in the soils within each region.

A recent study (Pavlin *et al.*, 2023) determined the concentrations of 18 elements in 173 honey samples covering 13 floral types and from five regions (Slovenia, Croatia, Bulgaria, Turkey, and Morocco). The main aim of the work was to assess the differences in the botanical sources, however the levels of Na, Mg, and Fe were found to be more influenced by environmental factors and so the authors summarised that these elements could be considered as markers of geographical origin.

Spark discharge-assisted laser-induced breakdown spectroscopy was used (Fechner *et al.*, 2021) to analyse 49 composite Argentinian honey samples. The samples were harvested across two years (2015 and 2016). Ca, K, Cu, Fe, and Mn were reported to be important indicators for the geographical origin of honey.

Bracatinga honeydew honey is exclusively produced in three regions of Brazil. Thirty nine elements were analysed in 34 Bracatinga honeydew honey samples (Silva *et al.*, 2021) obtained from across the three regions (Santa Catarina, Paraná, and Rio Grande do Sul). All samples were harvested in 2018. Statistical analysis (cluster analysis, principal components analysis, and linear discriminant analysis) allowed for the differentiation of the honeys from the different regions with 91% accuracy with Rb and Co highlighted as the primary elements of interest.

3.2.8 Meat

There are many reports where trace element data has been used to support the development of models for the determination of geographical origin of beef, pork and lamb. For example, the group of Fernandes analysed beef samples from various regions in Brazil for their elemental composition with the aim being to develop an internationally accepted traceability system for Brazilian beef products. Neutron activation analysis was used. Differences could be observed in the data sets obtained from the different regions using the three machine learning algorithms applied (Multilayer Perceptron, Random Forest and Classification and Regression Tree) and multivariate statistics (Fernandes *et al.*, 2020).

The Chinese Ministry of Agriculture identified Yanchi Tan sheep as an "Agro-product geographical indications" in 2008 and "Chinese Protected Designation of Origin" by the General Administration of Quality Supervision. Due to an increase in fraudulent activity methodology was needed to be able to identify the Tan sheep from the Yanchi region. Liu and co-workers (Liu *et al.*, 2021) collected lamb samples from across China and subjected them to multi-element and stable isotope analysis. Feed and soil samples were similarly analysed. The samples collected were from 2017 and 2018. Correlation was seen between the trace element concentration in the lamb samples and the soil. The authors reported differences in the stable isotope ratios as well as the elemental profiles among the different regions the samples were collected from. Concentration of Sr and Mo along with the stable isotope data could differentiate the Yanchi Tan lamb from other regions, however, further work is needed to consider lamb from other geographical locations to obtain a robust model.

Another study analysed pork belly fat samples from USA, Spain, Canada, Germany, Mexico, Chile and South Korea. The concentration of 19 elements (Al, Ca, Fe. P, K, Na, S, Zn, As, Cd, Cs, Cr, Li, Mn, Ni, Rb, Se, Sr and V) were determined along with the isotopes of the trace elements: ^{6}Li, ^{7}Li, ^{116}Cd, ^{50}Cr, ^{51}V, ^{52}Cr, ^{53}Cr, ^{55}Mn, ^{58}Niu, ^{75}As, ^{85}Rb, ^{82}Se, ^{84}Sr, ^{85}Sr, ^{87}Sr, ^{88}Sr and ^{133}Cs were measured. Linear discriminant analysis and principal component analysis were applied to the data sets for the 480 pork belly fat samples tested (350 domestic and 130 imported). Domestic samples could be differentiated from imported samples with a discrimination rate of 98%. The authors concluded that trace element concentrations and isotope ratios are "promising descriptors which may be used to reasonably discriminate the geographical origins of the pork samples" (Nho *et al.*, 2019).

The use of trace elements concentrations to determine the origin of pork from different regions of China was reported by Qi and collaborators. All seven regions studied (10 samples were collected from each region) had a distinct element profile. Data from 11 elements (Ca, Mn, Fe, Cu, Zn, Se, Rb, Sr, K, Na, and Mg) was used in the statistical analysis. The authors acknowledge that the data set was small and requires further input data to refine the models to generate a comprehensive method for origin determination (Qi *et al.*, 2021).

The group of Varrà reported the development of a "promising method to confirm the declared pig meat label attributes, deter potential complex fraud, and support meat traceability systems". 80 pig muscle and 80 pig liver samples were analysed by ICP-MS for 57 elements. Three sets of the samples consisted of meat from the PDO Parma Ham production process and a fourth group that were not compliant with the Parma Ham production process. Samples could be separated using the multivariate statistics applied but additional work is proposed to refine the models into useful traceability tools (Varrà *et al.*, 2023).

Further examples of the use of element profiling to successfully verify the geographical origin have been published. For example (Rees *et al.*, 2016), assessed poultry samples, mainly chicken, from Argentina, Brazil, Chile, China, Thailand and Europe (Austria, Czech Republic, Slovakia, Denmark, France, Germany, Hungary, Ireland, Italy, Netherlands, Poland and the United Kingdom. The study used stable isotope and elemental analysis, together with statistical processing of the resultant data to determine the geographical origin of the poultry. Interestingly the carbon stable isotope ratios of chicken meat indicated the quantity of maize in the diet leading to a useful discrimination between a large proportion of European poultry and poultry reared in locations such as South America, Thailand and China where maize feeding predominates. It was stated that the use of poultry carbon isotope values as a simple 'screening' parameter to differentiate European poultry meat from other major importers was not as reliable as for the differentiation of European and South American beef. However, carbon isotope ratios would be useful in most instances to corroborate suspicion of mislabelling of non-*corn-fed* European poultry. A study by Heaton and co-workers (Heaton *et al.*, 2008), analysed beef samples originating from the major cattle producing regions of the world (Europe, USA, South America, Australia and New Zealand) by IRMS and ICP-MS and was successfully able to discrimination between beef samples on the basis of the broad geographical areas (Europe, South America and Australasia). Although the authors included the caveat that the methodology, in its current state, could be used to provide reliable origin information, but that it was dependent upon the countries under investigation.

Kim and collaborators (Kim *et al.*, 2017) presented a study aimed at determining the concentrations of twenty-nine elements in 323 pork belly samples including 227 domestic (from Suncheon, Naju, Chungju, Gangjin and Yongin cities of South Korea), and 96 samples imported from USA, Germany, Austria, Netherlands and Belgium. The macro elements including Al, B, Ca, Fe, K, Mg, Na, P, S, and Zn were analysed by ICP-OES, whereas trace elements including Ba, Be, Bi, Cd, Co, Cr, Cu, Cs, Ga, Li, Mn, Ni, Pb, Rb, Se, Sr, U and V were analysed by ICP-MS, with multivariate data analyses of PCA and LDA. It was found that analysis of the trace elements were a promising approach for determining the geographical origins of the pork, where a discrimination index of 97% differentiated pork originating from different countries.

Finally, in a small single country study, (Sun *et al.*, 2011) researchers were able to successfully discriminate between mutton samples from agricultural and pastoral regions in China with 100% accuracy using ICP-MS analysis of Be, Na, Al, Ca, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, As, Se, Ag, Sb, Ba, Tl, Pb, Th and U (Sun *et al.*, 2011).

3.2.9 Olive oil

Another high value commodity subjected to fraudulent claims is olive oil. Two papers were identified that described the use of multi-element profiling to determine geographic origin.

In the first (Beltrán *et al.*, 2015), trace element profiles from Spanish olive oil and corresponding pomace from several different cultivars along with the corresponding soil samples were analysed for 34 elements. The olive oils were characterised by the levels of W, Fe, Mg, Mn, Ca, Ba, Li and Bi. The impact of fertiliser and fungicide used on the levels of the trace elements in the resulting olive oil needs to be investigated prior to standardising the methodology for the origin determination of Spanish olive oil.

In the second study (Telloli *et al.*, 2023), triple quadrupole ICP-MS analysis was used to characterise extra virgin olive oil samples from Italy (24 samples collected over a 6-year period). Concentrations of Be, B, Na, Al, P, K, V, Cr, Mn, Fe, Co, Cu, Se, Ag, Cs, Tl, Pb, Th, U, Mg, Ca, Ni, Zn, Ga, As, Rb, Sr, Cd and Ba were determined. PCA was applied to the datasets to provide a classification according to the region of origin, which the authors claim to be a starting point for continuing work with larger samples sizes to enhance the model.

3.2.10 Rice

Basmati rice is grown in a specific area of the Indo-Gangetic Plains. The properties and flavour of the rice cannot be replicated in other areas. As Basmati rice attracts a premium price there have been a number of issues with authentication over the years and methods have been developed to support the confirmation of the product. Arif *et al.* (2021) “assessed the application of elemental analysis for the authentication of the geographical origin of Basmati rice”. 64 samples (21 of known authenticity) were analysed by ICP-MS. Statistical analysis was performed using the concentrations measured for 35 of the elements tested (^{23}Na, ^{25}Mg, ^{26}Mg, ^{27}Al, ^{29}Si, ^{31}P, ^{33}S, ^{39}K, ^{43}Ca, ^{44}Ca, ^{45}Sc, ^{47}Ti, ^{52}Cr, ^{54}Fe, ^{55}Mn, ^{56}Fe, ^{59}Co, ^{60}Ni, ^{65}Cu, ^{66}Zn, ^{69}Ga, ^{72}Ge, ^{75}As, ^{76}Se, ^{79}Br, ^{81}Br, ^{85}Rb, ^{88}Sr, ^{89}Y, ^{95}Mo, ^{137}Ba, ^{139}La, ^{140}Ce, ^{197}Au and ^{206}Pb). Chemometric analysis (data driven soft independent modelling of class analogy - DD-SIMCA) identified eight elements (^{44}Ca, ^{66}Zn, ^{72}Ge,^{75}As, ^{76}Se, ^{88}Sr, ^{95}Mo, ^{206}Pb) as the key discriminators to confirm that the origin of the rice was the Basmati region. The authors reported that the “sensitivity and specificity of the one class DD-SIMCA model were 100% and 98%, respectively”. They also highlighted the value of coupling the elemental analysis with other techniques, for example isotope ratio mass spectrometry to further refine and improve the model (Arif *et al.*, 2021).

High resolution ICP-MS was used (Cheajesadagul *et al.*, 2013) to establish a model to confirm the geographical origin of Thai rice. 31 Thai jasmine rice and 5 other samples were tested with the concentrations of 21 elements measured. The multivariate analysis differentiated the samples and could also differentiate the Thai jasmine rice samples by region (northern, north-eastern or central regions of Thailand).

A recent study identified 14 elements as markers for the geographical origin of rice (Quinn *et al.*, 2022). Samples (n = 151) harvested in 2018 and 2019 in Vietnam, China and India were analysed by ICP-MS and the concentrations of the elements measured were subjected multivariate statistical assessment. The data showed the Chinese samples had higher concentrations of Ca, Al and Mn. Vietnamese samples had high Zn concentrations and low Ge concentrations and the Indian samples contained higher levels of B, Sr, Se, Cu, Mo, Co, W, Fe and Ti. The differences were “attributed to varying elemental compositions intrinsic to the soils from which they were grown”.

Bui and collaborators looked specifically at rice from Vietnam to establish if the different within country regions where the rice was grown could be differentiated. The authors described a linear discriminant analysis and partial least squares-discriminant analysis model that separated Sengcu rice from other regions. As, Ba, Sr, Pb, Se, Ca differentiated Sengcu rice from other regions with 100% accuracy (Bui *et al.*, 2022). A study investigating the geographical origin determination of hot pepper also reported that the $^{87}Sr/^{86}Sr$ ratios in rice could be linked to the those of the water and exchangeable fraction of soil (Song *et al.*, 2014).

Work by Maione and colleagues differentiated rice grown in the midwest and southern regions of Brazil using statistical models developed using the elemental composition of the rice 31 samples analysed. Cd, Rb, Mg and K concentrations were the primary discriminators (Maione *et al.*, 2016).

70 rice samples and 35 topsoil samples (from the same region of Brazil) were analysed by ICP-MS. Arsenic concentrations were determined in the rice samples by HPLC-ICP-MS to provide total and species concentrations. Others have reported that the variability in the data was linked to "the geographical area, to crop management, producers and in a lower extent to soil composition" (Lange *et al.,* 2019). There was no link between the data for the rice and for soil (applying PCA analysis). The authors suggested the inclusion of stable isotope data would provide a "more robust dataset for traceability".

In a 2016 study looking at the stable isotopes of Cs and Sr as markers for radioactivity Srinuttrakul and Yoshida reported that Cs and Sr concentrations in Thai rice were higher than those for Japanese rice (Srinuttrakul & Yoshida, 2016).

3.2.11 Saffron

As the most expensive spice on the market there have been a number of incidences of saffron fraud in recent times. One example of such fraud is the "rebranding of cheaper growing regions to be passed off as regions of higher quality" (Wakefield *et al.*, 2019). These authors determined the concentrations of 42 elements (Li, B, Na, Mg, Al, K, Ca, V, Cr, Mn, Fe, Co, Ni, Cu, Zn, As, Se, Rb, Sr, Y, Mo, Ag, Cd, Sn, Sb, Cs, Ba, La, Ce, Pr, Nd, Sm, Eu, Gd, Dy, Ho, Er, Tm, Yb, Lu, Pb, U) in 41 Iranian saffron samples and 9 Spanish saffron samples by ICP-MS. In addition, stable isotope analysis was performed. Application of LDA allowed the saffron from Spain to be differentiated from that grown in Iran. The authors also found differences in the profiles year on year and therefore, additional work is required to build on these databases and so better refine the models developed.

3.2.12 Tea

Elemental profiling of Indian tea has been demonstrated (Vinay Jain, 2022) to identify the growing region of the teas within India. Tea samples (n = 150) were obtained from eight regions and the concentrations of B, V, Cr, Co, Ni, Zn, Se, Rb, Sr, Mo, Cs, Ba, Mg, Mn, Al, La, Ce, and Nd were measured by ICP-MS. The application of principal component analysis showed significant variation using data derived from 18 of the elements tested. According to the PCA loading values, the separation between the geographical origin of tea was driven by Sr, Ba, and B for principal component 1; Cs, La and Rb for principal component 2; and Mo, Ce, and Nd for principal component 3. When testing the model with unknown teas, the origin of all 24 were correctly classified.

3.2.13 Tomatoes

As the major tomato producer in Europe, Italy has a commercial interest to be able to demonstrate the authenticity of the fruit grown on its land. A study from 2018 measured the concentrations of 26 elements (Li, Be, Na, Mg, Al, K, Ca, V, Cr, Mn, Co, Cu, Zn, Ga, As, Rb, Sr, Ag, Cd, In, Cs, Ba, Tl, Pb, Bi and U) in 183 tomato-based products originating from Italy, China, US and Spain. Sampling took place in 2013, 2015 and 2017. A total of 169 samples were analysed using Inductively Coupled Plasma orthogonal acceleration Time-of-Flight Mass Spectrometry (ICP-TOF-MS). Multivariate statistical analysis was applied to the data to generate a model to discriminate Italian tomatoes. Three element ratios were identified as being able to identify the Italian tomatoes (Li/Cu, Co/Rb, and Sr/Cd). The dataset was tested against 14 samples described by the authors as “Italian sounding” tomato passata of unknown origin obtained from American companies. Two of the 14 samples where well separated from Italian tomatoes in the statistical model (Fragni *et al.*, 2018)

3.2.14 Whisky

Most studies to determine the geographical origin of whisky have involved GC, LC and spectrophotometric techniques. The use of multi-element data to determine geographical origin of whisky has not been extensively reported. Some authors (Adam, *et al.*, 2002; Pawlaczyk *et al.*, 2019) using ICP-MS reported that their studies could not differentiate between samples of different geographical origins using multi-element data, although Irish whiskey could be distinguished from whisky samples from other countries as it was characterised by quite a high amounts of Ba and Ti. The authors suggested that their analyses were confounded by a small number of samples from each region and the type of whisky (single malt or grain vs blended malt or grain). The authors suggested that future studies should include a greater number of samples originating from different regions of Scotland and countries and from the same distillery.

Gajek and collaborators investigated the potential to use element concentrations to differentiate samples including according to their origin. 170 whisky samples from 11 countries (Scotland, the USA, Ireland, Poland, Japan, the UK, India, Azerbaijan, Slovakia, Wales, and Bulgaria) were analysed by ICP-MS for Ag, Al, B, Ba, Be, Bi, Cd, Co, Cr, Cu, Li, Mn, Mo, Ni, Pb, Sb, Sn, Sr, Te, Tl, U, and V, ICP-OES for Ca, Fe, K, Mg, P, S, Ti, and Zn and Cold Vapor-Atomic Absorption for Hg. The authors summarised the characteristics of the whiskies produced in USA, Ireland, Poland and Scotland. The authors reported that the whisky from the USA was characterised by the highest median values of Li, Be, V, Mn, Ag, Sb, Zn, P, Fe, and Ti when compared with samples from other countries. Scottish samples contained the highest median levels of Cu and Cd. They concluded that “the observed differences only prove that samples from various countries have completely different elemental fingerprints” (Gajek *et al.*, 2022).

Nelson *et al.* (2017)[22] analysed 69 samples (16 Bourbons, 8 Irish, 9 Japanese, 1 Rye, and 2 Tennessee whiskey products and 33 Scotch whisky products) by Microwave Plasma-Atomic Emission Spectroscopy (MP-AES). Concentrations of Al, Ba, Ca, Cu, Fe, K, Mg, Mn, Na, Rb, Si, Sr and Zn were determined to be statistically different. Differences in Cu

---

[22] Elemental Profiling of Whiskey using the Agilent 5100/5110 ICP-OES and MPP Chemometrics Software (agilent.com)

concentrations were proposed to be linked to processing conditions rather than the origin as copper is used in the distillery process.

3.2.15 Wine

As a commodity for which there have been numerous reports on counterfeit products then it is not surprising that the literature review identified the most publications on wine.

Cellier *et al.* reported the use of $^{87}Sr/^{86}Sr$ and $^{208}Pb/^{206}Pb$ ratios to differentiate between champagne and other sparkling wines from across the globe. Samples included Spanish Cava, Italian Prosecco, sparkling wines from USA, Brazil, Chile, China, Argentina, South Africa, England and Australia. Analysis was by Multicollector-ICP-MS which allows for other isotopes to be measured. The authors reported differences in the Sr an Pb isotopes in champagne compared to other sparkling wines and so proposed that this combination could be used to discriminate the geographical origins and so support authenticity testing in the future (Cellier *et al.*, 2021).

A recent article reported the use of ICP-OES and ICP-MS to determine trace element concentrations in soil, grapes grown on the soil and the wine produced by the grapes. Chemometric analysis was carried out on the data obtained. It was found that the use of bentonite clays in processing of wine (as fining agent) transfers rare earth element into the wine, influencing the trace element profile and so, in this study the geographical origin could not be differentiated based on the elemental profile (Temerdashev *et al.*, 2023).

Similarly, a previous study reported the influence of processing parameters as well as agricultural practices and pollution, on the elemental profile of, in this case, Turkish wines. ICP-AES and ICP-MS analysis of 13 wines was carried out and the limited sample set meant that it was not possible to fully assess the data statistically and so no conclusions as to the impact of geography on the element profiles could be ascertained (Sen & Tokatli, 2014).

Conversely, through the analysis of white wine throughout the production chain it was observed that there was no impact of the use of bentonite on the Sr ratios which remain consistent with those of the rocks and soils, suggesting that no further contribution is given by the addition of bentonite and yeast to the white wine Sr-isotope values. The Sr ratios in both the grapes and final wines preserved the isotope signature derived from the labile fraction of the soil where the vines were farmed (Tescione *et al.*, 2020).

Pepi and Vaccaro studied the elemental concentrations of Prosecco wines and compared the data with that of the soil on which the vines were grown and grapes harvested. Geochemical and statistical analyses were able to discriminate the vineyard soils according to the geo-lithological characteristics of each area and to identify geochemical “Prosecco” fingerprints. The authors claim that these fingerprints could be used against fraudulent use of DOC wine labels (Pepi & Vaccaro, 2018).

In a study of 639 Italian wine samples analysed for their elemental composition, Chianti and Chianti Classico wines from Tuscany in Italy were compared with another 18 geographical regions but it was not possible to completely discriminate between the samples tested due to the close proximity of the regions from which samples were taken (Bronzi *et al.*, 2020).

Martins and co-workers studied the Sr isotope ratios in Portuguese vineyard soils. Granite based soils showed higher $^{87}Sr/^{86}Sr$ ratios than the other soils (sedimentary formations). Although wines were not analysed in this study the authors proposed that an international databank of $^{87}Sr/^{86}Sr$ values should be set up to support geographical origin determination (Martins *et al.*, 2014).

Authentic Bordeaux wines were shown to have a narrow range of $^{87}Sr/^{86}Sr$ ratios and Sr concentrations and so were reported to be suitable indicators of the geographical origin of these prestigious wines. The data was compared to that derived for 17 red wines, purchased in China, which included some which were labelled as from Bordeaux, but with labels that led to questions around the authenticity of the products. The results demonstrated a relatively narrow span of variabilities for the $^{87}Sr/^{86}Sr$ ratio and Sr concentrations in authentic Bordeaux wines. In contrast, the Chinese wines results for $^{87}Sr/^{86}Sr$ ratios varied widely and contained approximately 4 times the concentration of Sr. The authors postulate that the unique Sr binary signature may detect imitated wines and trace genuine products from different regional wineries. Together with the soil, the authors highlighted that climatic conditions, agricultural practices, management of wineries and wine-making techniques may all also impact the Sr ratio and total concentrations in the wine. Again, the need to collate data to support further model development was proposed (Epova *et al.*, 2019).

Other research groups have published studies combining stable isotope and element concentration data to support the determination of geographical origin, highlighting the importance of combining datasets to achieve more suitable models to discriminate between locations. Rapa *et al.* (2023) described the analysis of seven Venetian wines for 63 elements and six isotope ratios with As, Ca, Cs, $\delta^{11}B$ and $^{87}Sr/^{86}Sr$ being the most informative. (Rapa *et al.*, 2023). Wu *et al.* (2021) looked at French red wines in the same way but identified Mg, Mn, Na, Sr, Ti and Rb as the most suited to French wine regional traceability (Wu *et al.*, 2021).

Saar de Almeida *et al.* (2023) reviewed the literature on the use of Sr isotope ratios to characterise wines according to their geographies. The method is based on the principle that the Sr isotope ratio in wine reflects that of the labile fraction of the vineyard soil from which the wine is produced and so can be used to determine geographical origin. The authors conclude that “Although limitations are evident when implemented at large (global) scales, we demonstrate that the $^{87}Sr/^{86}Sr$ isotope tracing technique remains a powerful and reliable tool for determining the geographical origin of wine when combined with detailed knowledge of the geological and soil characteristics of the substrata” (Saar de Almeida *et al.*, 2023).

A vendor’s application note from 2015[23] reported differences in the trace element profiles of wines from different regions, 75 red wines produced in different regions of Italy and from various grape types were analysed by ICP-MS measuring concentrations of 39 elements. PCA was performed on the data. Wines from Puglia were contained higher levels of Cu, Sb and Pb but it was not confirmed whether this was linked to soil type, grape type, or cultivation and production methods. The authors reported that the high levels of Cu maybe due to the use of copper compounds as mildewcides and fungicides or could be a consequence of the use of brass equipment during

---

[23] Geographical Differences of Trace Elements in Wines (resources.perkinelmer.com)

production and bottling. Wines from Tuscany had higher levels of Sr and Li. Trentino wines had higher levels of several rare earth elements. Overall differences were observed but the authors concluded that "further research will be needed to investigate whether these differences relate to soil and rainfall or are correlated to viniculture production differences".

An Agilent Technologies Application note authored by Nelson *et al.* (2015)[24] studied the trace element concentrations of Malbec wines from Argentina and the USA. Analysis was carried out by MP-AES. Sr, Rb, Mg, Ca, Na, and K concentrations in 41 wines from the two regions (26 from the Mendoza region of Argentina and 15 from California in the US) were determined. In-house chemometric software (Mass Profiler Professional) was applied to the data generated followed by Partial Least Squares – Discriminate Analysis (PLS-DA). The statistical analysis differentiated Malbec wines from Argentina and the US, with 14 out of 15 US samples and 25 out of 26 of the Argentinian wines correctly classified.

3.2.16 Cheese

As mentioned in Section 3.1.7, a working example of application of analytical techniques to authenticate PDO cheeses is reported by Camin *et al.*, 2015. In an international collaborative study and in combination with SIRA measurements, the content of Li, Na, Mn, Fe, Cu, Se, Rb, Sr, Mo, Ba, Re, Bi, U in cheese after acid microwave digestion using Inductively Coupled Plasma Mass Spectrometry or Optical Emission Spectrometry (ICP-MS or -OES) was measured in 13 different laboratories. For elemental data, the average RSD_r and RSD_R values ranged between 2 and 11% and between 9 and 28%, respectively, consistent with methods reported by the FDA and in the literature for cheese.

---

[24] Elemental profiling of Malbec Wines for geographical origin using an Agilent 4200 MP-AES (agilent.com)

3.2.17 Limitations and gaps in trace element analysis

Examples of the use of trace element data to determine the geographical origin of foodstuffs has been presented with over 50 papers and review articles being summarised above. As expected, the size of the sample set and confidence in the origin of the samples both influence the quality of the prediction models produced. In many of the examples considered it was a combination of techniques that allowed the geographical origin to be determined and so continuing to derive measurement data for authentic samples of high value foodstuffs and refining the models developed is essential to protect against future fraudulent claims as to the origin of these products. Many of the papers compared one region with another rather than being an indiscriminate tool to determine the origin of the sample when considering the potential worldwide source. While using trace element data, one needs to consider the relationship of geochemicals with the soil, and particularly the transfer factors and the pH of the soil which can influence the stability of a signal, as can the application of fertilisers. Finally, almost all the authors highlighted the need to continue to add trace element data to the databases developed to allow the impact of season by season and year on year variability to be fully addressed as well.

3.2.18 Conclusions for trace element analysis

In conclusion, the publications considered within this literature review demonstrate that trace element data can be used to support the determination of geographical origin. However, due to the limited data available the examples provided generally compare two distinct locations and would all benefit from more extensive data sets. At the present time trace element analysis alone cannot be considered to be an indiscriminate tool to determine the origin of the sample when considering the potential worldwide source.

3.2.18 Outlook for trace element analysis

Continued data collection to support the development of more extensive and robust databases covering more regions, addressing uncertainties of season by season and year on year variability is required to develop this tool further. In addition, the strength of using multiple complementary techniques to verify origin has been highlighted and so coupling trace element data with proteomic, genomic, stable isotope or other data is expected to improve the models developed to date.

3.3 Metabolomics and profiling techniques

Metabolomics is the large-scale study of small molecules, commonly known as metabolites, within biological materials. The papers cited in this section were selected through the filtering process to cover analytical technologies enabling metabolite profiling studies for the determination of geographical origin of food. The use of these technologies and their application to food commodities is discussed along with the application of data analysis techniques such as multivariate statistics, artificial intelligence (AI) and organisational workflows such as metabolomics. This section focusses on technologies that measure the molecular composition of foodstuff and use the variability in composition to assess geographical origin. This usually relates to specific compounds (markers) or groups of compounds (fingerprints). This section does not include isotope, protein or DNA based approaches which are covered elsewhere and largely focusses on small molecules, although some techniques discussed will register signals from more complex molecules such as oligosaccharides and triacyl-glycerols.

A comprehensive review of the analytical methods used for the determination of the geographical origin of food (Luykx & van Ruth, 2008) categorised methods into four groups: mass spectrometry techniques, spectroscopic techniques, separation techniques, and other techniques. The authors conclude that a combination of methods analysing different types of food compounds seems to be the most promising approach to establish geographical origin, recognising the need for statistical approaches to bring together the data.

However, separation techniques are frequently combined with mass spectrometry and sometimes with spectroscopic techniques such as NMR spectroscopy. Gas chromatography coupled with mass spectrometry has been extensively used for volatile profiling, whilst liquid chromatography coupled with high resolution mass spectrometry (LC-HRMS) provides a rich source of information for assessing the non-volatile components of food. Other hyphenated technologies such as gas chromatography coupled with ion mobility spectroscopy have shown recent promise

(Zhu *et al.*, 2023). The holy grail for origin determination is field portable and handheld devices that can be readily used by unskilled operators. Full integration of rapid online testing for food authenticity within the food supply chain has recently been discussed, capturing the most promising approaches for a fully integrated system, whilst highlighting challenges, particularly in developing countries (McVey *et al.*, 2021).

3.3.1 Targeted vs non-targeted

Analysis methods for food origin determination can be described as targeted or non-targeted the merits of which have been discussed (McGrath *et al.*, 2018). Targeted methods are those which seek to measure a specific component of a food, for example, pesticide residues or nutritional components such as iron or potassium. Targeted methods are often preferred as they do not rely entirely on databases of reference information to draw a conclusion, particularly in well studied food stuff. A good example is the use of markers from the nectar of specific flowers to determine the origin of honey (see honey later). Local flora is often specific to a particular country or region and it is often possible to use nectar markers to determine the origin of honey, with the most high profile example being the New Zealand governments use of 4 biomarkers to define authentic New Zealand Mānuka honey.

As origin determination is often complicated by factors that impact on the uniqueness of markers, non-targeted methods are also often employed. These methods measure the composition of a foodstuff through a “fingerprint” or profile. This can be a spectral profile which carries the signature of the product as a whole (for example UV or IR spectroscopy). These technologies rely on having an established and robust database or training set of samples that can be used to determine the success with which the origin of a test sample can be determined. As such, they rely heavily on the use of chemometrics and machine learning (Zhang *et al.*, 2021b) and often are referred to as black-box approaches, as the basis on which origin classification can be achieved is often unclear. More sophisticated non-targeted approaches such as LC-HRMS and NMR spectroscopy also provide data about components of a foodstuff that can be used for origin determination. This is usually a combination of several (often hundreds) of biomarkers that in combination are characteristic of a particular geographical source. The rationale for using these approaches may be based on local climate, geology or practices such as feedstuff and fertilisers.

Having access to meta data such as permissible additives, clarifying agents, yeasts and knowledge of permissible production processes is very valuable information, when performing targeted or non-targeted profiling of food. Taking the example of the EU wine databank, if information is available relating to the different approaches for populating the databank with authentic micro-vinified wine samples versus mass produced retail samples, signals can be detected to determine if using steel tanks or wood chips have been used during vinification, to support origin verification.

Whilst rarely conclusive on their own, non-targeted methods can provide highly indicative data to support a weight of evidence approach for origin determination, particularly when taken with the results of other analytical tests and data from traceability systems.

3.3.2 Multivariate analysis, data fusion and Omics

Data from both targeted and non-targeted approaches can be used in multivariate analysis and data fusion approaches. Multivariate analysis as the name suggests, relates to the statistical analysis of data with more than one variable. Often many thousands of variables are used as input to a multivariate analysis with data from many hundreds of samples resulting in complex data matrices. Tools that are often employed to simplify these data include principal components analysis and machine learning.

Data fusion seeks to take data from the same samples and join them to improve the information content of a dataset. This is a logical approach used routinely in assessing the origin of foods as most analysts will use multiple data sources when drawing conclusions. However, data fusion is usually used to cojoin data from different analytical techniques that are then further subjected to some kind of multivariate analysis. An example of this is the classification of salmon from 5 different regions using a dual-platform mass spectrometry data set. Eighteen robust lipid markers and nine elemental markers were found, which provided robust evidence of the provenance of salmon. This study showed that a data fusion - multivariate analysis strategy greatly improved the ability to correctly identify the geographical origin and production method of salmon (Hong *et al.*, 2023).

