DISCRIMINATION OF NEUTRAL SPIRITS AND GINS BY BOTANICAL AND GEOGRAPHICAL ORIGIN USING NTA LC-QTOF-HRMS

PATRICE HURTUBISE

Department of Food Science and Agricultural Chemistry
McGill University, Montréal

August 2023

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Masters of Science in Food Science and Agricultural Chemistry

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

With the rise in local alcoholic beverage popularity, fraudulent activities are an emerging threat. This study addresses this issue by developing a robust non-targeted (NTA) high-resolution mass spectrometry (HRMS) coupled with ultra-high-performance liquid chromatography (RP-UHPLC-Q-TOF) method to authenticate unaged spirits based on their chemical fingerprints. Sixty-three samples of gins (23) and neutral spirits (NS) (40) from various geographical and botanical origins were analyzed using the NTA LC-Q-TOF-MS method. This comprehensive approach, enhanced with multivariate analysis such as Principal Component Analysis (PCA), efficiently characterized the chemical fingerprint associated with origins in unaged spirits. An advanced tool, SIRIUS, assisted compound characterization, and compound identification followed levels of confidence (Schymanski, Jeon et al. 2014). The primary objectives are reviewing existing literature on alcoholic beverage fraud, establishing comprehensive fingerprints of NS and gins, and discerning potential compounds indicative of botanical and geographical origins, and assessing the impact of

botanical addition on authentication marker detection performance in gin production. Initial results demonstrated a discernible difference in the number of molecular features (MFs) among the samples. Specifically, the analysis yielded 76,447 MFs in ESI+ mode (48,345 for vodka samples and 28,102 for gin samples) and 42,206 MFs in ESI- mode (42,206 for vodka samples and 26,138 for gin samples). MFs demonstrating significant differences across categories and with peak intensity exceeding 100,000 were selected for MS2 analysis to initiate the identification procedure. This study fills a gap in gin fraud literature and contributes to the efforts of creating a protected geographical indication for “Quebec origin” spirits, underlining the efficacy of the NTA LC-Q-TOF-MS method in fighting alcoholic beverage fraud.

RÉSUMÉ

Avec l'augmentation de la popularité des boissons alcoolisées locales, les activités frauduleuses constituent une menace émergente. Cette étude aborde ce problème en développant une méthode robuste de spectrométrie de masse à haute résolution (HRMS) non ciblée (NTA) couplée à la chromatographie liquide à ultra-haute performance (RP-UHPLC-Q-TOF) pour authentifier les spiritueux non vieillis sur la base de leurs empreintes chimiques. Soixante-trois échantillons de gins (23) et de spiritueux neutres (SN) (40) d'origines géographiques et botaniques diverses ont été analysés à l'aide de la méthode NTA LC-Q-TOF-MS. Cette approche globale, complétée par une analyse multivariée telle que l'analyse en composantes principales (ACP), a permis de caractériser efficacement l'empreinte chimique associée aux spiritueux non vieillis. Un outil avancé, SIRIUS, a aidé à la caractérisation des composés, et l'identification des composés a suivi les niveaux de confiance proposés par Schymanski et al. (2014). Nos principaux objectifs sont une revue de la littérature existante sur la fraude en matière de boissons alcoolisées, l'établissement d'empreintes chimiques complètes des SN et gins et l'investigation des composés potentiels indiquant les origines botaniques et géographiques, ainsi que l'évaluation de l'impact de l'ajout d'aromates sur la performance de détection des marqueurs d'authentification dans la production de gin. Les premiers résultats ont montré une différence perceptible dans le nombre de caractéristiques moléculaires (CM) entre les échantillons. Plus précisément, l'analyse a produit 76 447 CM en mode ESI+ (48 345 pour les échantillons de vodka et 28 102 pour les échantillons de gin) et 42 206 CM en mode ESI- (42 206 pour les échantillons de vodka et 26 138 pour les échantillons de gin). Les CM présentant des différences significatives entre les catégories ont été

sélectionnés pour l'analyse MS2 afin d'entamer la procédure d'identification. Cette étude comble des lacunes dans la littérature sur la fraude des boissons alcoolisées et contribue aux efforts de création d'une indication géographique protégée pour les spiritueux "d'origine québécoise", soulignant l'efficacité de la méthode NTA LC-Q-TOF-MS dans la lutte contre la fraude sur les boissons alcoolisées.

CONTRIBUTION OF AUTHORS

The contribution of authors and their contribution to the different articles is as follows: Patrice Hurtubise is the MSc candidate who designed and carried out all experiments Chapter 3 and Chapter 4 in consultation with the supervisor. He performed all data analysis and prepared all the manuscripts for revisions by the supervisor. He performed the literature review in Chapter 2. Dr. Stéphane Bayen is the thesis supervisor under whose guidance all the research was conducted. He received and revised drafts of all the manuscripts for publication.

ACKNOWLEDGEMENTS

I would like to extend my heartfelt gratitude and deepest appreciation to my supervisor, Dr. Stéphane Bayen, for his indispensable mentorship and assistance in the formulation and execution of the experiment, as well as the meticulous data processing steps. Your assistance has been a beacon, guiding me to where I stand today. Your willingness to venture into a new project, conceived solely from one of my budding interests, demonstrated an immense trust in my capabilities. You've not only provided me with invaluable opportunities, but also opened doors for me to interact with distinguished individuals in the realm of alcoholic beverages and food authenticity.

I am tremendously grateful to my astute and supportive colleagues Lan, Zhi Hao, Shawn, Jingyun, Gabriel, Nancy, Karen, Shaghig, Lei, and Alexandra. Each one of you has imparted knowledge and shared experiences that enriched my journey immensely. I fervently wish you the utmost success in all your academic endeavours and life's ventures. I extend my thanks to Jean-François Roy for his time and adept assistance at the Agilent lab, particularly with instrumental analysis, data interpretation, and offering general insights which were indispensable. I also need to mention the incredible contribution from some local producers such as Ubald distillery without whom I could not have completed this project.

A special expression of gratitude goes out to my dear wife, whose unwavering encouragement was my stronghold during challenging times and whose celebrations with me during moments of triumph added colour to this journey. The road wasn't easy, yet your steadfast support for our family during this pursuit of mine was nothing short of inspirational. Thank you to my parents, whose emotional and financial support were my pillars of strength whenever the tides were rough. Your unfaltering faith and support throughout my life have been my guiding lights, and for that, I am eternally grateful. Lastly, thank you to my friends who have managed to keep me on track with small words of wisdom, reassurance, consolation, faith, or joy, or for those who have literally been by my side distracting me from stressful moments.

LIST OF ABBREVIATIONS

AAS: Atomic absorption spectroscopy
AES: Atomic emission spectroscopy
ANN: Artificial neural network
APCI: Atmospheric-pressure chemical ionization
A-TEEM: Absorbance transmittance and a fluorescence excitation emission Matrix
ATR: Attenuated total reflection
CART: Classification and regression trees
CE: Capillary electrophoresis
CFR: Code of Federal Regulations
CovSel: Covariance selection
CV: Cold vapour
CV-ANOVA: Cross-validated ANOVA
DA: Discriminant analysis
DAD: Diode array detector
EC: European Commission
EI: Electron ionization
EIC: Extracted ion Chromatogram
ESI: Electrospray Ionization
EUIPO: European Union Intellectual Property Office
FCM: Fuzzy c means
GC-MS: Gas chromatography mass spectrometry
GFSI: Global food safety initiative
HCA: Hierarchical cluster analysis
HPLC: High performance liquid chromatography
HR: High resolution
HS: Headspace
ICA: Independent component analysis
ICP: Inductively coupled plasma
ICR: Ion cyclotron resonance
IR: Infrared

IRMS: Isotope ratio mass spectrometry
JRC: Joint research centre
K-CM: Knowledge configuration model
kNN: k-Nearest neighbour
LC: Liquid chromatography
LDA: Linear discriminant analysis
LOD: Limit of detection
LOI: Limit of identification
m/z: Mass over charge ratio
MALDI: Matrix-Assisted Laser Desorption/Ionization
MANOVA: Multivariate analysis of variance
ME: Matrix effects
MF: Molecular feature
MIR: Mid infrared
MP: Microwave plasma
MS: Mass spectrometry
MVA: Multivariate analysis
n: number of samples
NIR: Near infrared
NIR: Near infrared
NMR: Nuclear magnetic resonance
NTA: Non-targeted analysis/Untargeted analysis
OES: Optical emission spectroscopy
OPLS: Orthogonal PLS
OPS: Optical plasma spectroscopy
PCA: Principal component analysis
(q)(dd)PCR: (quantitative) (double digest) Polymerase chain reaction
PDO: Protected designation of origin
PLS: Partial least square
(Q)TOF: (Quadrupole) time-of-flight
(dd)RAD: Double digest Restriction site Associated DNA
SIMCA: Soft Independent Modelling of Class Analogy
SNP: Single nucleotide polymorphism
SO: Sequential and orthogonalized
SPME: Solid phase microextraction
SVM: Support vector machine
TA: Targeted analysis
TIC: Total ion chromatogram
UHPLC: Ultra high-performance liquid chromatography
UV-Vis : Ultraviolet visible light spectrum
VOC : Volatile organic compound

1. INTRODUCTION

Alcoholic beverages hold a significant share of the food market worldwide, USD1,369.4 billion in 2020, representing slightly more than 10% of the global food and beverage market (Lu 2020). There has been a constant increase in demand in the last decades leading to alcoholic beverages sharing a progressively larger part of the food industry market. Even if adulteration is not a new concept for alcoholic beverages, this surge in market value has brought more opportunities for fraudsters. It is becoming apparent that important steps must be taken to protect consumers and industries. Indeed, for many years now, alcoholic beverages have been considered among the top food commodities to be adulterated for economic gain (Goodall, Harrison et al. 2018, Europol - OPSON IX 2020).

Food authenticity refers to compliance with all regulations and standards governing the truthful labelling and representation of a food product. Food fraud means to increase the economic gain by circumventing these regulations (GFSI 2018). Alcoholic beverages are subject to intrinsic and extrinsic types of adulteration. Intrinsic types are related to agricultural practices and processing (substitution, dilution, enhancement), and extrinsic types to finished product (mislabelling, origin, counterfeiting) (GFSI 2018). Adulteration of alcoholic beverages can have disastrous repercussions that are not only detrimental to the economy of the industry but can also endanger life (Holmberg 2010, Ellis, Muhamadali et al. 2019, Manning and Kowalska 2021). Despite laws regarding agricultural practices, production, processing, and labelling, there is a lack of the ability to detect adulterations more efficiently. Most traditional analytical tools have been commonly paired with targeted analysis (TA), using specific authentication markers, and comparing them against reference standards to evaluate authenticity (Cavanna, Righetti et al. 2018). Nowadays, non-targeted analysis (NTA), or untargeted analysis, where the analysis is done on the whole matrix, has emerged as a powerful tool to provide the exhaustive chemical composition of various matrices. NTA thus provides unique chemical fingerprints which increase the difficulty in adulterating food to a point where it becomes virtually impossible (Cajka and Fiehn 2016, Cavanna, Righetti et al. 2018). As local craft spirits are gaining popularity worldwide, there is an emerging need for authentication of artisanal qualities to prevent eventual fraud (van Ruth, Huisman et al. 2017). This problem has been addressed in part with the targeted analysis of certain categories of spirits, notably Scotch whisky, tequila and brandies, but very little has been

using untargeted analysis (Ng, Hupé et al. 1996, De León-Rodríguez, González-Hernández et al. 2006, Lukić, Banović et al. 2006, Wiśniewska, Boqué et al. 2017, Raičević, Popović et al. 2022). Many aspects of alcoholic beverages authentication remain to be explored, mainly the geographical and botanical origin of the raw materials, the agricultural practices, the fermentation and distillation processes, the identification and characterization of additives, the substitution and dilution with water or synthetic ethanol, to mention a few. The exploration of these complex matrices has often been conducted towards volatile compounds using gas chromatography, while non-volatile compounds have remained vastly understudied. The analysis of non-volatile compounds in spirits presents unique challenges, notably due to their expected lower concentrations due to the distillation process. Nonetheless, with sophisticated analysis, such as liquid chromatography paired with high resolution mass spectrometry, it is expected that the exploration of these non-volatile compounds can prove to be a rich source of chemical information for authentication. Furthermore, the use of volatile compounds as authentication markers can be challenging as well due to their inherent lower stability compared to non-volatile compounds, making the latter a potentially favorable alternative for authentication markers. In the context of this research, neutral spirits (NS) and gins were specifically selected for their unique chemical profiles and widespread popularity. Both spirits can be made from a diverse range of botanical ingredients and NS can often serve as base spirits for gins, offering an interesting contrasting comparison. The lack of literature in the authentication of both these spirits makes this research an important contribution in this field. The goal of this research project is to develop a robust authentication method using NTA LC-QTOF-HRMS to characterize the chemical fingerprint of NS and gins to potentially discover compounds serving as authentication markers for the botanical and geographical origins of these spirits. As mentioned, fraud of alcoholic beverages is an international issue challenging government bodies, industries, the food science community, and threatening consumers trust, resources, and health (GFSI 2018, Spink 2019). By developing this technique, it would reduce the consequences of unnoticed adulterated alcoholic beverages by increasing the detection efficiency of fraudulent products as well as reduce the incentive to commit the crimes in the first place by increasing the risk of getting caught. Therefore, this research aims at providing a solution to advance towards food fraud prevention.

1.1 Research Hypothesis

The present study was conducted with the hypothesis that:

Hypothesis 1: A robust RP-LC-QTOF-HRMS method can establish the unique chemical fingerprint for NS and gins. This fingerprint can identify and characterize key compounds related to specific spirit categories for authentication purposes.

Hypothesis 2: Investigating the fingerprint of NS and gins reveals key compounds corresponding to botanical and geographical, identifiable through MVA. The use of SIRIUS software allows the elucidation of the chemical formula and structure of previously detected molecular features of interest.

1.2 Research Objectives

The overall goal of this thesis is to improve the ability to authenticate gins based on the geographical origin of the raw materials by using specific markers or fingerprint patterns. More specifically, I will be addressing the following points:

Aim 1: Review the literature on food fraud, alcoholic beverage chemistry, and analytical tools and approaches used in food fraud detection to identify current knowledge gaps.

Aim 2: Develop a robust RP-LC-QTOF-HRMS method to establish the fingerprint of neutral spirits and gins to obtain a general understanding and potentially identify important patterns related to the chemical fingerprint of the spirits.

Aim 3: Investigate the fingerprint of neutral alcohols and gins to identify and characterize key compounds related to botanical and geographical origin, using multivariate analysis. Then, use MSMS data and SIRIUS to elucidate the chemical structure and formula of such candidate authentication markers. Use a classification model such as PLS-DA to establish the accuracy of classification of samples in their respective categories.

2. LITERATURE REVIEW OF ALCOHOLIC BEVERAGES AUTHENTICATION

Alcoholic beverages represent a significant portion of the global food market, accounting for slightly more than 10% in 2020 with a worth of USD 1,369.4 billion (Lu 2020). The European Commission estimates that food fraud costs the European Union 30 billion euros every year, with 1.3 billion euros related to alcoholic beverages fraud (Lecat, Brouard et al. 2017, European Commission 2018). This is particularly prominent in countries producing internationally renowned beverages like Scotch whisky, Cognac, Bourbon, and wine. Rising demand, coupled with production regulations, has led to a surge in the value of rarer or highly sought-after products, creating opportunities for fraudsters. Alcoholic beverages fraud, a practice dating back to the times of Pliny the Elder (1885), can occur in several ways such as substitution, dilution, enhancement, mislabelling, or counterfeiting. Fraudulent adulteration of these beverages has significant economic repercussions and can potentially endanger consumers' health. To counteract fraud, exhaustive analytical techniques should be incorporated into food control for early detection of adulterations, shifting from a reactive to a proactive approach. Legal measures are in place to ensure compliance with agricultural practices, production, processing, and labelling standards (Aylott 2013, Spink 2019). Strict regulations exist for beverages such as Scotch whisky, tequila, Cognac, bourbon, and most wines (27-CFR-5.22 2008, European Commission (EC) 2008, The Scotch Whisky Regulations 2009, Arvanitoyannis 2010, Comité Consultivo Nacional de Normalización de Seguridad al Usuario 2012, Bureau National Interprofessionnel du Cognac (BNIC) 2018). Typically, adulterated products are identified through investigations, whistle-blowers, or a lack of proper documentation (Europol - OPSON IX 2020, Europol - OPSON X 2021). However, these methods alone aren't always sufficient, and analytical tools are essential to provide intrinsic proofs, such as the detection of adulterants.

Analytical tools can offer spectroscopic, spectrometric, or isotopic data to authenticate a product and, in some cases, identify and characterize adulterants. Currently, there is an increased focus on NTA, which provides a unique chemical fingerprint of the sample for authenticity. This technique makes it virtually impossible to adulterate food without it being noticed. The resulting data can be analyzed with chemometrics to differentiate between authentic and fraudulent samples. Despite these efforts, fraud continues to plague the alcoholic beverages industry. Fraudsters adapt and find

new ways to pursue their illegal activities. One of the major challenges is the lack of standardization among the NTAs for authentication, which often leads to non-reproducible results and diminishes reliability (Cavanna, Righetti et al. 2018). In the case of alcoholic beverages, there is also a deficiency in understanding the complete chemistry of the products and how the environment and processing steps influence their compositions. A better understanding of the chemical composition, along with free access to fingerprint databases, would allow stakeholders to use alcoholic beverages authentication information anywhere along the supply chain. A robust authentication method, such as LC-MS, could provide powerful molecular information for use in authentication controls, offering an improvement over current basic analyses and sensory evaluation. In conclusion, it's essential to improve the scientific knowledge concerning the chemical fingerprint of alcoholic beverages, the correlation between authenticity traits and their chemical fingerprint, and the current state-of-the-art methods used for their authentication. This literature review will go over the definition and current state of food fraud, especially regarding alcoholic beverages. It will then explore the chemical composition of alcoholic beverages, the production methods, and the resulting fate of authentication markers. Finally, analytical methods used in alcoholic beverages authentication will be described. This would identify gaps in the current understanding and improve the accuracy of authentication methods, thereby protecting consumers and promoting authentic products.

2.1 DEFINITION OF FOOD AUTHENTICITY AND FOOD FRAUD

While food fraud is not a novel concept, the active scientific pursuit to detect and dismantle it has gained increased attention over the last few decades. This is largely due to several high-profile scandals, such as the poisoning related to diethylene glycol in wine in Austria in 1985 (Holmberg 2010), the 2008 Chinese milk scandal (Pei, Tandon et al. 2011), and more recently the appalling seafood fraud happening worldwide (Silva, Hellberg et al. 2021). These are only a few of the biggest scandals that occurred in the last decades, but are sufficient to highlight the fact that food fraud is a challenging problem of essential importance. The concept of Food Integrity developed by the Global Food Safety Initiative is divided into two categories: intentional adulteration which is a crime and unintentional adulteration which is considered accidental (Figure 2.1) (GFSI 2014, Goethem and Elliott 2018). The category of intentional adulteration includes food fraud, which is motivated by economic gain, and food defence, which is driven by the desire

to cause harm. The unintentional adulteration category includes food quality and food safety. The food industry and the scientific community have worked hard in the last decades to meet consumer demands for wholesome food, which includes food quality, food safety as well as food security (Elliott 2014, Manning 2017).

Food quality is based on the attributes and characteristics of food products that are valued by consumers, like taste, appearance, texture, and nutritional content (Petrescu, Vermeir et al. 2019). These factors are influenced by agricultural practices, production methods, processing techniques, and storage conditions. Promoting food quality means meeting consumer expectations by following the established industry standards. This ensures the preservation of the integrity of food products. Food security refers to the balance between food availability, consumption patterns, and production capacity at national and international levels (IFPRI). It is therefore not of importance in this review as it does not pertain to food fraud directly. However, in some cases of food fraud, like the Chinese milk scandal, some food becomes unavailable for a certain period following the exposure of the fraud due to recalls or can be fearfully avoided (Li, Wang et al. 2021). Food safety involves chemical, microbiological, and physical hazards that can be accidentally present in food due to various factors like agricultural practices, food processing, food packaging, and storage conditions and cause foodborne illnesses (Fung, Wang et al. 2018). Food authenticity can impact both food safety and food quality depending on the type of fraud or adulteration (IFPRI , Goethem and Elliott 2018, Spink 2019). Various factors are to be considered in food authenticity such as animal/botanical origin, geographical origin of ingredients, agricultural practices, production method, processing, packaging, and, most commonly, the presence of adulterants (Figure 2.1). Additionally, some food products are regulated by various guidelines and legislation to provide a ‘’standard of identity’’ (Goethem and Elliott 2018). It is possible to identify discrepancies from this reference standard as potential adulterations, which is food fraud. Food authenticity, on the other hand, is defined as a product of genuine origin, both biological and geographical, compliant with agriculture, production, and processing practices, free of intentional adulteration, as well as legally produced and sold. This review will consider all types of adulteration, with a particular emphasis on origin authenticity, including geographical and biological origins. The following review will endorse the same definition for food fraud as in the Food Integrity Handbook, which is the following: *’’an action intentionally causing a mismatch between food product claims and actual food product characteristics, either by deliberately*

making claims known to be false or by deliberately omitting to make claims that should have been made'' (Goethem and Elliott 2018).

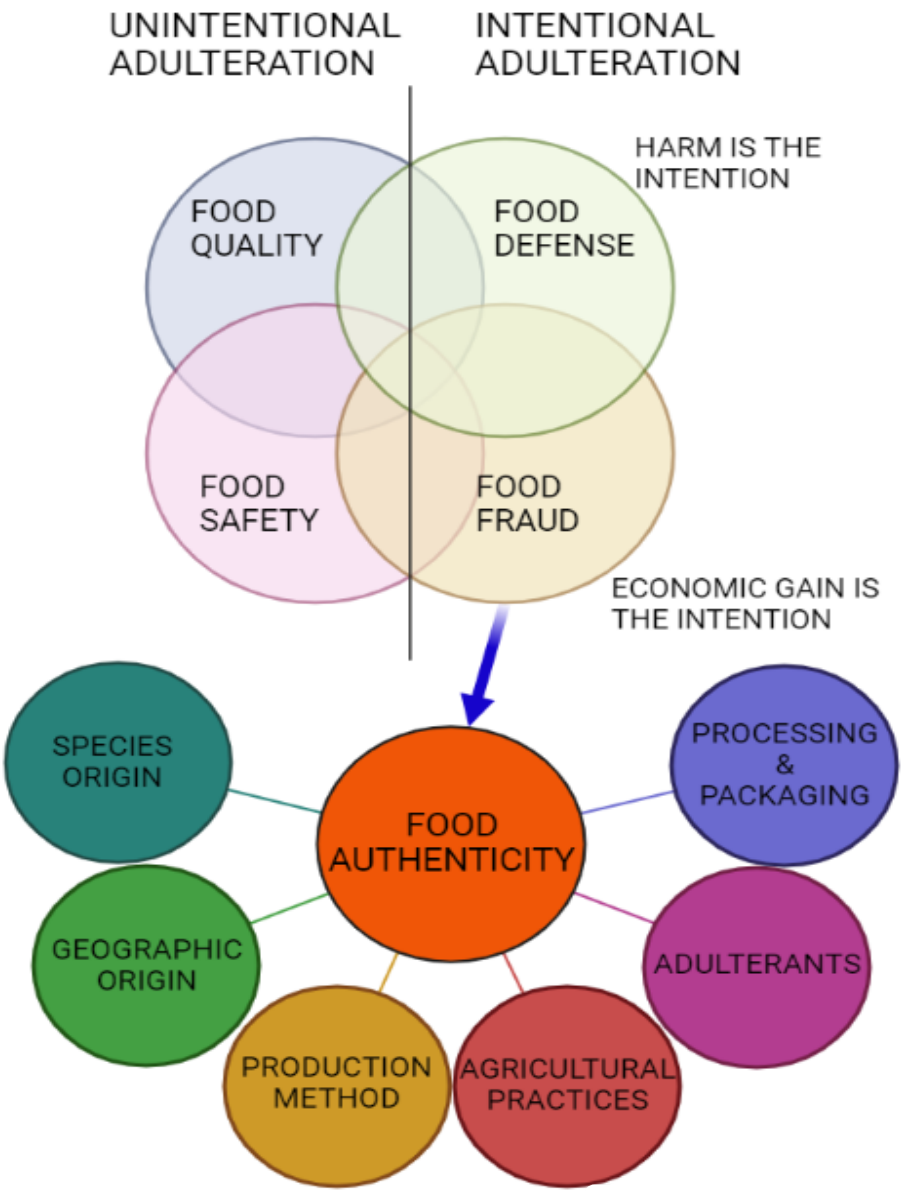


Figure 2.1. Food integrity englobes food quality and food safety under unintentional adulteration and food defence and food fraud under intentional adulteration. Food defence and food fraud are distinguished by the motivation behind adulteration, either intention to harm for the former or economic gain for the latter. Food authenticity is conditional on various factors such as animal/botanical origin, geographical origin of ingredients, method of production, agricultural practices, presence of adulterants, processing, and packaging (GFSI 2014, Hassoun, Måge et al. 2020).

2.1.1 Authenticity and Fraud Issues

To further define food fraud, the GFSI proposed dividing intentional adulterations into three subsets: adulteration, grey market, and counterfeiting. This section will discuss these subsets and provide examples to illustrate their characteristics. The GFSI categorizes adulteration into four types: substitution, dilution, unapproved enhancement, and concealment (GFSI 2014, GFSI 2018). Figure 2.2 briefly describes these types in the context of general food fraud. Substitution involves replacing a part of the food (nutrient or ingredient) with another, often of lower quality. Examples include substituting olive oil with seed oils, ethanol in alcoholic beverages with methanol, honey with cheaper sugars such as sucrose, or spices with sawdust, starch, sand, or chalk powder (Soares, Amaral et al. 2017, Ellis, Muhamadali et al. 2019, Kakouri, Revelou et al. 2021, Pacholczyk-Sienicka, Ciepielowski et al. 2021). Dilution is similar to substitution but occurs when a portion of liquid foods is mixed with lower quality or cheaper liquid, often water. Unlike substitution, dilution does not remove any part of the authentic food but merely dilutes it to increase volume and thus profits (Schelezki, Deloire et al. 2020). Common examples include adding water to milk or wine (Holmberg 2010, Nascimento, Santos et al. 2017). Unapproved enhancement involves adding undeclared compounds to enhance a food's quality. Examples include adding melamine to milk and infant powder to increase the nitrogen content, such as in the previously mentioned Chinese milk scandal, sugar to wine to increase alcoholic strength, and glycerol to wine to provide a thicker mouthfeel (Jung, Jaufmann et al. 2006, Holmberg 2010, Müller, Zhong et al. 2021). Concealment refers to hiding the low quality of food or ingredients, such as using hormones to make sick animals appear healthier or adding food colourings to fruits to cover bruises and blemishes. These categories often overlap, and fraud can remain undiscovered or have a minimal impact on health. However, fatal cases are not unheard of, particularly in the context of illicit alcoholic products where the addition of harmful adulterants like methanol to bypass religious restrictions, taxation, or for economic gains have led to several life-threatening incidents. Moreover, food allergies pose a significant public health concern as the substitution of ingredients, can result in severe allergic reactions or even death (Walker and Gowland 2015, Manning and Kowalska 2021). Grey market refers to the unauthorized sale of food products. Examples include diverting products from legitimate markets towards illegitimate ones (black markets, online platforms), not reporting overproduction and selling of excess food, and theft. In wine appellations, where production is highly regulated, or with fishing quotas, wineries or fishermen may sell the

regulated limit of food legally and sell the surplus without reporting it. Counterfeiting is a type of fraud that constitutes a crime on intellectual property by copying any aspects of a food product. It is common in contexts where specific or branded food products fetch high prices, such as high-end wines and spirits (Kuballa, Hausler et al. 2018, Fougere, Kaplan et al. 2020). The infamous Rudy Kurniawan scandal in the wine industry is one example (Fougere, Kaplan et al. 2020). Counterfeiting can also occur with regular brands, where counterfeit product are sold at a discount, relying on high-selling volume (Kuballa, Hausler et al. 2018). Mislabelling encompasses providing false information on labels to deceive consumers. Common forms of mislabelling include misrepresenting agricultural practices, geographical origin, biological origin, vintage, age, and processing techniques (Panossian, Mamikonyan et al. 2001, Holmberg 2010, Food Safety Authority of Ireland (FSAI) 2013, Khaksar, Carlson et al. 2015, Soares, Amaral et al. 2017). Although misrepresentations of geographical or biological origin may fall under substitution, mislabelling, or counterfeiting, they can also be considered as a separate entity labelled simply as origin adulteration (Esslinger, Riedl et al. 2014). Given the prevalence of fraudulent origin claims in alcoholic beverages, numerous studies have been conducted on origin authentication (Roullier-Gall, Lucio et al. 2014, Mannina, Marini et al. 2016, Roullier-Gall, Signoret et al. 2018, Karabagias, Karabagias et al. 2021, Phan and Tomasino 2021, Tzachristas, Dasenaki et al. 2021). Consequently, the following review will occasionally employ the term “origin fraud” over substitution, mislabelling, or counterfeit.

Distilled alcoholic beverages are frequently associated with fraud practices. One common fraud method is counterfeiting, which involves refilling empty bottles of high-end products with cheaper alternatives, a practice particularly prevalent in bars and similar establishments where regulatory oversight is less stringent than at import/export borders (Aylott 2013, Teodoro, Pereira et al. 2017, Kuballa, Hausler et al. 2018). This form of counterfeiting extends to premium beverages, similar to high-end wines, where inauthentic products are sold as genuine, sometimes fetching astronomical prices, ranging from hundreds of thousands to millions of dollars (Bennett 2017, Fougere, Kaplan et al. 2020). Other types of fraud may include the dilution of genuine spirits with cheaper ones, synthetic alcohol, or even methanol in extreme cases, posing several health risks and potential fatalities (Lachenmeier, Schoeberl et al. 2011). Fraudsters may also employ enhancement adulteration, using additives like colourings, flavourings, and sweeteners, to mimic the sensory attributes associated with a particular brand or beverage (Aylott 2013). These additives

often serve to replicate the effects of prolonged wood contact during maturation and aging, with caramel, wood, and tea extracts commonly used to simulate aging (Teodoro, Pereira et al. 2017).

Figure 2.2 Types of food fraud (GFSI 2018)

Mislabelling regarding geographical and botanical origins is another common authenticity issue (León-Rodríguez, Escalante-Minakata et al. 2007, Aylott 2013, Wiśniewska, Dymerski et al. 2015, Wiśniewska, Śliwińska et al. 2016, Wiśniewska, Boqué et al. 2017, Power, Néill et al. 2020). For instance, whiskies must adhere to stringent regulations to qualify for designations like Scotch whisky (The Scotch Whisky Regulations 2009), Irish whiskey (Food Safety Authority of Ireland (FSAI) 2019), or Bourbon whiskey (27-CFR-5.22 2008), necessitating specific ingredients, proportions, and country of origin. Similar regulations apply to other spirits such as tequila (Comité Consultivo Nacional de Normalización de Seguridad al Usuario 2012), gins (Aylott 2013), and brandy (Bureau National Interprofessionnel du Cognac (BNIC) 2018). Europol reported the seizure of over 50,000 litres of adulterated distilled spirits across multiple countries in 2016 during the OPSON V operation, and in 2021, approximately 50,000 litres of whiskey adulterated with colourants during the OPSON X operation (Europol - OPSON V 2016, Europol - OPSON X 2021). The extent of adulterated or counterfeited alcoholic beverages varies by country. In Russia, for example, spirits constituted half of the illicit alcohol market, equivalent to 10.2% of the total alcohol market (Euromonitor International 2015). A study on alcoholic beverages fraud

issues in South and Central America and Africa revealed that illicit alcohol accounted for 20-30% of the total alcohol market in most Central and South American countries and 50-65% in most African countries (International Alliance for Drinking Responsibly (IARD) 2018). Canada stands out with diverse unrecorded alcohol consumption, including home production, surrogate alcohol, smuggling, and cross-border shopping (Rehm, Kailasapillai et al. 2014). A study cited in a systematic review of these unrecorded alcohol consumption practices estimated that the illegal black market for alcoholic beverages in Canada resulted in $800 million in lost sales during the 1990s (Room and West 1998). Unfortunately, a comprehensive contemporary analysis of lost sales due to alcohol fraud in Canada appears to be lacking in the current literature. Ensuring authenticity is a significant challenge given the prevalence of fraud in the distilled spirits industry. The various types of fraud pose serious economic, social, and health consequences. The industry, regulatory authorities, and researchers must continue devising more effective strategies to detect and prevent fraud in this sector.

Wine, recognized globally as one of the most highly regulated food commodities, is at the same time one of the most sought-after beverages, leading to its unfortunate distinction as one of the most adulterated alcoholic beverages (Europol - OPSON IX 2020, Europol - OPSON X 2021). One prevalent adulteration is unapproved enhancement through chaptalization, where sugar or concentrated grape must is added to increase the wine's ethanol content, resulting in an inflated market value. Another common form of wine fraud is substitution to increase the volume produced. For instance, blending high-end wine with must coming from lower-quality grapes such as in the Brunellopoli case (Cavicchi and Santini 2011), or adding wine from inferior regions to reputable appellations. In regions cultivating non-*Vitis vinifera* species, these hardier, cheaper grapes, are sometimes substituted for approved grapes. Although not inherently illegal, this practice must be accurately disclosed on the label and generally precludes any appellation mention (Müller, Zhong et al. 2021). Substitution is often inextricably linked to mislabelling fraud issues. Mislabelling of origin is a prevalent form of wine fraud, driven by the intrinsic value associated with specific varietals or appellations. This incentivizes dishonest labelling practices, attributing higher-value names to the product (Takeoka and Ebeler 2011). Simultaneously, the burgeoning demand for organic products, stemming from ecological and health considerations, has instigated additional opportunities for fraudulent activities, especially considering the typically higher price points of organic wines (Abraben, Grogan et al. 2017, Vigar, Myers et al. 2020). Despite this,

studies focused on organic alcoholic beverage authentication remain limited, although MIR techniques have proven successful in distinguishing between conventional and organic Australian wines (Cozzolino, Holdstock et al. 2009). Vintage authentication is notoriously challenging, with radioisotopic dating offering approximate age determination, but sometimes falling short in detecting fraudulent mislabelling due to significant seasonal variations. Such vintage mislabelling is often complementary to origin mislabelling or counterfeiting (Godelmann, Fang et al. 2013, Geană, Ciucure et al. 2019, Merkytė, Longo et al. 2020). Concerning illegal additives, the use of aluminum to brighten a wine's colour, or berry extracts to darken red wines has been documented (Lachenmeier 2016). To imitate the effects of extensive cask aging, oak sawdust is often fraudulently used (Lachenmeier 2016). Historically, toxic lead salts were illicitly used to sweeten and clarify wines, a practice now virtually extinct due to health concerns (Lachenmeier 2016). Glycerol, naturally present in wine, contributes to a soft and full mouthfeel and signifies peak grape ripeness. Consequently, synthetic sources of glycerol are frequently added to wines, despite being an illegal additive in most winemaking regions (Müller, Zhong et al. 2021). These additives improve the perceived quality of the wine, permitting higher price points for cheaper wines. As with chaptalization, these practices are not universally illegal, reflecting the complex regulatory landscapes. The most high-profile aspect of wine fraud involves counterfeiting, particularly of premium or rare wines that command steep prices, presenting a lucrative opportunity for fraudsters (Fougere, Kaplan et al. 2020). Notably, fraudulent activities in the European Union's wine industry implicated more than 1,000,000 litres of wine and exceeded €1,200,000 in the latter half of 2020 alone (Popîrdă, Luchian et al. 2021). The implications of wine fraud are not solely economic as they also impact social and health aspects. Consumers who purchase high-quality wines expect transparency, and fraud activities are detrimental to the trust between producer and consumer. Health risks arise when hazardous unapproved additives are used, and their unknown or proven harmful effects on human health pose a threat. Therefore, the wine industry, regulatory agencies, and researchers must collaborate and improve current detection methods and prevention regarding wine fraud. This would ultimately benefit both the wine industry and wine consumers.