The combination of non-targeted analysis with multivariate techniques often attracts an “omics” suffix as is the case with when using metabolomics workflows (Mialon *et al.*, 2023). This has been suggested to be an effective way of marketing products with specific Geographical Indicators (GI) (Cassago *et al.*, 2021). The following sections describe published studies from the aforementioned literature review process. Each section describes how analytical methods have been combined with multivariate statistics and AI to address the reporting of the origin of food.

3.3.3 Edible Oils

Analytical methods that have been used for geographical origin determination of edible oils have been robustly reviewed (Tahir *et al.*, 2022). A systematic review of papers published between 1 January 2013 and 15 December 2020 identified sixty-six full text articles that met the selection criteria for inclusion in the review, which required the use of both analytical techniques and multivariate analysis. The authors concluded that geographical origin was a major source of the variation in oil composition. Targeted analysis of the following components was frequently carried out in oils: phenolic compounds, fatty acid profile, sterols, triacylglycerol (TAGs), volatile compounds and colour. The most popular approach (half of the papers) measured the elemental composition of the oils using, for example, inductively coupled plasma mass spectrometry (ICP-MS). A trend towards the use of NMR, IR spectroscopy and chromatography was noted. The predominant oil types reported to be analysed were olive oil (including virgin and extra virgin) mainly from Europe and North Africa. Several studies also considered palm oil from diverse locations in Asia and Africa (Pérez-Castaño *et al.*, 2015), (Ruiz-Samblás *et al.*, 2013), (Tres *et al.*, 2013), (Jolayemi *et al.*, 2018). Multivariate analysis techniques that were employed included principal components analysis (PCA) and partial least squares (PLS) regression, usually coupled with LDA or DA to predict geographical origin based on the input variables from analytical testing. A wide range of multivariate analysis techniques have been assessed to determine their applicability for the analysis of data from oils (Avramidou *et al.*, 2018) with novel AI approaches proposed (artificial neural networks, fuzzy logic, expert systems, decision trees, support vector machines). Similar approaches have been reported for geographical origin determination using FTIR applied to Greek olive oil (Soh *et al.*, 2023).

Using FT-NIR and headspace gas chromatography ion mobility spectrometry (HS-GC-IMS), methods were developed to reveal the authenticity of Slovenian olive oil. FT-NIR offers in-field testing capability, with high sample throughput, low operational costs, requires little or no sample preparation, and no need for chemicals or specialized laboratory facilities. The research carried out by the Joint FAO/IAEA laboratory was designed to verify the origin of Slovenian extra virgin olive oil from the Istria region, which has a protected designation of origin and is a high value product (Kelly & Midgley, 2024). More specifically, the project aimed to support 'made in Slovenia' branding regulation so a model was developed as Slovenian/non-Slovenian using OPLS-DA and SIMCA. A total of 64 authentic extra virgin olive oils were used in the study, collected over two years from Slovenia, Italy, Croatia, Greece, Tunisia and Spain as part of an IAEA collaborative research initiative with these countries. Scientists using the method were able to tell the difference between extra virgin olive oil from Slovenia and other countries with between 86 per cent and 93 per cent accuracy, after screening and processing the data obtained. Scientists at the FAO/IAEA laboratories also used Fourier transform infrared spectroscopy with attenuated total reflectance (FTIR ATR) to accurately discriminate olive oil between different regions of Lebanon, due to olive oil quality and value varying across different regions of Lebanon.

It was noted by Tahir and co-workers (Tahir *et al.*, 2022) that using multiple data sets from different analytical sources generally improved classification rates and therefore the robustness of origin determination, however all studies presented required bespoke reference data with sample numbers varying considerably.

The success of the discrimination between oils from different geographical locations was heavily dependent on the analytical technology employed and the sample set. Many technologies have been applied to determine the origin of oil including a wide range of spectroscopies and separation technologies. Some success has been achieved using technologies such as electronic noses and tongues although the nature of the technologies means that the data can be difficult to interpret due to low information content (Melucci *et al.*, 2016), (Haddi *et al.*, 2013), (Souayah *et al.*, 2017). The review concludes that the most successful methods for determining the origin of oils are multi-analyte methods when combined with multivariate analysis.

The use of NMR for the analysis of olive oil has been reviewed and points to this approach being amongst the most promising for geographical origin determination of extra virgin olive oil (Maestrello *et al.*, 2022). The authors highlight that the major advantage of the approach is the ability to detect different types of fraud and to perform quantitative measurement of oil components from the same analytical data. Liquid and gas chromatography coupled with TOF mass spectrometry has been evaluated for the discrimination of Mediterranean extra virgin olive oil (Olmo-García *et al.*, 2019). The study combined different data sets to produce statistical models which highlighted key variables that were associated with geographical origin and production year. The study focused on characterising compounds giving rise to those variables, showing the power of high-resolution mass spectrometry to discover new (bio)markers that can be transferred to lower technology platforms.

Official methods for the analysis of olive oil have been critiqued (Conte *et al.*, 2020) and suggestions for improvements made. There is a need to embrace more advanced technologies such as vibrational spectroscopy and NMR spectroscopy, MS, biosensors, and DNA-based approaches. These represent promising alternatives for the authentication and traceability of olive oil, because of their sensitivity, high-throughput, reproducibility and robustness in comparison with conventional methods currently used. However, the cost of the instrumentation may be prohibitive for routine analysis and there is also doubt over the robustness of sampling plans in published studies. Conte *et al.* (2020) conclude that tightening of (European) legislation and international harmonisation in relation to analytical approaches would help to prevent fraud in this sector.

3.3.4 Wine

Analytical methods applied to wine have recently been reviewed (Sun *et al.*, 2022). The authors broadly identified 4 types of approach to wine analysis for the determination of geographical origin as well as vintage and grape variety. Mass spectrometry, spectroscopic techniques, chromatography and other techniques were considered along with a range of multivariate statistical analysis methods. The importance was stressed that an initial critical step is undertaken to ensure that there is good information in relation to geographical origin and any compounding factors such as vintage, for reference samples. The authors conclude that the determination of wine origin using analytical methods is more mature than for the determination of age or grape variety.

Analytical methods are discussed along with their application. Raman spectroscopy and fluorescence spectroscopy are reported to be sparsely used in wine research. Mass spectrometry was believed to be more precise and perform better than spectroscopy and chromatography. However, stable isotope technology (SIRA, ICP-MS and NMR) was expected to continue to be the standard approach for wine traceability due to its high precision and relative ease of application. IR technology was thought to be easy to operate with low cost and size of the instrument aiding portability and uptake in the field. A study found that using FTIR spectra and quality parameters determined by chromatography and mass spectrometry could be combined using a novel multivariate method for the determination of wine origin (Dong *et al.*, 2023).

LC-MS, NMR and GC-MS are the three non-targeted analysis techniques commonly used for authenticity assessment of wine. Sun (Sun *et al.*, 2022) believed that the combination of multifactor analysis, instrument detection and chemometrics analysis is the development trend of the wine industry, and the technology for wine authenticity and traceability will become more sensitive, reliable and convenient.

The role of molecular spectroscopy (Chandra *et al.*, 2017) and NMR spectroscopy (Le Mao *et al.*, 2023) have also been discussed at length. The use of vibrational spectroscopy such as near and mid infra-red (NIR/MIR) was discussed alongside the use of Raman spectroscopy and sensor technologies. It was concluded that the use of even fairly simple techniques such as these were in their infancy in the wine sector for the determination of provenance and perhaps held back by a lack of technical knowledge about modern techniques within the sector. In relation to NMR spectroscopy, variations in sample preparation protocols were highlighted as a potential issue when comparing data from different studies. The impact of commercialising tools subsequently reducing transparency is also a barrier to future development.

Le Mao and collaborators (Le Mao *et al.*, 2023) suggested that the affordability and portability of NMR instruments should be improved to aid uptake, with benchtop NMR being one possible solution. This is particularly attractive for wine as some spectroscopies such as near infra-red have been shown to provide valuable information from unopened bottles of wine (Harris *et al.*, 2023), which is particularly pertinent to high value wines. Non-invasive technologies coupled with machine learning or artificial intelligence algorithms (Carneiro *et al.*, 2023) for the authentication of vintage wine would provide a useful advance in fraud prevention.

3.3.5 Coffee

Several techniques have been applied to coffee to determine geographical origin and often these are combined studies looking at botanical source (Arabica or Robusta). Spectroscopic techniques have been most widely applied to coffee and their use has been recently reviewed (Munyendo *et al.*, 2022). The authors discuss the application of vibrational spectroscopy to coffee for a wide range of applications including for the determination of geographical origin. The review suggests that the cited studies give highly variable results when classifying coffee based on its geographical origin, with some studies suggesting that NIR spectroscopy is a useful tool and others directly contradicting these findings. The validity of the studies will be highly dependent on the number and nature of the samples used as a reference collection for training the multivariate analysis models.

NMR spectroscopy has been used to differentiate 603 roasted arabica samples from Brazil, Ethiopia and Colombia (Gottstein *et al.*, 2024). Unlike wine and oil, coffee has significant lipophilic and hydrophilic components so two extractions were used to capture the profile of the coffee. The study showed that it was only possible to differentiate African from South American coffee using PCA and LDA to classify the NMR data.

An LC-HRMS study using coffee from Brazil and Mexico (Artêncio *et al.*, 2023) illustrated how the method could be combined with multivariate statistics to identify key components of the coffee that varied as a function of botanical and geographical origin. Whilst initial results looked promising, the study used only 21 samples with 19 of those from Brazil and 2 from Mexico.

A study of Colombian coffee (Arana *et al.*, 2016) compared a successfully implemented NMR method for verifying the protected origin of Colombian coffee with cheaper alternatives. The study used GCMS and GC-C-IRMS and compared the success of the techniques to NMR. The study found that identification of Colombian coffee when compared to coffee from neighbouring countries Brazil and Peru, was less successful using the GC based approaches than when using NMR.

In a study investigating Arabica and Robusta coffee from Africa, Asia, Central and South America (Mannino *et al.*, 2023), a combination of chemical (UV, HPLC-DAD–MS/MS, GC–MS, and GC-FID) and molecular fingerprinting (PCR-RFLP) was used to discriminate commercial green coffee accessions from different geographical origins. The authors concluded that using a combination of high-throughput metabolomics with phenolic compounds, fatty acids, xanthine derivatives, and melatonin, along with antioxidant power and DNA fingerprinting, they were able to discriminate the two coffee species and partition the individual accessions of the species according to their geographical origin. The study was somewhat limited in size as only 15 coffee samples were used for the analysis.

3.3.6 Rice

A review of the potential for rice fraud was published recently (Sliwinska-Bartel *et al.*, 2021). The lack of a coordinated approach is evident in the diversity of the methodologies applied in a range of studies summarised by the authors. Isolated studies have been used to exemplify a potentially simple set of issues, suggesting that there should be a standardised and international approach to address questions around rice authenticity, potentially at an inter-governmental or international standard body level. The review summarises well issues which may occur, such as blending of rice from different origins. The review also discusses the diversity and proliferation of academic studies in this area but provides no solid evidence that there is ongoing collaboration with stakeholders and reads largely as conjecture of the likelihood of mislabelling of rice origin.

Distinction of rice origin is achieved mainly by the application of ICP-MS and SIRA to determine trace element and stable isotope profiles respectively. This is described in other sections of this report. Other techniques include the use of GC-MS (with or without headspace (HS) capture) to differentiate rice from Asia (China, Korea and Malaysia), often using PCA-LDA to discriminate between volatile profiles (Ch *et al.*, 2021). LC-HRMS, Raman and NMR spectroscopy (Huo *et al.*, 2017) have all be used in separate studies but there is little by way of coordinated conclusions that can be discerned from the published data, short of the robust summary of activities provided in the Sliwinska-Bartel review (Sliwinska-Bartel *et al.*, 2021). The conclusions of this paper are necessarily somewhat vague, indicating that there is a need to coordinate international research to help to focus efforts. Similar conclusions can be drawn from the review of (Wadood *et al.*, 2022) which specifically highlights the effectiveness of IRMS and ICP-MS for the determination of rice origin.

Data analysis methods that have been applied for the authentication of rice have also been reviewed (Maione & Barbosa, 2019). The paper highlights that PCA-LDA is the most commonly applied multivariate analysis tool for the determination of rice origin, whilst highlighting the use of support vector machines (SVM) and artificial neural networks (ANN), bringing AI techniques into use. The paper also introduces the use of image analysis for the verification of rice, but this is largely for the determination of rice variety, which is not always linked to origin and can be robustly addressed using DNA based techniques.

3.3.7 Cocoa

Analytical approaches for the determination of the origin and authenticity of cocoa beans have recently been extensively and robustly reviewed (Fanning *et al.*, 2023). The authors highlight that the many steps in the supply chain and the price difference between ‘fine’ and ‘bulk’ cocoa present a risk in terms of food fraud. Methods to identify geographical origin using quality attributes were assessed, including spectrometry, spectroscopy and sensory studies. The paper suggested that integrating instrumental and sensory attributes will help to identify relevant and comprehensive geographical quality indications. A common theme for cocoa origin determination is a need to transition to more rapid, affordable and non-destructive analytical approaches. The use of advanced data analysis methods (for example AI) will also modernise traditional traceability methods. The need for harmonization of methods and the curation of authentic samples was highlighted as this is needed to produce robust geographical indications to establish cocoa terroir effectively. The authors concluded that it was abundantly clear that geographical origin plays a critical role in the quality of cocoa, with most studies reporting a significant difference in composition in cocoa from different geographical origins, a conclusion that is supported by data held at Fera (unpublished data).

Sensory methods (organoleptic) also effectively differentiated cocoa origin. However, more rapid NIR spectroscopy was able to reproduce the discrimination of cocoa samples obtained with a trained sensory panel. It was recommended that future studies explore more rapid and non-destructive techniques. Exploration of advanced machine learning algorithms (for example AI) to improve the origin classification and prediction would improve such analysis results with the caveat that transparency is a key component. To achieve this, methods and procedures must be harmonized and include reporting of verification of authentic samples used to draw conclusions.

Cocoa beans and cocoa bean products (CACBP) have been subjected to various analytical technologies to determine their origin. Practical application has focussed on spectroscopic techniques such as NIR spectroscopy (Anyidoho *et al.*, 2020), (Teye *et al.*, 2020) largely due to the requirement for field portability into often remote regions of for example West Africa and a need to verify origin close to source. The authors conclude that "more work needs to be done to move this technology from the laboratory applications to real usage in developing countries for optimum global benefits" and "Developing portable and affordable instruments is urgently needed particularly in West Africa. It would make onsite measurement application in developing countries possible and aid global traceability and production of high-quality cocoa beans". GC-MS successfully classified origin at the country level. Applying NMR spectroscopy as a non-targeted metabolomic approach was also successful. The need to supply metadata has been highlighted. Not all studies reported vital information on samples, such as the cocoa bean variety, which the authors say is necessary for the successful development of a database for cocoa traceability. Peripheral technologies such as Flow Infusion - Electro Spray Ionization - Mass Spectrometry have also been trialled with promising initial results (Acierno *et al.*, 2018) focussing on dark chocolate from a range of origins with the authors highlighting the promising steps made.

Others have also focussed on correlating the composition of cocoa bean products, such as chocolate, to the composition of the cocoa beans used in their manufacture. The polyphenol content of chocolate has been directly related to that in the source cocoa beans using HPLC-DAD-MS and PCA (Cambrai *et al.*, 2017) and shown to be a useful tool to differentiate chocolate containing cocoa beans from Madagascar, Caribbean, South American and African countries for the first time, where previous data has looked only to distinguish cocoa beans present in chocolate that is sourced from different continents, indicating that the development of technology is improving the spatial resolution of the analysis.

3.3.8 Honey

Methods for the determination of honey origin have been comprehensively reviewed by (Soares *et al.*, 2017, funded by the EU 7th Framework Programme Food Integrity: Ensuring the integrity of the European Food Chain) and more recently by (Danieli & Lazzari, 2022).

Both papers describe a range of methods that have been applied to determine the authenticity of honey and discuss botanical and geographical origin determination. Honey is somewhat unique in that its composition is determined by the variability in local flora and this can be used as a basis on which to determine origin. At the most rudimental level, pollen analysis (melissopalynological) using light microscopy can be used to determine the amount of pollen contributing to the overall pollen content by a range of different plant species. Whilst the approach is comparatively low-tech, the skill and experience required to differentiate accurately between pollen types should not be underestimated. This has resulted in a relatively small number of specialists who are able to determine geographical origin based on pollen. The approach also relies on pollen from different plants being distinguishable, which is usually the case. However, some pollen types are indistinguishable (for example Mānuka and Kanuka) and some pollen types will exist in many locations, so the pollen fingerprint of a honey may not uniquely identify its origin. In other cases when pollen native to a particular location can be identified, then pollen analysis is a powerful technique. To reduce the reliance on specialist pollen analysts, there have been recent attempts to use molecular biology techniques to analyse DNA from honey pollen. This can be hindered by a low copy number, but progress is being made with, for example, digital PCR which may allow the approach to become routine (You *et al.*, 2021).

Analytical approaches for the origin determination of honey are covered in different sections of this document, with IRMS and ICP-MS being the most routinely applied, to measure isotopic ratios and trace element composition in honey, which are both well known to be associated with geographical origin. However, the most recent rapid progress for the determination of honey origin is being made by applying a range of techniques to determine the composition of the honey. Targeted and non-targeted methods have been applied to determine biomarkers that are associated with origin. As with pollen analysis, these approaches can usually determine botanical origin, which then by proxy can be used to draw conclusions about geographical origin. Work carried out under by the EU 7th Framework Programme Food Integrity (Soares *et al.*, 2017) identified several classes of compounds that have been used to distinguish both the botanical and geographical origin of honey. Phenolic compounds have been extensively studied, along with volatile components, sugars, organic and amino acids. The technologies used to measure these groups are well established (for example GC-MS or GC-FID for volatiles) and the studies undertaken are described in the previously cited reviews, showing that each class of compound can be used to differentiate between honey, mainly by floral type. These analyses have resulted in a range of markers being identified that are specific to plants commonly associated with honey production. For example, the phenolic compound, hesperetin, has been identified as a marker of citrus honey.

An evaluation of the methods most used in the analysis of the geographical and botanical origin of honey, particularly in the last decade, has been published (Danieli & Lazzari, 2022). Current state-of-the-art technologies include metabolomic/genomic approaches and blockchain. The authors conclude that when methods are used in combination this usually leads to greater accuracy of origin determination, linking data through multivariate statistical or chemometric methods. Similar to other commodities discussed here, a range of techniques have been trailed and found to be able to classify honey by origin with NMR perhaps being the most accepted (Biswas *et al.*, 2023) but with rapid developments being made in the use of MS and other technologies such as Raman spectroscopy.

The development of data fusion and multivariate analysis methods specifically for application to honey is also noteworthy. The potential benefit of data fusion based on different complementary analytical techniques was investigated (Schwolow *et al.*, 2019). Sixty-four honey samples from three different origins were analysed by attenuated total reflection IR spectroscopy (ATR/FTIR) and HS-GC-IMS). The datasets obtained were combined in a low-level data fusion approach with a subsequent multivariate classification by PCA-LDA or PLS-DA. Validation results of the classification models were compared to the results that could be obtained by using the individual data blocks separately. A decreased cross-validation error rate and more robust model was obtained due to the low-level data fusion. The results show that data fusion is an effective strategy for improving the classification performance, particularly for challenging classification tasks such as determining the origin of honey.

The state-of-the-art for honey analysis also incorporates AI technologies for data analysis. For example, machine learning has been used to identify volatile components in citrus honey that are linked to origin (Karabagias & Nayik, 2023).

3.3.9 Meat

Recent food origin issues in the UK have largely been around meat, relating to the incorporation of horsemeat originating from elsewhere in Europe being found in beef labelled products, and more recently, the use of meat in processed products that were inaccurately declared to be from the UK. The scientific basis for these issues has been the subject of conjecture and to a lesser extent to scientific review. The confounding issues relating to meat origin mainly concern the movement of animals and meat products during rearing and supply, which, from an analytical perspective, makes it very difficult to identify the source of meat products. Analytical methods for the detection of meat origin also take little account of animal feed, which influences the biochemical profiles of all of the testing methodologies. For example, isotopic ratios and trace elements, when tested in a meat product will vary as a function of the feed used to produce the animals. Feed can be obtained from many different geographical locations and is a major factor in meat composition (Monahan *et al.*, 2018). Routine testing for origin and the establishment of databases for meat authenticity will fail if they do not account for feeding regimes and the movement of animals. Some success has been found in the use of technologies such as genomics which rely of the genetic code of the animals. However, movement still compromises this as germplasm is often acquired from gene banks (the UK is a leading supplier of beef germplasm, for example) and the animals may be reared elsewhere.

Profiling of the gut microbiome has been proposed, particularly for ruminant animals but this is highly speculative. Meat authenticity is perhaps the most challenging area covered here due to movement and feeding regimes. The lack of any confident review of meat origin determination indicates the need to consider all data (a weight of evidence approach) when assessing meat origin. Challenges pertaining to the interpretation of data, as they relate to assignment of dietary background or geographical origin, are discussed in the review of Monahan (Monahan *et al.*, 2018). They concluded that “among many factors such as the global nature of trade in meat, the complexities of the food chain associated with meat production, consumer demands for more information about the food they consume and the potential for fraud, there seems little doubt that meat authenticity will continue to be a subject of discussion and research into the future”. Furthermore, they say that “From the particular perspective of dietary background and geographical origin, it is clear that a single analyte is unlikely to be adequate and that measurement of multiple markers is required. Determination of the dietary background and geographical origin of meat brings specific challenges not least of which is the possibility for animals to move between different jurisdictions and to consume foods from different sources and from different geographical origins over their lifetime”. In authoring this section, it was felt that a deep dive into analytical technology was inappropriate as the major issues are around establishing traceability systems that can be subsequently tested in the market and may be an area where technology-based solutions such as RFID (radio frequency identification) will continue to allow for the monitoring of animal movement to support international trade. A recent review of beef testing methodologies is provided by (Bai *et al.*, 2023). Seven approaches were discussed in relation to their ability to determine the origin of beef. These are: stable isotope technique, DNA technology, spectroscopic technology, volatilomics technology, metabolomic analysis, fatty acid analysis and mineral element analysis. The main conclusion of the review was that “The supply chain for beef is highly complex”.

3.3.10 Limitations and gaps in using metabolomics and profiling techniques

- The validity of studies undertaken to date is highly dependent on the number and nature of the samples used as a reference collection. In many studies this is inadequate to reach firm conclusions and there is also doubt over the robustness of sampling plans in published studies.
- There is little evidence that there is ongoing collaboration to develop the skills/ experience and reference data required to differentiate accurately between origins.
- The lack of a coordinated approach is evident in the diversity of the methodologies applied in a range of studies.
- The affordability and portability of instruments should be improved to aid uptake as the cost of the instrumentation may be prohibitive for routine analysis.
- Commercialising tools subsequently reduces transparency and is a barrier to future development.
- The confounding issues relating to meat origin mainly concern the movement of animals and meat products during rearing and supply, which, from an analytical perspective, makes it very difficult to identify the source of meat products. Analytical methods for the detection of meat origin also take little account of animal feed. Routine testing for origin and the establishment of databases for meat authenticity will fail if they do not account for feeding regimes and the movement of animals.

3.3.11 Conclusions for using metabolomics and profiling techniques

- Using a combination of methods analysing different types of food compounds seems to be the most promising approach to establish geographical origin, recognising the need for statistical approaches to bring together the data. When methods are used in combination this usually leads to greater accuracy of origin determination, linking data through multivariate statistical or chemometric methods.

- Using multiple data sets from different analytical sources generally improved classification rates and therefore the robustness of origin determination. Data fusion is an effective strategy for improving the classification performance.
- An initial critical step must be undertaken to ensure that there is good information in relation to geographical origin and any compounding factors such as vintage, to ensure that robust metadata are collected.
- The determination of provenance is perhaps held back by a lack of technical knowledge about modern techniques within specific sectors.
- Integrating instrumental and sensory attributes will help to identify relevant and comprehensive geographical quality indications.
- Geographical origin plays a critical role in determining the quality of products such as cocoa.
- Non-targeted methods can provide highly indicative data to support a weight of evidence approach for origin determination, particularly when taken with the results of other analytical tests and data from traceability systems.

3.3.12 Outlook in using metabolomics and profiling techniques

Metabolomics and profiling techniques will continue to play a role supporting geographical origin verification, especially in combination with other methods. Linking data through multivariate statistical or chemometric methods is an effective strategy for improving the classification performance.

There is a need to transition to more rapid, affordable and non-destructive analytical approaches. Non-invasive profiling technologies coupled with machine learning or artificial intelligence algorithms would provide a useful advance in fraud prevention. Using AI techniques requires the caveat that transparency is a key component. For some commodities, particularly in developing countries, onsite measurement is highly desirable and would aid global traceability and production of high-quality products.

There is also a need to coordinate international research to help to focus efforts appropriately. There should be a standardised and international approach to address questions around food origin, potentially at an inter-governmental or international standard body level. The need for harmonisation of methods and the curation of authentic samples is highlighted as essential to produce robust geographical indications to establish terroir effectively. Tightening of (European) legislation and international harmonisation in relation to analytical approaches would help to prevent fraud.

Best practice requires that studies consider all data (a weight of evidence approach) when assessing origin.

Technology-based solutions such as RFID (radio frequency identification) should continue to be developed in conjunction with analytical methods, for the monitoring of the movement of products to support international trade.

3.4 Genomics

Genetic analysis for the assignment of geographical origin has not been widely used. The development of genetic tests, using the analysis of DNA, suitable for the assignment of geographic origin is inherently difficult. Such genetic tests would rely on the identification of private or near private DNA markers (Fontanesi, 2009). Animals and plants of different populations, but of the same species, can freely exchange genetic material by cross breeding, as there are no preventative biological barriers. Only selective breeding or genetic isolation can establish a unique sub-population of a species and only if then anchored to a discreet location, can geographical origin be assigned. To date the identification of these private genetic markers has been challenging, however, the development of next generation sequencing (NGS) which is able to scrutinise the genome in more detail, has enabled studies which may lead to tests able to assign geographical origin in the future.

3.4.1 Fish and seafood

An example of genetic isolation leading to 2 distinct identifiable sub-populations in Baltic Sea cod (*Gadus morhua*), was published by Hemmer-Hansen *et al.* in 2019. Western and eastern Baltic cod have overlapping geographical locations, with a mixing zone in the Arkona Basin waters, south of Sweden. These fish populations would ordinarily be expected to interbreed. However, the two populations have temporally distinct spawning times. While western Baltic cod spawning is restricted to a few weeks in early spring, eastern Baltic cod spawn over a prolonged period of time peaking in the summer months (Hüssy, 2011). Environmental conditions at spawning time also add to genetic isolation in that they rarely support the survival of eggs of the western Baltic cod after the spring spawning period.

This study developed a panel of SNP which had high levels of population differentiation between eastern and western Baltic Sea cod. The panel was then used to analyse 2000 fish tissue samples collected between 2011 and 2015 in the mixing zone. The study was able to assign fish to either the eastern or western population with a very high degree of confidence (Hemmer-Hansen *et al.*, 2019).

This study focused solely on the Baltic Sea cod population, however, it is known that there is mixing between the North Sea and western Baltic Sea cod, with individuals in that mixing zone of mixed genetic heritage (Berg *et al.*, 2015). It is not known how those fish would be classified using this panel of SNPs. This study was designed to facilitate the Baltic Sea cod fishery and was not designed to address the potential to assign GO of cod across the full geographical range. Nevertheless, the panel of SNPs developed during this study was highly informative and further work, using a greater number of samples from the full natural range of the Atlantic cod, may potentially reveal separate populations of cod based on geographical origin in the future. Spatial resolution in the marine environment is usually limited to large geographic distances because marine organisms generally show shallow population structure due to high dispersal capacity in the absence of strong barriers to gene flow. Next generation sequencing technologies, however, now allow the identification of hundreds/thousands of SNP using the genotyping-by-sequencing (GBS) approach and microbial community analysis using NGS-generated microbiome profiles (Milan *et al.*, 2019).

A study compared the analysis of foot tissue SNPs, with NGS-generated microbiome profiles of the hepatopancreas of the shellfish *Ruditapes philippinarum* (Manila clam), a shellfish of high commercial interest with worldwide distribution (Milan *et al.*, 2019). It was found that the clam place of origin could be located with high spatial resolution using the NGS-generated microbiome profiles of the hepatopancreas. Samples were collected in June and December/January in four areas of the Venice lagoon, over 2 years, although not all areas were sampled at each time point. Nevertheless, the spatial resolution for the origin of the clams remained over time. This initial study appears to show promise with regards to assigning geographical origin to Manila clams, however, the geographical area sampled was restricted to the Venice lagoon: an area with a distinctive marine environment. A wider study, using more samples from geographically diverse areas, is needed to confirm and extend this result, since work at Fera has shown that this type of analysis is not applicable to another type of shellfish: oysters (see Section 3.4.7).

A similar study was performed on the soft-shell clam species *Mya arenaria* collected in Canadian waters (Liu *et al.*, 2020). The NGS-generated microbiome profiles of the clams could be reliably differentiated by harvest site, which remained true over the 3-year sampling period. The microbial diversity of these freshly harvested clams was much higher when compared with batches of clams from retail samples. The processing of the retail samples was unknown, but likely involved depuration, a process in which shellfish are held in a tank of clean seawater prior to retail, which results in the expulsion of intestinal contents and improves the safety of the final product. In this study, the operational taxonomic unit (OTU) which is a measure of microbial diversity, fell from 2994 OTUs in the fresh samples to 149 OTUs in the retail samples. Additionally, the microbial community of the retail samples was heavily dominated by Proteobacteria, a typical spoilage organism for fresh seafood. It has previously been shown that the microbiome of shellfish become more similar as they spoil (Madigan *et al.*, 2014). Depuration and spoilage may well reduce the usefulness of this type of analysis for the attribution of geographical origin for shellfish. Interestingly, the study above, on Manila clams, was performed on fresh non-depurated samples (Milan *et al.*, 2019). It will be important to extend these analyses along the whole production chain, from sea to market, to better understand where the change in the microbiome community occurs. The depuration/spoilage-induced loss of microbial diversity makes determining the geographical origin of shellfish more difficult in retail samples than in freshly harvested samples.

3.4.2 Meat

A small study analysed the NGS-generated microbiome profiles of the Italian fermented sausage, Salame Piemonte (PGI) (Franciosa *et al.*, 2021). Three batches of the salame were manufactured in February, March and May and were then allowed to ferment and ripen under the usual temperature and humidity conditions used for this product. The microbiome profiles of the three batches clearly showed inter-batch variability with *Pediococcus pentosaceus, Latilactobacillus curvatus* and *L. sakei* associated with samples from February, March and May, respectively. Further, no consistent microbial profile was identified for the salame across the three batches. Therefore, inter-batch variability negates the use of this type of analysis for this product.

3.4.3 Honey

An example of genetic isolation facilitating the designation of geographical origin is that of Mānuka honey. A study by Chagné *et al.* published in 2023, used a panel of SNPs to investigate the population structure of *Leptospermum scoparium* J.R.Forst & G.Forst, called Mānuka by Māori, and the basis of a flourishing honey industry in Aotearoa New Zealand and Australia. A previous study had been able to genetically differentiate Australian and New Zealand *L. scoparium* but had only used a single sample from the Australian island of Tasmania (Koot *et al.*, 2022). The current study analysed 86 samples from Tasmania together with 418 samples from New Zealand using a high-density SNP array containing 9002 SNP for mānuka. SNP arrays are a widely used technology for genotyping DNA of many organisms, including animals and plants for selective breeding or pedigree analysis, as well as elucidating phylogeographic patterns and taxonomic structure (Montanari *et al.*, 2022). This study reported a strong genetic differentiation between the New Zealand samples and those from Tasmania. They also controversially stated that the differences were so large that the Australian *L. scoparium* trees be subject to taxonomic revision and that honeys marketed from *L. scoparium* growing in Australia, should not be called by the Māori term Mānuka (Chagné *et al.*, 2023).