While beers, ciders, meads, or other similar beverages may not be as highly valued as wine or spirits, they are not exempt from fraudulent practices. Beers have been found to be adulterated in several ways, including blending with inferior brands, mislabelling of origin, and substitution with methanol (Mattarucchi, Stocchero et al. 2010, Fernández Pierna, Duponchel et al. 2012,

Mannina, Marini et al. 2016, Pereira, Amador et al. 2016, Associated Press 2020). Fraud in ciders primarily involves the substitution of apple varieties approved for PDOs with cheaper alternatives, constituting a type of origin fraud (Alonso-Salces, Guyot et al. 2005, García-Ruiz, Moldovan et al. 2007). Mead, a beverage that is infrequently discussed in fraud-related literature compared to other alcoholic beverages, is nonetheless susceptible to fraudulent practices. Given that honey, a key ingredient in mead, is among the most adulterated food products, it is plausible to infer that mead may also be subject to various forms of adulteration (Soares, Amaral et al. 2017). Documented fraudulent practices include the substitution of honey with molasses or syrups, the addition of ethanol to honey without any fermentation taking place, and the unapproved enhancement of mead with flavouring agents such as vanillin to mask off flavours (Česlová, Pravcová et al. 2022). Furthermore, numerous local alcoholic beverages receive scant mention in the literature, suggesting a lack of research into the authenticity of these products. Data from the JRC's monthly food fraud reports show that, since 2016, wine and spirits together account for more than 85% of all reported cases related to alcoholic beverages fraud. This leaves less than 15% of reports related to beers, meads, ciders, and ready-to-drink beverages (European Commission 2022) (Figure 2.3). This distribution may underscore the need for increased attention and research on potential fraud in these lesser-studied alcoholic beverages.

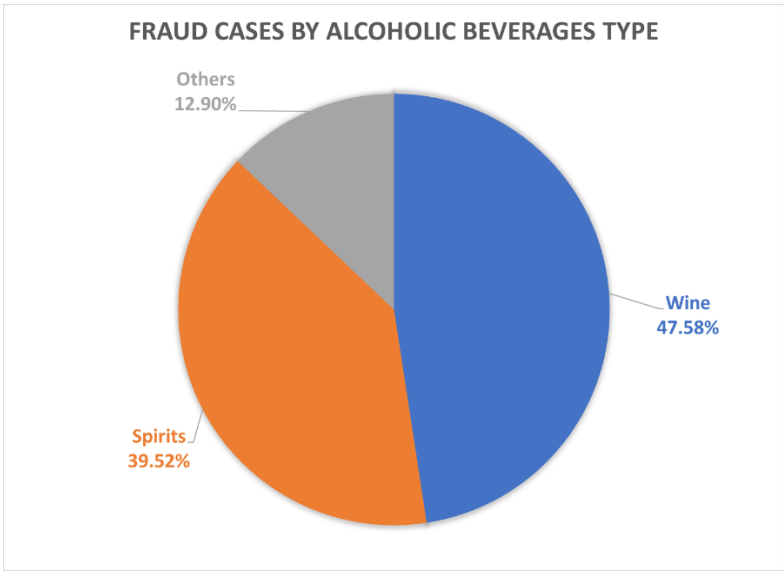


Figure 2.3: Reported food fraud cases related to alcoholic beverages from September 2016 to August 2022 in a pie chart categorized by affected alcoholic beverages. The categories are wine (blue, 49.14%), spirits (orange, 38.79%), and others (grey, 12.07%) which includes beers, ciders, and ready-to-drink beverages. All cases have been reported in the Knowledge Centre for Food

Fraud and Quality (KC-FFQ) a subgroup of the European Commission. The monthly food fraud report is drafted by the Joint Research Centre (JRC) Unit F.4 "Fraud Detection and Prevention" of Geel in Belgium, with the support of the JRC Unit I.3 "Text and Data Mining" of Ispra in Italy.

2.1.2 Consequences of Food Fraud

Fraudulent practices involving alcoholic beverages can have severe consequences, spanning economic, social, ethical, and health aspects. For instance, significant negative repercussions emerged from the infamous "Brunellopoli scandal", where Sangiovese grapes used to make Brunello di Montalcino DOCG wines were fraudulently substituted with cheaper grapes, tarnishing the appellation's reputation, diminishing consumer trust in Italian wines, and promoting unfair competition amongst producers (Cavicchi and Santini 2011). Further notable cases include the Rudy Kurniawan scandal starting early 2000s until his arrest in 2013 (Fougere, Kaplan et al. 2020), which involved a major counterfeiting scam where cheap wines were repackaged in famous vintage bottles and sold at premium prices, and the Glen Vodka scandal in the UK, where authorities seized 9,000 bottles of fake Vodka in 2019 (Lin and Salcido-Keamo 2021). A survey of industry stakeholders revealed complex view on the sharing of food fraud-related information. While most stakeholders see value information sharing, persistent drawbacks like cost, distrust of information shared by others, and increased workload make many reluctant to participate in such programs (Minnens, Lucas Luijckx et al. 2019). Instances of fraud have also had grave health implications. For instance, counterfeit alcohol has been known to contain dangerous amounts of harmful chemicals such as methanol, higher alcohols, ethylene glycol, and toxic metals, leading to numerous fatalities worldwide due to poisonings from illegally produced alcoholic beverages (Lachenmeier, Schoeberl et al. 2011, Pál, Muhollari et al. 2020, Manning and Kowalska 2021). In the Austrian Wine scandal of 1985, several wineries added diethylene glycol to their wines, resulting in significant harm to consumers and a near-collapse of the Austrian and parts of the German wine industry (Holmberg 2010). Premium appellations, such as Champagne and Cognac, have been implicated in multiple lawsuits over the years concerning the legal use of the appellation on labels, reflecting their importance in consumers' purchasing decisions (Pegan, Vianelli et al. 2020). Whiskey, particularly Bourbon and Scotch, have also been major targets of counterfeiting and fraud due to their high value (Collins, Zweigenbaum et al. 2014, Wiśniewska, Dymerski et al. 2015, Power, Néill et al. 2020). Operation OPSON, a Europol-INTERPOL joint initiative targeting food fraud, has identified alcoholic beverages as a priority, seizing millions of litres of counterfeit

or adulterated alcoholic beverages (Europol - OPSON IX 2020, Europol - OPSON X 2021). However, these represent only a fraction of all fraud cases involving alcoholic beverages, such as those illustrated in Figure 2.3. In summary, food fraud, particularly in the context of alcoholic beverages, presents itself in various forms and has significant economic, social, and health impacts on individuals, industries, and society.

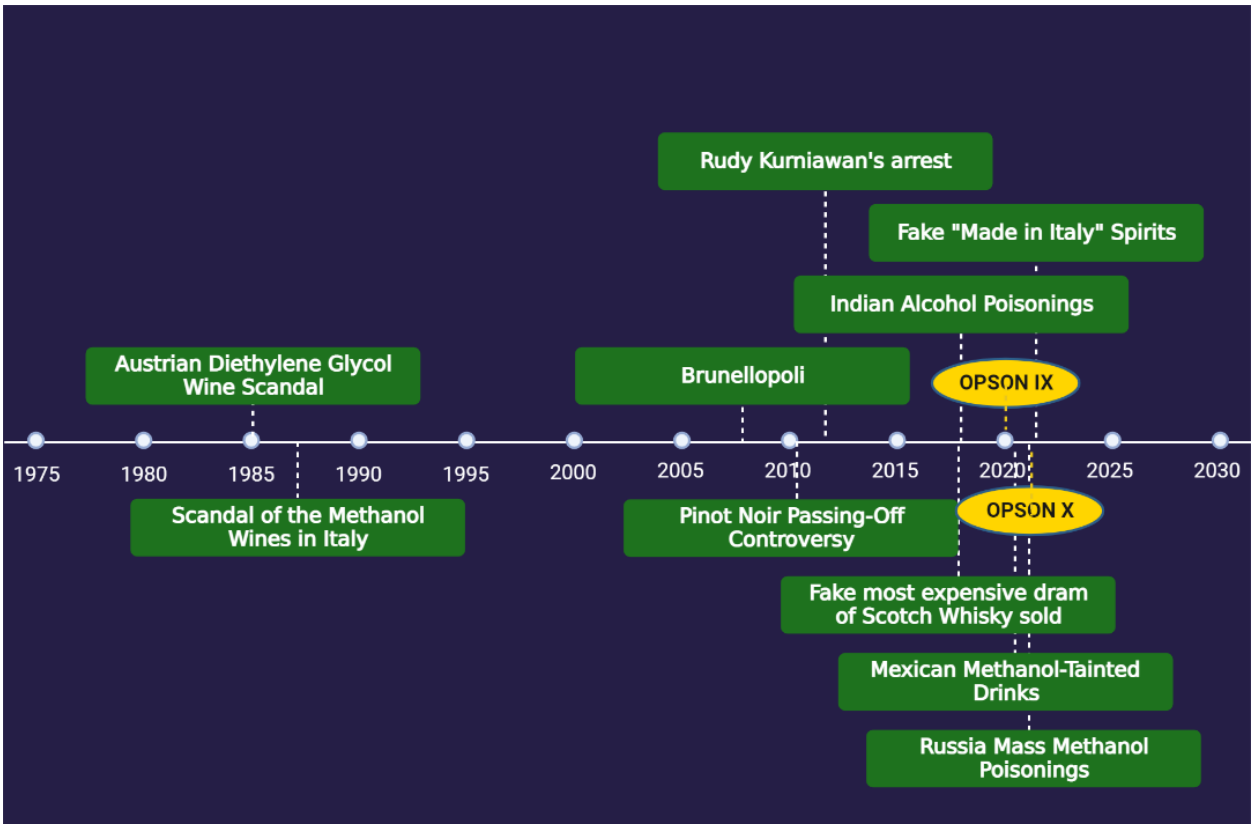


Figure 2.4 Timeline of key recent scandals or frauds involving alcoholic beverages (Holmberg 2010, Cavicchi and Santini 2011, Takeoka and Ebeler 2011, Bennett 2017, Associated local press 2020, Europol - OPSON IX 2020, Fougere, Kaplan et al. 2020, Europol - OPSON X 2021, Pomarici, Corsi et al. 2021, ANSA 2022, Balmforth 2022, Vaibhav Vikas 2022).

2.1.3 Consumer Perceptions and Expectations of Food Authenticity

Food fraud has wide-ranging repercussions beyond economic impact, leading to consumer distrust in the food industry and regulatory bodies, and often fostering feelings of anxiety and powerlessness (Charlebois, Schwab et al. 2016). Despite media attention, consumer awareness remains low, but a growing sentiment calls for stronger demand for food integrity (Charlebois, Juhasz et al. 2017). Surveys reveal that many consumers believe they have purchased fraudulent products, particularly in liquids like oils and alcoholic beverages. Trust in food authenticity varies,

and there is a desire for more information about food fraud issues (Ingrid Peignier 2017). Younger adults show more confidence in regulations but are less informed and less willing to change purchasing habits. While these findings are region-specific, a global knowledge gap exists and affects consumer willingness to pay, leading to economic implications (Spink 2019, McCallum, Cerroni et al. 2021). Increasing consumer knowledge and developing comprehensive techniques to prevent food fraud are essential. In understanding authenticity, consumers associate terms like accuracy, integrity, legitimacy, and originality with various meanings depending on the context (Nunes, Ordanini et al. 2021). Authetnic food is viewed as true to its origins, pure, compliant, and produced without solely economic motivations. Enhancing consumer knowledge can increase awareness and benefit both consumers and producers. Producers can also gain by presenting chemical fingerprints as proof of authenticity, promoting quality, and regional economic growth through origin certifications (Cavicchi and Santini 2011, Cassago, Artêncio et al. 2021). The sharing of information and more transparent practices could mitigate negative consequences associated with food fraud (Kendall, Clark et al. 2019).

Table 2.1 provides a comprehensive overview of authentication studies on various alcoholic beverages matrices, from the last five years. These studies serve a multitude of objectives, from the discrimination and classification of botanical or geographical origin to the detection of substitution and dilution fraud. Most researchers in this field have utilized analytical instruments such as spectrometry and spectroscopy. It's worth noting that the sample size (n) in these studies can vary widely, but very few have included over 200 samples. When it comes to data analysis the dominant methodologies often involve MVA, with PCA and DA being the two main statistical analyses employed.

MATRIX	OBJECTIVE	ANALYTICAL TECHNIQUE	N	DATA TREATMENT	
WINE	Varietal discrimination of Greek white wines by profiling and suspect screening	UHPLC-QTOF-MS	97	PCA, Random Forest	(Tzachristas, Dasenaki et al. 2021)
	Varietal discrimination of Douro red wines by profiling of anthocyanin profile	RP-HPLC-DAD	82	PCA, PLS-DA, CART, ANN	(Cosme, Milheiro et al. 2021)
	Geographical discrimination of Italian red wines by elemental fingerprinting	ICP-MS	639	K-CM (ANN and Fuzzy profiling)	(Bronzi, Brilli et al. 2020)
	Varietal identification assay development of Nebbiolo musts and wines using specific SNPs	SNP Genotyping	260	-	(Boccacci, Chitarra et al. 2020)
BEER	Discrimination of beers by style and production method using both a targeted and untargeted approach	NMR	31	PCA, OPLS-DA	(Palmioli, Alberici et al. 2020)
WHISK(E)Y	Untargeted analysis of diversely aged Bourbon whiskeys to identify potential authentication markers	FT-ICR-MS, HPLC-MS/MS	7	PCA	(Yang, Somogyi et al. 2020)
	Geographical discrimination of whisk(e)y using elemental fingerprinting	ICP-MS, MP-AES	68	MANOVA, CVA	(Hopfer, Gilleland et al. 2017)

RUM	Classification based on various production parameters and botanical origin using data fusion	NMR, HS-SPME-GC-MS, HRLC-MS	24	CV-ANOVA	(Belmonte-Sánchez, Romero-González et al. 2020)
FRUIT/MARC SPIRITS	Classification based on geographical and botanical origin using FT-Raman spectroscopy	Raman	97	Machine Llearning	(Magdas, David et al. 2022)
	Origin classification of Grappa and other Italian spirit using GC-MS and M/NIR spectroscopy	HS-SPME-GC-MS, FT-NIR, FT-MIR	75	SOPLS-DA, SO-CovSel-LDA	(Giannetti, Mariani et al. 2020)
VODKA	Comparison of NMR and IRMS for the detection of substitution fraud in Polish Vodka	NMR, IRMS	30	-	(Ciepielowski, Pacholczyk-Sienicka et al. 2019)
	Detection and quantification of dilution (water) and substitution (methanol) fraud	E-tongue	69	PCA, ANN	(Marenco, de Oliveira et al. 2021)

Table 2.1: Authentication Studies on Alcoholic Beverages.

2.1.4 Strategies for Detecting and Preventing Food Fraud

Fighting food fraud requires a multifaceted approach. The most common strategy involves investigations by food control authorities and other fraud-related law enforcement agencies. This approach is heavily dependent on legal documentation, whistle-blowers, surveillance programs, and traceability (Spink 2019). Examples of such initiatives include OPSON (Europol - OPSON IX 2020, Europol - OPSON X 2021), and the TRACES platform proposed by the European Commission (European Commission 2003). Investigative processes are particularly crucial for detecting frauds that can't be identified through analytical techniques, such as with many grey market frauds. In these cases, the product itself is often not adulterated; instead, the fraud involves the use of illegal markets, tax evasion, and unofficial distribution channels for economic gain. Since the chemical composition of the food remains unaltered, there is no adulteration signature to analyze. Therefore, investigative techniques are essential for preventing grey market fraud. While many analytical techniques require expertise, lengthy analysis time, or further development, investigations by authorities still take precedence over science. Analytical techniques are primarily used to provide evidence, confirm suspicions, or authenticate collections by private parties. These tools are invaluable for detecting fraudulent products and accurately identifying the adulterants involved. Analytical methods extract information from the chemical composition of the analyzed products, such as light absorption and emission, atomic vibrations, isotopic ratios, isotope decaying rates, and atomic and molecular masses. This raw information is then converted into measurable data to confirm product authenticity.