It should be noted however, that this report focused on genotyping leaves, whereas the commodity requiring geographical attribution is the honey from bees foraging on the *L. scoparium* blossom. Further work, therefore, to implement the use of this high-density SNP array for the verification of geographical origin of mānuka honey, would need to focus on samples of mānuka honey from New Zealand and Australia. In particular, the analysis of pollen from *L. scoparium* blossom found in mānuka honey would need to be assessed for compatibility with the use of high-density SNP arrays.

The traditional and most commonly used method for defining the regional origin of honey is melissopalynology, which aims to identify the plants from which bees have collected the nectar based on the morphological characteristics of the pollen found in the honey (Wirta *et al.*, 2021). However, although low cost, melissopalynology requires expert knowledge and reference collections. In recent years NGS has been proposed as a high-throughput alternative, analysing the pollen extracted from honey (Özkök *et al.*, 2023), or the pollen together with the DNA of plants, bacteria and fungi present in the honey (Wirta *et al.*, 2021).

A recent study (Özkök *et al.*, 2023) aimed to compare the melissopalynological method to NGS to determine honey's botanical and geographical origin. A total of 74 honey samples were collected from 16 different areas of Turkey. The samples were both subjected to melissopalynological analysis and NGS of two chloroplast genes, for which a database containing thousands of accessions is publicly available. The results showed that NGS analysis could detect more diversity in the plant species than could melissopalynological analysis and although the findings using both techniques were not identical, both were able to determine the honey's dominant components. Cluster analysis of the NGS data revealed clustering according to broad geographical location in Turkey. Notable exceptions were where samples were taken from geographically distant regions, but which had similar flora, resulting in them clustering together. This is a major drawback of studies such as this, in that the floral landscape is usually a continuum with poorly defined boundaries and regions with similar climate, soil and altitude could have similar floristic profiles. Additionally, neighbouring countries will have similar floral landscapes further confounding geographical attribution using pollen analysis.

A similar approach was taken in a study on Iranian honey (Khansaritoreh *et al.*, 2020). The majority of Iranian honey is produced in four provinces located in the Irano-Turanian floristic region. The climatic conditions in this region result in a short flowering season and poor nectar production. The beekeepers of colder cities migrate to warmer areas in early autumn and stay at that destination until mid-spring, making the assessment of pollen for geographical assignment complicated. This study concentrated on identifying the key species which represented the flora of a specific location using an extensive literature review of reference books on the flora of Iran. NGS analysis of two separate genes was used to determine the floral profile of pollen from 70 samples of honey. The analysis was able to accurately plot the migration of the hives from their summer to winter feeding grounds. It was particularly successful because of the substantial difference between the flora of the north compared to the west and south-west of Iran. It was a limited study and did not obtain any samples from neighbouring countries, each of which would also, as before, have a similar floral landscape as Iran.

The approach taken by Wirta and collaborators (Wirta *et al.*, 2021) was to use NGS to assess the presence of plant tissue, bacteria and fungi in honey from three neighbouring countries: Sweden, Finland and Estonia. This small study, analysing 48 commercial honey samples, was able to distinguish the country of origin using only the plant and fungal profiles of the honey, whereas the bacterial profiles did not separate the honey samples by country. Interestingly, filtering the honey, a common practice which makes honey remain liquid for longer, had only a minor effect on the taxonomic recovery of DNA. This approach, which does not rely on the pollen profile, is promising, however, the number of samples and the range of countries studied would need to be extended to confirm efficacy for the attribution of geographical origin.

3.4.4 Cereal

The genetic diversity and population structure of wheat has been the subject of many studies using DNA-based analyses. A small number of studies have attempted to link clusters of samples with their geographical origin. Combined analysis of random amplified polymorphic DNA and inter-simple sequence repeat markers was used to generate the population structure of wheat in India and Turkey. The study found that the Turkish hexaploid varieties were divided into two clusters, one group showed a close association with Indian hexaploid varieties and the other with Indian tetraploid varieties (Khan *et al.*, 2015). This exemplifies one of the largest draw backs for the attribution of geographical origin using DNA analysis of any kind, in crops which have undergone extensive breeding. This is particularly so for wheat because most modern cultivars can be traced back to a breeding program based in Foggia, southern Italy, under the direction of Nazareno Strampelli (Kabbaj *et al.*, 2017). Additionally, understanding wheat population structure as it pertains to geographical origin is hampered by the increase in germplasm exchange between breeding centers, causing changes to the historical structure of genetic diversity (Brbaklic *et al.*, 2015). A large study, using a high-density SNP array to analyse 370 samples from 32 countries, found that the greatest influence on clustering was from the breeding programmes in different centers (Kabbaj *et al.*, 2017) negating the influence of geographical origin. The study by Brbaklić used microsatellite analysis on 284 wheat varieties to cluster the samples in a slightly different way: into 6 sub-populations. Although none of the sub-populations were comprised of samples from a discreet geographical origin, most of the cultivars from the same country clustered together. Cultivars clustering with a different country, were most probably bred from a common or related ancestor. Nevertheless, clustering into six sub-populations does not offer much granularity for the attribution of geographical origin (Brbaklic *et al.*, 2015). Similarly, a study using a high-density SNP array was able to separate 726 samples broadly into seven groups: 5 are geographical (Asia, Australia, Canada, Europe or the USA), one is historical landraces, and one is durum wheat varieties. The authors found that there was a high number of shared alleles across the 5 geographical locations, reflecting the use of common ancestors for breeding and relatively little development of regional populations by out-crossing with locally adapted varieties (Wang *et al.*, 2014). Although this study was not designed for geographical origin assignment, the use of the high-density SNP array was able to start to separate some populations into geographical locations. It remains to be seen if this approach would be able to detect differences in commercial samples taken from different countries.

A SNP based study of 406 samples of spring barley, appeared to show some geographical origin association for sample origin (Genievskaya *et al.*, 2023). Samples were clustered into five sub-populations. Only two of these sub-populations contained samples from a single country, the USA (Sub-population 3) or Kazakhstan (Sub-population 2). The USA was also represented in two other sub-populations (4 & 5), whereas Kazakhstan was represented in three other sub-populations (1, 4 & 5). Europe and Africa were represented in the same two sub-populations (1 & 4), which also contained either Kazakhstan alone or Kazakhstan and the USA. Therefore, if a sample was assigned to Sub-population 4 it could originate from any of the four regions, if assigned to Sub-population 5 it could originate from the USA or Kazakhstan. This method, therefore, is not an appropriate approach for attributing geographical origin to spring barley samples and reflects the worldwide trade of spring barley.

3.4.5 Vanilla

Consideration of vanilla exemplifies some of the other issues around using genetic analysis for assignment of geographical origin. Produced from the cured seed pods of the tropical orchid genus *Vanilla*, it is the second most valuable spice in the world, after saffron. Originating in Mexico, it is now grown extensively in Madagascar, Indonesia and Mexico. *Vanilla* is only able to reproduce naturally within its native range in Mexico where it co-occurs with specialised pollinators and dispersers and so is usually propagated via stem cuttings from a plant which has not been allowed to flower (Lozano Rodríguez *et al.*, 2022). Due to the absence of natural genetic recombination through pollination and clonal vegetative propagation, the genetic diversity of globally cultivated *Vanilla* has become severely constrained, and is essentially genetically identical the world over (Ellestad *et al.*, 2022). This lack of genetic variability precludes any form of DNA based analysis.

3.4.6 Saffron

Similarly, saffron is propagated asexually through division and replanting corms. A study by Busconi and co-workers found that there was little genetic variability between saffron samples from Europe, Africa and Asia, reflecting clonal propagation methods. This study focused mainly on Spanish saffron but did include 39 samples from around the world. Using amplified fragment length polymorphism analysis, the study found that a predominant genotype (A1) was present in Spanish, French and Greek samples, which all showed a uniform genetic make-up. The A1 genotype was also present in all the other production areas studied. There were however eleven less frequent genotypes identified which were restricted to single locations, although the numbers of samples from each location varied only between three and five samples. The study however found high epigenetic variation between the samples which was consistent with their geographical origin. Epigenetic variations are influenced by environmental conditions and the study theorised that it would be very likely that samples from cultivation areas under different climates could be characterised by different epigenomes. The techniques used in this study are relatively insensitive with only 47 polymorphic loci identified (Busconi *et al.*, 2015). The use of more modern techniques, perhaps high-density SNP arrays or GBS could identify more geographically anchored epigenomes: a first step in geographical attribution of saffron.

The microbiome fingerprint of saffron corms was investigated in a small feasibility study across three production sites: Morocco, Kashmir (India) and Kishtwar (India) (Bhagat *et al.*, 2021) with a view to determine if the cultivation areas could be distinguished from each other. It was found that the cormosphere, the collection of all microbes (for example bacteria, fungi, viruses) associated with the corm, across the three sites harboured common phyla, with 24 genera found at all locations. However, there were some bacterial genera unique to each of Kashmir, Kishtwar and Morocco, which could be used to develop microbial markers for the geographical origin of corms from these regions. One significant drawback of this study is that it was the cormosphere which was studied and yet it is the stamen of the plant which enters the human food chain. It would be interesting to determine if the variation in the cormosphere extended to the stamen, otherwise, unfortunately this study is of little use for the attribution of geographical origin of saffron stamens.

3.4.7 Government Research Projects using genomic analysis

3.4.7.1 EU Framework 6 Project TRACE: Tracing (the Origin of) Food Commodities in Europe.(2005-2009)

An objective of Work Package 3 was the identification of PCR markers for the detection of plant species related to honey. This Work Package aimed to develop rapid, robust, accurate and cost-effective methods for determining the species/varietal origin of food using real-time PCR. The project focused on the development of methods which could authenticate ‘Miel de Corse’, a product of protected designation of origin (PDO). The study developed and validated real-time PCR systems for the detection of plant species commonly found on Corsica: sweet chestnut, lavender, eucalyptus, rockrose, oak and broom, which should be used to build a profile of the plants used by bees as forage during the production of “Miel de Corse”. Additionally, detection systems using real-time PCR for other plant species commonly identified in honey (acacia, linden, citrus, clover, heather, olive, rape, sunflower and rosemary) were also developed and validated. These real-time PCR systems were then used to distinguish Corsican honey from samples of honey from other geographic regions, including “Miel de Galicia”, a protected geographical indication (PGI) honey as well as German and English honeys were analysed.

This study showed that a combination of species specific systems, selected from the pool of real-time PCR systems developed during this project, was able to produce a plant species profile unique to Corsican honey when compared to honey from Galicia, Germany and England (Laube *et al.*, 2010), although the number of samples analysed was comparatively small and the locations relatively geographically remote from each other. Honey from Germany and England were easily distinguishable from Southern European honey by the detection of DNA from oilseed rape: a favourite foraging flower for bees, but not a crop grown in Southern Europe.

3.4.7.2 Development of Metagenomic Methods for Determination of Origin (Defra reference FA0160)

Between 2015 and 2019 this project piloted the application of microbial metabarcoding to trial the identification of microbial communities related to the origin of i) oysters, and ii) Stilton cheese.

The overall aim of this project was to use the microbial community associated with oyster gills as a blueprint for the application of metagenomics in food authenticity, and then to test the approach in a new commodity, blue Stilton cheese (a Protected Designation of Origin commodity). More specifically this involved the following objectives:

- Transfer methodology from the obsolete Roche 454 (Roche Diagnostic, Netherlands) DNA sequencing platform to the current MiSeq (Illumina, UK) DNA sequencer.
- Increase replication for the oyster analysis, including samples from multiple sites and seasons and a non-UK source (northern France).
- Attempt to identify regional signatures in the oyster bacterial community, and uncertainty in the approach.
- Apply the methodology to another commodity, PDO Blue Stilton cheese.

To address these aims, a total of 450 oyster samples were collected by the Centre for Environment, Fisheries and Aquaculture Science (Cefas) from around England, from sites in Cornwall, Dorset, Essex and Northumberland. Ten oysters were collected at each collection visit, and sampling took place from April 2015 to November 2016. A further 60 oysters were collected from two sites in northern France, between October 2016 and February 2017. Oyster gills underwent DNA extraction and PCR amplification at the 16S v4 variable region, followed by sequencing on an MiSeq (Illumina, UK) DNA sequencing instrument. A total of 504 samples were successfully sequenced for further analysis. Sequences were clustered into groups based on their sequence similarity, and representative members of these groups were compared to a database of known bacterial sequences, to determine which bacterial taxa were present in which samples. After sequence analysis was complete, Bayesian statistical methods were used to attempt to identify different groups of taxa that are associated with different geographical origins.

Once this methodology was established, it was rolled out to Stilton Cheese. Across 2016 and 2017, 101 cheese samples were taken, both directly from the various Stilton producers, or from retail sale in the UK. Samples were comprised of Blue Stilton cheese (n=62), and for comparison UK non-Stilton blue cheese (n=30) and non-UK blue cheese (n=9). Samples were processed as per the oyster samples, except that two genes were analysed; the 16S rRNA gene, for identification of bacteria, and the Internal Transcribed Spacer (ITS) region, for identification of fungi.

The bacterial communities in the oyster samples were highly variable, even among samples from the same time and location, and oysters could not be assigned to regions based on bacterial sequences. The blue cheese data were more promising, and samples could be assigned as Blue Stilton with 72-79% accuracy. There was also some ability to assign to individual producers (40-60% accuracy).

This work would benefit from further work in at least two areas – re-analysis of the existing data with more modern analytical techniques (for example de-noising and Amplicon Sequence Variant analysis, instead of Operational Taxonomic Unit clustering), and trialling on other foodstuffs with a microbial component to their production.

3.4.8 Limitations and gaps in using genomics techniques

In summary, this literature review did not locate any studies where the use of DNA-based analyses was mature enough to be put into practice. The majority of studies could at best, be called pilot studies, due to restricted numbers of samples, low numbers of locations and little sampling between years and seasons. Methods which assess the DNA of the target species are very difficult to anchor to a geographical location, for example the breed of Korean black pigs, which is normally restricted to Korea and was described using microsatellite markers (Oh *et al.*, 2014), could equally be reared in any other location. Swamp and river buffalo are usually found in South east Asia, and was described using a SNP panel, but could be reared in any comparably warm water body (Pérez-Pardal *et al.*, 2018). Putre's oregano, a subspecies of oregano recognised in Chile with a Seal of Approval (Contreras *et al.*, 2021) and described using sequencing and a SNP panel, could be grown in any other location where oregano grows.

3.4.9 Conclusions for using genomics techniques

In conclusion, DNA based analysis is not currently well placed for designating geographical origin. It can however add value, to traditional chemical-based analyses, in a combined approach, to verify the organism or product of interest, thereby increasing the veracity of the final geographical origin designation.

3.4.10 Outlook in using genomics techniques

What is undeniable is that DNA based analyses are unmatched in being able to identify species and to further differentiate subspecies, varieties and breeds. Where DNA can probably add the most value in ascribing geographical origin is when it is used in tandem with a more traditionally used technology for geographical origin determination. An example of this is a combined use of high-throughput sequencing using NGS with metabolomics using GC-TOF-MS analysis on Panxian ham, a traditional Chinese dry-cured ham protected by national geographical indication (Mu *et al.*, 2020). Other study (Fernandes *et al.*, 2015) combined a couple of DNA based methods to identify the variety of grape vine for Portuguese wine and was able to assess geographical origin using ^{87}Sr/^{86}SR isotope ratio data. The further development of SNP databases could provide a powerful tool to be used in tandem with other chemometric techniques.

3.5 Proteomics

Publications reporting the use of proteomic approaches for determination of geographical origin of food are scarce, with only a few studies identified during the review of published and grey literature performed for this project, as outlined below.

3.5.1 Honey

A recent review of proteomics for food authenticity (Afzaal *et al.*, 2022) included one article that described the application of proteomics to geographical origin determination of honey (Wang *et al.*, 2009). This work was based on fingerprinting of proteins by matrix assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) and protein barcoding generated using MALDI Biotyper 1.1 software to analyse the protein profiles. A database of protein profiles (barcodes) was created using mass spectra of 16 honey samples of known Hawaii origin. Commercially purchased honeys (n=38) with labels indicating the origin from different countries and various states of the USA, including Hawaii (n=15), were tested against the database. All the samples from Hawaii showed a good correlation with the barcodes from the authentic honeys in the database (correlation coefficient ≥ 0.75), although four samples from other countries or US states also showed correlation coefficients in the same range. Although these are preliminary data and they look interesting, a larger database and further analysis of many more samples from different regions would be needed to assess the performance of the approach.

3.5.2 Cocoa

A proteomics approach was used to compare the protein profiles of 25 cocoa samples from different continents and plant hybrid varieties (Kumari *et al.*, 2018). The study involved two-dimensional separation of proteins by gel electrophoresis followed by trypsin digestion, mass spectrometry analysis and protein identification. Statistical analysis of the results obtained with non-fermented beans (proteins detected and their abundances) showed a correlation with the geographical origin of the samples (at continental level) and not with the cocoa hybrid. The authors also observed a correlation between the total protein content of the non-fermented samples and the soil pH, and they speculate that the quality of the soil and the growth conditions may influence the protein accumulation in the plants. Although these results provide an interesting characterization of the cocoa beans studied, the approach is very laborious and not suitable for routine testing. The authors discussed the challenges of protein quantification as different methods produce different results due to the complexity of the cocoa matrix. This is likely to be the case for protein extraction also, which would need to be optimised and standardised to achieved robust data. Although the protein profile correlated with the continent of origin of the cocoa and not with the hybrid variety, there was one variety, CCN5, that was very different to the rest and made South American samples cluster with African samples in a PCA plot. This indicates that a sufficiently large number of samples from different locations and hybrid varieties would be required to fully assess the performance of this approach.

3.5.3 Limitations and gaps in using proteomics techniques

The lack of studies on the use of proteomics for geographical origin designation in food stuffs reflects the inherent problems with anchoring protein profiles to a unique geographical location. Proteins are ultimately derived from the DNA sequence of the target organism and, as with genomic analysis above, it is difficult to link genomic sequences to a geographical location when the intervention of man is taken into account. This is exemplified by the study on Hawai'ian honey. The Hawai'ian Islands are home to a large array of unique plants, due to its geographically isolated location. It has been estimated that there are approximately 1,400 vascular plant taxa (including species, subspecies, and varieties) native to the State of Hawai'i, of which, nearly 90 percent are found nowhere else in the world. It would be assumed, therefore that this geographic isolation would produce clearly distinct and unique barcodes for Hawai'ian honey, however, the arrival of humans brought new plant species to Hawai'i, not only agricultural plants, but also ornamental plants. The attribution of Hawai'ian honey to an Hawai'ian provenance was impressive, save for the fact that almost 10% of non-Hawai'ian honeys would have been wrongly classified as Hawai'ian (Wang *et al.*, 2009).

Although less advanced than other technologies, the use of proteomics for geographical origin determination is advocated by some due to certain advantages such as the higher stability of peptides to processing compared to DNA, or the effect of processing and production methods on specific chemical modifications of amino acids or formation of peptides that can be used as biomarkers (Ortea *et al.*, 2016). The Greek consortium FoodOmicsGR_RI, coordinating the development of Omics tools to support the agri-food sector are focusing their proteomics efforts on cheese and they are using nano-LC / HRMS to characterise the proteome and peptidome of Greek cheeses. The aim is to create a database of proteins and peptides identified in the original milk and in serial samples collected during the production and maturation processes. This consortium has also used MALDI-TOF MS to create a reference library of bacteria present in various Greek cheeses based on ribosomal proteins along with some housekeeping proteins. The use of Bruker MALDI Biotyper software enables pattern-matching to the reference library for identification of bacterial species present in different cheeses and their correlation with quality and safety traits. The authors claim that this approach can also be applied to PDO determination by studying the microbial diversity of different PDOs. Limitations that need to be addressed for further development of proteomic approaches for geographical origin determination include standardisation or detailed reporting of experimental details, the use of powerful statistical analysis and appropriate number of replicates, access to advanced MS equipment and bioinformatics tools. As for other technologies, extensive, well curated databases are critical to obtain robust data.

3.5.4 Conclusions for using proteomic techniques

In conclusion, as with DNA based analysis, the use of proteomic techniques is not currently well placed for designating geographical origin. They can, however, add value to traditional methodologies, in a combined approach, to verify the organism or product of interest, thereby increasing the confidence in the final geographical origin designation.

3.5.5 Outlook in using proteomics techniques

Due to the technical complexity and economic cost of proteomic approaches for geographical origin determination, it can be anticipated that developments in the short and mid-term will be targeted to specific commodities and geographical indication questions such as PDO and PGI to support marketability and brand protection.

3.6 Emerging techniques

This section focuses on the wide range of emerging technologies and computational tools which have been used in a research and development setting and which have been applied to a small number of samples and commodities. Some of these more promising technologies are discussed below, and a summary of other techniques identified in this literature search is presented in Table 4.

3.6.1 Corona Discharge Mass Spectrometry

Corona Discharge Mass Spectrometry (CD-MS) was applied for the first time to categorise black (n=11) and white (n=7) pepper seeds from different origins (Charoensumran *et al.*, 2021). This simple and rapid MS-based method combined with chemometric analysis demonstrated the ability to distinguish geographic origins from chemical profiles with discrimination efficiencies > 98%. The authors state that that this approach has potential to be used as part of a portable setup for remote and onsite analysis. Despite the low dataset numbers, this outcome suggests the methodology could be applied to a broader range of agricultural products for origin determination without any pre-treatment. Any future applications should focus on larger datasets and a broader range of samples.

3.6.2 Excitation-Emission matrix and Synchronous Fluorescence

Excitation-Emission matrix (EEM) and Synchronous Fluorescence was used for the first time to discriminate saffron from 3 different Moroccan provenances (n=18) (El Hani *et al.*, 2023). Moreover, global geographic discrimination between samples from Morocco, Afghanistan (n=10) and Iran (n=10) was achievable through PCA and LDA. Despite the low dataset numbers in this study and the limited geographic dataset, this non-destructive, simple, and reliable method could precede verification of saffron adulteration and quantification of fluorophore compounds for valorisation in applications such as cosmetics, food, and pharmacy.

3.6.3 Intelligent Sensory Technologies

PCA of Electronic nose (E-nose) data could rapidly distinguish samples of fresh instant rice (n=18) from three different Chinese provinces (Ren *et al.* 2023). The solid phase microextraction-gas chromatograph-mass spectrometer (SPME-GC-MS) results of samples from different provinces were clearly distinguished in PCA and hierarchical cluster analysis. Ten compounds were identified as potential markers of three Northeast Chinese fresh instant rice provenance. The authors conclude that the strategy of applying flavour profiles for determining geographical origin of fresh instant rice was an effective and non-destructive technique. Although this study is from a limited area in China on a specific product it stands as a proof of concept for future analysis of more diverse sample types

E-nose was also utilised in a study to discriminate the geographical origin of ginger (Yu *et al.*, 2022). In this study, HS-GC–MS and fast GC e-nose were used to successfully distinguish the varieties and geographical origins of dried gingers from seven major production areas in China. After chemometric analysis distinct separation of two different varieties of ginger was achieved on the HS-GC-MS data. However, this method was not effective for origin identification. Flavour profiles extracted by fast GC e-nose demonstrated better identification of ginger varieties and geographical origins based on pluralistic chemometrics and could be applied to trace source and region of ginger. This study provided evidence of applicability of these technologies to ginger authenticity within China.

A novel voltametric electronic tongue (VE-tongue) system based on three nanocomposites modified working electrodes was used for the discrimination of red wine from different geographical origins (Zheng *et al.*, 2022). Wine samples (n=120) from four different denominations of origin (France, Australia and two locations in China) were bought from a local supermarket and three working electrodes were applied for classification and prediction using PCA plot. The authors conclude that the three novel working electrodes delivered a VE-tongue system that can successfully discriminate different red wine samples by their denomination of origins, thus cutting down the detection cost of a versatile E-tongue system without interfering in the discrimination capacity of the system. The method would require further validation against a larger dataset to test and improve this model.

A paper comprising a comprehensive literature review addressing the authenticity determination of a variety of alcoholic beverages through intelligent sensory technology (IST) was presented by Wang *et al.* (2022). The techniques covered were E-nose, E-tongue and E-eyes. They concluded that IST have been successfully applied in quality assessments of alcoholic beverages, in terms of variety and geographical origins, monitoring production processes, detection of frauds and adulterations, discrimination of years of aging, distinction of brands and types, aroma analysis, detection of spoilage and off-flavours, and monitoring of the production process. However, the E-noses and E-tongues instruments still need improvement, especially the development of high sensitivity and selectivity bioelectronics sensor arrays aimed at improving accuracy and reliability of the analysis.

3.6.4 Sesquiterpene hydrocarbon fingerprinting by Headspace – Solid Phase Micro Extraction and Gas Chromatography-Mass spectrometry

Headspace – solid phase micro extraction (HS-SPME) and GC-MS was shown to be the fit-for-purpose tool for virgin olive oil geographical authentication (Quintanilla-Casas *et al.*, 2022). Virgin olive oil produced from EU (n = 246) and non-EU (n = 154) origins were correctly classified using chemometrics for 89.6% of samples during external validation experiments. The SH fingerprint provided a large amount of information, but the PLS-DA allowed discrimination of the most relevant variables according to the origin categories. Successful results were also obtained for classification models by country. Between EU countries the model correctly classified 92.2% and between non-EU countries 96.0%. These are remarkably high classification rates considering the natural heterogeneity of the oil and analytical variability. The group concluded that the proposed approach could be scaled down to authenticate the origin of oils obtained from smaller and closer areas of origin.

3.6.5 GC-Ion Mobility Spectrometry

Palm oil is one of the most economically important products in Malaysia with Malaysia and Indonesia being responsible for 85% of global palm oil production and the industry is predicted to expand due to demand. Palm oil is commonly traded as a global commodity with batches from different sources often being mixed at multiple stages during processing, shipment, refining, storage, and delivery to end users. The traceability of palm oil within the supply chain is a challenging issue. Several certification schemes, such as Malaysian Sustainable Palm Oil (MSPO) and the Roundtable on Sustainable Palm Oil (RSPO), have been set up to assure the sustainability of palm oil production and its traceability across the supply chain. The current measures in place for the traceability certification of palm oil production are based largely on paper trails and audits. Work has been undertaken by IAEA using gas chromatography ion mobility spectrometry (GC-IMS), coupled with principal component analysis (PCA) and orthogonal projections to latent structures discriminant analysis (OPLS-DA), applied for the geographical discrimination of crude palm oils from Malaysia (Food Safety and Control Newsletter Vol. 01 No. 1, July 2022, 2022). Crude palm oil samples were collected over 6 months (February–July 2019) at 4 different Malaysian locations. A supervised chemometric approach, OPLS-DA, was able to discriminate East Malaysia from the Peninsular Malaysia in most of the months. An example of the OPLS-DA score plot of the oil samples, collected in July 2019, is shown in Figure 1. The goodness of fit (R2X(cum), R2Y(cum)) and the predictive ability (Q2(cum)) values of the 7-fold cross-validated OPLS-DA model were 0.898, 0.899 and 0.745, respectively.

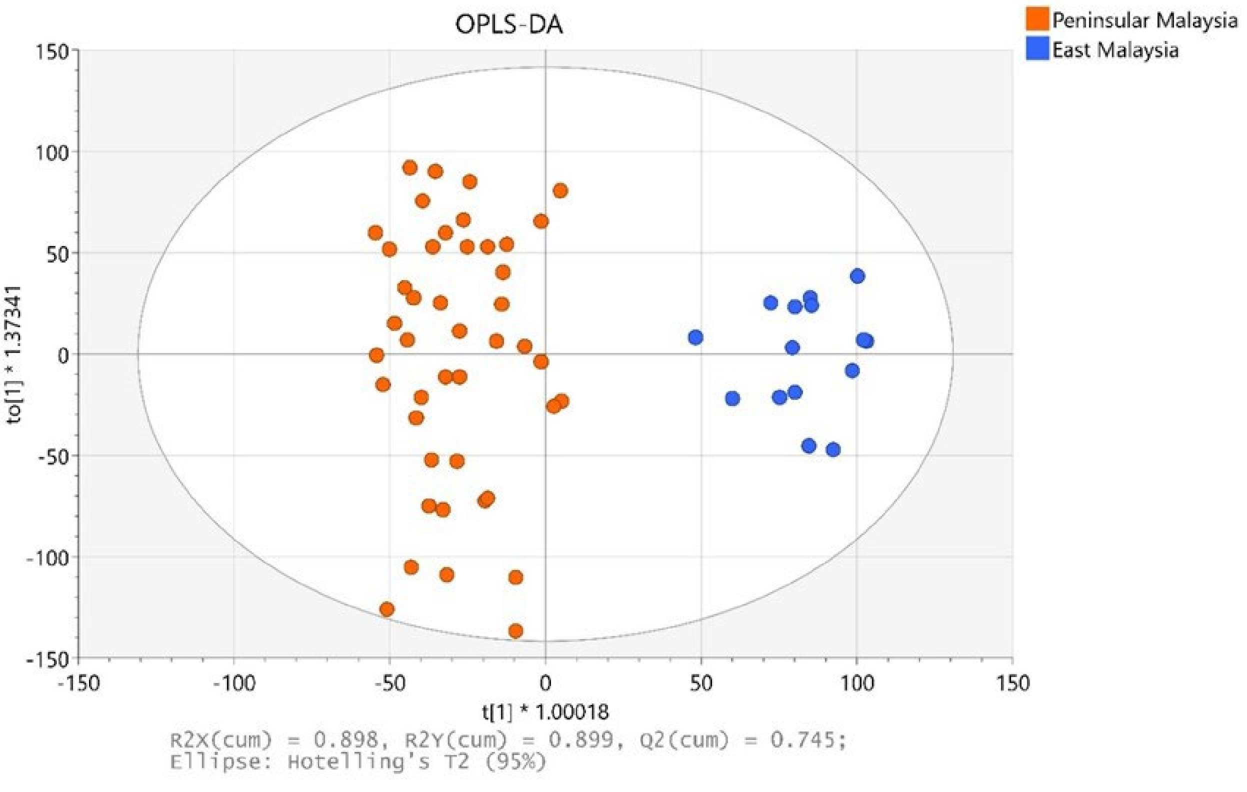

Figure 1: OPLS-DA score plot of the crude palm oil samples collected from Peninsular and East Malaysia in July 2019

As the next step, we assessed the ability of the OPLS-DA model to discriminate all 4 geographical locations. The discrimination of samples from Peninsular Malaysia was achieved in most of the months. Further work on the geographical discrimination of crude oil samples at IAEA has included the use of FT-NIR spectroscopy and stable isotope analysis.

3.6.6 Spectroscopy

Solid Phase microextraction-Gas chromatography-Mass spectrometry, IR and Raman techniques were reviewed as an application to authenticate the volatile profile of honey for botanical and geographic characterisation (Sotiropoulou *et al.*, 2021). Chemometric analysis was performed on the data from all techniques reviewed. SPME-GC-MS identified volatile compounds in honey that could be used as biomarkers for identifying botanical and geographic authentication. Numerous volatile components from a variety of botanical sources and geographical origins in mainly mono-floral honeys are listed. The authors note that the heating step to isolate the volatile components is the main disadvantage of this method as non-characteristic compounds are introduced into the samples. However, the technique is solvent free, inexpensive, rapid, and simple. They conclude that SPME-GC-MS based on volatile fraction was proved to provide reliable results to determine authenticity of honey in terms of botanical and geographic origin. It is not clear how effective this approach would be when challenged with multi-floral honey types. IR and Raman spectroscopies were determined to be suitable for the evaluation of botanical and geographical origin of honey when combined with chemometric analysis. The multi-spectroscopic approach using complementary techniques, and data interpretation through chemometric assessment, is required for the continued development and understanding of methodology related to honey authenticity.