In the context of alcoholic beverages, various other approaches to fraud detection and prevention exist. These include scrutinizing labels, logos, closure seals, production lot numbers, and other minute details, particularly for high-valued bottles that need to remain sealed to retain their value. Wineries and distilleries have also begun to develop protective measures for consumers against counterfeiting, such as smart labelling and holograms, invisible ink, and RFID, NFC, and QR communication technologies. These measures make authentic bottles more difficult to replicate and counterfeit ones easier to identify (Khalil, Doss et al. 2019, Lindley 2022). A relatively recent approach involves using blockchain technology to efficiently track the buying and selling of bottles and provide authentication information (Luzzani, Grandis et al. 2021). These measures lend credibility to the industry and provide some assurance to consumers that their

purchases are authentic. However, fraudsters are notoriously quick to adapt to new technologies and science, and it is only a matter of time before they devise new ways to deceive consumers. Therefore, the industry, authorities, and scientific community must remain vigilant and provide cutting-edge knowledge to protect consumers from fraudulent activities. The most effective strategy for tackling food fraud involves combining as many approaches as possible to cast the widest net. This is the only way to significantly improve the chances of preventing food fraud. When using analytical tools, the principle of measuring food and beverage authenticity involves comparing an experimental value or parameter to a reference or standard. If the experimental value falls within the control limit, the food is considered authentic or free from adulterants. These limits are set by analyzing authentic samples and represent the natural variation of a specific product. If the value falls outside these defined limits, it is considered problematic and may require additional investigation to confirm if it is fraudulent and to determine the source of adulteration (Esslinger, Riedl et al. 2014). Multiple methodologies paired with state-of-the-art analytical technologies are used to detect fraudulent products or to assess authenticity with accuracy beyond any reasonable doubt (Esslinger, Riedl et al. 2014). Traditional methods are based on TA, where selected analytes in suspicious samples are measured and compared to reference standards to evaluate if the analyzed analytes fall within or below certain limits (Cavanna, Righetti et al. 2018). NTA has shown promise in food authentication, using analytical technologies capable of providing exhaustive chemical fingerprints. The resulting analytical profile of the food provides authentication information on geographical origin, botanical or animal sources, potential adulterations, aging and spoilage, packaging material, contaminants, and many other parameters (Cajka and Fiehn 2016, Cavanna, Righetti et al. 2018). After analysis, other preventive measures come into play to deter fraud from happening, or at least to reduce its occurrence. From the perspectives of food science, social science, and business, it has been proposed that criminology can help understand the root cause of food fraud to further establish standards, certifications, and public policies to curb food fraud. Following these guidelines, supply chain management and enterprise risk management can act as the first line of defence in preventing food fraud (Spink 2019). The management of food fraud differs from that of food safety, in that it involves not so much risk analysis, but more of an evaluation and control over the vulnerabilities along the food supply chain (Spink 2019). Since fraudsters are actively trying to avoid detection due to the potential for significant economic gain, the goal of food fraud prevention is to reduce the opportunities for food fraud to a point where the

risk for the fraudsters is greater than the potential gain (Spink 2019). Promoting an interdisciplinary approach allows for the combination of the goals of every science field involved to simultaneously seize products, arrest fraudsters, judge and punish criminals, detect food fraud, identify fraudulent products, educate the public, and increase traceability and documentation. This enhances the risk for fraudsters since vulnerabilities in the food supply chain are significantly reduced, which in turn requires them to invest more money in their fraud scheme or jeopardize their criminal activities (Spink 2019). This literature review will highlight key aspects in the detection of adulteration in alcoholic beverages, an essential measure taken to protect consumers, but it is also important to realize that it only contributes slightly to preventing fraud from occurring initially. New strategies in analytical food science, notably untargeted analysis, permit the detection of adulteration with a broader spectrum, which decreases the probability of fraud going unnoticed. A by-product of this is that it makes it that much more difficult for fraudsters to succeed or to remain undiscovered. However, without a holistic approach targeted at identifying fraud opportunities and managing vulnerabilities, it appears impossible to eliminate food fraud, as it can't be expected to achieve conformity from fraudsters only from analytical testing (Spink 2019). The third section of this literature review will detail the various instruments and approaches to authenticate alcoholic beverages. The official techniques in the authentication of alcoholic beverages are isotope ratio analysis, elemental profiling, physicochemical analysis (ethanol content, pH), targeted profiling of authentication markers (shikimic acid, anthocyanins, congeners, sugars), and carbon dating (Elliott 2014). Emerging techniques are NTA mass spectrometry (MS) metabolomics, spectroscopic approaches, and DNA next generation sequencing (Elliott 2014). Non-targeted analysis (NTA) is relatively new in the context of food analysis and has proven very useful in authentication studies due to its broad and exhaustive scope of analysis. The advantage of using NTA is that it is harder for fraudsters to circumvent the analysis by modifying their adulteration process since many analytes are simultaneously measured, making it virtually impossible to avoid detection. Previously, food fraud prevention relied heavily on testing the physicochemical properties of products. However, these techniques have become outdated as fraudsters have found more sophisticated ways to adulterate food, making it necessary to implement new and improved authentication methods, such as fingerprinting (González-Pereira, Otero et al. 2021, Sarkar, Salauddin et al. 2022). In conclusion, preventing and detecting food fraud requires a comprehensive and multifaceted approach that combines investigative techniques,

advanced analytical methods, and proactive management strategies. By staying ahead of the curve and continually developing and implementing state-of-the-art techniques and strategies, the ability to protect consumers from fraudulent activities and ensure the integrity of the food supply can be significantly enhanced.

2.2 CHEMICAL COMPOSITION AND EFFECTS OF ADULTERATION ON SPIRITS

Understanding the chemical composition of authentic food and beverages is crucial to the task of identifying and quantifying potential authentication marker compounds. Prior knowledge of chemical composition can enhance the identification and characterization of adulterants in food matrices using analytical techniques like NTA, although it is not a requirement. (Gertsman and Barshop 2018). Adulteration has the potential to affect the physicochemical properties of other analytes, thus disturbing the equilibrium of the food matrix by favouring new chemical reactions to occur. Analytical tools are capable of detecting these chemical shifts (generation and degradation of compounds) and the data provided can reveal the type of adulteration present. (Gertsman and Barshop 2018). Table 2.1 highlights previous studies that characterized analytes affected by specific types of adulteration and demonstrated their potential as authentication markers.

2.2.1 Chemical Composition of Spirits: Influential Factors & Impact of Adulteration

Spirits are a type of alcoholic beverage defined by high alcohol content, usually above 15% v/v ethanol, and most commonly 35-50% v/v (Aylott 2013). Only ethyl alcohol of agricultural origin (EAAO), thus from products such as cereals, fruits, vegetables, and plants are allowed (European Commission (EC) 2008). Spirits have similar chemical compositions due to shared fermentation and distillation processes, with frequent reports of higher alcohols, esters, aldehydes, ketones, organic acids, fatty acids, terpenes, heterocyclic compounds, phenolic compounds, and nitrogen-, sulfur-, and oxygen-containing compounds. (Riu Aumatell 2012, Aylott 2013, Collins, Zweigenbaum et al. 2014, Tsakiris, Kallithraka et al. 2014, Quesada-Granados, Samaniego-Sánchez et al. 2016, Villanueva-Rodríguez, Rodríguez-Garay et al. 2016). As noted in Table 2.2, common cases of spirit adulteration are mislabelling (origin and ageing), counterfeiting, and substitution.

Whisky, with its roots in Scotland, Ireland, the US, Canada, Japan, and India, is made from cereal such as barley, corn, wheat, and rye (Aylott 2013, Roullier-Gall, Signoret et al. 2018). Its chemical profile comes from several steps including malting, kilning, mashing, fermentation, distillation, oak cask storage, and other steps like filtering. Each cereal introduces unique compounds, yet common classes of compounds such as carbohydrates, heterocyclic compounds, fatty acids, polyphenols, carbonyl compounds, and alcohols are often present (Lee, Paterson et al. 2001). Other factors like kilning duration, fermentation microorganisms, still types, and maturation conditions significantly impact the final product (Lee, Paterson et al. 2001, Power, Néill et al. 2020). Distillation is pivotal, providing specific compounds from the mash to the end product like ethanol and certain acids, esters, and higher alcohols (IARC 1988, Lee, Paterson et al. 2001). Whiskies from different geographical regions possess distinct chemical compositions due to terroir, an environmental factors that affect products (Roullier-Gall, Signoret et al. 2020). Regulations in major whisky-producing countries ensure quality and authenticity through specific requirements, indicated via labelling. For example, Irish whiskeys undergo triple distillation processes and must contain at least 25% malted barley (Power, Néill et al. 2020). A study highlighted the use of higher for authenticating Irish whiskeys, distinguishing them from American and Scotch whiskeys (González-Arjona, González-Gallero et al. 1999). American whiskeys, on the other hand, are distilled from specific grain mashes and aged in new charred oak (27-CFR-5.22 2008). A study identified key compounds unique to American whiskeys useful for brand identification and counterfeit detection (Collins, Zweigenbaum et al. 2014). Scotch whisky, made from 100% malted barley, is double-distilled in copper pot stills, and aged at least 2 years in oak, with the traditional use of peat bogs for drying malted barley, imparting unique compounds (Lee, Paterson et al. 2001). Factors like grain composition and aging time further influence the chemical profile of whiskies (Wiśniewska, Boqué et al. 2017, Roullier-Gall, Signoret et al. 2018, Kew, Goodall et al. 2019). The unique chemical composition of whiskies, thus stems from various factors including geographical origin, raw materials, processing, and aging.

Eau-de-vie, a fruit-distilled spirit, becomes brandy when aged in wooden barrels. This spirit can be named based on geographical origin or production method, such as with Cognac, Armagnac, and others. Pomace brandy is made from fruit pressing residue and includes examples such as Grappa and Orujo (Aylott 2013). Chemical compounds in brandies are mostly derived

from the fruits, with stone fruits exhibiting higher levels of benzaldehyde, phenolic compounds, acetals, and aldehydes compared to apples or grapes (Ledauphin, Le Milbeau et al. 2010, Śliwińska, Wiśniewska et al. 2015). For instance, the use of grapes results in the presence of aliphatic alcohols (1-hexanol), free, oxidated, and glycosidic terpenes (linalool, nerol, geraniol, limonene, citronellol, rose oxide, myrcene), as well as aldehydes (furfural, benzaldehyde) and ketones (β-damascenone) (Tsakiris, Kallithraka et al. 2016, Matijašević, Popović-Djordjević et al. 2019, Raičević, Popović et al. 2022). These compounds, known as congeners, can provide important information regarding the authenticity and origin of the spirit (Martí, Busto et al. 2004, Arvanitoyannis 2010, Hermosín-Gutiérrez, Castillo-Muñoz et al. 2011). During fermentation, various higher alcohols are produced, including 1-propanol, isoamyl alcohols, and isobutyl alcohols, as well as carboxylic acids such as propionic acid and butyric acid (Tsakiris, Kallithraka et al. 2016). Fermentation also leads to the formation of various fatty acids caused by the activity of different microorganisms (Tsakiris, Kallithraka et al. 2016). Additionally, over 160 esters have been reported to be produced during fermentation, such as ethyl acetate, ethyl lactate, and others (Tsakiris, Kallithraka et al. 2016). The distillation process typically involves two separate distillations. During the first distillation, the wine is brought up to around 24-30% alcohol by volume and many non-volatile compounds are present in wine, like organic acids, salts, polyphenolic compounds such as tannins, and some minerals, are removed. The second distillation separates and discards the head and tail fractions to keep only the heart fraction, resulting in a distillate of around 70% alcohol v/v (Tsakiris, Kallithraka et al. 2014). During distillation the heightened temperature results in the formation of carbonyl compounds, higher alcohols, Maillard reaction products, furans, and phenols (Tsakiris, Kallithraka et al. 2016).

MATRIX	ADULTERATION	MARKERS	REFERENCES
WINE	Dilution (Water)	Esters, higher alcohols, ethanol and reducing sugars.	(Karabagias, Karabagias et al. 2021)
	Mislabelling (Oak Ageing)	Organic acids, polyphenols, aldehydes	(Matějíček, Mikeš et al. 2005)
	Counterfeiting (Vintage)	137Cs	(Hubert, Perrot et al. 2009)
	Mislabelling (Origin)	Higher alcohols, esters, ketones, fatty acids, polyphenols, phenols, aldehydes	(Chávez-Márquez, Gardea et al. 2022)
	Mislabelling (Origin, Vintage)	Na, Mg, Si, P, S, Cl, K, Y, U, Cr, Ni, Mn, Cd, As, Pb, Zn, Cs, Rb, Ca, Fe, Cu, V, Ba, Li, Sr, Co, Al, Be, Sb	(Capron, Smeyers-Verbeke et al. 2007, Forina, Oliveri et al. 2009, Martin, Watling et al. 2012)
SPIRITS	Counterfeiting	Alcohols, glycerol, organic acids	(Kuballa, Hausler et al. 2018)
	Substitution	Methanol, formic acid, ethyl formate	(Kuballa, Hausler et al. 2018)
BRANDY	Counterfeiting, Mislabelling (Oak Ageing)	Sinapaldehyde, syringaldehyde, coniferaldehyde	(Panossian, Mamikonyan et al. 2001)
	Substitution	Caramel	(Markechová, Májek et al. 2014)
ABSINTH	Substitution, Concealment	Absinthin, thujone, tartrazine, patent blue V, brilliant blue FCF	(Lachenmeier 2007)
WHISK(E)Y	Counterfeiting	Fatty acids, polyphenols, higher alcohols, ethyl esters	(Garcia, Vaz et al. 2013) (Teodoro, Pereira et al. 2017)
	Mislabelling (Origin), Counterfeiting	Aldehydes, fatty acids, polyphenols, phenols, higher alcohols, organic acids	(Collins, Zweigenbaum et al. 2014)
BEERS	Counterfeiting	Na+/K+ adducts of malto-oligosaccharides	(Pereira, Amador et al. 2016)
	Origin (Trappists)	Organic acids, alanine, adenosine, isopentanol, propanol	(Mannina, Marini et al. 2016)
	Mislabelling	Sugars, higher alcohols, organic acids	(Duarte, Barros et al. 2004)
AGAVE SPIRITS	Mislabelling (Oak Ageing)	Organic acids, polyphenols, aldehydes	(Muñoz-Muñoz, Grenier et al. 2008)
	Mislabelling (Origin)	Esters, higher alcohols, organic acids, fatty acids, ketones, aldehydes	(León-Rodríguez, Escalante-Minakata et al. 2007)
	Mislabelling (Origin + Oak Ageing)	Na, K, Cu, Ca, Mg, Fe, S, Zn, Sr	(Ceballos-Magaña, Jurado et al. 2009)
RUMS	Mislabelling (Ageing)	Organic acids, 5-HMF, aldehydes	(de Aquino, Rodrigues et al. 2006)

Table 2.2: Authentication and fraud markers by fraud type on various alcoholic beverages

Rum is mostly made from sugarcane derivatives, such as molasses and sugarcane honey, but can also be made from other sources such as beets (Aylott 2013). The spirit is primarily produced in Central and South America, but also has production sites in Europe, North America, and various islands including Madagascar, Hawaii, and the Canary Islands. These regions employ various production methods, ranging from processing raw sugarcane juice to using blackstrap molasses (Quesada-Granados, Samaniego-Sánchez et al. 2016). Sugarcane juice and molasses comprise a range of chemical compounds influencing the chemical composition of rum. These compounds can undergo chemical reactions during various stages of rum production, including fermentation, distillation, and aging, leading to highly complex compositions (Quesada-Granados, Samaniego-Sánchez et al. 2016). The fermentation mostly produces glycerol, ethanol, and higher alcohols (such as isoamyl and isobutyl alcohols), lactic and acetic acids, acetaldehyde, various esters, and fatty acids (Quesada-Granados, Samaniego-Sánchez et al. 2016). Factors that are most influential to their chemical composition include still type and distillations passes (Sampaio, Reche et al. 2008, Quesada-Granados, Samaniego-Sánchez et al. 2016). Pot still distillates typically contain more ethyl acetate and 5-HMF, while column still distillates are characterized by higher concentrations of benzaldehyde, acetaldehyde, isoamyl alcohol, ethyl carbamate, higher alcohols, and acetone (Reche, Leite Neto et al. 2007, Sampaio, Reche et al. 2008).

Agave spirits are distilled from various types of succulent plants from the genus Agave, including mezcal, tequila, and others. Mezcal can be made using a wide range of agaves, while tequila must be made only with *Agave tequilana* Weber v. azul (Cisneros 2001, Aylott 2013). Agave spirits are almost exclusively produced in Mexico, and Tequila is only made in the state of Jalisco, and restrictively in other states (Villanueva-Rodríguez, Rodríguez-Garay et al. 2016). Mexican laws specify that it must be made from at least 51% *A. tequilana*, the remaining being any other carbohydrates sources except for other agaves (Comité Consultivo Nacional de Normalización de Seguridad al Usuario 2012, Villanueva-Rodríguez, Rodríguez-Garay et al. 2016). Studies on Tequila have focused on its fructans and only limited literature describes its chemical composition (Lopez, Mancilla-Margalli et al. 2003, Waleckx, Gschaedler et al. 2008, Ávila-Fernández, Rendón-Poujol et al. 2009). The degradation of lignin during production forms compounds serving as authentication markers such as vanillin and syringaldehyde, pectin, fatty acids, homoisoflavanones, and various terpenes (Peña-Alvarez, Díaz et al. 2004, Peña-Alvarez,

Capella et al. 2006, Morales-Serna, Jiménez et al. 2010, Villanueva-Rodríguez, Rodríguez-Garay et al. 2016). Tequila production begins with cooking and juicing the piñas (the stem and basal parts of the agave leaves) to release the fermentable carbohydrates, which are then hydrolyzed to glucose and fructose by cooking. The cooking generates multiple compounds, including organic acids, higher alcohols, furans, aldehydes, ketones, norisoprenoids, aromatic compounds, terpenes, and others (Villanueva-Rodríguez, Rodríguez-Garay et al. 2016). Fermentation is then induced with indigenous or cultured yeasts, the former resulting in a greater diversity of chemical compounds (such as esters and terpenes), albeit to the detriment of ethanol production (Villanueva-Rodríguez, Rodríguez-Garay et al. 2016). The by-products of fermentation, in addition to ethanol, are higher alcohols, phenylethanol, ethyl esters, ketones, and aldehydes (Lachenmeier, Sohnius et al. 2006, Villanueva-Rodríguez, Rodríguez-Garay et al. 2016).

Vodka can be produced from a variety of fermentable carbohydrate sources, with the most common being grains such as wheat, corn, and rye. Vodka and most spirits before being aged or flavoured are considered to be neutral spirits due to their neutral flavour profile. Although vodkas made from other raw materials such as potatoes, cassava, honey, maple, whey, and fruits also exist, they have limited representation in the international market (Aylott 2013). Compared to other spirits, the raw materials used in the production of vodka have a lesser impact on its chemical composition due to the extensive distillation process made to increase the purity of ethanol, which leaves only minimal residual congeners (Aylott 2016). Artisanal vodka producers sometimes lower the distillation proof to increase the presence of congeners, thus enhancing the sensory profile related to the raw materials used. Similarly, the impact of the fermentation process on the final chemical composition also follows these principles. Due to the extensive distillation and filtration processes, vodka has a simpler chemical composition when compared to other spirits (Aylott 2016). For this reason, vodkas often require preconcentration to allow the detection of trace analytes. Adulteration can be more easily detected due to a lower variety of compounds, although geographical and botanical origins are harder to detect for the same reasons (Lachenmeier, Attig et al. 2003, Legin, Rudnitskaya et al. 2005, Ciepielowski, Pacholczyk-Sienicka et al. 2019, Marenco, de Oliveira et al. 2021). After distillation, most spirits are filtered but few are as filtered as vodka is. Producers claim various filtration methods to promote the purity of their spirit, ranging from the use of lava rock, coconut carbon, charcoal, and quartz crystals to gold, platinum, micron paper, pearls, and diamonds. These methods have been recognized by expert tasters to affect the

sensory properties of vodka, but the scientific literature on their impact on chemical composition is scarce. One study observed that higher-end filtration methods using deionization such as reverse osmosis and ion exchange resulted in a lower concentration of various anions. The analysis of the anionic composition has been demonstrated to be sufficient in the brand identification of various Russian and German vodkas (Lachenmeier, Attig et al. 2003). Besides filtration, another post-distillation step is the addition of colourings, flavourings, or sweeteners, which is more common in vodkas compared to other spirits and can significantly affect its chemical composition, leading to the requirement of mentioning flavoured vodka on its labelling in some countries (Aylott 2016). The most common approved additives are sugars, glycerol, citric acid, and propylene glycol (Ng, Hupé et al. 1996, Aylott 2016). Vodka is produced worldwide and regulations vary by country. In Russia, vodka must be made from grain or potatoes and the alcohol content must be 38-45% or exactly 50% or 56% (Legin, Rudnitskaya et al. 2005). In Canada and the US, it is defined as a neutral spirit distilled or treated after distillation, as to be without distinctive character, aroma, taste, or colour (27-CFR-5.22 2008, Regulations 2022). However, unlike Canadian regulations, the US allows the addition of sugar and citric acid without mentioning it on labels (27-CFR-5.22 2008, Regulations 2022). This has been demonstrated by the presence of 5-hydroxymethylfurfural (5-HMF) and triethyl citrate (TEC) in American vodkas, which result from sugar degradation and the reaction between citric acid and ethanol, respectively (Ng, Hupé et al. 1996).

Gin is a unique category of spirits because the type of raw ingredients used to produce ethanol is less important than the requirement to infuse juniper berries according to legal regulations (*J. communis*) (Aylott 2013). Foremost, gin is produced by infusing juniper berries in a neutral spirit (i.e., vodka), along with other botanicals, spices, and fruits (Riu Aumatell 2012). Different terms and designations, such as London Dry Gin, Distilled Gin, Compound Gin, Navy Gin, Plymouth Gin, or Genever, describe specific subtypes of gin and are governed by legal regulations. The European Commission defines gin as a spirit flavoured with *J. communis* and made from ethyl alcohol of agricultural origin (EAAO) with a minimum alcoholic strength of 37.5% alc./volume. Distilled gins must be produced by redistilling EAAO in the presence of *J. communis*, and London Dry Gin, a type of distilled gin, must not contain added sweeteners exceeding 0.1 g /L, and only water can be added after distillation (European Commission (EC) 2008). Compound Gins are produced by flavouring EAAO with *J. communis* and other ingredients using a maceration step. Plymouth Gin and Genever are spirits made with the infusion of *J. communis*, both of which have

PDO (Riu Aumatell 2012). While there is limited data on the differences in chemical compositions between distilled and compound gins, distilled gins are often considered to be of higher sensory quality (Riu Aumatell 2012). The use of a variety of botanicals in gin production results in the creation of a diverse range of flavours, reflecting cultural influences, regional flora, and specific aroma profiles. However, certain ingredients are predominantly used in gin production. *J. communis* is known to provide around 80 compounds to gin, mostly monoterpenes, oxygenated monoterpenes, diterpenes, sesquiterpenes, oxygenated sesquiterpenes, aldehydes, and alcohols (Vichi, Riu-Aumatell et al. 2007, Riu Aumatell 2012)». Of the monoterpenes, terpinene-4-ol, *p*-cymene, β-myrcene, γ-terpinene, α-pinene, and limonene represented more than 70% of the volatile fraction of gins (Vichi, Riu-Aumatell et al. 2007). Various factors affect the relative abundance of these compounds in gin, such as geographic location, altitude, plant age, and berry ripeness (Barjaktarović, Sovilj et al. 2005, Vichi, Riu-Aumatell et al. 2007). The chemical composition of *J. communis* essential oil has been characterized as containing possibly close to 200 compounds, therefore, many more compounds might be present in gins, albeit in trace amounts (Barjaktarović, Sovilj et al. 2005). Otherwise, coriander (*Coriandrum sativum*), orange (*Citrus sinensis*), cassia (*Cassia fistula*), orris root (*Iris florentina*), cardamom (*Elettaria cardamomum*), angelica root (*Angelica archangelica*), cinnamon (*Cinnamomum zeylandicum*), calamus (*Acorus calamus*), fennel (*Foeniculum vulgare*), aniseed (*Pimpinella anisuum*), lemon (*Citrus limon*), cumin (*Cuminum cynimum*), almond (*Prunus amygdalus*), and licorice root (*Glycyrrhiza glaba*) are all common ingredients in gin production (Riu Aumatell 2012). The individual ingredients used in gin production have vastly different chemical compositions, meaning each gin recipe will have a unique chemical profile. Additionally, these ingredients are influenced by their environment, and are susceptible to seasonal variations and geographical differences, resulting in fluctuation between batches of the same gins. Despite the importance of understanding the chemical composition of gins, few studies have been conducted in this area, and there is currently no established method for their authentication. Different infusion techniques for botanicals are used depending on the type of gin produced and the desired sensory profile, such as maceration in the distillate during redistillation, vapour infusion with hanging baskets in the still, and vacuum distillation (Riu Aumatell 2012, Hodel, Pauley et al. 2019). Maceration is the most traditional method and lets the botanicals steep in the EAAO while it is slowly being heated for redistillation (Riu Aumatell 2012). A study comparing vapour infusion and maceration techniques observed a

higher abundance of compounds using vapour infusion, the only exception being linalool, which was present at lower concentrations (Hodel, Pauley et al. 2019). On the other hand, vacuum distillation increases compound oxygenation and reduces the presence of undesirable heat-generated compounds (Greer, Pfahl et al. 2008).

2.2.2 Chemical Markers in Spirits as Indicators of Authenticity and Fraud

The data presented in Figure 2.5 highlights the fact that half of the reported fraud cases in spirits are associated with the grey market, with the majority being smuggling and contraband issues. Since the food matric often remains unaltered, it is practically impossible to identify grey market activity using analytical instruments. The remaining half of the fraud cases are dominated by substitution, which accounts for 28%, followed by counterfeiting at 15%, mislabelling at 6%, and a combination of dilution, concealment, and enhancement at 1% (JRC 2023).

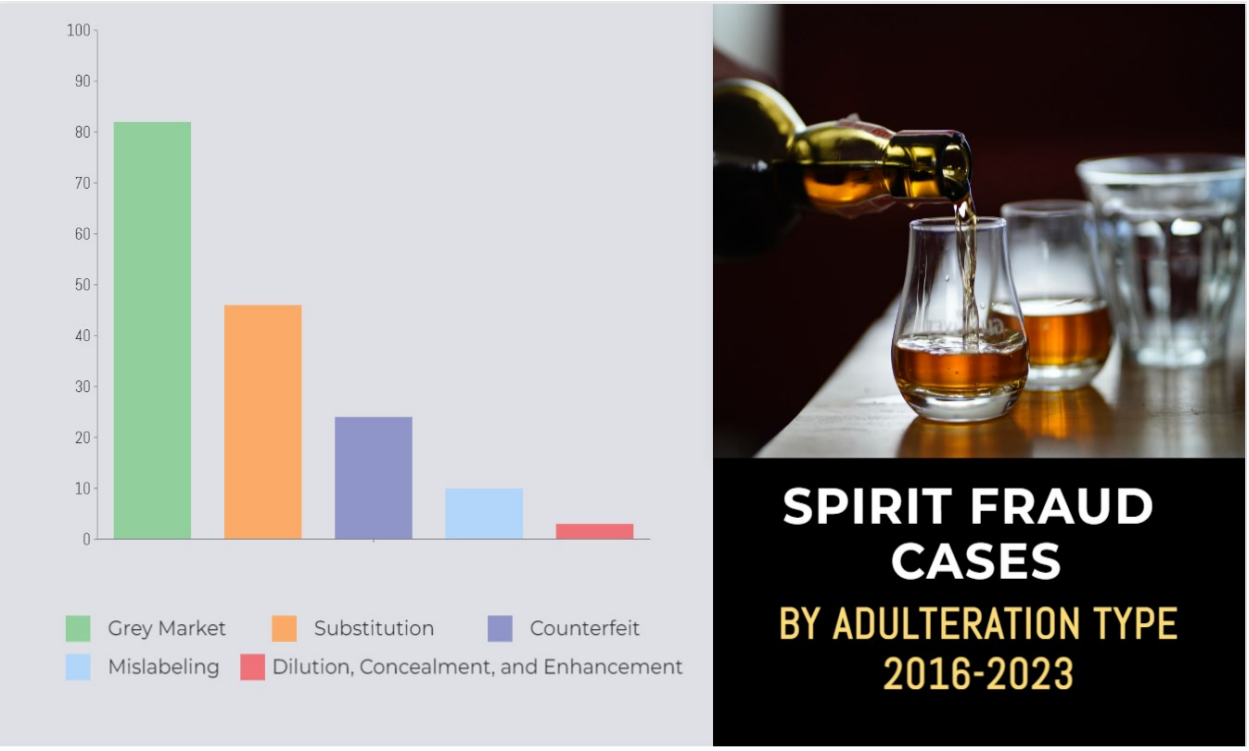


Figure 2.5 Breakdown of reported food fraud cases related to spirits, from September 2016 to January 2023 (*n = 165*). These cases are divided into categories: Grey market (50%, *n = 82)*, substitution (28%, *n = 46*), counterfeit (15%, *n = 24*), mislabelling (6%, *n = 10*), and a category combining dilution, concealment, and enhancement (2%, *n = 1* for each). These findings are sourced from the Knowledge Centre for Food Fraud and Quality (KC-FFQ), a subgroup of the European Commission. The monthly food fraud reports are prepared by the JRC Unit F.4 “Fraud

Detection and Prevention" based in Geel, Belgium, and supported by the JRC Unit I.3 "Text and Data Mining" located in Ispra, Italy.

Substitution is a common type of fraud in spirits, often involving the addition of methanol or synthetic ethanol (Tsakiris, Kallithraka et al. 2014). Synthetic ethanol can be produced using various methods, such as the hydration of ethene or the reaction of CO and H_2 under high pressure or with catalysts (Mohsenzadeh, Zamani et al. 2017, Kang, He et al. 2020). One way to detect synthetic ethanol is by analyzing its different oxygen isotopic ratios compared to ethanol of agricultural origin, using techniques such as IR-MS (Perini and Camin 2013). Other markers that can help detect synthetic ethanol include acetone, 2-butanol, crotonaldehyde, and degradation products resulting from the production of synthetic ethanol (Tsakiris, Kallithraka et al. 2014). Counterfeiting occurs when criminals produce fake versions of a specific spirit or brand. This type of fraud is often uncovered because of dubious labels, incomplete documentation, or when illegal production sites are investigated. In some cases, the authenticity of rare spirits sold at auctions or held in private collections may be questioned, necessitating scientific analysis to verify their origin. Mislabelling, such as falsely claiming geographical or botanical origin fraud or age statements, is a prime target for analytical methods as it is often difficult to detect by other means. To verify the identity of a spirit, specific markers mentioned in the previous sections on chemical composition can be used such as organic acids, higher alcohols, aldehydes, ketones, lactones, flavonoids, terpenes, phenolic compounds, and coumarins (Riu Aumatell 2012, Aylott 2013, Collins, Zweigenbaum et al. 2014, Tsakiris, Kallithraka et al. 2014, Quesada-Granados, Samaniego-Sánchez et al. 2016, Villanueva-Rodríguez, Rodríguez-Garay et al. 2016). For example, brandies made with the Muscat variety have higher linalool content, while those made with Chardonnay produce more 1-hexanol, and those made from Pinot Blanc have much lower concentrations of ethyl esters (Lukić, Banović et al. 2006). Similarly, apple, grape, or plums brandies have been authenticated using aldehydes, acetals, benzaldehyde, higher alcohols, and furanic compounds (Ledauphin, Le Milbeau et al. 2010). Spirits that have false age statements will mostly have discrepancies in the compounds extracted from wood by ethanol and their relative abundances, such as vanillin, sinapinaldehyde, coniferaldehyde, and syringaldehyde and their derivatives (Panossian, Mamikonyan et al. 2001, de Aquino, Rodrigues et al. 2006, Muñoz-Muñoz, Grenier et al. 2008, Collins, Zweigenbaum et al. 2014, Tsakiris, Kallithraka et al. 2014, Espinosa-Vega, Belio-Manzano et al. 2019). Diluting spirits with water or with neutral spirits can be relatively

easily detected due to a lower abundance of all chemicals normally present. Enhancement or concealment through the use of illegal additives such as colourings, flavourings, and sweeteners is usually prohibited, such as in Bourbon production (Collins, Zweigenbaum et al. 2014). Known additives can also be easily detected with the use of traditional targeted analysis techniques. For example, a study on brandies has characterized the furfural:5-HMF ratio as a potential authentication marker for spirit caramel E150, a commonly added colouring agent (Granados, Mir et al. 1996). Dilution, concealment, and enhancement, have been the least reported adulteration types in the JRC fraud report since 2016 (JRC 2023).

2.3 ANALYTICAL TECHNIQUES TO AUTHENTICATE ALCOHOLIC BEVERAGES AND DETECT FRAUD

Authenticating the origin and detecting adulterants in alcoholic beverages is challenging due to their complex chemical compositions and the wide range of adulterants, many of which are still unknown. Two major analytical strategies, or workflows, can be applied to authenticate alcoholic beverages and food in general. The first one is based on targeted analysis (TA), also called profiling. It focuses on the analysis of a specific set of known metabolites (Cavanna, Righetti et al. 2018). TA is usually applied on samples where the goal is to quantify or detect the presence of known compounds. The second one, non-targeted analysis (NTA) is a method that analysis at once a massive part of the chemical composition of a sample, without prior knowledge of the matrix composition, or the need for specific targets (Broadhurst, Goodacre et al. 2018). NTA is usually applied with the objective of discovering new compounds or markers or establishing a chemical fingerprint. Therefore, both TA and NTA are used with different objectives in mind and, accordingly, produce vastly different results.

TA in food authentication is precise but limited to known compounds, making it insufficient for detecting adulterants or markers (Esslinger, Riedl et al. 2014, Broadhurst, Goodacre et al. 2018). NTA, on the other hand, measures numerous analytes semi-quantitatively without prior knowledge, providing a broader view (Broadhurst, Goodacre et al. 2018). Ta analyzes a few specific compounds, while NTA detects thousands, offering a fingerprint of the sample's chemical composition (Esslinger, Riedl et al. 2014, Broadhurst, Goodacre et al. 2018). NTA requires minimal sample preparation to enhance the non-discriminatory approach (Uttl, Bechynska et al. 2023). It identifies patterns and relative metabolite abundances rather than

specific metabolites (Esslinger, Riedl et al. 2014). Due to the vast amount of data generated in NTA, statistical analyses are required to interpret relationships between variables (Esslinger, Riedl et al. 2014). Figure 2.6 depicts common pipelines of TA and NTA. The integration of both NTA and TA offers the most effective approach for future food authentication studies (Cajka and Fiehn 2016, Chen, Zhong et al. 2020). The main advantage of using NTA over TA is in the amount of information extracted from the samples and the simplicity of the sample preparation step, making it a quick and exhaustive method in authentication studies. Food fingerprinting has many applications besides authentication, such as food quality (Cevallos-Cevallos, Reyes-De-Corcuera et al. 2009), food safety and food processing (He and Bayen 2020), as well as in food microbiology (Chao and Krewski 2008). NTA has also been implemented for PGI registration purposes (Cassago, Artêncio et al. 2021). Overall, food fingerprinting using NTA has demonstrated significant potential for the authentication and quality control of alcoholic beverages, as will be further explored in the following sections (Aylott 2013, Lachenmeier 2016, Arslan, Tahir et al. 2021, Lin and Salcido-Keamo 2021).