Fluorescence spectroscopy in conjunction with parallel factor analysis (PARAFAC), PCA and SIMCA (soft independent modelling of class analogy) was used for the development of geographic and botanical discrimination models to differentiate among distinct honey classes (Ramona-Crina *et al.*, 2022). Honey samples (n=96) from seven botanical sources and two geographical origins (France and Romania) were studied. The study proved the efficiency of the association between EEM fluorescence spectroscopy in conjunction with chemometric methods for honey discrimination according to geographical and botanical origin. A differentiation group of 95.8% was achieved for geographic origin and 94.5% for botanical varieties. No meaningful distinctions were recorded for the honeydew, linden and acacia honeys produced in the two countries. It was also noted that differentiations for colza and sunflower honeys might be caused by the distinct agricultural practice from Romania and France.

Honey samples (n=1040) from 6 different botanical origins were collected from 4 different locations within Indonesia and analysed by ultraviolet (UV) spectroscopy to determine authenticity (Suhandy and Yulia 2021). A SIMCA classification method was applied to the data. This technique demonstrated to be a simple and low-cost analytical method for authenticating Indonesian honey for botanical, entomological and geographic origin. This study was limited in terms of geographic and botanical origin capability, and it is questionable as to the suitability of this methodology for larger and more complex datasets.

Various molecular and atomic spectroscopic techniques (^{1}H NMR, portable NIR, benchtop NIR, Fourier transform infrared spectroscopy on the middle infrared region (ATR-FTIR-MIR), and Function-as-a-Service (FAAS) were used to characterize and discriminate Brazilian Canephora coffees of specific producers, including two with geographical indication, and also to differentiate them from the Arabica (Baqueta *et al.,* 2023). The main objective was to compare and evaluate the feasibility of discrimination of different analytical techniques to determine a role in real time applications. The sample set comprised 100 Canephora samples of different geographical origins in Brazil (Conilon from Espírito Santo, Amazonian Robusta from indigenous and non-indigenous producers of Rondônia, and Conilon from Bahia) and Arabica coffee (25 samples). The authors concluded that although there was a contribution from all the different techniques for characterisation, the multi-block discrimination showed that NIR spectroscopy dominated for this purpose. Due to this finding, comparisons between benchtop NIR and portable NIR were tested. Portable NIR provided slightly inferior results to benchtop NIR through variable selection.

Given that benchtop FT-NIR is a highly accessible and novel rapid screening technology, work is underway by IAEA to develop methods for the geographical discrimination of green coffee from Costa Rica and of crude palm oil from Malaysia, and the verification of authenticity of organic strawberries. A training programme with support Member State efforts to improve their food safety and authenticity control systems and raising awareness of this highly accessible technology.

Spectroscopic techniques (NIR, MIR, Raman and UV-vis) were reviewed for rapid compositional analysis, authenticity, and traceability in beer and wine (Chapman *et al.*, 2019). They record that all these techniques are rapid and easy to set up with very little or no sample pre-treatment. However, there are drawbacks such as requirements for application specific calibrations and overlapping signals. They concluded that these analytical platforms provide new opportunities for compositional analysis of beer and wine samples as well as help with authenticity, regional discrimination, and traceability. There are issues to resolve such as the availability of data mining tools, spectroscopic databases and difficulties combining the methods when large datasets are generated. There is also limited knowledge of what molecular changes take place during production of beer and wine. The paper concludes that future developments in vibrational spectroscopy will require a new approach in the analysis and interpretation of the data generated. The application of these techniques will determine that food science analytics is moving away from data being a discreet range of numbers reflecting what has happened, to more dynamic space where data mining, modelling, and big data together will provide relevant information that can be revealed to better understand complex interrelationships, processes, and functionality.

A study to review the application of handheld and portable spectroscopy-based devices, for the determination of food authenticity monitoring and traceability verification was conducted by McVey *et al.* (2021). Specifically, NIR, MIR, Raman, Vis-NIR, Visible and hyperspectral imaging (HIS) were reviewed as established technologies in this area. The largely commercially available NIR spectroscopy are miniaturised handheld devices with a low cost. These devices have a good classification capability when the resulting data is analysed by with chemometric tools. MIR are portable systems that are limited in size due to moving parts and detector requirements. These factors also make them less cost effective and can compromise performance of hand-held versions. However, they do have advantages over NIR due to the ability to identify and characterise structure and isolate mixtures for quantification. The Raman devices are a favourite technology for development in food adulteration due to low cost, long life, portability, and high sensitivity. Although negatively impacted by noise and other interfering signals, they are rapid in their experiment time. The development of portable devices in the Vis wavelength region has proven successful due to the availability of high performing, low-cost photodetector arrays. Portable, handheld Vis and Vis-NIR devices share many advantages seen in NIR as they are cost effective, non-destructive, rapid and can be used for chemical characterisation of foods. Although these devices are not as established as the NIR and MIR technologies they have shown promising developments in this area. HIS is emerging as a powerful tool for food authenticity analysis. Collecting both spectral and spatial information from a sample allows HIS to characterise complex heterogeneous mixtures as well as identifying surface and sub-surface constituents. This technique is the most flexible as it can analyse numerous samples at the same time. However, the cost for these instruments is high.

White wine samples (n=180) from 3 north-eastern Italian varieties were analysed by Surface Enhanced Raman Spectroscopy to determine their classification (Zanuttin *et al.*, 2019). Using the ratios of three chemical characteristics, discrimination could be made between different wine varieties and producers. The results show that the quality features of these wines are due to local environmental characteristics and wine making processes related to the winery. The portable nature of this methodology is an advantage but proof of the concept on a larger, more complex dataset is required for it to be utilised as a global authentication tool.

Laser-induced Breakdown Spectroscopy has also been demonstrated as a promising emerging technology for the geographical origin of commodities such as olive oil (Gazeli *et al.*, 2020, Gyftokostas *et al.*, 2020).

3.6.7 Other techniques

Samples of dark chocolate originating from Africa (n=15), Asia (n=11) and South America (n=31) were analysed by Flow Infusion-Electrospray Ionization- Mass Spectrometry (FI-ESI-MS) to assess the geographical origin of cocoa beans (Acierno *et al.*, 2018). The results of chemometric assessment confirmed separation between African and Asian origin chocolate but no clear trend for South America. The inability to separate all three continents was linked to brand related factors such as formulations, climate growing conditions and industrial processing.

Absorbance-transmission and fluorescence excitation-emission matrix (A-TEEM) was used for the first time to discriminate Shiraz wines (n=186) from sub-regional level within the Barossa and Eden Valleys (Ranaweera *et al.*, 2023). Clear vintage variations were seen between the samples. A 98.4% accuracy in classification was achieved when modelling both regional and vintage classifications. The method would require further validation against a larger dataset to test and improve this model.

Table 4: Other technologies and computational tools that have been applied to the geographical origin determination of food

Technique	**Commodity**	**Outcome**	**Reference**
APCI-MS	Apple juice	Correct GO classification Accuracy: 94.2%	(Zielinski *et al.*, 2014)
FTIR ATR	Coffee	GO classification of Arabica coffee Accuracy: 100%	(Zaukuu *et al.*, 2022)
FTIR ATR	Olive oil	Accuracy 100% when genetic analysis added	(Kakouri *et al.*, 2021)
FTIR	Turmeric	Unsuccessful at attributing GO	(Zaukuu *et al.*, 2022)
FTIR	Coffee	GO classification of green coffee Accuracy: 100%	(Zaukuu *et al.*, 2022)
FTIR	Coffee	GO classification of Arabica coffee Accuracy: 100%	(Zaukuu *et al.*, 2022)
FTIR	Olive oil	Cluster samples according to GO	(Rohman *et al.*, 2020)
FT-NIR	Coffee	GO classification of Arabica and Robusta coffee Accuracy: >93%	(Zaukuu *et al.*, 2022)

Technique	**Commodity**	**Outcome**	**Reference**
FT-NIR	Coffee	GO classification of Arabica coffee Accuracy: 100%	(Zaukuu *et al.*, 2022)
HPLC	Cocoa	Bioactive compounds - "some studies showing more success than others"	(Fanning *et al.* 2023)
IR	Honey	"suitable for the evaluation of botanical and geographical origin of honey when combined with chemometric analysis"	(Sotiropoulou *et al.*, 2021)
Instrumental neutron activation analysis	Lemon juice	Correct classification of origin in Argentina	(Zielinski *et al.*, 2014)
LC-MS	Cocoa	Bioactive compounds - "some studies showing more success than others"	(Fanning *et al.*, 2023)
LC-MS	Olive oil	Good discrimination	(Dou *et al.*, 2023)
LC-MS	Honey	Good GO discrimination	(Dou *et al.*, 2023)
MIR	Olive oil	Classified samples to GO correctly	(Kakouri *et al.*, 2021)
NIR	Saffron	Pattern recognition for Iranian samples 100%	(Zaukuu *et al.*, 2022)

Technique	Commodity	Outcome	Reference
^{1}H-NMR	Turmeric	GO classification of saffron Accuracy: 100%	(Zaukuu *et al.*, 2022)
^{1}H-NMR	Saffron	GO classification Accuracy: 100%	(Zaukuu *et al.*, 2022)
^{1}H-NMR	Sweet cherry juice	Poor GO classification Accuracy: 21% using principle component analysis, but >83.7% using SIMCA	(Zielinski *et al.*, 2014)
Particle-Induced X-ray Emission, FTIR and Accelerator MS – Radiocarbon Analysis (AMS-^{14}C)	Coffee	Combined technologies illustrate how different cutting-edge techniques can synergistically determine GO	(Chytry *et al.*, 2022)
PTR-MS	Coffee	GO classification of Arabica and Arabica/Robusta coffee mixtures Accuracy: 69%	(Zaukuu *et al.*, 2022)
Pyrolysis mass spectrometry	Orange juice	Separation of Brazilian and Israeli juice	(Zielinski *et al.*, 2014)
Raman spectroscopy	Wine	Accuracy = 76.6%	(Xu *et al.*, 2020)

Technique	**Commodity**	**Outcome**	**Reference**
Raman spectroscopy	Honey	Accuracy = 100%	(Xu *et al.*, 2020)
Raman spectroscopy	Olive oil	Accuracy = 86.6%	(Kakouri *et al.*, 2021)
Raman spectroscopy in combination with IRMS	Olive oil	Accuracy = 82%	(Kakouri *et al.*, 2021)
Raman spectroscopy in combination with IRMS	Rice	Identification error = 4.1%	(Xu *et al.*, 2020)

3.6.8. Blockchain

Several papers and review articles have been published which assess the use of digital ledger Blockchains for traceability and monitoring of food products from farm to shelf. Blockchain technology was studied by Peng *et al.* (2022) as a mechanism for collection of data in the rice supply chain. They concluded that theoretical verification of this emerging technology for resource sharing in the food supply chain provides digital ideas transferable to the food industry.

Blockchain, electronic tagging, and the combination of these traceability efforts are relevant for growing number of specifications. Emerging information and technology play a vital role in electronic agriculture (Saurabh & Dey, 2021). The technology adoption factors for grape wine supply chain using conjoint analysis were studied and presented. When designing a modular information system architecture for grape wine supply using Blockchain technology, Saurabh and Dey found that perceived disintermediation was the most important factor for users followed by traceability and then price. The direction of their future research is to expand to a system for deep traceability of bottled wine to wineries and vines.

A traceability prototype system was established to verify a case analysis of grain and oil traceability (Xu *et al.*, 2022). A whole chain of grain and oil quality and safety was constructed, and a classification table of key information was built. The traceability system proposed had the advantages of data tamper resistance through hash encryption, data traceability through multiple trade links and data sharing through nodes within the network capable of receiving and sending message to maintain the ledger. They found that further theoretical, practical, and quantitative indicators are required to study the benefits of Blockchain technology based on consumer demand for complete and correct information about the goods they buy.

Lavazza, an Italian roast coffee company, describe their case study (Gazzola *et al.*, 2023) introducing a Blockchain-tracked product to the market. Sensors were set up at the farm to monitor air and soil conditions. However, this was problematic due to reliance on continuous cellular connectivity. A user interface was created to detail the growing, harvesting, drying, quality, transportation, and roasting stages. They found that the strengths of the blockchain included data integration, tangible traceability and transparency. Whereas the weaknesses were manual applications for data input, limited to one batch ID at a time, high costs, difficult to escalate and possible falsification. The paper offers an example of how blockchain technology can be applied to increase traceability for stakeholders while maintaining transparency for consumers. However, this analysis aimed to be illustrative rather than definitive and only considered a single case study.

In a magazine article summarising the advantages and disadvantages of blockchain software in food supply (McKenzie, 2018) the author cited examples of software development by Walmart and IBM, aimed at addressing food safety and food traceability issues. Blockchain has the potential to support food safety. Almost 28 million people fall ill due to foodborne illness each year (Center for Disease Control and Prevention, USA). Should it become easily accessible to trace the supply chain for individual products linked to a food safety issue, contamination could be quickly identified and contained and products could be removed from the supply chain at a faster rate, potentially protecting consumer health, reducing the costs linked to food recalls and increasing consumer trust in certain products. A second area in which blockchain could provide protection relates to food fraud. Blockchain has the potential to protect food supply chains by increasing transparency and accessibility to companies. In addition to providing a manner to quickly trace products and their ingredients, blockchain may provide customers with more confidence in the authenticity of premium products and increase trust and thus increase spending on these products.

As a proof-of-principle example, Walmart asked their team to trace a package of sliced mangoes from their US stores back to their source, a 30-day chain involving 16 farms, two packing houses, three brokers, two import warehouses and one processing facility. This tracking process took the team six days, 18 hours and 26 minutes, a task completed by tracing food safety audits and certificates. Walmart then used the software developed by IBM and tracked the mangoes in only 2.2 seconds.

While this scenario shows impressive time savings due to blockchain software, Mitchell Weinberg, founder of food fraud protection and prevention firm Inscatech, argues that it is unrealistic to expect that all food can be tracked by blockchain, citing the example that, for a pack of spices, blockchain cannot determine whether the correct herb or spice was harvested in the first instance. Walmart acknowledge that tracking the entire list of ingredients which are contained within a single product is complex, due to the fragmentation of the food supply system, with much lack of digitisation and reliance on paperwork. While blockchain can be accurately implemented to track money by creating an accounting trail, Weinberg cites two major impediments to the success of blockchain as a food traceability solution. The first is that it requires for every step of the supply chain to be included in the data. The second is that it requires honest participation, in terms of, for example, correctly declaring the commodity, the authenticity of the handling party and where they are located. At present, there is no means of validating that information logged is genuine.

IBM acknowledge that data will not necessarily be guaranteed to be more accurate due to concerns over dishonesty and incorrect information declaration. However, the company hopes that the implementation of blockchain will improve data quality over time by removing the anonymity of those entering the data which may disinhibit unscrupulous behaviours. Making data more accessible should also mean that the discovery of both accidental and deliberate errors is facilitated. However, the notion that removal of anonymity will result in less unscrupulous behaviour is disputed by Weinberg, arguing that those adulterating food are sophisticated criminals who make much effort to cover their crime. Also, if information such as what is being harvested is not the declared commodity, there is no way of identifying this once incorrect source data are entered into blockchain. Therefore, while blockchain technology may facilitate efficiency in supply, such as quickly identifying delays and where efficiencies in supply can be made, the human element required to generate the data may mean that blockchain is unsuitable for identifying unscrupulous activity in food supply.

In conclusion, although there is the possibility to address the demand for tangible and transparent traceability systems throughout food production systems, there are flaws to Blockchain systems. While blockchain could aid paperchain traceability in a weight-of-evidence approach, a common opinion is that implementation costs of these systems are high. This is mainly due to required investment in the manual setup of the technology at a variety of different and complex stages within the food production process. Furthermore, there are various supply chains from raw ingredients to final product vendors within the chain. Another common concern is that much of the data is entered manually and is therefore subject to possible falsification and vulnerable to security attacks.

3.6.9 CEN activities, AOAC/OIV methods and other collaborative opportunities relevant to country of origin determination

Activities for food and liquid commodities by the European Committee for Standardisation (CEN), and for wine by the Association of Analytical Chemists (AOAC) and the International Organization of Vine and Wine (OIV), have been underway to facilitate the standardisation of methods to determine geographical origin. Unfortunately, standardisation activities can take much time.

The European Committee for Standardisation (Food Authenticity Working Group, WG6) has focused, since 2020, on the standardisation of methods for stable isotope analysis (reference CEN/TC460) in the following areas:

a) Determination of C and/or N isotope ratios in food by Elemental analyser – Isotope Ratio Mass Spectrometry (EA-IRMS)

b) Determination of $^{18}O/^{16}O$ isotope ratios in liquid aqueous food matrices by Equilibration – Isotope Ratio Mass Spectrometry (Eq-IRMS)
c) Sample preparation for isotope ratio analysis of the different fractions of fruit and vegetable juices and related products

These methods include sample preparation, usage of standards, normalisation and processing of raw data, which are vital aspects for comparability of data sets from different laboratories and consequently country of origin determination.

The finalisation of the standards will be accompanied by inter-laboratory validations which are planned by the Food Authenticity Working Group in the near future.

AOAC INTERNATIONAL's Official Methods of Analysis℠ program is the organisation's premier program for consensus method development. Methods approved in this program have undergone rigorous scientific and systematic scrutiny and are deemed to be highly credible and defensible. Concerning wine, Table 5, produced by the Association of Analytical Chemists (AOAC) and the International Organization of Vine and Wine (OIV), summarises the validated and officially approved analytical methods for wine using SIRA and ^{2}H-NMR for the determination of stable isotope ratios in wine constituents (Christoph *et al.*, 2015). The methods are also published in the "Compendium of International Methods of Wine and Must Analysis" of the International Organization of Vine and Wine (OIV). These are used for proof of adulteration in cases of chaptalisation, water and sugar addition, vintage but also for mislabelling of geographical origin.

Finally, other collaborative opportunities could be capitalised on to progress origin verification methods. Twenty-three countries are currently involved on setting up and implementing databases, alongside initiatives from NASIR (National Association of Security and Investigative Regulators, Germany) and NIST (USA). Should the UK become involved, there would be opportunity to share these data. There is also opportunity to link up with Commonwealth countries to share data to expand on previous UK geographical origin projects to re-visit and expand on the data UK hold on beef, salmon and wine.

In order to form and strengthen collaborations and data sharing, with an aim to provide origin databases scaling large areas, the IAEA is supporting countries (from low to high premium member states) for geographical origin database initiatives, providing training to standardise sampling strategies, sample analyses, data evaluation and leading to routine testing.

Outside of food, the IAEA functions as a third-party auditor, testing 5% of the total samples collected by World Forest ID activities to cross-check data to support timber origin. World Forest ID initially focussed on timber, holding centralised data, which gets updated continuously by accredited laboratories and are also expanding to other commodities, for example to verify that soya and palm oil have not been grown in areas cleared for this purpose by unpermitted deforestation and on mapping shrimp origin. More information relating to timber origin methodologies, showing an example of a successful working system to trace geographical origin, is included in stakeholder interview.

Table 5: Summary of the validated and officially approved analytical methods for wine using SIRA and 2H-NMR

Analytical method	**Molecule (OIV Method)**	**Isotope ratio**
^{2}H-NMR, (SNIF®-NMR)	Sugar, ethanol *(OIV-MA-AS311-05),* *AOAC 995.17*	$(D/H)_I$ ppm $(D/H)_{II}$ ppm
^{13}C/^{12}C, SIRA	Sugar, ethanol *(OIV-MA-AS312-06),* *AOAC 2004.01* Glycerol *(OIV-MA-AS312-07)* Carbon dioxide *(OIV-MA-AS314-03)*	δ^{13}C, ‰ V-PDB

Analytical method	**Molecule (OIV Method)**	**Isotope ratio**
$^{18}O/^{16}O$, SIRA	Water *(OIV-MA-AS2-12)*	$\delta^{18}O$, ‰ V-SMOW

Key: D, Deuterium (2H); V-SMOW: Vienna Standard Mean Ocean Water, $\delta^{18}O = 0$ ‰ V-SMOW; V-PDB: Vienna Pee Dee Belemnite, $\delta^{13}C = 0$‰ V-PDB. OIV AOAC

Finally, there is a future requirement for the standardisation of statistical analysis to prepare data for geographical origin elucidation.

3.6.10 Outlook for emerging technologies

It is envisaged that further development will be made by manufacturers of handheld and portable devices to offer increased sensitivity, selectivity, and availability. These devices have been shown to be a proof of concept for the rapid determination of geographic origin on local levels. As the technology develops the expertise in operation and interpretation will require development to facilitate the building and understanding of larger more robust datasets on a global scale. A final consideration for these technologies as they develop will be the accreditation and standardisation of the methodologies.

Digital leger blockchain technology could be a useful tool for geographic origin traceability of farm produce and processed foods. Large amounts of information can be stored in one place. However, the success of these systems is reliant on the co-operation of every member of the supply chain to invest in and apply the technology which could be a challenge. For this technology to be applied in field, issues around security and integrity will need to be addressed.

3.7 Summary of most promising techniques per commodity

The progress of various technologies in attributing geographical origin has been discussed in detail above. While emerging technologies may come to the fore in the future, from this information, the list below summarises the methods shown to be most developed and showing the most demonstrated capability and potential.

- **Cereals:** Relatively little research has been completed for this commodity type and it may be worth investigating with the combination of SIRA and trace element analysis.
- **Cocoa:** A common theme for cocoa origin determination is a need to transition to more portable, rapid, affordable and non-destructive analytical approaches to support testing in remote areas. Spectroscopy techniques show great promise, combining NIR and sensory techniques with AI.
- **Coffee:** SIRA analysis, particularly for oxygen isotope ratios, in combination with trace element analysis.
- **Fish and Seafood:** Gaining analytical data relating to the geographical origin of commodities cultivated in the marine environment is particularly challenging. Trace element profiles vary greatly due to a combination of natural and anthropogenic activities and depending on harvest time. It is therefore likely that the combination of several technologies will provide the most informative models overall, including trace element analysis, NIR and REIMS study of lipid markers. Due to the wide range of varying factors, it will be imperative that databases are constantly updated to account for these variations.
- **Fruit juice:** Trace element analysis is the most mature analysis which has shown promise.
- **Garlic:** The research for this commodity is immature, comprising a small dataset over a restricted geographical range. Methods to consider include combining trace element, volatile compounds and metabolite data.
- **Honey:** Pollen analysis using light microscopy can be used to determine the relative amounts of pollen contributed to honey by different plant species. However, this requires high levels of skill and expertise from a limited number of analysts who are currently able to determine geographical origin based on this method. The study of volatile compounds, sugars, organic acids and amino acids has been used to differentiate between floral type of honey. SIRA, particularly of H, O and C isotopes shows promise, and trace element analysis can add value in terms of determining environmental factors, such as trace elements in water sources, which relate to where the pollen originated. Metabolomic and genomic approaches in combination with blockchain provide current state-of-the-art technologies. These could be coupled with analysis of

volatile compounds. Limitations relating to the low copy numbers which in the past have challenged the success of digital PCR are starting to be overcome and this technology may also be of value in the future.

- **Meat:** Livestock can be moved across geographical origins during their lifetime and dietary background (which can vary due to geography and also due to feed type and origin). SIRA (H, C, N, S and O) is a well-established technique to infer geographical origin with carbon isotopes beneficial to discriminate diet and deuterium and oxygen to support the discrimination of geographical origin. Combining SIRA with trace element analysis, and also considering fatty acid profiling can improve the confidence of the data, along with RFID to monitor livestock movement.
- **Olive oil and other edible oils:** Since geographical origin is a major source of the variation in oil composition, a multivariate approach seems to offer the most potential. SIRA technology (particularly for C, H and O) has shown great promise for this commodity group, and it is suggested that this could be combined with NMR and profiling of phenolic compounds, fatty acid profile (including FAMEs), sterols, triacylglycerol (TAGs), volatile compounds and colour. The additional inclusion of trace element data could be considered but the impact of fertilisers and fungicide on trace element levels must be understood. The potential of FTIR ATR should also be considered.
- **Rice:** A combination of SIRA (particularly C, H and O) with trace element analysis seems the most promising application.
- **Saffron:** There has been insufficient work on this commodity to conclude on the most promising method, but SNP arrays and GBS appear to be worth investigating further.
- **Tea:** Trace element profiling is the most mature method that has offered the most promise.
- **Tomato:** Studies are limited but trace element profiling appears to show promise, for example Li/Cu, Co/Rb and Sr/Cd ratios.
- **Vanilla:** Much more research is required here across various technologies. It appears that genomics technologies are not applicable due to the absence of genetic variation.

- **Whisky:** There has been much focus on GC, LC and spectroscopy, and it appears that trace element analysis could add to a weight-of-evidence approach.
- **Wine:** SIRA and SNIF-NMR methodologies are the most mature and are applicable, as exemplified by their application in the EU Wine Databank. These could be combined with trace element profiling to compare the trace element of the wine with important components of the soil, especially Sr and Pb. However, the impact of processing must be understood including the contribution of Cu and Al from the fermentation vats. The impact of trace elements from fining agents, agricultural practices and pollution must also be considered.

3.8 Challenges and limitations to geographical origin determination

This literature review clearly highlights many challenges remaining in the determination of geographical origin across all technologies. Firstly, there is a lack of available databases, with databases not existing, not available for public access or not incorporating a sufficient scope of samples. Keeping databases up-to-date with regular inclusion of contemporary samples will allow for natural and seasonal variation so that the methods are ready for use to respond to issues as soon as they arise. Without regular maintenance, delays in testing will be incurred, the length of which will depend on the level of variation which has occurred in the interim period.

In order for methods to be considered as ready for use, methods must be standardised and validated and must have undergone an inter-laboratory validation. There has been a lack of reference materials and calibration standards (Camin *et al.*, 2017) used during the determination of geographical origin. The AOAC/OIV have published methods for the determination of the geographical origin of wine which have undergone rigorous scientific and systematic scrutiny and are deemed to be highly credible and defensible. Work is underway by the European Committee for Standardisation's Food Authenticity Working Group to standardise testing of the geographical origin of solid food matrices (raw or processed) and also, separately, liquid matrices, with inter-laboratory trials planned for the future. Reference materials for these methods being considered by CEN are available for purchase and then necessitate intra-laboratory validation by each individual user laboratory. Once this work is complete, proficiency testing schemes could be co-ordinated in the future to support laboratories in achieving the quality of analysis demanded of them by their customers, accreditation bodies and managers.

Data in Appendix 3 show the cost to purchase and run a range of the main technologies currently applied to address geographical origin analyses, including time requirements for training. This is to inform on investments required to support official control in the future. As shown in this appendix, most of the technologies involved require significant investment to purchase and run the equipment and require high levels of training. The estimated charge per sample assumes that a sufficient number of samples are analysed by the required method each year to cover the costs of method accreditation including participation in any future proficiency testing rounds, instrument maintenance, annual servicing and also the costs of gases where applicable. It is therefore recommended that it may be most cost effective for Official Control Laboratories to sub-contract the analyses involving the instruments which they do not routinely own rather than investing in new instruments and training. The technology which requires the lowest cost is the portable NIR which requires much less investment and training and produces data rapidly.

4. Stakeholder engagement activities

In order to capture information and views from stakeholders relating to the geographical origin of food, and to better-understand the current outlook regarding geographical origin verification of food and feed, stakeholders from enforcement, geographical origin testing laboratories, representatives of food traceability networks and trade bodies, independent experts in food authenticity, those working in database and technology development along with stakeholders in the supply chain were contacted either to be interviewed or to complete a questionnaire in relation to their involvement in geographical origin verification and to determine potential future actions to support geographical origin. In addition to these stakeholders, an independent consultant, Dr Simon Kelly was contacted, alongside an organisation which is running a successful working model of traceability of food and timber commodities (World Forest ID). Dr Kelly is knowledgeable in origin verification, particularly specialising in SIRA. The outputs of this stakeholder engagement are included in Appendix 4.

World Forest ID provided an example of a working model for verification of geographical origin, albeit mainly for timber, with a more minor element for food and feed commodities. This model has a large dataset, dispersed over an almost global area, the methods of collecting samples and analysing data have been standardised and the database is curated centrally. The datasets are constantly expanded, allowing for variation according to commodity type (including tree species), season and year.

4.1 Summary of key aspects highlighted by stakeholders

- It is understood across stakeholders that there is no ‘silver bullet’ technology to determine the accurate origin of food and feed commodities and that testing data need to be checked to determine if they are consistent or not with the declared origin in order to make a judgement regarding the accuracy of food labels or paper trails. While SIRA and trace element testing are the most commonly used technologies across commodity types, stakeholders acknowledged that testing using other technologies can add weight-of-evidence.
- Geographical origin of food is not a priority testing area. Food safety is the main focus, followed by species identification so research and capabilities for origin verification are lagging behind. Origin verification tends to be a question for those higher in the supply chain who tend to be the parties who instigate testing and the point was raised that it is difficult to engage with those involved earlier in the supply chain as origin questions are regarded as a problem rather than as a solution to food authenticity and value.
- Traceability documents are an easy target for counterfeit activities. Large companies tend to have a simpler supply chain than smaller companies, as well as technical staff to support traceability. Traceability is therefore more difficult for smaller companies, which would benefit from more support in this area.
- SIRA and trace element analysis are the main methodologies used to verify origin, often in combination. These technologies provide data which can be

used to determine whether this aligns or not with the declared origin of the food, but don't provide unequivocal origin determination.

- The main analytical challenge highlighted was the need for extensive and robust databases, with large numbers of authentic samples and being updated continuously. Almost all stakeholders highlighted this.
- Issues with databases:
  - Databases need to be representative of seasonal and annual variations, including taking account of climate change impacts,
  - Databases should be built with assured authentic samples, collected and tested against standardised protocols to provide robust databases, with no chance that fraudulent samples could be included.
  - Metadata should be recorded to assist data interpretation, e.g., extreme temperatures in a specific year / season.
  - Databases should be challenged regularly using authentic and non-authentic samples.
  - Datasets are usually owned by private companies and not publicly available. This is due to the large investments required to build databases and the commercial value that they have.
  - Sharing of data and transparency about numbers / types of samples and uncertainty associated with specific testing methods / databases / data analysis interpretation were flagged as areas for improvement.
  - Currently, many UK testing customers access databases which hold only UK data and therefore verification of origin of non-UK produce is a challenge. The point was raised that it would be ideal for the databases to be curated centrally with public access and that the lack of transparency from database holders relating to database design and data interpretation was a challenge to food origin verification.
  - Due to the lack of transparency regarding database design and composition, points were raised that, when a customer receives an unexpected test result relating to origin declarations, they can approach an alternative provider and receive an alternative outcome without understanding which test may provide the more representative result for their query.

- Testing laboratories tend to produce their own reference materials, which usually have been subjected to inter-laboratory validation and have been produced from authentic samples.
- There is agreement regarding the preference for accredited laboratories and methods, where available. Also, it would be beneficial if laboratories supported the initiation of proficiency testing or similar, in which they could participate to demonstrate their competency in methods.

5. HorizonScan™ data relating to geographical origin

Data were sourced to inform on the number of global incidents relating to geographical origin issues in food and feed over the last decade. HorizonScan™ provides food safety professionals with a comprehensive, scientific basis for assessing supply chain risks. The tool is a web-based information service with a suite of proprietary risk assessment and supplier check tools tracking current and historical global food fraud and contamination issues in near real time which are reported via official means. HorizonScan™ collects daily data for 550 commodities, from over 110 food safety agencies and 180 countries. Data were taken from HorizonScan™ to detail incidents recorded relating to geographical origin issues over the last ten years. The data are shown in Figures 2a and 2b.

As shown in Figure 2a, the commodities for which the largest numbers of incidents which have been notified over the last decade relating to geographical origin are honey and wine, followed by fish, meat, olive oil and food/dietary supplements. In Figure 2a, each of the honey notifications relate to honey of incorrectly declared country of origin and often also of (related) incorrectly declared botanical origin in the same product notifications. Along with incorrect origin declarations for a given product, some notifications also include adulteration of these honeys with undeclared or unauthorised dyes and sweeteners. The wine notifications in Figure 2a relate to the grapes having been harvested in a different country of origin to that declared and the fish notifications relate to incorrect or insufficient information regarding country of origin.

From the HorizonScan™ data, 68% of the notifications for exports from Czech Republic (Figure 3) relate to honey being of the incorrect geographical origin (relating to incorrect botanical origin) plus issues in the same honeys relating to addition of colourings and sweeteners. These notifications were all reported in the period January 2015 to December 2023. Relating to wine exported in Czech Republic, 27% of the notifications relate to the grapes having been harvested in a different country of origin to that declared.

Of the issues shown in Figure 3 for exports from Slovak Republic, 85% of these relate to wines containing grapes harvested from an alternative country of origin to that declared. Concerning the high number of issues shown on Figure 3 relating to unknown exporting nations (exporting nation not declared in Rapid Alert System for Food and Feed (RASFF) notification), these data relate to a wide range of commodities of unknown origin. The issues relating to Spain are also for a range of different commodities.

In addition to this officially recorded data, issues were captured in the grey literature (trade journals) relating to the faking of PGI balsamic vinegar of Modena (prepared with suspected lower-grade grapes)[25], falsification relating to the vintage and origin of wine in New Zealand[26] and concerns for the geographical origin of meat labelled as British[27].

---

[25] ‘Fake balsamic vinegar’ scandal as Italy uncovers major fraud case (thegrocer.co.uk)
[26] Wine fraud: Southern Boundary Wines fined $1.7m (rnz.co.nz)
[27] Tesco ‘foreign’ pork chop sold as British a one off, industry claims (thegrocer.co.uk)

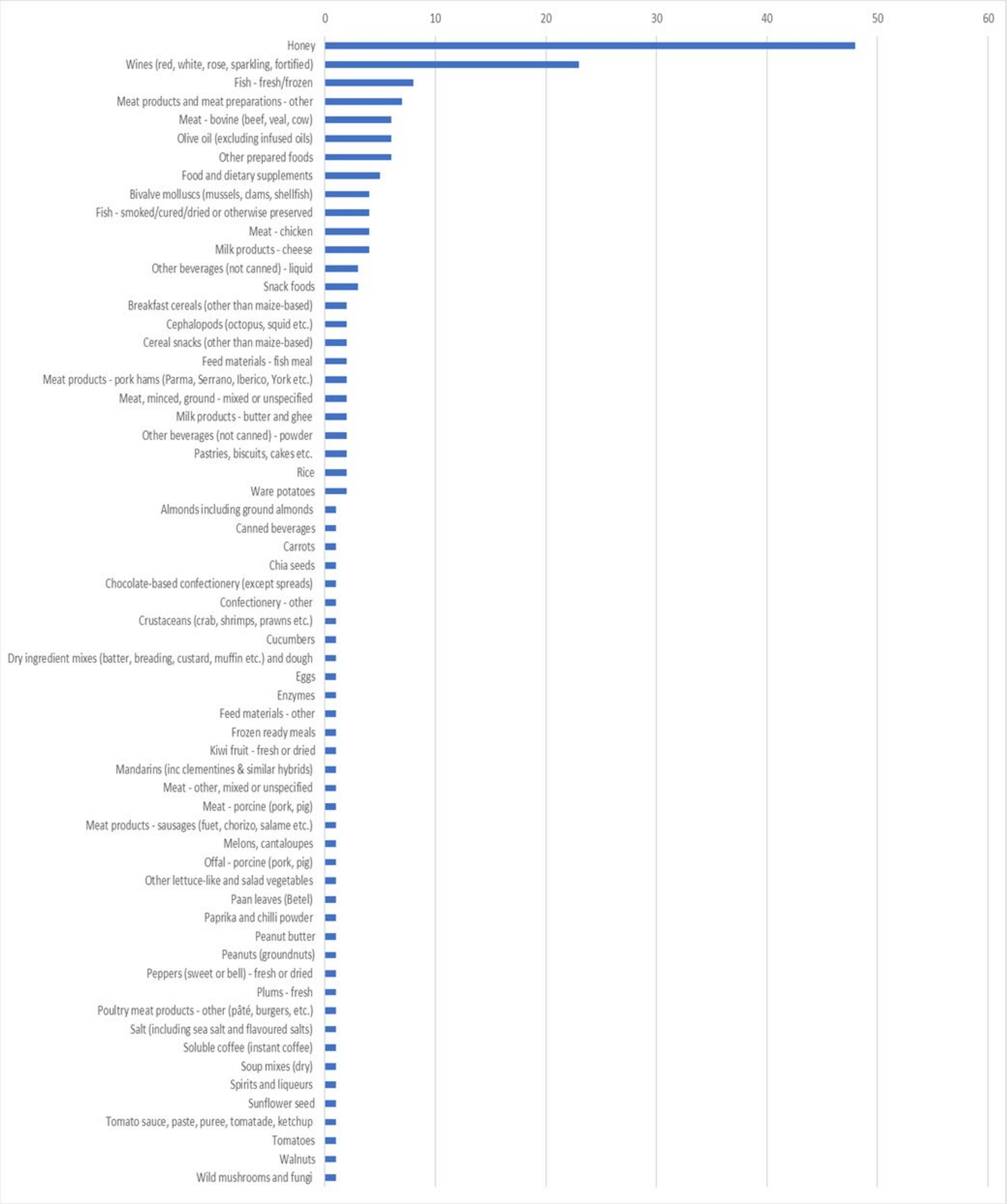

Figure 2a: Data showing the number of notifications (x-axis) per commodity type (y-axis) from the EC RASFF data plus alerts/recalls published by the relevant food safety agency in each country over the decade preceding this report

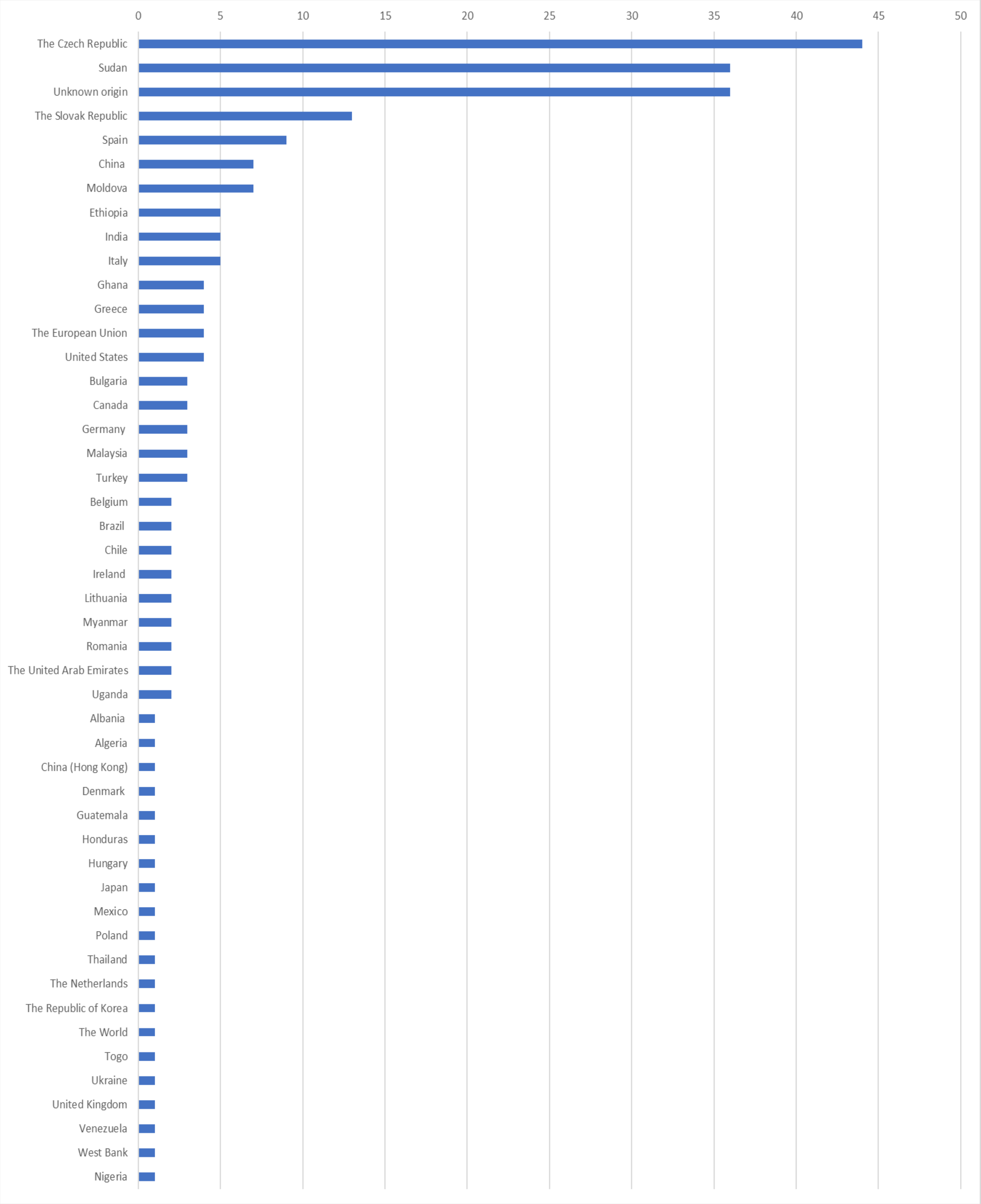

Figure 3: Data showing number of notifications from RASFF data (x-axis) per country (y-axis) including alerts/recalls published by the relevant food safety agency in each country over the decade preceding this report

6. Conclusions and Challenges: Review of methods for the verification of geographical origin of food and feed

There is no unequivocal single technique with which to verify the country of origin of food and feed. Testing methods do not recognise country borders and there can be wide variations in data profiles across countries. Technologies can be applied but, rather than the data providing an absolute geographical origin, the data must be evaluated to determine whether they are consistent or not with the declared origin of a product.

The maturity of the testing for origin varies greatly between commodities and geographical locations. The validity of studies undertaken to date is highly dependent on the number and nature of the samples used as a reference collection. In many studies this is inadequate to reach firm conclusions and there is also doubt over the robustness of sampling plans in published studies. Obtaining authentic reference materials is a critical problem. While combined SIRA and trace element analyses are the most widely used technologies and often provide the most pertinent data for provenance especially in combination also with Strontium analysis, depending on the commodity and scenario, other technologies such as genomics and spectroscopy are applicable or can add value to SIRA and TE data.

Using multiple data sets from different analytical sources generally improves classification rates and therefore the robustness of origin determination. Data fusion is an effective strategy for improving the classification performance. There is scope to investigate the inclusion of other data such as that from contaminants such as pesticides, dioxins etc to improve the outcomes of testing methods. If successful, applying multivariate data in this way provides a large suite of evidence to substantiate testing results. A point to note is that using multivariate methods may require more frequent monitoring compared to using, for example, SIRA and TE data only.

There are barriers to the implementation of methods for geographical origin determination. The main challenges relate to quality, size, geographical range and curation of the databases. The lack of continuous expansion, and therefore lack of current relevance in terms of seasonal and annual natural variation, size and robustness of databases used for verification of origin, is discussed throughout this report.

Datasets are fragmented, with small datasets available for certain constrained locations, for example neighbouring regions of one country with data generated for a single food commodity. Data sharing is poor. Due to investment and IP concerns, data tend to be held in private databases and the UK can mostly access UK data. This makes the origin verification of imports challenging. Before robust geographical origin testing methods can be achieved, large representative databases containing authentic samples must be prepared, incorporating high numbers of samples for which datasets span large areas of the globe. The number of samples and extent of geographical range should be stated to facilitate accurate interpretation and uncertainty judgements of results. Investment is required since databases must be continuously expanded to incorporate natural season-to-season and annual variability. Increasing the number of samples and the geographical area within a dataset would add confidence to the outputs and would influence on the quality of the prediction models, to the benefit of all users. The involvement in the food industry in providing samples would increase the representation of natural variability within datasets.

Databases must be continuously expanded to account for season-to-season and annual variation. They should be challenged with new/additional samples to ensure continued relevance.

Ideally, in the future, databases would be curated centrally, an approach successfully adopted for the provenance of timber. Data should only be accepted into the database one validity has been demonstrated by use of accepted/standardised methods for collection and analysis of authentic samples. Standardisation activities are under development by AOAC, CEN, CODEX, however this takes time. Transferability of data from one lab to another must be demonstrated by proficiency trials.

Other barriers to geographical origin verification relate to:

- Costs – the main methods used require mass spectrometric analysis which requires high-cost instruments and skilled analysts.
- Contacts are required to source authentic samples in foreign countries in a standardised manner, to build and expand databanks of analytical data.
- Lack of food-specific matrix-matched reference materials for SIRA.

- Demand for proficiency testing is lacking for geographical origin determination and this is impeding the trustworthiness of the data.
- Requirement for future standardisation of statistics applied to geographical origin data.
- There is a challenge to present multivariate data in court. Only limited methods have been considered ready to be used in court, for example single isotope analysis data for wine and for German butter trade.
- IP considerations of databases and related accessibility.
- Due to the nature of the methods, should a product be adulterated and contain a commodity from a mixture of origins, this is problematic for origin verification and the achievable limits of testing for adulterated samples must be understood.
- Few methods to verify origin are accredited and the accreditation process can take twelve months. Means to fast-track methods for accreditation are needed when a method is specifically developed to address and investigate a known issue in the supply chain.

The confounding issues relating to meat origin mainly concern the movement of animals and meat products during rearing and supply, which, from an analytical perspective, makes it very difficult to identify the source of meat products. Analytical methods for the detection of meat origin also need to consider different types and geographical sources of animal feed which can impact on the chemical profile. Routine testing for origin and the establishment of databases for meat authenticity will fail if they do not account for feeding regimes and the movement of animals.

In the cases for which there are many variables which contribute to the authenticity of a product, for example wine, an initial critical step must be undertaken to ensure that there is good information in relation to geographical origin and any compounding factors such as vintage, to ensure that robust metadata are collected.

While the most substantial approach of data/techniques for country of origin determination to date has involved the combination of SIRA, trace element and chemometrics methodologies which have been used for the long standing application for PDO/PGI foods, there is scope for the emergence of other methods, either for initial, fast screening or more detailed analysis. While hand-held spectroscopy technologies have been used in small proof-of-principle studies, multivariate analysis may be considered as a natural next step in the quest to verify origin, fusing data generated by genomics (including environmental genomics), isoscapes (for example hydrology) and metabolomics. In the future, it would also be interesting to combine other existing datasets to determine their value in verifying geographical origin. Datasets which could provide valuable information include environmental data, pesticide residue data, dioxin and polychlorinated biphenyl levels (particularly relevant to foods including salmon), per- and polyfluoroalkyl substances (PFAS), viral screen data, fungi species presence. Other existing data which have not been applied to origin determination, such as various data held by the British Geological Survey, could be incorporated into datasets.

While many of the larger members of the food supply chain are aware of the importance of accurate origin labelling and perform testing to challenge label information, there is opportunity to engage and educate parties at the lower end of the supply chain in order to mitigate subsequent issues. Projects at IAEA such as the Implementation of Nuclear Techniques for Authentication of Foods with High-Value Labelling Claims (INTACT Food) project are going some way to filling this knowledge gap.

Finally, there are opportunities for public, private and inter-nation funding to improve methods in geographical provenance which could be capitalised upon to protect our supply chain.

6.1 Future Direction

Based on the information gathered in this project, key areas of future work which we recommend are detailed below:

- All stakeholders highlighted challenges with a lack of quality data or up-to-date data in the various databases which have been prepared over the years in support of food provenance testing. In the future, investment could be made to support the building of robust datasets, necessitating the harmonised collection and analysis of high number of authentic samples from a large (e.g.

global) geographical area. Efforts should focus on commodities which are vulnerable to origin fraud.

- These datasets should be curated in a centralised hub to facilitate harmonisation and the datasets should be regularly expanded with samples to account for seasonal and annual variation and so that the models are ready for immediate use to respond to issues in the supply chain. The hub should be funded so as to allow continued expansion and curation of the datasets over time. The suitability of the datasets should be challenged each year by the testing of additional samples (i.e. not the samples included in the prediction models) and all testing facilities involved should take part in proficiency trials and achieve acceptable data, demonstrated through statistical measures, for example z-scores.
- Few methods to verify origin are accredited and the accreditation process can take twelve months. Means to fast-track methods for accreditation are needed when a method is specifically developed to address and investigate a known issue in the supply chain.
- Relevant matrix-matched certified reference materials should be made available.
- Demand for proficiency testing in origin analyses should be encouraged and certified testing schemes should be initiated. Currently no proficiency testing schemes are detailed for geographical origin on the global EPTIS database.
- There is much existing data from other testing exercises which are not yet/rarely used in origin verification, and which could have value for origin determination. This could be explored. Investment could be made to initiate the collation of these existing data and metadata and investigate its utility for origin verification purposes. Prediction models could be built to determine the relevance of these data for addressing origin issues. Such existing data could include environmental and geological data, pesticide residue data, dioxin and polychlorinated biphenyl levels (particularly relevant to foods including salmon), viral screen data and fungi species presence information.
- Government funding and private and EU funding opportunities in origin determination could be considered to expand collaborations between those working on building datasets for a commodity (or range of commodities) across

different geographical locations and, in turn, to expand the geographical range of datasets.

7. Acknowledgements

Fera Science gratefully acknowledges the joint funding from Food Standards Agency and Defra to conduct this work. We are indebted to all stakeholders who engaged with this project via interviews and completion of questionnaires to inform on issues relating to the current status of geographical origin of food. We also gratefully acknowledge Dr Simon Kelly who was a specialist consultant for the project, particularly specialising in SIRA.

This report has been prepared by Fera Science Limited ("Fera") for the for the sole benefit of Food Standards Agency and Defra. This document, and all the information, images and intellectual property rights in it belong to Fera (or its licensees). No part of the text or graphics may be reproduced without the prior written permission of Fera. Except as otherwise advised in writing by Fera, this information is confidential in nature must be treated by the receiver with at least the degree of care that it applies to its own confidential information (and always with at least a reasonable standard of care).

Fera shall not be liable for any claims, losses, demands or damages of any kind whatsoever (whether such claims, losses, demands or damages were foreseeable, known or otherwise and whether direct, indirect or consequential) arising out of or in connection with: (i) any advice given by Fera or its representatives; and/or (ii) the preparation of any technical or scientific reports. Fera makes no representation as to the suitability of using any particular goods in any manufacturing processes or scientific research, nor as to their use in conjunction with any other materials. Fera shall not be liable for any reliance placed on, nor for any recommendations, interpretation, analysis, guidance, suggestions, proposals or endorsements made in connection with, the services and/or the commercial or scientific activities carried out by Fera or its representatives.

Appendix 1. Country of origin food labelling rules – (UK)

Food labelling rules are maintained largely by EU retained legislation, in particular the 1169/2011 Regulation on the provision of Food Information to Consumers and its related Implementing Regulations. These include that origin labelling is required in the following situations:

1. Where failure to indicate this might mislead the consumer as to the true country of origin or place of provenance of the food, in particular if the information accompanying the food or the label as a whole would otherwise imply that the food has a different country of origin or place of provenance.

2. Where meat falls within the Combined Nomenclature ('CN') codes listed in Annex XI (1169/2011)*:

- Meat of swine, fresh, chilled or frozen.
- Meat of sheep or goats, fresh, chilled or frozen.
- Meat and edible offal of chickens, turkeys, duck, geese, or guinea fowls, fresh, chilled or frozen.

*Details of this are governed by Commission Implementing Regulation (EU) 1337/2013 – this stipulates that there must be 'origin information' displaying place of 'rearing' and 'slaughter' for the above meats. Use of the word 'origin' only applies where the animal was born, reared and slaughtered within one country.

3. Where the country of origin or the place of provenance of a food is given and where it is not the same as that of its primary ingredient, the country of origin or place of provenance of the primary ingredient in question shall also be given or indicated as being different to that of the food. (Commission Implementing Regulation (EU) 2018/775.)

4. Where the country of origin or place of provenance is given mandatorily in accordance with the Regulation or voluntarily through any indication such as statements, terms, pictorial presentation or symbols.

Separate regulations guide the mandatory labelling of beef, beef products, fish and shellfish, fruit and vegetables, honey, olive oil and wine.

Where origin information is not mandatory, it may be provided voluntarily as long as it is accurate, verifiable and does not mislead. This includes statements, terms, pictorial presentation or symbols.

1.1 PDO and PGI schemes in the EU

PDO and PGI products represent the excellence of European agricultural food production and are both the result of a unique combination of human and environmental factors that are characteristic of a specific territory[28].

The European Union dictates precise regulations for their protection, creating specific quality standard regimes that protect the good faith of the consumers and with the purpose of giving producers concrete instruments for identifying and promoting products with specific characteristics in a better way, as well as protecting them from unfair practices.

Regulation (EU) no. 1151/2012 (article 5) scrupulously describes the meaning of the acronyms PDO and PGI, specifying that:

The label PDO (Protected Designation of Origin) identifies a product that originates in a specific place, region or country, the quality or characteristics of which are essentially or exclusively due to a particular geographical environment with its inherent natural factors (raw materials, environmental characteristics, location) and human factors (traditional and craft production) and the production, transformation and elaboration phases of which all take place in the defined geographical area, in

---

[28] European Certification System (pdopgi.eu)

respect of rigid production regulations established in the procedural guidelines of production.

The label PGI (Protected Geographical Indication) indicates a product that originates in a specific place, region or country, whose given quality, reputation or other characteristics are essentially attributable to its geographical origin, and for which at least one of the production steps takes place in the defined geographical area.

Only those production processes that demonstrate a consolidated and coded production tradition, an inseparable bond with the territory of origin, a suitable social-business fabric and which manage to reach high quality standards that are certified by third party control organisations can hope to obtain and preserve the coveted recognition of the community and contemporary entrance into the European register of PDO and PGI products.

1.2 UK Geographical Indication Schemes

As with the EU, in the UK food, drink and agricultural products with a geographical connection, together with those that are made using traditional methods, can be registered and protected as intellectual property[29]. This protection is called a geographical indication (GI).

GI protection guarantees a product’s characteristics or reputation, authenticity and origin. It protects the product name from misuse or imitation.

The UK GI schemes protect registered product names when they are sold in Great Britain (England, Scotland and Wales).

---

[29] Guidance: Protected geographical food and drink names: UK GI schemes (gov.uk)

The EU GI schemes protect registered products names when they are sold in Northern Ireland and the EU.

All product names protected in the EU on 31 December 2020 following successful applications to the EU GI schemes are protected under the UK and EU GI schemes.

1.3 UK GI scheme logos and product labelling

There are UK GI logos that represent the 3 designations of GI product. The use of these labels is strictly controlled, and a raft of rules must be followed before a logo can be applied to a product.

Figure 4: PDO, PGI and TSG logos published 31 December 2020

Appendix 2. SIRA databases from the literature

Table 6: Table of SIRA databases from the literature

Food	Country of origin	Number of samples	Year	Stable isotopes	Analytical technique	Other parameters	Chemometrics	Outcome	Reference
Honey	Italy (botanical varieties)	265	1999 to 2005	δ13C (ethanol, honey and protein), δ15N (protein), (D/H)I	EA-IRMS and SNIF-NMR	TE (Al, B, Ba, Ca, Cr, Cu, Fe, K, Mg, Mn, Na, Ni, Pb, Rb, Sr, Zn) by ICP-OES	Analysis of Variance and Tukey's test; PCA	Separation of Northern and Southern chestnut honey	(Bontempo *et al.*, 2017)
Honey	Turkey (pine)	373	2015 to 2017	δ13C (honey and protein)	EA-IRMS	C4% sugar, sugars (fructose, glucose, sucrose, and maltose) and a multitude of physicochemical properties	PCA	Three different groups (South Anatolia, North Agean and high-altitude locations)	(Uçurum *et al.*, 2023)
Beef	13 regions in 8 European countries	224	2005 to 2009	H, C, N, S (protein)	T/C and EA-IRMS	N/A	Analysis of Variance and Tukey's test; PCA	Three different groups (coastal area, low latitude, C4 based diet; inland area, low latitude, mixed diet; coastal area, higher latitude, C3 based diet)	(Bontempo *et al.*, 2023)

Food	**Country of origin**	**Number of samples**	**Year**	**Stable isotopes**	**Analytical technique**	**Other parameters**	**Chemometrics**	**Outcome**	**Reference**
Wheat	Regions within Federal State of Hesse (Germany)	132 (including 24 reference samples from elsewhere)	2013 to 2014	H, C, N, O, S (protein) and H, C, O (lipids)	EA-IRMS	N/A	Analysis of Variance and boxplots	Differentiation of non-Hessian from Hessian samples	(Gatzert *et al.*, 2021)
Olive Oil (EVOO)	Greece	210	2005 to 2006, 2010, 2015 to 2016	C, O (EVOO and irrigation)	EA and Py-IRMS	N/A	Boxplots and correlation of C over O (EVOO)	Differentiation of Ionian Island olive oils from olive oils from Crete and Chalkidiki	(Karalis *et al.*, 2020)
Olive Oil (VOO)	Italy (9 regions)	387	2009 to 2011	C, O	EA and Py-IRMS	N/A	Analysis of Variance and Post Hoc Fisher test	O vs C strong correlation, driven by latitude, temperature, rain and xerothermic index	(Portarena *et al.*, 2014)
Rice	China (17 provinces)	900	2017	H, C, N, O	EA and Py-IRMS	TE (K, Mg, Ca, Na, Fe, Zn, Mn, Cu, Ni, Cr and Mo) by ICP-MS	Machine learning (BPNN model)	Classification of Japonica (97.2%) and Indica (77.9%) rice	(Li *et al.*, 2022)

Food	**Country of origin**	**Number of samples**	**Year**	**Stable isotopes**	**Analytical technique**	**Other parameters**	**Chemometrics**	**Outcome**	**Reference**
Pork	Korea (non-Korea)	252 (94)	Not provided	C, N	EA-IRMS	TE profile, including 27 isotopes of 19 elements by ICP-MS	multivariate chemometric linear discriminant analysis (LDA)	Cross validation classification rate 88 to 100%	(Park *et al.*, 2018)
Wine	Austria (10 sites in 4 regions)	349	1997 to 2004, no 2000 data	C (ethanol), O (wine water)	EA-IRMS and CO2-eq-IRMS	(D/H)I, (D/H)II, R-Wert by SNIF-NMR	Box plots, Canonical discrimination	North to South trend for (D/H)I, information of vintage recommended for classification	(Philipp *et al.*, 2018)
Wine	Venetian (7 regions)	2019	Not provided	δ11B, 208Pb/206Pb, 207Pb/206Pb, 206Pb/204Pb, 208Pb/207Pb and 87Sr/86Sr)	MC-ICP-MS	TE (n= 63) by HR-ICP-MS	seven classification methods (LDA, QDA, k-NN, NB, RF, ANN, SVM)	Red and white wine clusters; good results for Amarone, Bardolino, Pinot Grigio and Recioto PDO.	(Rapa *et al.*, 2023)
Wine	Global (France, Italy, Spain, USA, Chile, South Africa, Australia and China)	600 (red, imported)	1999 to 2014	C (ethanol and glycerol), O (wine water)	GC-C-IRMS and CO2-eq-IRMS	metals (Ca, K, Mg, Na, Fe, B, Sc, Cr, Cu, Zn, Rb, Sr, Al, Ba, Mn, Ti) by ICP-OES and MS	Analysis of Variance and Tukey's test, ANN	Overall discrimination accuracy of 83.9%	(Wu *et al.*, 2019)
Wine	France (4 regions)	240 (red)	1999 to 2014	C (ethanol and glycerol), O (wine water)	GC-C-IRMS and CO2-eq-IRMS	TE (B, Sc, Cr, Cu, Zn, Rb, Sr, Al, Ba, Mn, Ti) by ICP-MS	PLS-DA and ANN	Verification accuracy of 98.2%	(Wu *et al.*, 2021)

Food	**Country of origin**	**Number of samples**	**Year**	**Stable isotopes**	**Analytical technique**	**Other parameters**	**Chemometrics**	**Outcome**	**Reference**
Chicken	Global (17 countries)	384 (including 5% turkey)	No indication	H, O, C, N and S of defatted dry mass	EA-IRMS	TE (n= 53; 26 used) by ICP-MS and strontium ratios by TIMS	CDA	Multivariate verification accuracy of 88% for all samples, without strontium (global average across all samples)	(Rees *et al.*, 2016)
Scallops	China (7 coastal regions)	575 (Patinopecten yessoensis, Chlamys farreri, and Argopecten irradians)	2016 (spring and autumn)	C, N	EA-IRMS	N/A	Analysis of Variance and Scheffe's or Tamhane's T test; LDA of 150 samples	Average 92% classification for origin prediction and 98.3% accuracy for species prediction	(Zhang *et al.*, 2019)

Appendix 3. Costs to purchase and run technologies to respond to geographical origin issues

Table 7: Table of costs to purchase and run technologies to respond to geographical origin issues

Technology	Instrument type	Manufacturer	Model	Estimated cost of instrument in 2024, including software and database access if applicable	Approx. cost of reference material for a batch of 20 samples	Approx. cost of consumables for a batch of 20 samples	Approx. time required for training, to become competent, e.g. for a science graduate	Approx. time requirement to analyse, interpret and report a batch of 20 samples	Approx. service costs and frequency	Approx. cost of gases if applicable	Approx. cost for a commercial testing lab to analyse a batch of 20 samples
1	Stable Isotope Ratio Mass Spectrometer	Elementar	Isoprime presicION IRMS	£174,600 (excludes VAT)	N/A	N/A	6 months	N/A	£9,000	£9,000	N/A

Technology	Instrument type	Manufacturer	Model	Estimated cost of instrument in 2024, including software and database access if applicable	Approx. cost of reference material for a batch of 20 samples	Approx. cost of consumables for a batch of 20 samples	Approx. time required for training, to become competent, e.g. for a science graduate	Approx. time requirement to analyse, interpret and report a batch of 20 samples	Approx. service costs and frequency	Approx. cost of gases if applicable	Approx. cost for a commercial testing lab to analyse a batch of 20 samples
1	Elemental analyser (EA)	Elementar	Vario Pyrocube	£68,274 (excludes VAT)	£500 for each isotope (H,O, C, N, S) of interest	£200-£275, depending on isotopes analysed	6 months	23 hours for H/O or C/N/S mode; assuming EA is fully set up, following overnight equilibration of reactor packing	£9,000	£12,000	Samples preparation and analysis for H, O and S, N, S - £6000

Technology	**Instrument type**	**Manufacturer**	**Model**	**Estimated cost of instrument in 2024, including software and database access if applicable**	**Approx. cost of reference material for a batch of 20 samples**	**Approx. cost of consumables for a batch of 20 samples**	**Approx. time required for training, to become competent, e.g. for a science graduate**	**Approx. time requirement to analyse, interpret and report a batch of 20 samples**	**Approx. service costs and frequency**	**Approx. cost of gases if applicable**	**Approx. cost for a commercial testing lab to analyse a batch of 20 samples**
1	Gas/Liquid equilibration unit	Elementar	Isoprime MultiFlow-Bio	£59,967 (excludes VAT)	£500	£40 - £600, depending on isotopes analysed	6 months	14 hours (H), 18 hours (O)	£4,500 per annum	£11,000	£3,600 - £7,200 for H and O
2	ICP-MS (Single quad)	Agilent	7700x	£181,000 (excludes VAT)	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Technology	**Instrument type**	**Manufacturer**	**Model**	**Estimated cost of instrument in 2024, including software and database access if applicable**	**Approx. cost of reference material for a batch of 20 samples**	**Approx. cost of consumables for a batch of 20 samples**	**Approx. time required for training, to become competent, e.g. for a science graduate**	**Approx. time requirement to analyse, interpret and report a batch of 20 samples**	**Approx. service costs and frequency**	**Approx. cost of gases if applicable**	**Approx. cost for a commercial testing lab to analyse a batch of 20 samples**
2	ICP-MS (Triple quad)	Agilent	8900	£285,000 (excludes VAT)	N/A	N/A	N/A	N/A	N/A	N/A	N/A
2	Microwave Digestor (High pressure)	Milestone	UltraWave 3	£89,250 (excludes VAT)	£70	£30	5 to 15 days	2 days	~£240 per sample (multi-element screen)	£10,000 per annum	£9,000 per annum