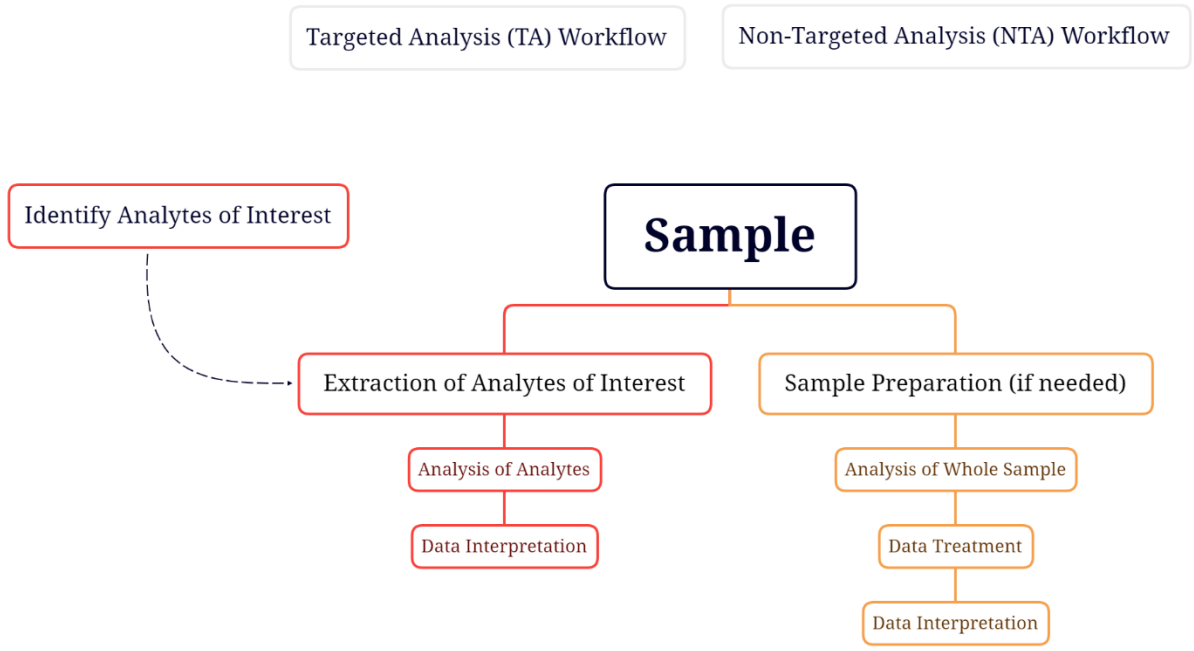


Figure 2.6 Workflows of targeted analysis and non-targeted analysis.

2.3.1 Analytical Instruments Used to Authenticate Alcoholic Beverages

Many analytical instruments and methods exist to detect fraud in food and beverages. Table 2.3 lists many of the analytical tools used for alcoholic beverages in recent authentication studies. GC-MS, LC-MS, and spectroscopic methods are dominant, but many other techniques are gaining popularity in food authentication such as elemental analyses, isotopic ratio analyses, radioisotope dating, DNA sequencing, and electronic devices (E-nose, E-tongue). Most analytical methods include common steps such as sample preparation, separation of analytes, and detection of analytes. Sample preparation is used to extract compounds from complex matrices, enhance the signals of certain analytes, and allow the analysis of matrices unfit for analytical instruments due to their physical properties. Various organic solvent extraction can also be used to target compounds from the bulk of the matrix. Extractions are used sparingly in NTA to provide a comprehensive portrait of the sample. Samples can also be concentrated or purified resulting in a reduced matrix interference and an improved recovery, which increases the concentration of analytes and their signal intensity. Sometimes, the food matrices must first be treated to be able to analyze properly, as is the case for sticky or fatty substances like honey or milk in LC-MS. When that is the case, dilution, or the use of solvent for extraction will allow converting the matrix into a more appropriate form, fit for analysis. Separation techniques often rely on chromatography in food authentication studies, such as GC and LC. GC aims at separating volatile particles with a gas carrier, representing the mobile phase, and an adsorbent or liquid-coated packing material, representing the stationary phase (Marriott 2005). LC separates compounds with a liquid mobile phase carrying the analytes and a solid stationary phase contributing to the retention of the analytes through ionic activity or by obstruction with particles of various sizes that impedes the flow of analytes. In both techniques, parameters such as polarity, viscosity, size, and volatility, will force analytes to eluate at different rates to separate them from one another, and depending on the interactions with the stationary phase, the analytes will then reach the detector at different moments (Palamareva 2005). GC is predominantly used with highly polar and volatile compounds since the analytes must be carried by a gas, while LC is mainly used with non-volatile or nonpolar compounds, although variations, such as reverse phase LC (RP-LC), allow for a wider range of compatible analytes (Marriott 2005, Palamareva 2005). Capillary electrophoresis (CE) is another technique that separates ions based on their electrophoretic mobility. It can easily analyze polar and charged compounds and is mostly paired with mass spectrometry, although its presence in

authentication studies is still uncommon (Mamani-Huanca, de la Fuente et al. 2021). The detection part is dominated by mass spectrometry (MS) and spectroscopy detectors. MS is based on the measurement of the analytes' fragment ions' mass-to-charge ratios (m/z) in relation to their relative abundance in a sample (Haag 2016). MS is commonly used in conjunction with GC and LC (GC-MS, LC-MS). On the other hand, spectroscopy is based on the measurement of the absorption and the emission of radiation, and the most frequent spectroscopic methods are nuclear magnetic resonance (NMR) and vibrational spectroscopy (IR, RAMAN). The importance of each method and its advantages and disadvantages will be discussed in the following sections.

MS is a commonly used analytical detector in the field of food authentication studies, often in combination with GC and LC. MS measures the m/z of analytes in a sample and plots them based on their relative abundances and retention times. MS has three essential components: ionization source, mass analyzer, and detector (Haag 2016). The ionization sources used to create atomic and molecular ions often rely on electron beams (Siuzdak 2004). Numerous ionization sources exist, such as ESI, EI, MALDI, ICP, and APCI (Siuzdak 2004, Feider, Krieger et al. 2019). The selection of ionization methods depends on various factors, such as the mass of the analytes, the type of ionization needed, the required energy level for fragmentation, and the desired sensitivity (Siuzdak 2004). Fragmentation occurs after ionization when the absorbed energy from the beam atomizes compounds. The molecular fragments serve as a specific molecular fingerprint of a compound, unique to the chosen analytical method. The mass analyzer is a critical component of the mass spectrometer responsible for separating ionized fragments based on their m/z, selecting only charged analytes of a specific mass suitable for detection (Haag 2016). The separation of ions is achieved through the use of electromagnetic fields, allowing for the retention of only ions with a predetermined m/z range (Jennings and Dolnikowski 1990). The selection of a mass analyzer depends on the range of studied m/z, the mass of analytes, the required resolving power, and the limit of detection (Haag 2016). Commonly used mass analyzers include Quadrupole, TOF, ion trap, FT-ICR, and Orbitrap (Haag 2016). Combining multiple mass analyzers can greatly enhance the specificity of the analysis, resulting in tandem MS, $MS^{2,}$ or MS/MS (Haag 2016). Examples of these include triple quadrupole (QqQ), quadrupole TOF (QTOF), or TOF/TOF (Haag 2016).

MATRIX	FRAUDS	ANALYTICAL TOOL	REFERENCES
WINE	Dilution (water)	HS-SPME-GC-MS	(Karabagias, Karabagias et al. 2021)
	Mislabelling (Oak Ageing)	HPLC-UV-DAD	(Matějíček, Mikeš et al. 2005)
	Counterfeiting (Vintage)	Gamma-Ray Spectroscopy	(Hubert, Perrot et al. 2009)
	Mislabelling (Botanical)	HPLC-QTOF-MS	(Vaclavik, Lacina et al. 2011)
BRANDY	Counterfeiting, Wood Ageing	Capillary Electrophoresis	(Panossian, Mamikonyan et al. 2001)
	Substitution	Fluorescence Spectroscopy	(Markechová, Májek et al. 2014)
ABSINTH	Substitution, Concealment	HPTLC	(Lachenmeier 2007)
WHISK(E)Y	Counterfeiting, Mislabelling (Age)	Paper spray-MS, ESI-FT-ICR-MS	(Garcia, Vaz et al. 2013) (Teodoro, Pereira et al. 2017)
	Mislabelling (Origin), Counterfeiting	UHPLC-QTOF-MS	(Collins, Zweigenbaum et al. 2014)
BEERS	Counterfeiting	Paper spray-MS	(Pereira, Amador et al. 2016)
	Origin (Trappists)	NMR	(Mannina, Marini et al. 2016)
	Mislabelling	NMR, FTIR	(Duarte, Barros et al. 2004)
AGAVE SPIRITS	Dilution, Substitution, Enhancement, Counterfeit	UV-Vis Spectroscopy	(Contreras, Barbosa-García et al. 2010)
	Mislabelling	HPLC-DAD	(Muñoz-Muñoz, Grenier et al. 2008)
	Mislabelling (Origin)	GC-MS	(León-Rodríguez, Escalante-Minakata et al. 2007)
RUM/SUGARCANE SPIRITS	Mislabelling (Ageing)	HPLC-UV	(de Aquino, Rodrigues et al. 2006)
	Mislabelling (Origin)	ICP-AES, GC-MS, GC-FID, HPLC-UV-Vis	(Cardoso, Andrade-Sobrinho et al. 2004)

Table 2.3: Studies reporting the detection of frauds using various analytical tools.

The detector detects the charged mass fragments and converts them into a measurable signal (David W. Koppenaal 2005). Common detectors include electron multipliers, a serial connection of dynodes amplifying the current of ions; Faraday cups, measuring the potential drop from the result of the electrons passing through a resistor after the ions hit a collector; photomultiplier conversion dynodes, which is similar to electron multiplier, but photons are amplified instead of electrons; and array detectors, which measures the spatial distribution of the ions hitting a sensor, allowing simultaneous detection of multiple analytes (David W. Koppenaal 2005). In summary, the analytes enter the MS after being separated by chromatography. They are then ionized and fragmented, and the resulting fragments are then separated by a mass analyzer. The flow of ions is finally converted into a measurable signal by the detector to be further analyzed by data treatment tools. MS is incredibly valuable in authentication studies due to its sensitivity and its common pairing with chromatographic separation instruments (GC, LC) (Cajka and Fiehn 2016). The authentication studies cited in Table 2.3 illustrate well the dominant presence of MS compared to other analytical tools (Cajka and Fiehn 2016). Indeed, MS has found many applications in alcoholic beverage authentication with various matrices and for different purposes, such as the detection of adulterants, and authentication of origin (See Table 2.3). GC-MS has been used in the origin authentication of tequila and Cabernet Sauvignon (León-Rodríguez, Escalante-Minakata et al. 2007, Chávez-Márquez, Gardea et al. 2022). LC-MS has been used to profile the nonvolatile compounds in American whiskeys (Bourbon, Rye, Tennessee, and non-classified) and has allowed the identification of 7600 compounds in total, further allowing to classify whiskeys based on raw materials, age, and producers (Collins, Zweigenbaum et al. 2014). It is also commonly used for the discrimination of wine based on varietals (Vaclavik, Lacina et al. 2011). In summary, MS has both strengths and drawbacks compared to spectroscopy, the second most used method. MS offers high sensitivity, a wide detection range, numerous databases, and the ability to couple it with advanced separation techniques, making it a versatile tool in a wide range of applications. However, it also has limitations such as low quantitation, low reproducibility, destructive nature, and significant sample volumes (Bujak, Struck-Lewicka et al. 2015). Despite these challenges, MS remains an essential analytical tool, providing researchers insights into the chemical composition of complex samples such as alcoholic beverages.

Spectroscopy is a well-established technique in food science and authentication. Spectroscopic techniques are considered the main techniques for some food matrices, such as fruit juices, oil, and wine (Su, Arvanitoyannis et al. 2018). Spectroscopic techniques are useful for analyzing high-value alcoholic beverages, as they can be conducted non-destructively (Arslan, Tahir et al. 2021). Spectroscopy works by analyzing the absorption and emission of radiation through wavelengths isolation. The pattern of these wavelengths provides a spectrum that is used to fingerprint samples (Schermann 2008). Common spectroscopic methods include IR-related ones (NIR, MIR), nuclear magnetic resonance (NMR), Raman, and UV-Vis. NMR is the predominant analysis technique used in food authentication (Cajka and Fiehn 2016). It can analyze several matrices with little sample preparation required and its measurements are highly reproducible and repeatable. However, NMR is less sensitive than other techniques like GC/LC-MS or even IR (Le Gall and Colquhoun 2003). It has been used for authentication purposes with cases of extension with various sugars (Bertelli, Lolli et al. 2010), wine origin (Son, Kim et al. 2008, Alonso-Salces, Héberger et al. 2010, Godelmann, Fang et al. 2013), rum classification (Belmonte-Sánchez, Romero-González et al. 2020), vodka authentication (Ciepielowski, Pacholczyk-Sienicka et al. 2019), and tracing beer origins (Monakhova, Schäfer et al. 2011). IR was among the first tools for food fingerprinting (Lai, Kemsley et al. 1994). FTIR, combined with MVA, like PCA and DA, quickly authenticates food products and detects chemical shifts (Lai, Kemsley et al. 1994, Su, Arvanitoyannis et al. 2018). MIR has evolved as a swift screening tool for alcoholic beverage authentication (Fernández Pierna, Duponchel et al. 2012, Parpinello, Ricci et al. 2019). Raman spectroscopy is used for classifying wines and other products but is considered to have weaker resolution (Kizil and Irudayaraj 2018, Magdas, David et al. 2022). Fluorescence spectroscopy relies on measuring light emitted by exciting molecules and has been applied to authenticate various alcoholic beverages, with the ability to detect trace analytes (Hassoun, Måge et al. 2020, Xagoraris, Revelou et al. 2021). UV-Vis is allows the detection of wavelengths in organic compounds' visible light or ultraviolet light range (Passos, Sarraguça et al. 2019). UV-Vis coupled with CE has been applied in wine to authenticate grape cultivars with the use of proteins in must or various phenolic acids (Popîrdă, Luchian et al. 2021). Spectroscopic techniques play an important role in food authentication, with each method having its advantages and limitations. Current research is aimed at improving sensitivity and resolution, while also exploring new applications for food authentication such as portable instruments (Chaudhary, Kajla et al. 2022).

Elemental analysis is effective in identifying the geographical origin of alcoholic beverages (García-Ruiz, Moldovan et al. 2007, Magdas, Cristea et al. 2021). Techniques like ICP-OES, ICP-MS, and AAS are used in authentication. ICP, known for its low detection threshold and ability to analyze multiple elements simultaneously, is especially useful in NTA (Eyring 2003, Thomas 2019). ICP-OES detects photon wavelengths, while ICP-MS measures relative abundance based on m/z (Fernández-Sánchez 2019). AAS, less expensive but limited in resolution and multi-analyte detection, is less preferred (Eyring 2003). Recent advancements include coupling NTA ICP-MS with LC for elemental speciation analysis, providing detailed insights into molecular forms and isotopic composition (Lorenc, Hanć et al. 2022). Elemental analysis, traditionally part of TA, is evolving with speciation analysis (Feldmann, Raab et al. 2018). Techniques like LIBS, though less sensitive than ICP, are useful for infield analyses and have been applied in wine authentication (Cremers 2013, Tian, Yan et al. 2017). These methods have diverse applications in authenticating alcoholic beverages. For instance, they differentiate tequila from mezcal, various ages of tequila, rums from cachaça, and different beer styles (Bellido-Milla, Moreno-Perez et al. 2000, Cardoso, Andrade-Sobrinho et al. 2004, Ceballos-Magaña, Jurado et al. 2009). They are also extensively used to determining the geographical origins of wine, while their effectiveness in identifying grape varietals or vintage remain inconsistent (Capron, Smeyers-Verbeke et al. 2007, Forina, Oliveri et al. 2009, Martin, Watling et al. 2012, Đurđić, Pantelić et al. 2017). Despite these challenges, elemental analysis holds significant potential in authenticating alcoholic beverage.

Isotope ratio analysis is a valuable tool in alcoholic beverage authentication, particularly in the detection of adulteration and identifying geographical and botanical origins of substances (Perini and Camin 2013, Magdas, Cristea et al. 2021, Magdas, David et al. 2022). The ratio of $^{13}C/^{12}C$ ($\delta^{13}C$) is useful in botanical authentication due to differences in photosynthetic pathways between plants. It can distinguish between C3 and C4 plants and detect sugar adulteration in spirits like brandy and whiskey (Evert, Eichhorn et al. 2013)«(Rhodes, Heaton et al. 2009). Associated with water sources in alcoholic beverages, $\delta^{18}O$ ($^{18}O/^{16}O$) is instrumental in detecting fraudulent dilution and varying production methods. It can also detect synthetic ethanol but has limitations in detecting sugar adulteration (Perini and Camin 2013). The hydrogen isotopic ratio $\delta^{2}H$ (D/H) is widely used to determine geographical origin and distinguish between fermentation and synthetic ethanol. It has authenticated the origin of Polish vodka and detected sugar adulteration

(Ciepielowski, Pacholczyk-Sienicka et al. 2019, Martinelli, Nardoto et al. 2020). The nitrogen isotope $\delta^{15}N$ ($^{15}N/^{14}N$) shows promise in authenticating agricultural, varying based on fertilizers and pesticides used (Kelly, Heaton et al. 2005, Nietner, Haughey et al. 2014, Kelly, Brodie et al. 2018, Spangenberg and Zufferey 2018). In summary, isotopes ratios are influenced by a plant's metabolism and geographical location, leading to unique profile ratios for authentication studies. Challenges include determining vintages and the minimal discussion of non-traditional isotopes like B, Cu, Zn, Hg, Cr, and Fe (Rossmann 2001, Adam, Duthie et al. 2002, Cellier, Berail et al. 2021). Measurements techniques include IRMS, SNIF-NMR, and ICP with advanced instruments like MC-ICP-MS achieving high precision.