Technology	Instrument type	Manufacturer	Model	Estimated cost of instrument in 2024, including software and database access if applicable	Approx. cost of reference material for a batch of 20 samples	Approx. cost of consumables for a batch of 20 samples	Approx. time required for training, to become competent, e.g. for a science graduate	Approx. time requirement to analyse, interpret and report a batch of 20 samples	Approx. service costs and frequency	Approx. cost of gases if applicable	Approx. cost for a commercial testing lab to analyse a batch of 20 samples
3	NMR spectroscopy	Bruker	500 MHz Avance Neo Onebay System	£658,492 (excludes VAT)	£100	£50	30 days	1 working day	£500, annually	£12,000 per annum	£3,000
4	FTIR-ATR (Bench-top)	Perkin Elmer	Spectrum 2 plus ATR and adulterant screen software	£60,000 (excludes VAT)	N/A	negligible	5 days	1.5 hours	£3,000/annual preventative maintenance	N/A	£200

Technology	Instrument type	Manufacturer	Model	Estimated cost of instrument in 2024, including software and database access if applicable	Approx. cost of reference material for a batch of 20 samples	Approx. cost of consumables for a batch of 20 samples	Approx. time required for training, to become competent, e.g. for a science graduate	Approx. time requirement to analyse, interpret and report a batch of 20 samples	Approx. service costs and frequency	Approx. cost of gases if applicable	Approx. cost for a commercial testing lab to analyse a batch of 20 samples
4	FT-NIR (Bench-top)	Bruker	MPA II plus an additional autosampler (30 positions) for reflection mode plus OPUS software	£90,000 (without an autosampler: £60,000) (excludes VAT)	N/A	negligible	5 days	Solid samples: analysis in reflection mode using the autosampler: 10 mins; Liquid samples: single sample analysis in transmission mode: 20 mins	£3,000/annual preventative maintenance	N/A	£100

Technology	**Instrument type**	**Manufacturer**	**Model**	**Estimated cost of instrument in 2024, including software and database access if applicable**	**Approx. cost of reference material for a batch of 20 samples**	**Approx. cost of consumables for a batch of 20 samples**	**Approx. time required for training, to become competent, e.g. for a science graduate**	**Approx. time requirement to analyse, interpret and report a batch of 20 samples**	**Approx. service costs and frequency**	**Approx. cost of gases if applicable**	**Approx. cost for a commercial testing lab to analyse a batch of 20 samples**
4	NIR (portable)	VIAVI Solutions	MicroNIR 1700 ES plus MicroNIR PRO software (perpetual licence)	£25,000 (excludes VAT)	NA	negligible	2 days	30 mins	N/A	N/A	£150
5	Raman spectroscopy	Thermo	First defender RMX	£96,870 (excludes VAT)	Included with instrument	None required	0.5 day	1 hour	£6000, annually	N/A	£3,000

Technology	Instrument type	Manufacturer	Model	Estimated cost of instrument in 2024, including software and database access if applicable	Approx. cost of reference material for a batch of 20 samples	Approx. cost of consumables for a batch of 20 samples	Approx. time required for training, to become competent, e.g. for a science graduate	Approx. time requirement to analyse, interpret and report a batch of 20 samples	Approx. service costs and frequency	Approx. cost of gases if applicable	Approx. cost for a commercial testing lab to analyse a batch of 20 samples
6	Next Generation (DNA) Sequencing (NGS)	Illumina	NextSeq1000 or 2000	£175,000 or £279,000, depending on model (excludes VAT)	N/A	£600	2 months to generate data on sequencer. Specialist bioinformatician required to interpret data	4 working days	£10,000 annually	N/A	£2,000

Appendix 4. Specialist Consultation Interview

4.1 Dr Simon Kelly, Introduction

Dr Simon Kelly is an internationally recognised specialist in the field of geographical origin verification of food, particularly in the fields of SIRA and trace element analysis, but Dr Kelly also possesses a broad range of expertise and over 30 years of experience in food authenticity origin analysis. Dr Kelly was given permission to participate in this project as part of his work-related duties for the FAO/IAEA (Food and Agricultural Organisation of the United Nations/International Atomic Energy Agency) joint centre for Nuclear Techniques in Food and Agriculture. However, it should be noted that his views do not necessarily represent those of the FAO or IAEA. Dr Kelly's involvement was provided free-of-charge and included two consultations. The first was at project inception to consult on the current status of global origin testing capabilities, the challenges, limitations and evidence gaps and to discuss his views on these matters. The second consultation was towards the end of the project to discuss Dr Kelly's opinion on the outcomes of the project and to add insight into the practical next steps proposed by Fera Science to take forward for origin verification.

4.2 First consultation - current status of global origin testing capabilities, the challenges, limitations and evidence gaps and to discuss his views on these matters

Glossary

SIRA = stable isotope ratio analysis

TE = trace elements

FAO = Food and Agriculture Organisation of the United Nations

IAEA= International Atomic Energy Agency

PDO = protected designated origin

PGI = protected geographical indication

FT = Fourier transformed

Isoscape = a map representing the natural variation in stable isotope composition of a food

MID-IR = mid infrared spectroscopy

NIR = near infrared spectroscopy

NMR = nuclear magnetic resonance

MS = mass spectrometry

EUDR = European Deforestation Regulations

FIRMS = Forensic Isotope Ratio Mass Spectrometry

Q1. What is the current status of geographical origin testing?

According to Dr Simon Kelly, arguably the best hypothesis-driven approach ('gold standard') is to combine stable isotope ratio analysis (SIRA) and trace element (TE) data with chemometrics and/or isoscape mapping for determining country of origin of a food produce. These techniques are closely associated with foods of protected designated origin (PDO) and protected geographical indication (PGI). This is based on the understanding that isotopic signatures (light stable isotope ratios of e.g. hydrogen and oxygen in rain/ground water or heavy element isotope ratios of e.g. strontium and lead) and elemental fingerprints (based on soil chemistry and underlying geology) are linked to climate and geography, transferred to plants and animals from the environment in which they were cultivated or reared.

A number of full working systems for country-of-origin determination exist for food, e.g. EU wine databank, AHDB pork, Italian cheese consortia (Grana Padano, Parmigiano Reggiano), but also in the textile industry, for due diligence checking and monitoring the cotton origin to avoid countries with human rights issues related to production.

From an economical point of view, PDO/PGI schemes from Europe are accepted in China and vice versa; e.g. Jinxiang garlic (China, PGI). These schemes assist in promotion of foods (e.g. Blue mountain coffee - PDO, Jamaica; Thai jasmine (Hom Mali) rice - PGI, Thailand; Basmati rice – PGI, India and Pakistan) and their associated qualities. The eAmbrosia database holds the EU geographical indications register (europa.eu).

There are many complementary techniques to SIRA. These include spectroscopy techniques (e.g. FT-NIR, FT-MID-IR, Raman, GC- ion mobility and benchtop NMR), which benefit from reduced running costs, rapidity, and non-destruction of the sample, and also targeted and non-targeted analysis by chromatographic separation and MS detection.

Q2. Please describe the latest developments in geographical origin testing.

Multivariate analysis is progressing, using approaches such as (i) data fusion, e.g. combining genetic information with chemical data of foods, or hydrology-based isoscapes with food chemistry, and (ii) metabolomics (determination of unknowns). This goes hand in hand with the growing use of artificial Intelligence and machine learning for multivariate model building.

The reliability of rapid screening techniques such as infrared spectroscopy for geo-origin analysis is being demonstrated with increasing frequency.

Products of interest wanting to establish databases are Sri Lankan/Ceylon tea and Maynmar rice. For the latter MID-IR data from three harvests was successfully challenged with the 4th harvest. These promising results will be followed-up with monitoring the stability of the rice (consistency of spectra over time might be influenced by the changes to the protein composition).

Many of our food commodities originate from ‘developing countries.’ The International Atomic Energy Association (IAEA) is at the forefront of training these nations in various geographical authenticity methods, emphasising the importance of quality assurance and proficiency testing when building databases.

With new EUDR legislation [(EU) 2023/1115] in place for deforestation-free products, new databases are needed so governments and consumers can make informed choices, if they wish to avoid goods for which agricultural practices are linked to deforestation activities and impact on climate change, e.g. soya, coffee, cocoa, palm oil or farmed shrimp.’ World Forest ID’ (a US registered not-for profit organisation) aim to record data from 1000s of timber samples and forest risk commodities to combat illegal logging.

Q3. Are there any promising emerging methods in this field?

The use of ancillary SIRA data and chemical contaminant data is an interesting area to explore as an alternative or to supplement SIRA/TE data for country of origin determination.

Potentially studying existing environmental data (e.g. global network of isotopes in precipitation for ancillary data after demonstrating correlation between input water and food commodity of interest, OR DNA), pesticide residues, dioxins and PCBs (especially of interest for salmon), viral screens and fungi presence, not necessarily published in the context of geographical origin, might lead to valuable links to geographical origin information. There would be added benefit here as vast quantities of data exist which, once amassed and their metadata verified, may have the potential to inform on geographical origin.

Q4. Which are the factors of study design and database design we should place most and least emphasis on?

The following factors are equally important to allow for a robust database:

1. Validity of the data should be demonstrated by use of standardised methods (AOAC, CEN, CODEX) and accepted methods e.g. SIRA (exchangeable hydrogen correction in carbohydrates/proteins).

2. Database has been curated over multiple years and seasons and should stand up to non-database challenge samples, ideally with anonymised sample testing through independent auditors to assess reliability of the database in question.

3. Large datasets (or as a minimum stating the size of the database) so that interpretation and uncertainty of results can be evaluated, and evidence weighted.

4. Transferability from one lab to another, e.g by means of using certified reference materials, which is standard for SIRA/TE operation, but can be challenging for spectroscopic databases.

5. The samples used to build databases should ideally be representative of any permitted technological processing of the food to account for potential effects of processing on isotopic and chemical composition.

Q5. What are the most pertinent challenges, limitations and evidence gaps you are aware of at present in geographical origin testing?

Measuring of hydrogen isotopes in protein and carbohydrates, lack of progress and the reactive (rather than proactive) nature of development of standardised methods in general.

There is an array of barriers limiting the implementation of methods for country of origin determination:

a) cost of authenticity testing (including sample analysis, curating of database and need for verification samples) in addition to safety testing,

b) lack of existing databases for specific commodities, often premium products for example single origin tea, coffee, chocolate, etc.,

c) lack of reliable contacts to source authentic samples in foreign countries and need for time-consuming import licences,

d) challenge of obtaining industry assistance and support with assessing effect of permitted food industry technological processes on isotopical / chemical variation,

e) lack of food specific reference materials for SIRA and other techniques,

f) lack of proficiency testing schemes for country of origin, no standardisation of applying statistics to the data (e.g. R- freely available software package or other statistical software packages),

g) lack of standardised statistical analysis and statistical packages,

h) acceptance of methods by CEN/AOAC/Codex Alimentarius takes a long time,

i) remaining challenge to present multivariate data at court (single isotopes have been used in the past successfully for wine and German butter trade, see below for further detail),

j) disconnection of enforcement and evidence provider; e.g. corned beef (quick ^{13}C test to detect Brazilian beef in UK corned beef, no extensive database needed)

k) lack of data generated between laboratories, especially for spectroscopy methods. Need to gather the data over several years, and most of all

l) the intellectual property of the databases. **Ideally databases for food origin determination need to be open access and held by independent and trusted curators.** An identified gap are papers or guidance documents which illustrate how to get from ‘zero to a working system’ for authenticating a sample (FIRMS network is currently drafting one) and having a robust testing scheme in place. Such a document should cover four phases; a) sampling (including minimum of sample numbers needed), b) implementation (including year on year variation, production method, variation of processing technology, shelf life/storage), c) guidance on best statistical techniques to use e.g., when a laboratory has only access to UK local/national domestic samples for food origin analysis, then the use of Soft Independent Modelling by Class Analogy (SIMCA) to identify ‘alien’ (non-UK) food samples requires consideration, d) testing (in situ or specialist laboratory, how to test the samples) and e) routine (anonymised testing and proficiency testing schemes).

Q6. Other aspect of interest in context of country of origin determination

Recommending a two-tier system of techniques for determining the country of origin, using a Tier 1 rapid and ‘cheap’ screening method (e.g. IR), and then Tier 2 confirmatory (often more expensive) techniques such as SIRA plus other complementary techniques e.g. trace elements.

To note, suitable Tier 1 techniques have shown to be reliable for country of origin screening (for example rice), but need 100s and preferably 1000s of samples per year to monitor variation over time closely.

A pre-requisite of any testing laboratory performing tests to verify authenticity of samples, so that producers/traders can be taken to court, is that the laboratory is accredited to ISO17025, uses standardised or nationally recognised analytical methods and participates successfully in proficiency testing schemes. Whilst in a court case scenario historically only single isotope analysis data has been permitted for wine authentication (e.g. $\delta^{13}C$ for C_4-sugar or $\delta^{18}O$ for water addition), during the German circular butter trade investigation in 1998, isotope ratios of $^{18}O/^{16}O$ in the water fraction, $^{13}C/^{12}C$ in the bulk butter and $^{13}C/^{12}C$, $^{15}N/^{14}N$, $^{34}S/^{32}S$ and $^{87}Sr/^{86}Sr$ in the butter whey protein were permitted (Roßmann *et al.*, 2000)[30].

For the authentication of Grana Padano PDO cheese (Camin *et al.*, 2017)[31] 3-4 isotopes (H, C, N and occasionally S) and 13 most prominent elements: Sr, Cu, Re, Mo, U, Na, Bi, Ni, Mn, Fe, Se, Li and Ga are considered.

4.3 Second consultation -Discussion linked to the findings of the literature review

SIRA

Regarding SIRA technologies, the literature review highlighted that the following challenges for application of SIRA remain; costs to set up databases and to curate them is high, therefore the majority of them are not open access. There is still a lack of SIRA food -matrix reference materials, although the United States Geological Survey released 10 new materials in 2021 with assistance from IAEA. Proficiency tests have been available since 1995 focusing on data accuracy and not data interpretation (country of origin verification). The future will be in data fusing of various parameters which may have derived from different analytical techniques. A ‘2-tier’ approach of screening and confirmatory methods is envisaged, e.g. vibrational spectroscopy rapid screening with SIRA confirmation.

---

[30] Rossmann A., Haberhauer G., Hölzl S., Horn P., Pichlmayer F., Voerkelius S. (2000). The potential of multielement stable isotope analysis for regional origin assignment of butter. European Food Research Technology, 211, 32–40. https://doi.org/10.1007/s002170050585

[31] Camin, F., Boner, M., Bontempo, L., Fauhl-Hassek, C., Kelly, S. D., Riedl, J., & Rossmann, A. (2017). Stable isotope techniques for verifying the declared geographical origin of food in legal cases. Trends in Food Science & Technology, 61, 176-187. https://doi.org/10.1016/j.tifs.2016.12.007

In response to this, Dr Kelly added further points of consideration, for example the lack of opportunities to work with industry on country-of-origin verification which would ensure the capturing of natural variability of pure and processed foods to ensure any isotopic variability associated with permissible technological processes was understood. Also, the challenge of implementing the use of databases into a working system, monitoring supply chains, especially at points of highest vulnerability. It can question the feasibility/robustness of a database, when mixtures of edible oils are concerned, e.g. when mixtures of olives oils from more than one EU country of origin are declared on the label.

Dr Kelly emphasised the uniqueness of the EU wine databank, which started off for general authenticity purposes (detecting adulterations with sugar or water) using SIRA and SNIF-NMR. The $(D/H)_I$ ratio of the methyl group in the ethanol of a wine is linked to botanical origin (C3 plant), while the $(D/H)_{II}$ ratio of the methylene group gives indication of the fermentation water and therefore the geographical origin of a wine. This latter data, as well as the stable oxygen isotope ratio of wine water, is being used to verify origin. Isoscapes are maps representing the natural variation in stable isotope composition and have their value when no food sampling has happened in a particular region/location yet, but the relationship between the isotope composition of the food and input water is characterised and correlated. In this situation, for example, ground water hydrogen and oxygen isotope values at a particular location can be used to infer the isotope composition of food cultivated at the same location. There are significant amounts of water isotope data collected across the world by the IAEA’s Global Network of Isotopes in Precipitation (GNIP) and this may be used to underpin isoscape maps for food.

Metals = trace elements (TE)

In relation to the areas covered by trace element detection techniques, Dr Kelly pointed out that, while it is very common to find TE data published in conjunction with SIRA data, one needs to consider the relationship of geo-chemicals with the soil, and particularly the metal transfer factors, the pH of the soil, and fertiliser treatments, which can influence the stability of a TE signals in an origin control system. In relation to working models for geographical origin verification, working database systems for which TE analyses are currently used include commodities such as cotton and PDO

Italian cheeses (Grana Padano and Parmigiano Reggiano). These implemented systems are routinely challenged with anonymised authentic and controlled “alien” samples to monitor their robustness.

There are a number of ‘key’ elements, when considering the geographical origin of food, such as rubidium, strontium and calcium which, are associated with underlying geology (rock formation, bottom of soil layers). Other elements are more linked to soil chemistry, occurring in the top layer of soil. For statistical evaluation of TE data in general, one would acquire data from macro, micro and trace TEs and by using, for example, stepwise canonical discrimination narrow down the parameters needed. The identified elements can then be linked to soil chemistry maps, when available, to supplement the understanding of the variability in TE composition of foods.

Metabolomics

Having access to meta data such as permittable additives, clarifying agents, yeasts and knowledge of permissible production processes is very valuable information, when performing targeted or non-targeted profiling of food. Taking the example of the EU wine databank, one needs to know the different approaches for populating the databank with authentic micro-vinified wine samples versus mass produced retail samples using steel fermentation vessels and presence of flavour enhancing additives, such as wood chips. In order to be able to provide intrinsic solutions for geographical origin verification via chemical profiles, it is essential to monitor annual and seasonal variations and regularly challenge the built models with anonymised authentic and non-authentic control samples.

Other techniques

Blockchain and electronic tagging, and the combination of these techniques are relevant for growing number of food traceability applications. However, they do rely on full and honest information inputs, meaning the blockchain will only be as strong as its weakest link.

The various parties along the supply chain would appreciate the potential for geographical origin testing laboratories to be able to measure 100 samples from a region within a week. But to achieve this scenario the testing laboratory would need to know which methods were used by others to benefit from database sharing in an international system. There is also a plethora of chemical contaminant data that is gathered through routine testing of food around the world. Using multivariate statistics, and potentially Artificial Intelligence, it may be possible to take advantage of this data to help characterise the geographical origin of food. This approach of modelling patterns in official food control data has recently been applied to dioxin and PCB data for authentication of Baltic Sea salmon.

IAEA involvement

The FAO/IAEA Joint Centre is supporting 17 low- and middle-income countries to develop their food fraud control systems by building capacity in food authenticity and geographical origin analysis through its Coordinated Research Project "INTACT FOOD". This is done to both protect and promote many premium foods that are exported from these IAEA member States, for example, Jamaican Blue Mountain and Costa Rican coffee, Thai Jasmine rice, Chinese Jinxiang Garlic (PGI), and Moroccan Taliouine saffron. The IAEA, with the help of institutes from developed countries, supplies advice for geographical origin database initiatives, with sampling strategies, analytical protocols, and multivariate data evaluation software. The IAEA Technical Cooperation programme also supplies SIRA, and other complementary testing, equipment to low- and middle-income countries, along with training and expert missions with the aim of establishing reliable food authenticity and geographical origin control systems.

Outlook

CEN efforts to standardise methods in food authenticity used for geographical origin verification are progressing slowly. Dr Kelly highlighted the opportunity to link up all the progressive work of many competent authorities in countries working on food origin databases initiatives, for example, NIST (USA) and NARO (Japan), as part of a new UK government led initiative or funded project. There would be further opportunity to link up country of origin verification efforts throughout the UK's Commonwealth. Furthermore, previous UK geographical origin projects such as those on beef, wine, and salmon could be revisited and compiled into a significant single database resource. All of these activities would require a consistent long-term commitment from the FSA and/or Defra to underpin geographical origin analysis of food like the setup of the National Reference Centre for Authentic food (NRZ-Authent) in Germany.

4.4 World Forest ID - Example of a working model in geographical origin identification

Background

World Forest ID is a US registered not-for-profit international organisation which verifies and determines the global geographical origin of timber and Forest Risk Commodities. Timber and Forest Risk Commodities must be marked according to their geographical origin. EU rules require that the products EU citizens consume do not contribute to deforestation or forest degradation worldwide. Reasons to verify timber origin include enforcement, environmental regulations including protection of forests from clearance to grow crops such as cacao, soy, coffee and palm oil, and to verify sanctions are being adhered. The method used Gaussian Process modelling of chemical profiles against latitude and longitude. The methods used by World Forest ID can be applied to verify origin if a certain origin is suspected, or to predict origin if there is little or no additional intelligence relating to its location.

Verification is performed by scientific analysis as paper trail methods have been shown to be inefficient, paper-based and open to fraud. World Forest ID's initial consortium, before becoming a not-for-profit, comprised US Forest Service International Programs, Royal Botanical Gardens Kew, World Resources Institute, AgroIsoLab, the Forest Stewardship Council. The organisation was developed since much international work is undertaken in timber origin identification and it was recognised that a centralised hub was required to incentivise, facilitate and manage the sharing of samples, data and protocols, hence the development of this not-for-profit organisation World Forest ID focusses on the harvest location verification and determination of both timber and Forest Risk Commodities (especially with the adopted EUDR).

Testing

World Forest ID has contributed to databases focussing on timber species identification, such as wood anatomy and chemical composition by DART-TOFMS which can provide indications of species. DNA analysis is currently not actively pursued due to the destruction of DNA within processed wood. However, the main focus of World Timber ID is to verify or determine geographical origin of timber and forest risk commodities. Large amounts of pre-existing data have been pulled together and much further data has been generated through the analysis of timber samples, mostly analysing by SIRA, X-ray fluorescence (to determine elemental composition) and ICP-MS trace element analysis, to build more databases to verify or determine geographical origin. SIRA data can be used to predict geographical on a larger spatial scale, especially when combining data from several isotopes whereas trace element profiles can change over much smaller areas depending on soil variability. The resolution to determine the harvest location becomes much finer when both SIRA and trace element data are combined, and this is the preferred way that World Forest ID collect their data. Since the composition of trace elements and stable isotopes will often change with season for temperate species, data are collected over a number of seasons.

These data are then incorporated into a machine learning model, which allows to predict over large spatial scales, moving away from the limited classification models which are currently used by other actors. To date, samples from 290 species and more than 35 counties have been collected, and World Forest ID have partnered with labs, governments, non-government organisations and universities in localities across the globe to recruit and train specialist collectors. SIRA and ICP-MS are relatively expensive testing methods to employ but, unlike for initial testing of some foods (applicable as a first step prior to confirmatory testing by SIRA and/or ICP-MS), lower cost spectroscopy methods are not powerful enough for robust timber geographical origin verification or determination. Each project undertaken by World Forest ID includes a replicability study, taking a sample set of 5-10% of the collected samples, to confirm whether multiple labs can get the same results.

To a lesser extent, other approaches are being explored by World Forest ID, including a pilot genomic identification project for fungal, viral and bacterial endophyte contamination of trees. This method is especially applicable in the tropics where these microbes are more prevalent on timber and their presence is variable spatially, but stable isotope and trace element profile determination is applicable on a global scale.

Regarding the testing of food commodities, World Forest ID has recently started a trial on whether palm oil origin can be determined using chemical measurements. All results are at the preliminary stage.

Success stories

Data from the library of references that World Forest ID is building can be accessed by companies and prosecutors around the world. Large datasets have been captured, spanning large areas and across a variety of seasons, and the datasets continue to grow with time. World Forest ID encourages the sharing of methods between global laboratories to verify the origin of timber products. World Forest ID also implement inter-laboratory trials to verify that a wide range of laboratories around the world can perform the same methods to the same standards and with high level of reproducibility. With high levels of reproducibility, the data can stand up to scrutiny in a court of law. World Forest ID data have been used to settle origin cases and the vast majority of cases are settled outside of court.

Challenges for timber origin verification

Regarding challenges and gaps in knowledge, World Forest ID state that challenges for timber origin identification lie in the fact that such a wide range of timber products exist and that more research is needed as to whether the chemical signal in those products is similar to the chemical signal of reference samples. There are also a great number of species to consider which can add variation to datasets. Also, despite World Forest ID investing to develop large databases, a full dataset will never be complete for any one commodity. For example, although oak grows across much of the globe, a full dataset will be almost impossible to capture. However, World Forest ID targets key countries that are known to trade a specific species and where illegal laundering is possible. Not every country where a species grows, has a risk of illegal laundering. When a timber product is traded, the year of harvest (or the age of that particular piece of timber/location within rings of the trunk) is not known. World Forest ID state that there are complications when assessing origin in samples containing a mixture of origins, for example in a particleboard comprising woods from a range of origins.

Perceived challenges for food analysis compared to timber analysis

Some of the challenges faced during the identification of timber origin differ from those applicable to food. Timber is a spatially continuous commodity, often spanning countries and borders. Unlike food, timber does not move when growing (as food can, moving for example at times from farm to farm) and also is not grown within the boundaries of separate farms or ponds. Unlike timber, two similar foods that have been grown or raised in close geographical range can have different feeding regimes, selected by individual farmers, which will complicate trace element data. These differences present challenges to the determination of geographical origin for food compared to for timber. Given the complications described above when determining geographical origin of timber for particleboard (comprising mixtures of different timber sources) it is likely that similar complications would be observed if these methods were applied to adulterated food, containing mixtures of food sources.

Conclusions

The work of World Forest ID demonstrates a working model of verifying and determining geographical origin of a group of commodities. Although dealing with different commodities, there are many areas described above against which any future models developed for the determination of food origin could be benchmarked.

World Forest ID have created large datasets, with sample sets spanning the globe. For food, many studies have been completed to develop methods to determine the geographical origin of certain food types. As demonstrated in the literature review, these studies are usually restricted with respect to the time span of the data collection. Samples are collected across one season, or a small number of seasons and for certain commodities only and tend to consider only a limited geographical area, thus discounting other areas of the globe which may yield similar data, potentially leading to false positive identifications. In addition, these studies usually focus on using classification models (“Country X vs Country Y”), as such ignoring any potential spatial variability. Once a project in food origin analysis has been completed, funding is usually no longer available and the dataset is rarely added to at future dates. Therefore, changes in other environmental conditions over time are not accounted for, which leads to a reduced confidence in the data and the potential for false positives increases. With continuous funding, World Forest ID take the approach of continuing to collect samples as time progresses to continue to expand the dataset. World Forest ID perform reproducibility studies (known to World Forest ID as replicability studies) to verify that multiple labs achieve the same result for the same sample, and also test blind samples. This provides increased confidence in the data.

In general, as demonstrated in the literature review, where projects have been undertaken for a particular food commodity, very little publicly available data exists. Most data are stored in proprietary databases, often held by companies who offer commercial origin testing services. This reduces the opportunity for collaboration and for inter-laboratory trials to determine the reproducibility of the data. This lack of collaborative data tends to result in the robustness of data being questioned and data generated by methods used in a single laboratory would be unlikely to be considered in a court proceeding. The result is poor enforcement support which can result in a reduced deterrent against food origin fraud and therefore we recommend that the approaches undertaken to challenge methods in inter-laboratory trial in timber origin identification are also undertaken for methods used in food analysis to improve the integrity of methods to verify the origin of food.

References

Acierno, V., Alewijn, M., Zomer, P., & van Ruth, S. M. (2018). Making cocoa origin traceable: Fingerprints of chocolates using Flow Infusion - Electro Spray Ionization - Mass Spectrometry. *Food Control*, *85*, 245-252. https://doi.org/10.1016/j.foodcont.2017.10.002

Adam, T., Duthie, E., & Feldmann, J. (2002). Investigations into the Use of Copper and Other Metals as Indicators for the Authenticity of Scotch Whiskies. *Journal of the Institute of Brewing*, *108*(4), 459-464. https://doi.org/10.1002/j.2050-0416.2002.tb00576.x

Afzaal, M., Saeed, F., Hussain, M., Shahid, F., Siddeeg, A., & Al-Farga, A. (2022). Proteomics as a promising biomarker in food authentication, quality and safety: A review. *Food Science & Nutrition*, *10*(7), 2333-2346. https://doi.org/10.1002/fsn3.2842

Ahn, S. J., Lee, A., Min, S. S., In, S., Kim, E., & Kim, H. J. (2019). Comparison of Physicochemical Characteristics of Garlic Produced from South Korea and China. *Journal of Food Science*, *84*(7), 1806-1811. https://doi.org/10.1111/1750-3841.14684

Anyidoho, E. K., Teye, E., & Agbemafle, R. (2020). Nondestructive authentication of the regional and geographical origin of cocoa beans by using a handheld NIR spectrometer and multivariate algorithm. *Analytical Methods*, *12*(33), 4150-4158. https://doi.org/10.1039/d0ay00901f

Arana, V. A., Medina, J., Esseiva, P., Pazos, D., & Wist, J. (2016). Classification of Coffee Beans by GC-C-IRMS, GC-MS, and ¹H-NMR. *Journal of Analytical Methods in Chemistry*, *2016*, Article 8564584. https://doi.org/10.1155/2016/8564584

Arif, M., Chilvers, G., Day, S., Naveed, S. A., Woolfe, M., Rodionova, O. Y., Pomerantsev, A. L., Kracht, O., Brodie, C., Mihailova, A., Abrahim, A., Cannavan, A., & Kelly, S. D. (2021). Differentiating Pakistani long-grain rice grown inside and outside the accepted Basmati Himalayan geographical region using a 'one-class' multi-element chemometric model. *Food Control*, *123*, Article 107827. https://doi.org/10.1016/j.foodcont.2020.107827

Artêncio, M. M., Cassago, A. L. L., da Silva, R. K., Giraldi, J. D. E., & da Costa, F. B. (2023). Untargeted Metabolomic Approach Based on UHPL-ESI-HRMS to Investigate Metabolic Profiles of Different <i>Coffea</i> Species and Terroir. *Revista Brasileira De Farmacognosia-Brazilian Journal of Pharmacognosy*. https://doi.org/10.1007/s43450-023-00454-y

Asfaha, D.G., Quetel, C.R.;Thomas, F., Horacek, M., Wimmer, B., Heiss, G., Dekant, C., Deters-Itzelsberger, P.,Hoelzl, S., Rummel, S., Brach-Papa, C., Van Bocxstaele, M., Jamin, E., Baxter, M.,Heinrich,K., Kelly, S., Bertoldi, D., Bontempo, L., Camin,F., Larcher, R., Perini, M., Rossmann, A., Schellenberg, A., Schlicht, CC., Froeschl, H ., Hoogewerff, J., Ueckermann, H., (2011) Combining isotopic signatures of n((87)Sr)/n((86)Sr) and light stable elements (C, N, O, S) with multi-elemental profiling for the authentication of provenance of European cereal samples. *Journal of Cereal Science*, 53, 170-177. https://doi.org/10.1016/j.jcs.2010.11.004

Avramidou, E. V., Doulis, A. G., & Petrakis, P. V. (2018). Chemometrical and molecular methods in olive oil analysis: A review. *Journal of Food Processing and Preservation*, *42*(11), Article e13770. https://doi.org/10.1111/jfpp.13770

Bai, S., Qin, D., Chen, Z., Wu, S., Tang, S., Gao, L., & Wang, P. (2022). Geographic origin discrimination of red swamp crayfish Procambarus clarkii from different Chinese regions using mineral element analysis assisted by machine learning techniques. *Food Control*, *138*, 109047. https://doi.org/10.1016/j.foodcont.2022.109047

Bai, Y., Liu, H. J., Zhang, B., Zhang, J. K., Wu, H., Zhao, S. S., Qie, M. J., Guo, J., Wang, Q., & Zhao, Y. (2023). Research Progress on Traceability and Authenticity of Beef. *Food Reviews International*, *39*(3), 1645-1665. https://doi.org/10.1080/87559129.2021.1936000

Bajoub, A., Bendini, A., Fernández-Gutiérrez, A., & Carrasco-Pancorbo, A. (2018). Olive oil authentication: A comparative analysis of regulatory frameworks with especial emphasis on quality and authenticity indices, and recent analytical techniques developed for their assessment. A review. *Critical Reviews in Food Science and Nutrition*, *58*(5), 832-857. https://doi.org/10.1080/10408398.2016.1225666

Baqueta, M. R., Valderrama, P., Mandrone, M., Poli, F., Coqueiro, A., Costa-Santos, A. C., Rebellato, A. P., Luz, G. M., Rocha, R. B., Pallone, J. A. L., & Marini, F. (2023). 1H NMR, FAAS, portable NIR, benchtop NIR, and ATR-FTIR-MIR spectroscopies for characterizing and discriminating new Brazilian Canephora coffees in a multi-block analysis perspective. *Chemometrics and Intelligent Laboratory Systems*, *240*, Article 104907. https://doi.org/10.1016/j.chemolab.2023.104907

Baqueta, M. R., Valderrama, P., Mandrone, M., Poli, F., Coqueiro, A., Costa-Santos, A. C., Rebellato, A. P., Luz, G. M., Rocha, R. B., Pallone, J. A. L., & Marini, F. (2023). 1H NMR, FAAS, portable NIR, benchtop NIR, and ATR-FTIR-MIR spectroscopies for characterizing and discriminating new Brazilian Canephora coffees in a multi-block analysis perspective. *Chemometrics and Intelligent Laboratory Systems*, *240*, Article 104907. https://doi.org/10.1016/j.chemolab.2023.104907

Beltrán, M., Sánchez-Astudillo, M., Aparicio, R., & García-González, D. L. (2015). Geographical traceability of virgin olive oils from south-western Spain by their multi-elemental composition. *Food Chemistry*, *169*, 350-357. https://doi.org/10.1016/j.foodchem.2014.07.104

Berg, P. R., Jentoft, S., Star, B., Ring, K. H., Knutsen, H., Lien, S., Jakobsen, K. S., & André, C. (2015). Adaptation to Low Salinity Promotes Genomic Divergence in Atlantic Cod (Gadus morhua L.). *Genome Biology and Evolution*, *7*(6), 1644-1663. https://doi.org/10.1093/gbe/evv093

Berriel, V., Barreto, P., & Perdomo, C. (2019). Characterisation of Uruguayan Honeys by Multi-Elemental Analyses as a Basis to Assess Their Geographical Origin. *Foods*, *8*(1), Article 24. https://doi.org/10.3390/foods8010024

Bertoldi, D., Barbero, A., Camin, F., Caligiani, A., & Larcher, R. (2016). Multielemental fingerprinting and geographic traceability of <i>Theobroma cacao</i> beans and cocoa products. *Food Control*, *65*, 46-53. https://doi.org/10.1016/j.foodcont.2016.01.013

Bhagat, N., Sharma, S., Ambardar, S., Raj, S., Trakroo, D., Horacek, M., Zouagui, R., Sbabou, L., & Vakhlu, J. (2021). Microbiome Fingerprint as Biomarker for Geographical Origin and Heredity in <i>Crocus sativus</i>: A Feasibility Study. *Frontiers in Sustainable Food Systems*, *5*, Article 688393. https://doi.org/10.3389/fsufs.2021.688393

Biswas, A., Naresh, K. S., Jaygadkar, S. S., & Chaudhari, S. R. (2023). Enabling honey quality and authenticity with NMR and LC-IRMS based platform. *Food Chemistry*, *416*, Article 135825. https://doi.org/10.1016/j.foodchem.2023.135825

Bontempo, L., Perini, M., Pianezze, S., Horacek, M., Roßmann, Kelly, S., Thomas, F. Heinrich, K., Schlicht, C., Schellenberg, A., Hoogewerff, J., Heiss, G., Wimmer, B. and Camin, F. (2023). Characterisation of beef coming from different European Countries through stable isotope (H, C, N and S) ratio analysis. *Molecules*, *28*(6). https://doi.org/10.3390/molecules28062856

Brbaklic, L., Trkulja, D., Kondic-Spika, A., Mikic, S., Tomicic, M., & Kobiljski, B. (2015). Determination of Population Structure of Wheat Core Collection for Association Mapping. *Cereal Research Communications*, *43*(1), 22-28. https://doi.org/10.1556/crc.2014.0027

Bronzi, B., Brilli, C., Beone, G. M., Fontanella, M. C., Ballabio, D., Todeschini, R., Consonni, V., Grisoni, F., Parri, F., & Buscema, M. (2020). Geographical identification of Chianti red wine based on ICP-MS element composition. *Food Chemistry*, *315*, Article 126248. https://doi.org/10.1016/j.foodchem.2020.126248

Bui, M. Q., Quan, T. C., Nguyen, Q. T., Tran-Lam, T. T., & Dao, Y. H. (2022). Geographical origin traceability of Sengcu rice using elemental markers and multivariate analysis. *Food Additives & Contaminants Part B-Surveillance*, *15*(3), 177-190. https://doi.org/10.1080/19393210.2022.2070932

Busconi, M., Colli, L., Sánchez, R. A., Santaella, M., Pascual, M. D., Santana, O., Roldán, M., & Fernández, J. A. (2015). AFLP and MS-AFLP Analysis of the Variation within Saffron Crocus (<i>Crocus sativus</i> L.) Germplasm. *Plos One*, *10*(4), Article e0123434. https://doi.org/10.1371/journal.pone.0123434

Camargo, A. B., Resnizky, S., Marchevsky, E. J., & Luco, J. M. (2010). Use of the Argentinean garlic (Allium sativum L.) germplasm mineral profile for determining geographic origin. *Journal of Food Composition and Analysis*, *23*(6), 586-591. https://doi.org/10.1016/j.jfca.2010.01.002

Cambrai, A., Marchioni, E., Julien-David, D., & Marcic, C. (2017). Discrimination of Cocoa Bean Origin by Chocolate Polyphenol Chromatographic Analysis and Chemometrics. *Food Analytical Methods*, *10*(6), 1991-2000. https://doi.org/10.1007/s12161-016-0744-7

Camin, F., Bontempo, L., Heinrich, K., Horacek, M., Kelly, S.D., Schlicht, C., Thomas, F., Monahan, F.J., Hoogewerff, J., Rossmann, A. (2007) Multi-element (H,C,N,S) stable isotope characteristics of lamb meat from different European regions. Analytical and Bioanalytical Chemistry, 389, 309-320. https://doi.org/10.1007/s00216-007-1302-3

Camin,F., Larcher, R., Nicolini,G., Bontempo,L., Bertoldi,D., Perini,M., Schlicht,C., Schellenberg,A., Thomas,F., Heinrich,K., Voerkelius,S., Horacek,M., Ueckermann,H., Froeschl,H., Wimmer,B., Heiss,G., Baxter,M., Rossmann,A., Hoogewerff,J. (2010) Isotopic and Elemental Data for Tracing the Origin of European Olive Oils. *Journal of Agricultural and Food Chemistry, 58*, 570-577 https://doi.org/10.1021/jf902814s

Camin, F., Bertoldi, D., Santat, A. Bontempo, L., Perini, M., Ziller, L., Stroppa, A. and Larcher, R. (2015). Validation of methods for H, C, N and S stable isotopes and elemental analysis of cheese: results of an international collaborative study. *Rapid Communications in Mass Spectrometry*, *29*(5), 415-423. https://doi.org/10.1002/rcm.7117

Carneiro, C. N., Gomez, F. J. V., Spisso, A., Silva, M. F., Santos, J. L. O., & Dias, F. D. (2023). Exploratory Analysis of South American Wines Using Artificial Intelligence. *Biological Trace Element Research*, *201*(9), 4590-4599. https://doi.org/10.1007/s12011-022-03529-4

Cassago, A. L. L., Artêncio, M. M., Giraldi, J. D. E., & Da Costa, F. B. (2021). Metabolomics as a marketing tool for geographical indication products: a literature review. *European Food Research and Technology*, *247*(9), 2143-2159. https://doi.org/10.1007/s00217-021-03782-2

Cellier, R., Berail, S., Barre, J., Epova, E., Claverie, F., Ronzani, A. L., Milcent, S., Ors, P., & Donard, O. F. X. (2021). Analytical strategies for Sr and Pb isotopic signatures by MC-ICP-MS applied to the authentication of Champagne and other sparkling wines. *Talanta*, *234*, Article 122433. https://doi.org/10.1016/j.talanta.2021.122433

Ch, R., Chevallier, O., McCarron, P., McGrath, T. F., Wu, D., Le, N. D. D., Kapil, A. P., McBride, M., & Elliott, C. T. (2021). Metabolomic fingerprinting of volatile organic compounds for the geographical discrimination of rice samples from China, Vietnam and India. *Food Chemistry*, *334*, Article 127553. https://doi.org/10.1016/j.foodchem.2020.127553

Chagné, D., Montanari, S., Kirk, C., Mitchell, C., Heenan, P., & Koot, E. (2023). Single nucleotide polymorphism analysis in <i>Leptospermum scoparium</i> (Myrtaceae) supports two highly differentiated endemic species in Aotearoa New Zealand and Australia. *Tree Genetics & Genomes*, *19*(4), Article 31. https://doi.org/10.1007/s11295-023-01606-w

Chandra, S., Chapman, J., Power, A., Roberts, J., & Cozzolino, D. (2017). Origin and Regionality of Wines-the Role of Molecular Spectroscopy. *Food Analytical Methods*, *10*(12), 3947-3955. https://doi.org/10.1007/s12161-017-0968-1

Chapman, J., Gangadoo, S., Truong, V. K., & Cozzolino, D. (2019). Spectroscopic approaches for rapid beer and wine analysis. *Current Opinion in Food Science*, *28*, 67-73. https://doi.org/10.1016/j.cofs.2019.09.001

Charoensumran, P., Rauytanapanit, M., Sricharoen, N., Smith, B. L., Wongravee, K., Maher, S., & Praneenararat, T. (2021). Rapid geographical indication of peppercorn seeds using corona discharge mass spectrometry. *Scientific Reports*, *11*(1), Article 16089. https://doi.org/10.1038/s41598-021-95462-0

Charoensumran, P., Rauytanapanit, M., Sricharoen, N., Smith, B. L., Wongravee, K., Maher, S., & Praneenararat, T. (2021). Rapid geographical indication of peppercorn seeds using corona discharge mass spectrometry. *Scientific Reports*, *11*(1), Article 16089. https://doi.org/10.1038/s41598-021-95462-0

Cheajesadagul, P., Arnaudguilhem, C., Shiowatana, J., Siripinyanond, A., & Szpunar, J. (2013). Discrimination of geographical origin of rice based on multi-element fingerprinting by high resolution inductively coupled plasma mass spectrometry. *Food Chemistry*, *141*(4), 3504-3509. https://doi.org/10.1016/j.foodchem.2013.06.060

Christoph, N., Hermann, A., & Wachter, H. (2015). 25 Years authentication of wine with stable isotope analysis in the European Union - Review and outlook.*BIO Web of Conferences* [38th world congress of vine and wine (part 1)]. 38th World Congress of Vine and Wine, Mainz, GERMANY. http://dx.doi.org/10.1051/bioconf/20150502020

Chytry, P., Souza, G. M. S., Debastiani, R., dos Santos, C. E. I., Antoine, J. M. R., Banas, A., Banas, K., Calcagnile, L., Chiari, M., Hajdas, I., Molnar, M., Pelicon, P., Barradas, N. P., Quarta, G., Romolo, F. S., Simon, A., & Dias, J. F. (2022). The potential of accelerator-based techniques as an analytical tool for forensics: The case of coffee. *Forensic Science International*, *335*, Article 111281. https://doi.org/10.1016/j.forsciint.2022.111281

Chytry, P., Souza, G. M. S., Debastiani, R., dos Santos, C. E. I., Antoine, J. M. R., Banas, A., Banas, K., Calcagnile, L., Chiari, M., Hajdas, I., Molnar, M., Pelicon, P., Barradas, N. P., Quarta, G., Romolo, F. S., Simon, A., & Dias, J. F. (2022). The potential of accelerator-based techniques as an analytical tool for forensics: The case of coffee. *Forensic Science International*, *335*, Article 111281. https://doi.org/10.1016/j.forsciint.2022.111281

Conte, L., Bendini, A., Valli, E., Lucci, P., Moret, S., Maquet, A., Lacoste, F., Brereton, P., García-González, D. L., Moreda, W., & Toschi, T. G. (2020). Olive oil quality and authenticity: A review of current EU legislation, standards, relevant methods of analyses, their drawbacks and recommendations for the future. *Trends in Food Science & Technology*, *105*, 483-493. https://doi.org/10.1016/j.tifs.2019.02.025

Contreras, R., Calle, I., Oses, R., Aguayo, F., Porcile, V., & Arias, M. (2021). Genetic identification of the Putre's oregano (<i>Origanum vulgare</i> L.) by ITS and microsatellites, a species recognized in Chile with Seal of Origin. *Boletin Latinoamericano Y Del Caribe De Plantas Medicinales Y Aromaticas*, *20*(2), 177-194. https://doi.org/10.37360/blacpma.21.20.2.14

Danieli, P. P., & Lazzari, F. (2022). Honey traceability and authenticity. Review of current methods most used to face this problem. *Journal of Apicultural Science*, *66*(2), 101-119. https://doi.org/10.2478/jas-2022-0012

D'Archivio, A. A., Foschi, M., Aloia, R., Maggi, M. A., Rossi, L., & Ruggieri, F. (2019). Geographical discrimination of red garlic (<i>Allium sativum</i> L.) produced in Italy by means of multivariate statistical analysis of ICP-OES data. *Food Chemistry*, *275*, 333-338. https://doi.org/10.1016/j.foodchem.2018.09.088

Dong, R. A., Sun, J. H., Cui, P. Y., Shi, J. Q., Li, Y. K., & Ma, X. Y. (2023). Quantitative and classification analysis of red wine by infrared spectra and gas chromatography-mass spectrometry data coupling with a new variable selection method. *Journal of Food Composition and Analysis*, *120*, Article 105325. https://doi.org/10.1016/j.jfca.2023.105325

Dou, X., Zhang, L., Yang, R., Wang, X., Yu, L., Yue, X., Ma, F., Mao, J., Wang, X., Zhang, W., & Li, P. (2023). Mass spectrometry in food authentication and origin traceability. *Mass Spectrom Rev*, *42*(5), 1772-1807. https://doi.org/10.1002/mas.21779

Driscoll, A. W., Howa, J. D., Bitter, N. Q., & Ehleringer, J. R. (2020). A predictive spatial model for roasted coffee using oxygen isotopes of α-cellulose. *Rapid Communications in Mass Spectrometry*, *34*(7), Article e8626. https://doi.org/10.1002/rcm.8626

El Hani, O., García-Guzmán, J. J., Palacios-Santander, J. M., Digua, K., Amine, A., Gharby, S., & Cubillana-Aguilera, L. (2023). Geographical Classification of Saffron (<i>Crocus Sativus</i> L.) Using Total and Synchronous Fluorescence Combined with Chemometric Approaches. *Foods*, *12*(9), Article 1747. https://doi.org/10.3390/foods12091747

El Hani, O., García-Guzmán, J. J., Palacios-Santander, J. M., Digua, K., Amine, A., Gharby, S., & Cubillana-Aguilera, L. (2023). Geographical Classification of Saffron (<i>Crocus Sativus</i> L.) Using Total and Synchronous Fluorescence Combined with Chemometric Approaches. *Foods*, *12*(9), Article 1747. https://doi.org/10.3390/foods12091747

Epova, E. N., Bérail, S., Sèby, F., Vacchina, V., Bareille, G., Médina, B., Sarthou, L., & Donard, O. F. X. (2019). Strontium elemental and isotopic signatures of Bordeaux wines for authenticity and geographical origin assessment. *Food Chemistry*, *294*, 35-45. https://doi.org/10.1016/j.foodchem.2019.04.068

Fallah, A. A., Zeynali, F., Saei-Dehkordi, S. S., Rahnama, M., & Jafari, T. (2011). Seasonal bioaccumulation of toxic trace elements in economically important fish species from the Caspian Sea using GFAAS. *Journal für Verbraucherschutz und Lebensmittelsicherheit*, *6*(3), 367-374. https://doi.org/10.1007/s00003-011-0666-7

Fanning, E., Eyres, G., Frew, R., & Kebede, B. (2023). Linking cocoa quality attributes to its origin using geographical indications. *Food Control*, *151*, Article 109825. https://doi.org/10.1016/j.foodcont.2023.109825

Fanning, E., Eyres, G., Frew, R., & Kebede, B. (2023). Linking cocoa quality attributes to its origin using geographical indications. *Food Control*, *151*, 109825. https://doi.org/https://doi.org/10.1016/j.foodcont.2023.109825

Fechner, D. C., Silva, T. V., Nespeca, M. G., Vieira, A. L., Gomes, J. A., Santos, D., Pellerano, R. G., & Ferreira, E. C. (2021). Protected Geographical Identification of Honey by Spark Discharge-assisted Laser-induced Breakdown Spectroscopy. *Atomic Spectroscopy*, *42*(3), 148-153. https://doi.org/10.46770/as.2021.022

Fernandes, E. A. D., Sarriés, G. A., Bacchi, M. A., Mazola, Y. T., Gonzaga, C. L., & Sarriés, S. R. V. (2020). Trace elements and machine learning for Brazilian beef traceability. *Food Chemistry*, *333*, Article 127462. https://doi.org/10.1016/j.foodchem.2020.127462

Fernandes, J. R., Pereira, L., Jorge, P., Moreira, L., Gonçalves, H., Coelho, L., Alexandre, D., Eiras-Dias, J., Brazao, J., Clímaco, P., Baleiras-Couto, M., Catarino, S., Graça, A., & Martins-Lopes, P. (2015, Jul 05-10). Wine fingerprinting using a bio-geochemical approach. *BIO Web of Conferences* [38th world congress of vine and wine (part 1)]. 38th World Congress of Vine and Wine, Mainz, GERMANY. http://dx.doi.org/10.1051/bioconf/20150502021

Fontanesi, L. (2009). Genetic authentication and traceability of food products of animal origin: new developments and perspectives. *Italian Journal of Animal Science*, *8*, 9-18. https://doi.org/doi.org/10.4081/ijas.2009.s2.9

Fragni, R., Trifirò, A., Nucci, A., Seno, A., Allodi, A., & Di Rocco, M. (2018). Italian tomato-based products authentication by multi-element approach: A mineral elements database to distinguish the domestic provenance. *Food Control*, *93*, 211-218. https://doi.org/10.1016/j.foodcont.2018.06.002

Franciosa, I., Ferrocino, I., Giordano, M., Mounier, J., Rantsiou, K., & Cocolin, L. (2021). Specific metagenomic asset drives the spontaneous fermentation of Italian sausages. *Food Research International*, *144*, Article 110379. https://doi.org/10.1016/j.foodres.2021.110379

Fu, X. S., Hong, X. Z., Liao, J. Y., Ji, Q. G., Li, C. F., Zhang, M. Z., Ye, Z. H., & Yu, X. P. (2021). Fingerprint Approaches Coupled with Chemometrics to Discriminate Geographic Origin of Imported Salmon in China's Consumer Market. *Foods*, *10*(12), Article 2986. https://doi.org/10.3390/foods10122986

Gajek, M., Pawlaczyk, A., Józwik, K., & Szynkowska-Józwik, M. I. (2022). The Elemental Fingerprints of Different Types of Whisky as Determined by ICP-OES and ICP-MS

Techniques in Relation to Their Type, Age, and Origin. *Foods*, *11*(11), Article 1616. https://doi.org/10.3390/foods11111616

Gazeli, O., Bellou, E., Stefas, D and Couris, S. (2020). Laser-based classification of olive oils assisted by machine learning. Food Chemistry, *302*, Article 125329. https://doi.org/10.1016/j.foodchem.2019.125329

Gazzola, P., Pavione, E., Barge, A., & Fassio, F. (2023). Using the Transparency of Supply Chain Powered by Blockchain to Improve Sustainability Relationships with Stakeholders in the Food Sector: The Case Study of Lavazza. *Sustainability*, *15*(10), Article 7884. https://doi.org/10.3390/su15107884

Genievskaya, Y., Almerekova, S., Abugalieva, S., Abugalieva, A., Sato, K., & Turuspekov, Y. (2023). Identification of SNPs Associated with Grain Quality Traits in Spring Barley Collection Grown in Southeastern Kazakhstan. *Agronomy-Basel*, *13*(6), Article 1560. https://doi.org/10.3390/agronomy13061560

Gottstein, V., Lachenmeier, D. W., Kuballa, T., & Bunzel, M. (2024). 1H NMR-based approach to determine the geographical origin and cultivation method of roasted coffee. *Food Chemistry*, *433*, Article 137278. https://doi.org/10.1016/j.foodchem.2023.137278

Grainger, M. N. C., Klaus, H., Hewitt, N., Gan, H., & French, A. D. (2023). Graphical Discrimination of New Zealand Honey from International Honey Using Elemental Analysis. *Biological Trace Element Research*. https://doi.org/10.1007/s12011-023-03680-6

Gyftokostas, N., Stefas, D. and Couris, S. (2020). Olive oils classification via laser-induced breakdown spectroscopy. *Applied Sciences*, *10*(10), 3462. https://doi.org/10.3390/app10103462

Harris, N., Viejo, C. G., Barnes, C., & Fuentes, S. (2023). Non-Invasive Digital Technologies to Assess Wine Quality Traits and Provenance through the Bottle. *Fermentation-Basel*, *9*(1), Article 10. https://doi.org/10.3390/fermentation9010010

Heaton, K., Kelly, S. D., Hoogewerff, J., & Woolfe, M. (2008). Verifying the geographical origin of beef: The application of multi-element isotope and trace element analysis. *Food Chemistry*, *107*(1), 506-515. https://doi.org/10.1016/j.foodchem.2007.08.010

Hemmer-Hansen, J., Hüssy, K., Baktoft, H., Huwer, B., Bekkevold, D., Haslob, H., Herrmann, J. P., Hinrichsen, H. H., Krumme, U., Mosegaard, H., Nielsen, E. E., Reusch, T. B. H., Storr-Paulsen, M., Velasco, A., von Dewitz, B., Dierking, J., & Eero, M. (2019). Genetic analyses reveal complex dynamics within a marine fish management area. *Evolutionary Applications*, *12*(4), 830-844. https://doi.org/10.1111/eva.12760

Herald, N. Z. (2016). Riddle of how 1,700 tons of manuka honey are made... but 10,000 are sold. Riddle of how 1,700 tons of manuka honey are made... but 10,000 are sold - NZ Herald

Hong, Y. H., Birse, N., Quinn, B., Li, Y. C., Jia, W. Y., McCarron, P., Wu, D., da Silva, G. R., Vanhaecke, L., van Ruth, S., & Elliott, C. T. (2023). Data fusion and multivariate analysis for food authenticity analysis. *Nature Communications*, *14*(1), Article 3309. https://doi.org/10.1038/s41467-023-38382-z

Huo, Y. Q., Kamal, G. M., Wang, J., Liu, H. L., Zhang, G. N., Hu, Z. Y., Anwar, F., & Du, H. Y. (2017). ¹H NMR-based metabolomics for discrimination of rice from different geographical origins of China. *Journal of Cereal Science*, *76*, 243-252. https://doi.org/10.1016/j.jcs.2017.07.002

Hüssy, K. (2011). Review of western Baltic cod (Gadus morhua) recruitment dynamics. *ICES Journal of Marine Science*, *68*, 1459–1471. https://doi.org/10.1093/icesjms/fsr088

Kabbaj, H., Sall, A. T., Al-Abdallat, A., Geleta, M., Amri, A., Filali-Maltouf, A., Belkadi, B., Ortiz, R., & Bassi, F. M. (2017). Genetic Diversity within a Global Panel of Durum Wheat (<i>Triticum durum</i>) Landraces and Modern Germplasm Reveals the History of Alleles Exchange. *Frontiers in Plant Science*, *8*, Article 1277. https://doi.org/10.3389/fpls.2017.01277

Kakouri, E., Revelou, P. K., Kanakis, C., Daferera, D., Pappas, C. S., & Tarantilis, P. A. (2021). Authentication of the Botanical and Geographical Origin and Detection of Adulteration of Olive Oil Using Gas Chromatography, Infrared and Raman Spectroscopy Techniques: A Review. *Foods*, *10*(7). https://doi.org/10.3390/foods10071565

Karabagias, I. K., & Nayik, G. A. (2023). Machine Learning Algorithms Applied to Semi-Quantitative Data of the Volatilome of Citrus and Other Nectar Honeys with the Use of HS-SPME/GC-MS Analysis, Lead to a New Index of Geographical Origin Authentication. *Foods*, *12*(3), Article 509. https://doi.org/10.3390/foods12030509

Karabagias, V. K., Karabagias, I. K., Louppis, A., Badeka, A., Kontominas, M. G., & Papastephanou, C. (2019). Valorization of Prickly Pear Juice Geographical Origin Based on Mineral and Volatile Compound Contents Using LDA. *Foods*, *8*(4), Article 123. https://doi.org/10.3390/foods8040123

Katerinopoulou, K., Kontogeorgos, A., Salmas, C. E., Patakas, A., & Ladavos, A. (2020). Geographical Origin Authentication of Agri-Food Products: A Review. *Foods*, *9*(4), 489. https://doi.org/10.3390/foods9040489

Khan, M. K., Pandey, A., Thomas, G., Akkaya, M. S., Kayis, S. A., Ozsensoy, Y., Hamurcu, M., Gezgin, S., Topal, A., & Hakki, E. E. (2015). Genetic diversity and population structure of wheat in India and Turkey. *Aob Plants*, *7*, Article plv083. https://doi.org/10.1093/aobpla/plv083

Khansaritoreh, E., Salmaki, Y., Ramezani, E., Akbari Azirani, T., Keller, A., Neumann, K., Alizadeh, K., Zarre, S., Beckh, G., & Behling, H. (2020). Employing DNA metabarcoding to determine the geographical origin of honey. *Heliyon*, *6*(11), e05596. https://doi.org/10.1016/j.heliyon.2020.e05596

Kim, J. S., Hwang, I. M., Lee, G. H., Park, Y. M., Choi, J. Y., Jamila, N., Khan, N., & Kim, K. S. (2017). Geographical origin authentication of pork using multi-element and multivariate data analyses. *Meat Science*, *123*, 13-20. https://doi.org/10.1016/j.meatsci.2016.08.011

Koot, E., Arnst, E., Taane, M., Goldsmith, K., Thrimawithana, A., Reihana, K., González-Martínez, S. C., Goldsmith, V., Houliston, G., & Chagné, D. (2022). Genome-wide patterns of genetic diversity, population structure and demographic history in mānuka (Leptospermum scoparium) growing on indigenous Māori land. *Horticulture Research*, *9*. https://doi.org/10.1093/hr/uhab012

Kumari, N., Grimbs, A., D'Souza, R. N., Verma, S. K., Corno, M., Kuhnert, N., & Ullrich, M. S. (2018). Origin and varietal based proteomic and peptidomic fingerprinting of <i>Theobroma</i> <i>cacao</i> in non-fermented and fermented cocoa beans. *Food Research International*, *111*, 137-147. https://doi.org/10.1016/j.foodres.2018.05.010

Lange, C. N., Monteiro, L. R., Freire, B. M., Franco, D. F., de Souza, R. O., Ferreira, C. S. D., da Silva, J. J. C., & Batista, B. L. (2019). Mineral profile exploratory analysis for rice grains traceability. *Food Chemistry*, *300*, Article 125145. https://doi.org/10.1016/j.foodchem.2019.125145

Le Mao, I., Da Costa, G., & Richard, T. (2023). 1H-NMR metabolomics for wine screening and analysis. *Oeno One*, *57*(1), 15-31. https://doi.org/10.20870/oeno-one.2023.57.1.7134

Liu, H. C., You, C. F., Chen, C. Y., Liu, Y. C., & Chung, M. T. (2014). Geographic determination of coffee beans using multi-element analysis and isotope ratios of boron and strontium. *Food Chemistry*, *142*, 439-445. https://doi.org/10.1016/j.foodchem.2013.07.082

Liu, H. Y., Qin, Y. C., Ma, Q., Zhao, Q. Y., Guo, X. Q., Ma, L. N., Gou, C. L., Xia, Y., Gan, R. Y., & Zhang, J. M. (2021). Discrimination the geographical origin of Yanchi <i>Tan Lamb</i> with different muscle sections by stable isotopic ratios and elemental profiles. *International Journal of Food Science and Technology*, *56*(6), 2604-2611. https://doi.org/10.1111/ijfs.14900

Liu, H. Y., Wei, Y. M., Zhang, Y. Q., Wei, S., Zhang, S. S., & Guo, B. L. (2017). The effectiveness of multi-element fingerprints for identifying the geographical origin of wheat. *International Journal of Food Science and Technology*, *52*(4), 1018-1025. https://doi.org/10.1111/ijfs.13366

Liu, T. S., Lin, J. N., & Peng, T. R. (2018). Discrimination of Geographical Origin of Asian Garlic Using Isotopic and Chemical Datasets under Stepwise Principal Component Analysis. *Journal of Forensic Sciences*, *63*(5), 1366-1373. https://doi.org/10.1111/1556-4029.13731

Liu, X. J., Teixeira, J. S., Ner, S., Ma, K. S. V., Petronella, N., Banerjee, S., & Ronholm, J. (2020). Exploring the Potential of the Microbiome as a Marker of the Geographic

Origin of Fresh Seafood. *Frontiers in Microbiology*, *11*, Article 696. https://doi.org/10.3389/fmicb.2020.00696

Lozano Rodríguez, M. Á., Rodríguez, M.L., Canché, J.M.P., García, R.A.M., & Cabrera, C.R.C. . (2022). Visit frequency of Euglossine bees (Hymenoptera: Apidae) to mature fruits of Vanilla planifolia (Orchidaceae). *Acta Botanica Mexicana* https://doi.org/10.21829/ABM129.2022.2001

Luykx, D. M. A. M., & van Ruth, S. M. (2008). An overview of analytical methods for determining the geographical origin of food products. *Food Chemistry*, *107*(2), 897-911. https://doi.org/10.1016/j.foodchem.2007.09.038

Madigan, T. L., Bott, N. J., Torok, V. A., Percy, N. J., Carragher, J. F., de Barros Lopes, M. A., & Kiermeier, A. (2014). A microbial spoilage profile of half shell Pacific oysters (Crassostrea gigas) and Sydney rock oysters (Saccostrea glomerata). *Food Microbiol*, *38*, 219-227. https://doi.org/10.1016/j.fm.2013.09.005

Maestrello, V., Solovyev, P., Bontempo, L., Mannina, L., & Camin, F. (2022). Nuclear magnetic resonance spectroscopy in extra virgin olive oil authentication. *Comprehensive Reviews in Food Science and Food Safety*, *21*(5), 4056-4075. https://doi.org/10.1111/1541-4337.13005