Carbon dating is a technique widely applied in various scientific fields to determine the age of an item or substance by calculating the decay rate of ^{14}C, a carbon isotope. Due to the nuclear weapon testing in the 1950s and 1960s, the level of ^{14}C in the atmosphere almost doubled the level it was before the tests. This drastic change in atmospheric ^{14}C can be used to date much more accurately than by using the traditional radioactive decaying of ^{14}C (Asenstorfer, Jones et al. 2011). In the context of food authentication, carbon dating has been used to date vintage in wines or year of production in old bottles of Scotch whiskies, proving its usefulness in detecting counterfeit bottles (Asenstorfer, Jones et al. 2011, Cook, Dunbar et al. 2020). A major limitation of this technique used to be the requirement of a physical sample from the wine, however, a non-destructive technique has been developed to authenticate a wine's age by applying a vacuum on the cork of a bottle and cryo-trapping the molecules that migrate from the inside of the bottle to the outside, mainly ethanol and other gases (Fahrni, Fuller et al. 2015). By analyzing these molecules using carbon dating, the authors that developed the method were able to verify the vintage of the wines without opening the bottle, thus protecting the quality of the content and the value of the product. This technique has the potential to be used with other techniques such as chemical or elemental analysis and is considered superior to other non-destructive methods such as the fingerprinting of the bottle, the cork, or the label (Fahrni, Fuller et al. 2015). In addition to carbon dating, gamma rays from ^{40}K, ^{226}Ra, ^{210}Pb, ^{228}Ra, and ^{137}Cs can also be calculated to distinguish alcoholic beverages made before and after the nuclear testing without even opening the bottle (Hubert, Perrot et al. 2009, Médina, Salagoïty et al. 2013).

DNA sequencing is used in the authentication of alcoholic beverages to identify biological origins and detect food fraud, such as the substitution of high-quality ingredients with cheaper ones. This method has been employed for verifying the variety of grapes in wine and the botanical source in beer. However, it has limitations in authenticating geographical origin, especially when the same raw materials are used across different regions (Böhme, Calo-Mata et al. 2019, Cusa, St John Glew et al. 2022). PCR is the main DNA-based authentication method, with modern variations like qPCR and ddPCR providing more precise quantification (Taylor, Laperriere et al. 2017, Böhme, Calo-Mata et al. 2019). Challenges include the low amount of DNA extractable from alcoholic beverages, especially spirits, and the interference from other compounds (Anca, Adriana BĂ et al. 2021). Emerging techniques, such as double digest restriction enzyme associated DNA, are allowing authentication without reference databases, broadening the scope of genetic material analysis in food fingerprinting (Peterson, Weber et al. 2012). In summary, DNA sequencing tools like qPCR, ddPCR, and ddRAD are useful to precisely authenticate botanical origin in alcoholic beverages, although there are still limitations to using them for other fraud issues, namely the requirement of genetic diversity and low DNA concentration in finished products.

E-nose and e-tongue are analytical tools that have emerged in the last few decades (Zou, Wan et al. 2015). They function by detecting compounds in a mixture with an array of partially selective electronic chemical sensors with varied sensitivity to recognize, identify, classify, and quantify these compounds. E-noses are used with volatile compounds, while e-tongues are used with non-volatile compounds (Zou, Wan et al. 2015). The compiled data detected by the sensors provide a chemical fingerprint of the sample (Zou, Wan et al. 2015). E-noses uses a range of sensors for improved efficiency, such as conducting polymers, optical sensors, and piezoelectric sensors (Freund and Lewis 1995, Johnson, Sutter et al. 1997, Wu 1999, Zou, Wan et al. 2015, Sierra-Padilla, García-Guzmán et al. 2021). A study has been conducted using e-nose to characterize the aging fingerprint of beer (Ghasemi-Varnamkhasti, Mohtasebi et al. 2011), while another used HS-MS to authenticate wines' geographical origins, grape varietals, and aging (Martí, Busto et al. 2004). Polish vodkas could also be classified by botanical origin and substitution with maize could be detected (Wiśniewska, Śliwińska et al. 2016). E-tongues also has multiple sensors, such as electrochemical, optical and mass sensors. Electrochemical sensors measure various electrical current parameters (voltage, amperage, conductivity, resistance) (Ciosek and

Wróblewski 2007, Zou, Wan et al. 2015), while optical sensors measure the response to the light activity of non-volatile compounds. Mass sensors accumulate a charge when a mechanical stress is applied. E-tongues have been used to evaluate the type of wood and charring level in aged red wines (Parra, Arrieta et al. 2006), discriminate the geographical origin of Italian wines down to the specific vineyard (Legin, Rudnitskaya et al. 2003) and classify wines based on botanical origin and vintage year (Geană, Ciucure et al. 2020). Similarly, e-tongues were used to classify brandies by vintage and quality (Cetó, Llobet et al. 2013) and classify whiskies by brand (Novakowski, Bertotti et al. 2011). Despite the potential of these technologies in the analysis of alcoholic beverages, important challenges such as susceptibility to the environment and loss of sensitivity in the presence of high concentrations of ethanol remain (Harper 2001, Baldwin, Bai et al. 2011).

2.3.2 Analytical Tool Issues in Alcoholic Beverage Authentication

Alcoholic beverage authentication using analytical tools still presents several challenges, including issues with small sampling sizes, workflow standardization, and experiment reproducibility. Studies with small sample sizes, which are common in the literature, often below 50 samples and rarely above 200 samples, can introduce biases in the data obtained and decrease the value of the studies. In Table 2.4, studies with relatively large sample sizes were selected to ensure data accuracy and reduce bias, but it can still be seen that most don't have large sample sizes. Another issue is the lack of standardization and reproducibility of the analytical methods used in different laboratories. Many laboratories develop their protocol for detecting adulterants in alcoholic beverages or authenticating origin, but these protocols may not provide the same results in different environments or with slightly different sample sets (Cavanna, Righetti et al. 2018, Goethem and Elliott 2018). Recent studies have shown that the lack of standardization and reproducibility can significantly impact the value of untargeted analytical studies for food authentication purposes. To address these challenges, researchers are exploring new methods to increase the value and accuracy of untargeted analytical studies for food authentication purposes. These methods are mainly focused on standardizing analytical methods across the food authentication research field and enhancing reproducibility. Future research is aimed at implementing the potential for machine learning and artificial intelligence to improve accuracy and reproducibility (Gabrieli, Muszynski et al. 2022, Goyal, Kumar et al. 2022, Mavani, Ali et al.

2022). Overall, addressing these challenges is essential to ensure the authenticity of alcoholic beverages and protect consumers and honest producers.

2.3.3 Data Treatment in Alcoholic Beverage Authentication Using NTA

Data treatment is necessary for any research that relies on the analysis of data through statistical means. In the NTA framework, the process begins with collecting raw data, which is then both preprocessed and processed to convert it into clean and usable data for analysis. Statistical methods are then employed to extract meaningful correlations and causation from the data. MVA is a key feature of NTA data analysis, allowing multiple different measurements to be made on each experimental dataset, which helps in prediction analysis, sample classification, clustering, and the isolation of outliers. Once all the results of the analysis have been obtained and various relationships have been identified and characterized, it can be interpreted to identify authentic or fraudulent products. If the result is uncertain, questionable samples may undergo further investigation using different analytical approaches or instruments. Reporting on all data treatment steps employed in a study is essential for reproducibility and standardization in NTA approaches.

Data preprocessing is required after the acquisition of the raw data for its ‘’correction’’. Although each analytical approach requires its own data preprocessing procedure, several common steps are necessary (Esslinger, Riedl et al. 2014). These steps include cleaning, transformation, reduction, and wrangling (Fan, Chen et al. 2021, Simonnet-Laprade, Bayen et al. 2021).

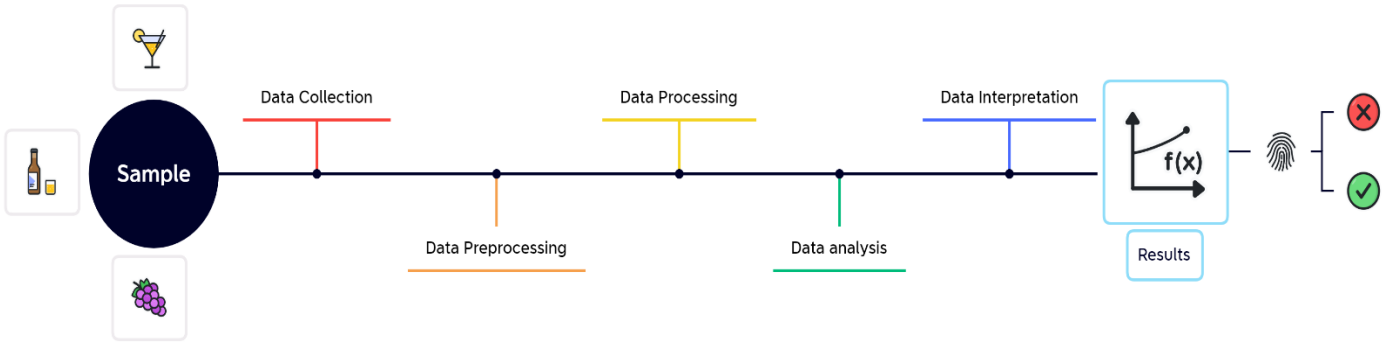


Figure 2.7 Workflow from collection of data from samples to end results.

Data cleaning involves several tasks, such as removing or filling missing values and handling noisy data through binning, regression, or clustering. It also includes resolving of

inconsistencies, such as null values, outliers, or irrelevant data (Fan, Chen et al. 2021) Data transformation aims to transform data into suitable forms by converting numerical data values into categorical ones. Data transformation steps include normalization, discretization, and concept hierarchy generation. Normalization, often called data scaling, will adjust all the data into defined limits. Discretization is the transformation of continuous data values into discrete forms, which is usually more suitable for the numerical evaluation of statistical models. Concept hierarchy generation involves converting lower-tiered concepts into higher-tiered ones (Fan, Chen et al. 2021). Data reduction is useful in big data analysis, such as NTA, to reduce the amount of data. This can be achieved by data cube aggregation, feature selection, numerosity reduction, and dimensionality reduction (Fan, Chen et al. 2021). Data cube aggregation simply summarizes the data by aggregating similar concepts into broader ones. Feature subset selection involves the removal of less significant data, which can be done by discarding data with *p*-values greater than the accepted level of significance or selecting a subset of relevant attributes from the original features to optimize model construction. Numerosity reduction is the conversion of the data into simpler representation models. Dimension reduction involves transforming data from a high-dimensional space to a lower-dimensional space using compression and encoding mechanisms. It may result in a loss of information but aims to retain as many properties as possible from the original data set. A common dimension reduction tool is the principal component analysis (PCA) which will be described later (Fan, Chen et al. 2021). Data wrangling is usually performed last and involves transforming the preprocessed raw data into structurally appropriate values suitable for analysis (Azeroual 2020). The more multidimensional a technique is, the more complex the data pre-processing has to be. An increasing number of software is being developed to respond to the demand for these types of analytical instruments' data pre-processing steps (Esslinger, Riedl et al. 2014). Finally, data processing commonly includes initial m/z detection by the selection of m/z retention time pairs, retention time alignment based on information from reference compounds, removal of shoulder peaks, noise filtering, and reduction, removal of undesired peaks, isotopologues grouping, peak alignment by comparison between samples, gap-filling, feature filtering, removal or filtering of duplicates, normalization and thresholding of abundance based on spikes and blanks, searching libraries for known compounds, and annotation and identification of features (Katajamaa and Orešič 2005, Fisher, Croley et al. 2021). This list is representative of the commonly featured data processing steps. Researchers using NTA or screening approaches should

report all data processing steps employed in their publications to ensure reproducibility and promote standardization.

Developing a robust method development for authentication studies using NTA requires careful consideration of several key aspects. These include extraction replicates, quality controls (QC), performance validation, marker identification, MVA or machine learning model validation, receiver operating characteristic (ROC) curves, and sample accuracy in terms of representativeness of authentic products (Cavanna, Righetti et al. 2018). It is crucial to consider that chemical variations may not arise from the food matrices themselves but rather from the analytical techniques or the sample preparation. Applying validation parameters is essential to reduce the occurrence of these biased results (Esslinger, Riedl et al. 2014). QC measures in NTA approaches are complex, as they must take into account not only individual analytes but also the complete food matrix (Esslinger, Riedl et al. 2014). To ensure the reliability of QC measures, it is important to use homogenous, stable pooled samples that cover all signals in the matrix (Esslinger, Riedl et al. 2014). Injecting these pooled samples at regular intervals throughout the experiment helps maintain rigorous measurements and prevent deviations (Cavanna, Righetti et al. 2018). Model and performance validation are crucial steps in developing a robust authentication method. Internal validation using cross-validation approaches can provide insight into the robustness of the model by removing parts of the samples and recreating the model with the remaining data, a procedure repeated multiple times until all data points have been included in the test at least once. Additionally, external validation is conducted by evaluating an independent set of samples not used for model building to demonstrate the model's validity through introduction of variability (Cavanna, Righetti et al. 2018). Marker identification and validation are essential components of authentication studies using NTA. The accuracy in the identification of compounds relies on the data available. Recently, researchers have adopted the Identification Confidence Levels system, as shown in Figure 2.8. (Schymanski, Jeon et al. 2014, Schrimpe-Rutledge, Codreanu et al. 2016). These levels range from level 5, where accurate m/z and retention time are known, to level 1, which involves comparing experimental data with reference standards. Intermediate levels include unequivocal molecular formula, substituent and class elucidation using libraries and databases, and matching of experimental MS/MS spectra with those in libraries and databases (levels 4,2, and 2, respectively) (Schymanski, Jeon et al. 2014). The objective of using these confidence levels is to establish a standardized approach to compound identification. Data fusion, a new approach that

combines data from different analytical instruments, can further improve the validity of a marker if the signal is detected using multiple instruments, making it relevant across platforms (Cavanna, Righetti et al. 2018).

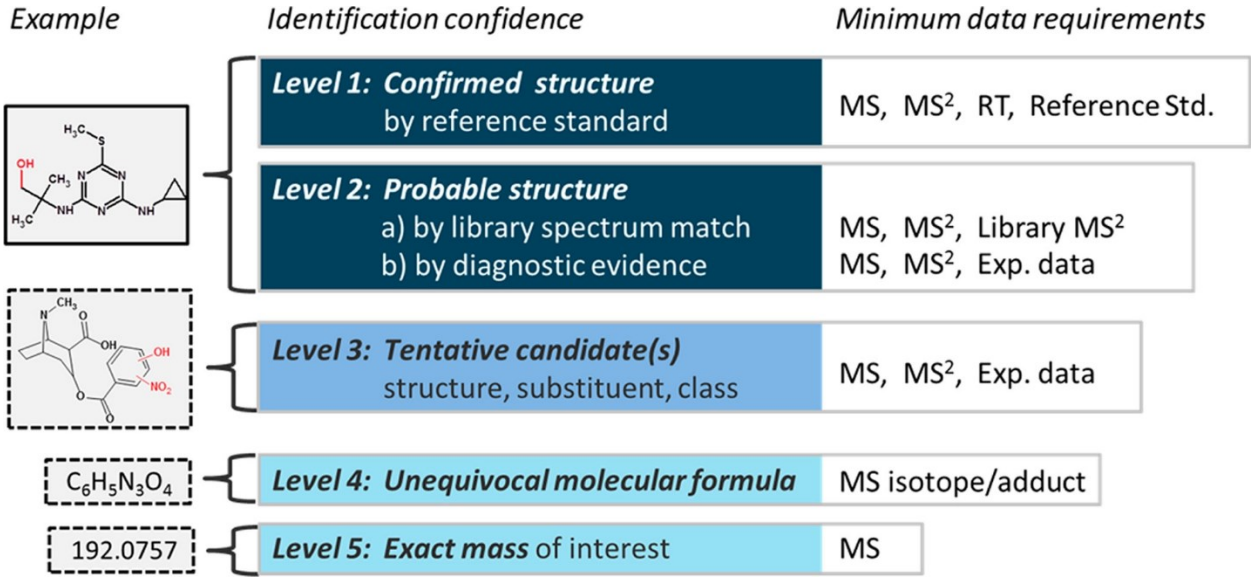


Figure 2.8 Identification confidence levels using HRMS analysis (Schymanski, Jeon et al. 2014).

Sample accuracy presents a significant challenge in authentication studies, as it is difficult to confidently assess whether a provided sample is truly authentic. A review paper on the challenges in model validity for NTA revealed that among the presented papers (49 articles), more than 25% did not perform any validation study on their methods or models (Cavanna, Righetti et al. 2018). In summary, a robust method for authentication studies using NTA involves rigorously applying validation procedures, including QC measures, model validation, marker identification, and sample accuracy.

2.3.4 Data Analysis in Alcoholic Beverage Authentication Studies

Data analysis is essential to authentication studies, particularly in the context of NTA. The vast amount of information generated by these analytical strategies necessitates the use of statistical classification models, dimension reduction and pattern recognition methods, and more recently, machine learning approaches. MVA, or chemometrics, is commonly employed to interpret the data obtained. MVA is used to reduce the complexity of datasets and cluster information, enabling the differentiation of samples based on their intrinsic chemical composition.

This process allows for the discrimination or classification of samples according to their chemical fingerprint. MVA can also serve as a predictive tool to assess the analytical parameters required for specific fingerprints, thus determining to which class a sample belongs, without prior knowledge of its chemical composition, origin, or production methods (López-Ruiz, Romero-González et al. 2019). MVA works by analyzing multiple measurements on each data point and is divided into two approaches: supervised and unsupervised MVA, which are not interchangeable. Supervised and unsupervised MVA are both valuable in assessing complex groups of variables, such as those involved in the authentication or detection of adulteration in alcoholic beverage matrices. Supervised MVA methods, such as PLDS-DA, LDA, SVM, kNN, and SIMCA, evaluate variables in the data to determine combinations that best represent causality. They establish classification thresholds based on previous sample sets and classify samples using predefined parameters. Supervised methods can work effectively with predetermined parameters, due to their exposure to correctly labelled training sets. In authentication studies, the chemical composition and relative abundance of compounds serve as the primary variables, given the significant information they provide (Cubero-Leon, Peñalver et al. 2014). Contrastingly, unsupervised MVA, such as HCA, PCA, and FCM, seek to condense data to understand its structure and detect underlying patterns, without focusing on a specific end goal (Scott and Crone 2021). It excels in detecting differences among samples and employs techniques such as clustering to group samples based on similarities and sequence associations to identify patterns (He and Bayen 2020).