Maione, C., & Barbosa, R. M. (2019). Recent applications of multivariate data analysis methods in the authentication of rice and the most analyzed parameters: A review. *Critical Reviews in Food Science and Nutrition*, *59*(12), 1868-1879. https://doi.org/10.1080/10408398.2018.1431763

Maione, C., Batista, B. L., Campiglia, A. D., Barbosa, F., & Barbosa, R. M. (2016). Classification of geographic origin of rice by data mining and inductively coupled plasma mass spectrometry. *Computers and Electronics in Agriculture*, *121*, 101-107. https://doi.org/https://doi.org/10.1016/j.compag.2015.11.009

Mannino, G., Kunz, R., & Maffei, M. E. (2023). Discrimination of Green Coffee (<i>Coffea arabica</i> and <i>Coffea canephora</i>) of Different Geographical Origin Based on Antioxidant Activity, High-Throughput Metabolomics, and DNA RFLP Fingerprinting. *Antioxidants*, *12*(5), Article 1135. https://doi.org/10.3390/antiox12051135

Martins, P., Madeira, M., Monteiro, F., De Sousa, R. B., Curvelo-Garcia, A. S., & Catarino, S. (2014). 87Sr86Sr ratio in vineyard soils from Portuguese denominations of origin and its potential for origin authentication. *Journal International Des Sciences De La Vigne Et Du Vin*, *48*(1), 21-29. http://dx.doi.org/10.20870/oeno-one.2014.48.1.1652

McGrath, T. F., Haughey, S. A., Patterson, J., Fauhl-Hassek, C., Donarski, J., Alewijn, M., van Ruth, S., & Elliott, C. T. (2018). What are the scientific challenges in moving from targeted to non-targeted methods for food fraud testing and how can they be addressed? - Spectroscopy case study. *Trends in Food Science & Technology*, *76*, 38-55. https://doi.org/10.1016/j.tifs.2018.04.001

McVey, C., Elliott, C. T., Cannavan, A., Kelly, S. D., Petchkongkaew, A., & Haughey, S. A. (2021). Portable spectroscopy for high throughput food authenticity screening: Advancements in technology and integration into digital traceability systems. *Trends in Food Science & Technology*, *118*, 777-790. https://doi.org/10.1016/j.tifs.2021.11.003

McVey, C., Elliott, C. T., Cannavan, A., Kelly, S. D., Petchkongkaew, A., & Haughey, S. A. (2021). Portable spectroscopy for high throughput food authenticity screening: Advancements in technology and integration into digital traceability systems. *Trends in Food Science & Technology*, *118*, 777-790. https://doi.org/10.1016/j.tifs.2021.11.003

Mendil, D., Demirci, Z., Tuzen, M., & Soylak, M. (2010). Seasonal investigation of trace element contents in commercially valuable fish species from the Black sea, Turkey. *Food and Chemical Toxicology*, *48*(3), 865-870. https:/doi.org/10.1016/j.fct.2009.12.023

Mi, S., Zhang, X. N., Wang, Y. H., Yan, F. L., Sang, Y. X., Gong, H. Y., & Wang, X. H. (2021). Geographical discrimination and authentication of Chinese garlic based on multi-element, volatile and metabolomics profiling combined with chemometrics. *Food Control*, *130*, Article 108328. https://doi.org/10.1016/j.foodcont.2021.108328

Mialon, N., Roig, B., Capodanno, E., & Cadiere, A. (2023). Untargeted metabolomic approaches in food authenticity: A review that showcases biomarkers. *Food Chemistry*, *398*, Article 133856. https://doi.org/10.1016/j.foodchem.2022.133856

Milan, M., Maroso, F., Dalla Rovere, G., Carraro, L., Ferraresso, S., Patarnello, T., Bargelloni, L., Cardazzo, B., & Fariselli, P. (2019). Tracing seafood at high spatial resolution using NGS-generated data and machine learning: Comparing microbiome versus SNPs. *Food Chemistry*, *286*, 413-420. https://doi.org/10.1016/j.foodchem.2019.02.037

Monahan, F. J., Schmidt, O., & Moloney, A. P. (2018). Meat provenance: Authentication of geographical origin and dietary background of meat. *Meat Science*, *144*, 2-14. https://doi.org/10.1016/j.meatsci.2018.05.008

Monahan, F. J., Schmidt, O., & Moloney, A. P. (2018). Meat provenance: Authentication of geographical origin and dietary background of meat. *Meat Science*, *144*, 2-14. https:/doi.org/10.1016/j.meatsci.2018.05.008

Montanari, S., Deng, C., Koot, E., Bassil, N. V., Zurn, J. D., Morrison-Whittle, P., Worthington, M. L., Aryal, R., Ashrafi, H., Pradelles, J., Wellenreuther, M., & Chagné, D. (2022). A multiplexed plant-animal SNP array for selective breeding and species conservation applications. *bioRxiv*, 2022.2009.2007.507051. https://doi.org/10.1101/2022.09.07.507051

Morrison, L., Bennion, M., Gill, S., & Graham, C. T. (2019). Spatio-temporal trace element fingerprinting of king scallops (<i>Pecten maximus</i>) reveals harvesting period and location. *Science of the Total Environment*, *697*, Article 134121. https://doi.org/10.1016/j.scitotenv.2019.134121

Mu, Y., Su, W., Mu, Y. C., & Jiang, L. (2020). Combined Application of High-Throughput Sequencing and Metabolomics Reveals Metabolically Active Microorganisms During Panxian Ham Processing. *Frontiers in Microbiology*, *10*, Article 3012. https://doi.org/10.3389/fmicb.2019.03012

Munyendo, L., Njoroge, D., & Hitzmann, B. (2022). The Potential of Spectroscopic Techniques in Coffee Analysis-A Review. *Processes*, *10*(1), Article 71. https://doi.org/10.3390/pr10010071

Nho, E. Y., Choi, J. Y., Lee, C. M., Dang, Y. M., Khan, N., Jamila, N., & Kim, K. S. (2019). Origin Authentication of Pork Fat via Elemental Composition, Isotope Ratios, and Multivariate Chemometric Analyses. *Analytical Letters*, *52*(9), 1445-1461. https://doi.org/10.1080/00032719.2018.1545133

Oh, J. D., Song, K. D., Seo, J. H., Kim, D. K., Kim, S. H., Seo, K. S., Lim, H. T., Lee, J. B., Park, H. C., Ryu, Y. C., Kang, M. S., Cho, S., Kim, E. S., Choe, H. S., Kong, H. S., & Lee, H. K. (2014). Genetic Traceability of Black Pig Meats Using Microsatellite Markers. *Asian-Australasian Journal of Animal Sciences*, *27*(7), 926-931. https://doi.org/10.5713/ajas.2013.13829

Olmo-García, L., Wendt, K., Kessler, N., Bajoub, A., Fernández-Gutiérrez, A., Baessmann, C., & Carrasco-Pancorbo, A. (2019). Exploring the Capability of LC-MS and GC-MS Multi-Class Methods to Discriminate Virgin Olive Oils from Different Geographical Indications and to Identify Potential Origin Markers. *European Journal of Lipid Science and Technology*, *121*(3), Article 1800336. https://doi.org/10.1002/ejlt.201800336

Özkök, A., Bilgiç, H. A., Kosukcu, C., Arik, G., Canli, D., Yet, I., & Karaaslan, C. (2023). Comparing the melissopalynological and next generation sequencing (NGS) methods for the determining of botanical origin of honey. *Food Control*, *148*, Article 109630. https://doi.org/10.1016/j.foodcont.2023.109630

Pavlin, A., Kocar, D., Imperl, J., Kolar, M., Marolt, G., & Petrova, P. (2023). Honey Origin Authentication via Mineral Profiling Combined with Chemometric Approaches. *Foods*, *12*(15), Article 2826. https://doi.org/10.3390/foods12152826

Pawlaczyk, A., Gajek, M., Jozwik, K., & Szynkowska, M. I. (2019). Multielemental Analysis of Various Kinds of Whisky. *Molecules*, *24*(7). https://doi.org/10.3390/molecules24071193

Peng, X. Z., Zhang, X., Wang, X. Y., Li, H. S., Xu, J. P., Zhao, Z. Y., & Wang, Y. H. (2022). Research on the Cross-Chain Model of Rice Supply Chain Supervision Based on Parallel Blockchain and Smart Contracts. *Foods*, *11*(9), Article 1269. https://doi.org/10.3390/foods11091269

Pepi, S., & Vaccaro, C. (2018). Geochemical fingerprints of "Prosecco" wine based on major and trace elements. *Environmental Geochemistry and Health*, *40*(2), 833-847. https://doi.org/10.1007/s10653-017-0029-0

Pérez-Pardal, L., Chen, S., Costa, V., Liu, X., Carvalheira, J., & Beja-Pereira, A. (2018). Genomic differentiation between swamp and river buffalo using a cattle high-density single nucleotide polymorphisms panel. *Animal*, *12*(3), 464-471. https://doi.org/10.1017/s1751731117001719

Qi, J., Li, Y. Y., Zhang, C., Wang, C., Wang, J. Q., Guo, W. P., & Wang, S. W. (2021). Geographic origin discrimination of pork from different Chinese regions using mineral elements analysis assisted by machine learning techniques. *Food Chemistry*, *337*, Article 127779. https://doi.org/10.1016/j.foodchem.2020.127779

Quinn, B., McCarron, P., Hong, Y., Birse, N., Wu, D., Elliott, C. T., & Ch, R. (2022). Elementomics combined with dd-SIMCA and K-NN to identify the geographical origin of rice samples from China, India, and Vietnam. *Food Chemistry*, *386*, 132738. https:/doi.org/10.1016/j.foodchem.2022.132738

Quintanilla-Casas, B., Torres-Cobos, B., Guardiola, F., Servili, M., Alonso-Salces, R. M., Valli, E., Bendini, A., Toschi, T. G., Vichi, S., & Tres, A. (2022). Geographical authentication of virgin olive oil by GC-MS sesquiterpene hydrocarbon fingerprint: Verifying EU and single country label-declaration. *Food Chemistry*, *378*, Article 132104. https://doi.org/10.1016/j.foodchem.2022.132104

Quintanilla-Casas, B., Torres-Cobos, B., Guardiola, F., Servili, M., Alonso-Salces, R. M., Valli, E., Bendini, A., Toschi, T. G., Vichi, S., & Tres, A. (2022). Geographical authentication of virgin olive oil by GC-MS sesquiterpene hydrocarbon fingerprint: Verifying EU and single country label-declaration. *Food Chemistry*, *378*, Article 132104. https://doi.org/10.1016/j.foodchem.2022.132104

Rainbow, P. S. (2017). *Heavy Metals in the Marine Environment* https://doi.org/10.1201/9781351073158

Ramona-Crina, S., Guyon, F., & Magdas, D. A. (2022). Application of emission - excitation matrices in parallel with factor analysis with other chemometric techniques for honey classification. *Journal of Food Composition and Analysis*, *107*, Article 104401. https://doi.org/10.1016/j.jfca.2022.104401

Ramona-Crina, S., Guyon, F., & Magdas, D. A. (2022). Application of emission - excitation matrices in parallel with factor analysis with other chemometric techniques for honey classification. *Journal of Food Composition and Analysis*, *107*, Article 104401. https://doi.org/10.1016/j.jfca.2022.104401

Ranaweera, R. K. R., Bastian, S. E. P., Gilmore, A. M., Capone, D. L., & Jeffery, D. W. (2023). Absorbance-transmission and fluorescence excitation-emission matrix (A-TEEM) with multi-block data analysis and machine learning for accurate intraregional classification of Barossa Shiraz wine. *Food Control*, *144*, Article 109335. https://doi.org/10.1016/j.foodcont.2022.109335

Ranaweera, R. K. R., Bastian, S. E. P., Gilmore, A. M., Capone, D. L., & Jeffery, D. W. (2023). Absorbance-transmission and fluorescence excitation-emission matrix (A-TEEM) with multi-block data analysis and machine learning for accurate intraregional classification of Barossa Shiraz wine. *Food Control*, *144*, Article 109335. https://doi.org/10.1016/j.foodcont.2022.109335

Rapa, M., Ferrante, M., Rodushkin, I., Paulukat, C., & Conti, M. E. (2023). Heavy metals, rare earth elements and isotopic fingerprint of Venetian Protected Designation of Origin sparkling wines. *British Food Journal*. https://doi.org/10.1108/bfj-06-2022-0483

Rees, G., Kelly, S. D., Cairns, P., Ueckermann, H., Hoelzl, S., Rossmann, A., & Scotter, M. J. (2016). Verifying the geographical origin of poultry: The application of stable isotope and trace element (SITE) analysis. *Food Control*, *67*, 144-154. https://doi.org/10.1016/j.foodcont.2016.02.018

Ren, X., Li, S. X., Zhang, M., Guan, L. A., & Han, W. X. (2023). Geographical discrimination of fresh instant rice according to non-destructive analysis of flavor profiles. *Cereal Chemistry*, *100*(2), 414-423. https://doi.org/10.1002/cche.10621

Ren, X., Li, S. X., Zhang, M., Guan, L. A., & Han, W. X. (2023). Geographical discrimination of fresh instant rice according to non-destructive analysis of flavor profiles. *Cereal Chemistry*, *100*(2), 414-423. https://doi.org/10.1002/cche.10621

Rodrigues, C. I., Maia, R., Miranda, M., Ribeirinho, M., Nogueira, J. M. F., & Máguas, C. (2009). Stable isotope analysis for green coffee bean: A possible method for geographic origin discrimination. *Journal of Food Composition and Analysis*, *22*(5), 463-471. https:/doi.org/10.1016/j.jfca.2008.06.010

Rodrigues, C., Brunner, M., Steiman, S., Bowen, G. J., Nogueira, J. M., Gautz, L., Prohaska, T., & Máguas, C. (2011). Isotopes as tracers of the Hawaiian coffee-producing regions. *J Agric Food Chem*, *59*(18), 10239-10246. https://doi.org/10.1021/jf200788p

Rodrigues, C., Máguas, C., & Prohaska, T. (2011). Strontium and oxygen isotope fingerprinting of green coffee beans and its potential to proof authenticity of coffee. *European Food Research and Technology*, *232*(2), 361-373. https://doi.org/10.1007/s00217-010-1362-z

Rohman, A., Ghazali, M. A. I. B., Windarsih, A., Irnawati, Riyanto, S., Yusof, F. M., & Mustafa, S. (2020). Comprehensive Review on Application of FTIR Spectroscopy Coupled with Chemometrics for Authentication Analysis of Fats and Oils in the Food Products. *Molecules*, *25*(22), 5485. https://doi.org/10.3390/molecules25225485

Saar de Almeida, B., Fedele, L., D'Antonio, M., Morra, V., Mercurio, M., Stevenson, R., & Widory, D. (2023). Characterizing wine terroir using strontium isotope ratios: a review. *Isotopes Environ Health Stud*, 1-22. https://doi.org/10.1080/10256016.2023.2245122

Salla, V., & Murray, K. K. (2013). Matrix-assisted laser desorption ionization mass spectrometry for identification of shrimp. *Analytica Chimica Acta*, *794*, 55-59. https://doi.org/https://doi.org/10.1016/j.aca.2013.07.014

Saurabh, S., & Dey, K. (2021). Blockchain technology adoption, architecture, and sustainable agri-food supply chains. *Journal of Cleaner Production*, *284*, Article 124731. https://doi.org/10.1016/j.jclepro.2020.124731

Schellenberg,A., Chmielus,S., Schlicht,C., Camin,F., Perini,M., Bontempo,L., Heinrich,K., Kelly,S.D., Rossmann,A., Thomas,F., Jamin·,E., M. Horacek,M. (2010) Multielement Stable Isotope Ratios (H, C, N, S) of Honey from different European Regions. (2010) *Food Chemistry, 121,* 770-777. https://doi.org/10.1016/j.foodchem.2009.12.082

Schwolow, S., Gerhardt, N., Rohn, S., & Weller, P. (2019). Data fusion of GC-IMS data and FT-MIR spectra for the authentication of olive oils and honeys-is it worth to go the extra mile? *Analytical and Bioanalytical Chemistry*, *411*(23), 6005-6019. https://doi.org/10.1007/s00216-019-01978-w

Sen, I., & Tokatli, F. (2014). Characterization and Classification of Turkish Wines Based on Elemental Composition. *American Journal of Enology and Viticulture*, *65*(1), 134-142. https://doi.org/10.5344/ajev.2013.13081

Serra, F., Guillou, C. G., Reniero, F., Ballarin, L., Cantagallo, M. I., Wieser, M., Iyer, S. S., Héberger, K., & Vanhaecke, F. (2005). Determination of the geographical origin of green coffee by principal component analysis of carbon, nitrogen and boron stable isotope ratios. *Rapid Commun Mass Spectrom*, *19*(15), 2111-2115. https://doi.org/10.1002/rcm.2034

Silva, B., Gonzaga, L. V., Maltez, H. F., Samochvalov, K. B., Fett, R., & Costa, A. C. O. (2021). Elemental profiling by ICP-MS as a tool for geographical discrimination: The case of bracatinga honeydew honey. *Journal of Food Composition and Analysis*, *96*, 103727. https:/doi.org/10.1016/j.jfca.2020.103727

Simpkins, W. A., Louie, H., Wu, M., Harrison, M., & Goldberg, D. (2000). Trace elements in Australian orange juice and other products. *Food Chemistry*, *71*(4), 423-433. https:/doi.org/10.1016/S0308-8146(00)00150-3

Sliwinska-Bartel, M., Burns, D. T., & Elliott, C. (2021). Rice fraud a global problem: A review of analytical tools to detect species, country of origin and adulterations. *Trends in Food Science & Technology*, *116*, 36-46. https://doi.org/10.1016/j.tifs.2021.06.042

Soares, S., Amaral, J. S., Oliveira, M., & Mafra, I. (2017). A Comprehensive Review on the Main Honey Authentication Issues: Production and Origin. *Comprehensive Reviews in Food Science and Food Safety*, *16*(5), 1072-1100. https://doi.org/10.1111/1541-4337.12278

Soh, C. G., Zhu, Y., & Toh, T. L. (2023). A regularised logistic regression model with structured features for classification of geographical origin in olive oils.

Chemometrics and Intelligent Laboratory Systems, *237*, Article 104819. https://doi.org/10.1016/j.chemolab.2023.104819

Song, B. Y., Ryu, J. S., Shin, H. S., & Lee, K. S. (2014). Determination of the Source of Bioavailable Sr Using ⁸⁷Sr/⁸⁶Sr Tracers: A Case Study of Hot Pepper and Rice. *Journal of Agricultural and Food Chemistry*, *62*(38), 9232-9238. https://doi.org/10.1021/jf503498r

Sotiropoulou, N. S., Xagoraris, M., Revelou, P. K., Kaparakou, E., Kanakis, C., Pappas, C., & Tarantilis, P. (2021). The Use of SPME-GC-MS IR and Raman Techniques for Botanical and Geographical Authentication and Detection of Adulteration of Honey. *Foods*, *10*(7), Article 1671. https://doi.org/10.3390/foods10071671

Sotiropoulou, N. S., Xagoraris, M., Revelou, P. K., Kaparakou, E., Kanakis, C., Pappas, C., & Tarantilis, P. (2021). The Use of SPME-GC-MS IR and Raman Techniques for Botanical and Geographical Authentication and Detection of Adulteration of Honey. *Foods*, *10*(7), Article 1671. https://doi.org/10.3390/foods10071671

Srinuttrakul, W., Yoshida, S., & Iop. (2016, Aug 04-06). Determination of stable cesium and strontium in rice samples by inductively coupled plasma mass spectrometry.*Journal of Physics Conference Series* [International nuclear science and technology conference 2016]. International Nuclear Science and Technology Conference (INST), Thammasat Univ, Bangkok, THAILAND. https://doi.org/10.1088/1742-6596/860/1/012013

Suhandy, D., & Yulia, M. (2021). The Use of UV Spectroscopy and SIMCA for the Authentication of Indonesian Honeys According to Botanical, Entomological and Geographical Origins. *Molecules*, *26*(4), Article 915. https://doi.org/10.3390/molecules26040915

Suhandy, D., & Yulia, M. (2021). The Use of UV Spectroscopy and SIMCA for the Authentication of Indonesian Honeys According to Botanical, Entomological and Geographical Origins. *Molecules*, *26*(4), Article 915. https://doi.org/10.3390/molecules26040915

Sun, S., Guo, B., Wei, Y., & Fan, M. (2011). Multi-element analysis for determining the geographical origin of mutton from different regions of China. *Food Chemistry*, *124*(3), 1151-1156. https://doi.org/10.1016/j.foodchem.2010.07.027

Sun, X. Y., Zhang, F., Gutiérrez-Gamboa, G., Ge, Q., Xu, P. K., Zhang, Q. W., Fang, Y. L., & Ma, T. T. (2022). Real wine or not? Protecting wine with traceability and authenticity for consumers: chemical and technical basis, technique applications, challenge, and perspectives. *Critical Reviews in Food Science and Nutrition*, *62*(24), 6783-6808. https://doi.org/10.1080/10408398.2021.1906624

Tahir, H. E., Arslan, M., Mahunu, G. K., Mariod, A. A., Hashim, S. B. H., Zou, X. B., Shi, J. Y., El-Seedi, H. R., & Musa, T. H. (2022). The use of analytical techniques coupled with chemometrics for tracing the geographical origin of oils: A systematic review (2013-2020). *Food Chemistry*, *366*, Article 130633. https://doi.org/10.1016/j.foodchem.2021.130633

Telloli, C., Tagliavini, S., Passarini, F., Salvi, S., & Rizzo, A. (2023). ICP-MS triple quadrupole as analytical technique to define trace and ultra-trace fingerprint of extra virgin olive oil. *Food Chemistry*, *402*, Article 134247. https://doi.org/10.1016/j.foodchem.2022.134247

Temerdashev, Z., Bolshov, M., Abakumov, A., Khalafyan, A., Kaunova, A., Vasilyev, A., Sheludko, O., & Ramazanov, A. (2023). Can Rare Earth Elements Be Considered as Markers of the Varietal and Geographical Origin of Wines? *Molecules*, *28*(11), Article 4319. https://doi.org/10.3390/molecules28114319

Tescione, I., Casalini, M., Marchionni, S., Braschi, E., Mattei, M., & Conticelli, S. (2020). Conservation of ⁸⁷Sr/⁸⁶Sr During Wine-Making of White Wines: A Geochemical Fingerprint of Geographical Provenance and Quality Production. *Frontiers in Environmental Science*, *8*, Article 153. https://doi.org/10.3389/fenvs.2020.00153

Teye, E., Anyidoho, E., Agbemafle, R., Sam-Amoah, L. K., & Elliott, C. (2020). Cocoa bean and cocoa bean products quality evaluation by NIR spectroscopy and chemometrics: A review. *Infrared Physics & Technology*, *104*, Article 103127. https://doi.org/10.1016/j.infrared.2019.103127

Vadalà, R., Mottese, A. F., Bua, G. D., Salvo, A., Mallamace, D., Corsaro, C., Vasi, S., Giofrè, S. V., Alfa, M., Cicero, N., & Dugo, G. (2016). Statistical Analysis of Mineral Concentration for the Geographic Identification of Garlic Samples from Sicily (Italy), Tunisia and Spain. *Foods*, *5*(1). https://doi.org/10.3390/foods5010020

Varrà, M. O., Husàkovà, L., Zanardi, E., Alborali, G. L., Patocka, J., Ianieri, A., & Ghidini, S. (2023). Elemental profiles of swine tissues as descriptors for the traceability of value-added Italian heavy pig production chains. *Meat Science*, *204*, Article 109285. https://doi.org/10.1016/j.meatsci.2023.109285

Vinay Jain, S. D. a. B. K. (2022). Authenticating Geographical Origin of Tea Using ICP-MS and Agilent Mass Profiler Professional Software. *Agilent Application Notes*.

Wadood, S. A., Nie, J., Li, C. L., Rogers, K. M., Khan, A., Khan, W. A., Qamar, A., Zhang, Y. Z., & Yuwei, Y. (2022). Rice authentication: An overview of different analytical techniques combined with multivariate analysis. *Journal of Food Composition and Analysis*, *112*, Article 104677. https://doi.org/10.1016/j.jfca.2022.104677

Wakefield, J., McComb, K., Ehtesham, E., Van Hale, R., Barr, D., Hoogewerff, J., & Frew, R. (2019). Chemical profiling of saffron for authentication of origin. *Food Control*, *106*, Article 106699. https://doi.org/10.1016/j.foodcont.2019.06.025

Wang, A. L., Zhu, Y. Y., Qiu, J., Cao, R. G., & Zhu, H. (2022). Application of intelligent sensory technology in the authentication of alcoholic beverages. *Food Science and Technology*, *42*, Article e32622. https://doi.org/10.1590/fst.32622

Wang, A. L., Zhu, Y. Y., Qiu, J., Cao, R. G., & Zhu, H. (2022). Application of intelligent sensory technology in the authentication of alcoholic beverages. *Food Science and Technology*, *42*, Article e32622. https://doi.org/10.1590/fst.32622

Wang, J., Kliks, M. M., Qu, W., Jun, S., Shi, G., & Li, Q. X. (2009). Rapid Determination of the Geographical Origin of Honey Based on Protein Fingerprinting and Barcoding Using MALDI TOF MS. *Journal of Agricultural and Food Chemistry*, *57*(21), 10081-10088. https://doi.org/10.1021/jf902286p

Wang, S. C., Wong, D. B., Forrest, K., Allen, A., Chao, S. M., Huang, B. E., Maccaferri, M., Salvi, S., Milner, S. G., Cattivelli, L., Mastrangelo, A. M., Whan, A., Stephen, S., Barker, G., Wieseke, R., Plieske, J., Lillemo, M., Mather, D., Appels, R., . . . Int Wheat Genome, S. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnology Journal*, *12*(6), 787-796. https://doi.org/10.1111/pbi.12183

Wieser, M. E., Iyer, S. S., Krouse, H. R., & Cantagallo, M. I. (2001). Variations in the boron isotope composition of Coffea arabica beans. *Applied Geochemistry*, *16*(3), 317-322. https://doi.org/10.1016/S0883-2927(00)00031-7

Wirta, H., Abrego, N., Miller, K., Roslin, T., & Vesterinen, E. (2021). DNA traces the origin of honey by identifying plants, bacteria and fungi. *Scientific Reports*, *11*(1), Article 4798. https://doi.org/10.1038/s41598-021-84174-0

Wu, H., Lin, G. H., Tian, L., Yan, Z., Yi, B. Q., Bian, X. H., Jin, B. H., Xie, L. Q., Zhou, H. C., & Rogers, K. M. (2021). Origin verification of French red wines using isotope and elemental analyses coupled with chemometrics. *Food Chemistry*, *339*, Article 127760. https://doi.org/10.1016/j.foodchem.2020.127760

Xu, J. P., Han, J. Q., Qi, Z. B., Jiang, Z. X., Xu, K., Zheng, M. Z., & Zhang, X. (2022). A Reliable Traceability Model for Grain and Oil Quality Safety Based on Blockchain and Industrial Internet. *Sustainability*, *14*(22), Article 15144. https://doi.org/10.3390/su142215144

Xu, Y., Zhong, P., Jiang, A., Shen, X., Li, X., Xu, Z., Shen, Y., Sun, Y., & Lei, H. (2020). Raman spectroscopy coupled with chemometrics for food authentication: A review. *TrAC Trends in Analytical Chemistry*, *131*, 116017. https://doi.org/https://doi.org/10.1016/j.trac.2020.116017

Yu, D. X., Zhang, X., Guo, S., Yan, H., Wang, J. M., Zhou, J. Q., Yang, J., & Duan, J. A. (2022). Headspace GC/MS and fast GC e-nose combined with chemometric analysis to identify the varieties and geographical origins of ginger (<i>Zingiber officinale</i> Roscoe). *Food Chemistry*, *396*, Article 133672. https://doi.org/10.1016/j.foodchem.2022.133672

Yu, D. X., Zhang, X., Guo, S., Yan, H., Wang, J. M., Zhou, J. Q., Yang, J., & Duan, J. A. (2022). Headspace GC/MS and fast GC e-nose combined with chemometric analysis to identify the varieties and geographical origins of ginger (<i>Zingiber officinale</i> Roscoe). *Food Chemistry*, *396*, Article 133672. https://doi.org/10.1016/j.foodchem.2022.133672

Yu, G., Wang, F., Zhang, B., & Fan, J. (2016). In vitro inhibition of platelet aggregation by peptides derived from oat (Avena sativa L.), highland barley (Hordeum vulgare Linn. var. nudum Hook. f.), and buckwheat (Fagopyrum esculentum Moench) proteins. *Food Chem*, *194*, 577-586. https://doi.org/10.1016/j.foodchem.2015.08.058

Zanuttin, F., Gurian, E., Ignat, I., Fornasaro, S., Calabretti, A., Bigot, G., & Bonifacio, A. (2019). Characterization of white wines from north-eastern Italy with surface-enhanced Raman spectroscopy. *Talanta*, *203*, 99-105. https://doi.org/10.1016/j.talanta.2019.05.024

Zanuttin, F., Gurian, E., Ignat, I., Fornasaro, S., Calabretti, A., Bigot, G., & Bonifacio, A. (2019). Characterization of white wines from north-eastern Italy with surface-enhanced Raman spectroscopy. *Talanta*, *203*, 99-105. https://doi.org/10.1016/j.talanta.2019.05.024

Zaukuu, J.-L. Z., Benes, E., Bázár, G., Kovács, Z., & Fodor, M. (2022). Agricultural Potentials of Molecular Spectroscopy and Advances for Food Authentication: An Overview. *Processes*, *10*(2), 214. https://www.mdpi.com/2227-9717/10/2/214

Zhang, T. W., Wang, Q., Li, J. R., Zhao, S. S., Qie, M. J., Wu, X. L., Bai, Y., & Zhao, Y. (2021). Study on the origin traceability of Tibet highland barley (<i>Hordeum vulgare</i> L.) based on its nutrients and mineral elements. *Food Chemistry*, *346*, Article 128928. https://doi.org/10.1016/j.foodchem.2020.128928

Zhang, X. L., Yang, J., Lin, T., & Ying, Y. B. (2021). Food and agro-product quality evaluation based on spectroscopy and deep learning: A review. *Trends in Food Science & Technology*, *112*, 431-441. https://doi.org/10.1016/j.tifs.2021.04.008

Zheng, Z. W., Qiu, S. S., & Wei, Z. B. (2022). A Novel Voltammetric Electronic Tongue Based on Nanocomposites Modified Electrodes for the Discrimination of Red Wines from Different Geographical Origins. *Chemosensors*, *10*(8), Article 332. https://doi.org/10.3390/chemosensors10080332

Zheng, Z. W., Qiu, S. S., & Wei, Z. B. (2022). A Novel Voltammetric Electronic Tongue Based on Nanocomposites Modified Electrodes for the Discrimination of Red Wines from Different Geographical Origins. *Chemosensors*, *10*(8), Article 332. https://doi.org/10.3390/chemosensors10080332

Zhu, H., Zhu, D. Z., & Sun, J. M. (2023). Application of GC-IMS coupled with chemometric analysis for the classification and authentication of geographical indication agricultural products and food. *Frontiers in Nutrition*, *10*, Article 1247695. https://doi.org/10.3389/fnut.2023.1247695

Zielinski, A., Haminiuk, C., Nunes, C., Schnitzler, E., Ruth, S., & Granato, D. (2014). Chemical Composition, Sensory Properties, Provenance, and Bioactivity of Fruit Juices as Assessed by Chemometrics: A Critical Review and Guideline. *Comprehensive Reviews in Food Science and Food Safety*, *13*. https://doi.org/10.1111/1541-4337.12060

Food Standards Agency
food.gov.uk

For enquiries about this publication, contact the Food Standards Agency.

Project reference: FS900435

Follow us on Twitter: @foodgov

Find us on Facebook: facebook.com/FoodStandardsAgency