Matrix	Analytical Technique	n	Data Treatment	Reference
Wine	UHPLC-QTOF-MS	97	PCA, Random Forest	(Tzachristas, Dasenaki et al. 2021)
	RP-HPLC-DAD	82	PCA, PLS-DA, CART, ANN	(Cosme, Milheiro et al. 2021)
	HS-SPME-GC-MS	45	MANOVA, LDA, PLS	(Karabagias, Karabagias et al. 2021)
	ICP-MS	639	K-CM (ANN and Fuzzy profiling)	(Bronzi, Brilli et al. 2020)
	A-TEEM, ICP-MS	86	PCA, SVMDA, PLS, XGBDA	(Ranaweera, Gilmore et al. 2021)
	ART-FTIR	84	PCA, LDA	(Kruzlicova and Gruberova 2022)
	NIR	88	PCA, PLS, OPLS-DA	(Nardi, Petrozziello et al. 2020)
	SNP Genotyping	260	-	(Boccacci, Chitarra et al. 2020)
	NMR	201	PCA, PLS-DA	(Ehlers, Horn et al. 2022)
Cider, Radler	ICP-MS, ICP-OES, CVAAS	73	PCA	(Gajek, Pawlaczyk et al. 2021)
Beer	NMR	31	PCA, OPLS-DA	(Palmioli, Alberici et al. 2020)
	ATR-MIR	24	PCA, PLS-DA	(Gordon, Chapman et al. 2018)
	LC-MS	232	PCA, OPLS-DA	(Mattarucchi, Stocchero et al. 2010)
Baiju	Fluorescence Spectroscopy	30	LDA, PCA	(Burns, Alexander et al. 2021)
Whisk(e)y	FT-ICR-MS, HPLC-MS/MS	7	PCA	(Yang, Somogyi et al. 2020)
	ICP-MS, MP-AES	68	MANOVA, CVA	(Hopfer, Gilleland et al. 2017)
	NMR	148	PCA, OPLS-DA	(Kew, Goodall et al. 2019)
Rum	NMR, HS-SPME-GC-MS, HRLC-MS	24	CV-ANOVA	(Belmonte-Sánchez, Romero-González et al. 2020)
Fruit/Marc Spirits	Raman	97	Machine Learning	(Magdas, David et al. 2022)
	HS-SPME-GC-MS, FT-IR	75	SOPLS-DA, SO-CovSel-LDA	(Giannetti, Mariani et al. 2020)
Vodka	NMR, IRMS	30	-	(Ciepielowski, Pacholczyk-Sienicka et al. 2019)
	E-tongue	69	PCA, ANN	(Marenco, de Oliveira et al. 2021)

Table 2.4: Recent authentication studies conducted on various alcoholic beverages matrices using different statistical analyses strategies.

In effect, both supervised and unsupervised analyses offer distinct strengths: while supervised models are particularly apt for establishing classification parameters, unsupervised models are adept at uncovering previously unseen patterns in data (Pereira, Amador et al. 2016). Most MVA statistical algorithms can be categorized into four distinct usages: predictive, descriptive, generative, and discriminative modelling (Cevallos-Cevallos, Reyes-De-Corcuera et al. 2009). Predictive modelling forecasts future outcomes with new data, potentially detecting new adulteration issues in alcoholic beverages or classifying new samples based on existing databases (Malek, Hui et al. 2019). Descriptive, or informative, modelling establishes relationships between data and outcomes, helping identify outliers in datasets to detect fraudulent products in authentication studies (Cevallos-Cevallos, Reyes-De-Corcuera et al. 2009). Generative models create probability distributions for individual classes, representing data placement across dimensions. In authentication studies, samples are submitted to these classes to evaluate their fit and classification. Discriminative models classify data points by establishing boundaries between classes and determining key differences that separate one class from the others. Both generative and discriminative models are widely used for classifying samples, making them essential statistical models in authentication studies (Sallans, Bruckner et al. 2006, Cevallos-Cevallos, Reyes-De-Corcuera et al. 2009, Ji 2020).

In terms of supervised methods, PLS-DA is used for predictive, descriptive, and discriminative modelling. It combines dimension reduction (PLS), a technique common to multiple statistical tools, and discriminative modelling (DA) (Lee, Liong et al. 2018). Although PLS-DA was not initially designed for sample classification, it frequently succeeds in determining whether data fits a class by comparing it to a training set (Barker and Rayens 2003). A study analyzing Brazilian beer brands, which are often counterfeited through substitution with cheaper brands, used PLS-DA as a classification tool to detect MS patterns. The researchers achieved a 98% reliability rate in discriminating between cheap and premium brands. By enhancing PLS-DA with variable selection, in this case using OPS, they reached a 100% reliability rate (Pereira, Amador et al. 2016). LDA is another supervised classification technique that focuses on reducing dimensions and maximizing variance among samples by creating a pooled covariance matrix (Næs, Isaksson et al. 2004). LDA can then distribute samples into predetermined classes. It is frequently used in authentication studies, for example in the classification of Italian wines (Parpinello, Ricci et al.

2019). New methods involving LDA have aimed at creating semi-supervised and unsupervised algorithms (Un-LDA) to address challenges faced by supervised techniques, such as the required labelling of classes (Wang, Wang et al. 2021). SVM, a supervised machine-learning method, is used for descriptive modelling and discriminative classification models (Wilson 2008). SVM operates by constructing a line or plane, known as a hyperplane, that separates all data into two categories while maximizing the distance between each point in the set and the hyperplane (Wilson 2008). As SVM classifies samples into two groups separated by a hyperplane, it is ideal for identifying outliers, such as fraudulent products in routine analysis. However, SVMs struggle to establish a hyperplane in large datasets, often encountered in NTA. To address this issue, the kernel trick is commonly applied, adding dimensions to the dataset to find a hyperplane that separates data in a higher dimension (Wilson 2008). kNN is a supervised technique that classifies new data points by associating them with their nearest neighbours in the dataset (Neath and Johnson 2010). The data point is then assigned to the class with the highest number of nearest neighbours. kNN is particularly useful in predictive modelling, as demonstrated in a study that used Pinot Noir lipidomics to determine whether a sample belongs to a specific geographical origin (Phan and Tomasino 2021). SIMCA, a supervised classification technique, is based on PCA, an unsupervised technique discussed below (Hopke 2003). As an unsupervised technique, PCA does not consider predetermined class information and is therefore not helpful for classifying data into predetermined classes. SIMCA was developed to use PCA with predetermined class information. It operates by developing a class model for each group in the training set and then adding new data points to these classes based on their relative distance from the existing information (Hopke 2003, Ballabio and Todeschini 2009).

Delving into unsupervised methods, HCA is a straightforward unsupervised method that constructs hierarchies of clusters, ensuring that each cluster is distinct while maintaining related data within clusters. HCA creates a dendrogram by clustering data from top to bottom or vice versa (Pezoulas, Exarchos et al. 2020). In a study analyzing red monovarietal wines produced from Zweigelt and Rondo grapes, researchers found that HCA could classify the wines into two groups based on their phenolic compound composition (Stój, Kapusta et al. 2020). PCA aims to maximize variation between samples while reducing the dataset's dimensionality. It identifies the fewest linear combination that can summarize as much data as possible without losing significant information (Mardia, Kent et al. 1979, Parpinello, Ricci et al. 2019). PCA is useful for determining

correlations between samples and variables (Næs, Isaksson et al. 2004). In a study, researchers used high-resolution NMR paired with chemometrics techniques such as ICA, PCA, and OPLS-DA, to successfully model various whisky production parameters like malt status, presence of peated malt, alcoholic strength, and maturation wood type. PCA was primarily used to discriminate between blended and single-malt whiskies (Kew, Goodall et al. 2019). FCM operates similarly to previously described clustering analysis, assigning data points to clusters such that they are as alike as possible. However, FCM allows data points to belong to more than one cluster (Theodoridis, Pikrakis et al. 2010, Subasi 2020). While HCA works in a top-down (or bottom-up) manner, FCM uses means to indicate the degree to which specific data points belong to a cluster. Although many fuzzy logic tools have been applied to authenticate alcoholic beverages, FCM has not yet been used for this purpose, despite its potential for clustering (Raptis, Siettos et al. 2000, Petropoulos, Karavas et al. 2017). FCM offers advantages over other unsupervised methods, such as partial membership, which could detect subtle chemical shifts caused by adulteration, identify subcategories of alcoholic beverage types, or authenticate beverages containing multiple ingredients, like gins or wine blends (Tanatavikorn and Yamashita 2015).

Statistical learning, which combines data science, statistics, and machine learning, is increasingly being used in authentication studies, particularly in the era of big data. It employs machine learning tools, such as deep learning with DT and NN, for pattern recognition (Pezoulas, Exarchos et al. 2020). These tools can be supervised, unsupervised, or semi-supervised, and they are a subset of artificial intelligence designed to perform tasks without being explicitly programmed for them (El Naqa and Murphy 2015). DTs are supervised techniques used for classification and regression models. They work by dividing datasets into nodes until only terminal subsets remain, with these terminal nodes representing clusters (Krzywinski and Altman 2017). Newer and improved algorithms, such as bagging trees and Random Forest, are more stable but also provide more nodes, increasing the complexity of the models (Breiman 1996, Breiman 2001, Nuti, Jiménez Rugama et al. 2021). NNs can be both supervised and unsupervised. They function by mimicking biological neural networks to model complex datasets statistically or structurally, offering high flexibility in their approaches (Jain, Jianchang et al. 1996, Abiodun, Jantan et al. 2019). Fuzzy logic is a legacy machine learning method behind FCM and is based on the premise that all data points simultaneously belong, to varying degrees, to all possible clusters in the dataset. As such it assigns a degree of membership to clusters and then proceeds to map nonlinearly the

data points with these degree values (Mendel 1995, Kayacan and Khanesar 2016). SVM and SVC, two previously mentioned approaches, are also machine learning techniques. Deep learning involves representation-learning methods that use multiple levels of representation, composed of simple, non-linear modules. Deep learning models in chemometrics analyze large chemical or spectroscopic datasets by learning data representations with multiple abstraction levels. This enables automatic feature extraction from raw data like chromatograms or spectra. Large datasets and robust computational resources are necessary for training deep learning models (López-Monroy and García-Salinas 2022). Despite its potential, deep learning is underexplored in alcoholic beverage authentication, presenting an exciting opportunity for future research. In all cases, machine learning has been developed to help overcome challenges in statistical analyses, such as pattern recognition, clustering, functions estimation, prediction power, self-optimization, and adaptiveness to new data, all of which are highly desired characteristics for understanding data (Jain, Jianchang et al. 1996).

2.3.5 Data Treatment and Data Analysis Issues in Alcoholic Beverages Authentication Studies

The current scientific challenges in developing a proficient and widely available authentication workflow include a lack of method and model validation and reproducibility, which leads to difficulties in confidently assessing the classification or discrimination abilities of proposed approaches. Other challenges include the absence of responsible sharing of widely available databases of mass spectra and chemical fingerprints of food products, issues in asserting the authenticity of collected samples, and incoherence in data processing steps, as well as a deficiency in the transparency of these steps in published articles (Riedl, Esslinger et al. 2015, Cavanna, Righetti et al. 2018, Gertsman and Barshop 2018). These challenges are common to all NTA approaches. Resolving these issues would enable the validation of workflows and their establishment in various parts of the food industry to maintain strict control over authentication, thus reducing fraudulent products from reaching markets. The use of standardized methods would unify measurements made from similar instruments across multiple brands and laboratories. Providing a shared database would allow the scientific community to compare their results and gain access to all analyzed products based on geographical origin, botanical origin, harvest year, and production processes. This would facilitate authentication controls and aid in identifying and

characterizing unknown compounds, one of the most significant bottlenecks in NTA (Gertsman and Barshop 2018).

2.4 CONCLUSION

Food fraud, notably in alcoholic beverages, poses significant risks to consumers, industries, and public health. Authenticating these beverages requires verifying their origin, production methods, and ingredients. Chemical composition analysis and identifying adulteration-related chemical shifts are key in addressing fraud. Techniques like LC-QTOF-HRMS based NTA with MVA are effective for detecting fraud and differentiating between geographical and botanical origins, but research in this area, especially for spirits, is still developing. There is a need to expand these methods to new spirits and origins. Challenges include identifying knowledge gaps, particularly in how adulteration affects chemical composition. Developing methods to establish fingerprints and databases for various alcoholic beverages is crucial, as is standardizing these methods across laboratories. Current research focuses on distilled spirits using GC, with limited LC use, and doesn't fully explore the correlation between authenticity traits and chemical fingerprints. High water and ethanol content in beverages pose specific challenges, masking important trace analytes (Esslinger, Riedl et al. 2014). Many studies have small sample sizes, highlighting the need for larger-scale research for reliable findings. Neutral grain spirits and gins, despite their increasing popularity and complexity due to the inclusion of various botanicals, suffer from a lack of research, especially concerning their chemistry and the identification of authentication markers. This gap in research, coupled with the rising popularity of gins, presents potentially profitable opportunities for fraudsters. Currently, there are no existing methods capable of certifying the authenticity of gins, highlighting the need for comprehensive research in this area. In conclusion, addressing these challenges and preventing fraudulent practices in the alcoholic beverage industry through improved authentication methods is vital. By increasing detection efficiency and reducing incentives for committing crimes, it is possible to protect consumers and ensure they receive authentic products that meet their expectations. The complexity of alcoholic beverages' chemistry, influenced by various processing steps, necessitates the development of robust analytical tools for authentication. This knowledge will contribute significantly to the fight against food fraud, ensuring the integrity of the food systems.

CONNECTING TEXT

In the preceding chapter, a survey was performed on the complex and diverse literature of food authenticity and fraud, with a particular focus on alcoholic beverages. This exploration revealed a critical role of chemical composition analysis in authenticating these products. Chapter 2 emphasized the significance of techniques like LC-QTOF-HRMS based NTA coupled with MVA in identifying adulteration and verifying geographical and botanical origins. However, it also highlighted the emergent state of this approach in the realms of spirits, indicating an urgent need for further research in this area. Key challenges such as standardization of methods and the necessity for larger-scale studies were identified, underscoring the gaps in the current understanding of alcoholic beverage fraud, especially in relation to spirits like neutral grain spirits and gins.

Building on these foundations to transition to Chapter 3, an innovative study that pioneers the use of NTA RP-LC-QTOF-HRMS tailored to the chemical composition of neutral spirits and gins. This significant advancement in the literature of alcoholic beverages introduces a novel approach to the characterization of the complex chemical fingerprinting of these spirits. By integrating advanced techniques such as ESI and APCI ionization, and utilizing MVA, this study not only distinguishes between different categories of spirits but also uncovers unique markers, such as methyl cinnamate in gins, which have not been reported previously. This marks a substantial contribution to the field of alcoholic beverages authentication, addressing the highlighted challenges of the previous chapter and paving the way for further exploration in the chemical characterization of spirits.

3. CHARACTERIZING THE CHEMICAL FINGERPRINTS OF NEUTRAL SPIRITS AND GINS USING A ROBUST RP-LC-QTOF-HRMS METHOD AND MULTIVARIATE ANALYSIS

3.1 Abstract

This study innovatively employs a NTA approach utilizing RP-LC-QTOF-HRMS to establish the chemical fingerprints of neutral spirits (NS) and gins. Through the application of both ESI+ and ESI-, as well as APCI+ ionization modes, a robust and rapid method for NS and gins fingerprinting was developed and validated. Analysis using Venn diagrams allowed for the isolation of unique entities while PCA demonstrated that the spirits categories represented the main source of variability among the dataset with clear separation between both spirits. With the unique entities identified and the help of a screening list, three compounds were considered carefully as potentially being authentication markers for gins when compared with NS. Those entities underwent MSMS analysis to obtain detailed mass spectra details. These results along with the mass-to-charge ratio and retention time values obtained from MS data were used to compare with reference standards and imported in SIRIUS to proceed with chemical structure and formula elucidation. One compound, methyl cinnamate, emerged with significant similarity to its reference standard, proving to be a strong candidate in the authentication of gins. These results significantly advanced the current understanding of the chemical composition of NS and more specifically gins, paving the way for improved authentication and quality control measures within the beverage industry. The findings corroborate the hypothesis that NTA RP-LC-Q-TOF-HRMS is a viable technique for deciphering the complex chemical fingerprints of alcoholic beverages, and propose the exploration of this methodology in discovering new compounds of interest in other beverage categories. Future work should extend this validated method to a broader spectrum of spirits and alcoholic beverages to further explore the chemical diversity and potential applications in food and beverage authenticity studies.

3.2 Introduction

The authenticity and quality of alcoholic beverages, particularly spirits, have been the subject of increasing interest in recent years. This is attributed to the economic, health, and social implications of counterfeit and adulterated products. The complexity of alcoholic beverages, which

are composed of a myriad of volatile and non-volatile compounds, presents a unique challenge. This complexity is further enhanced by the fact that the chemical composition of alcoholic beverages is influenced by numerous factors, including the type of raw materials used, the fermentation process, and the aging conditions (Basalekou, Kyraleou et al. 2022). Recent progress in analytical techniques have made it possible to characterize in depth the chemical composition of spirits, such as high-resolution mass spectrometry (HRMS). These techniques, coupled with multivariate analysis (MVA), can provide a robust method for the authentication of spirits. However, these undertakings are not without challenges. Issues related to the validity of the samples labelled as authentic, and the sample size needed to prove that these approaches are applicable for establishing the fingerprint of neutral alcohols are expected. Moreover, the diversity and intricacies of the alcoholic beverage matrices complicate things further, resulting in gaps in the current scientific literature aimed at understanding their chemical composition. Recent studies have demonstrated the potential of these techniques in the discovery of relevant compounds in various spirits. For example, a study demonstrated the potential of UHPLC-QTOF-MS in the authentication of spirits. Through the analysis of the non-volatile composition of various whiskeys and by using a discriminant analysis model, the researchers were able to differentiate whiskey samples by type and age based on 40 compounds, of which only 8 were previously mentioned in the literature (Collins, Zweigenbaum et al. 2014). To the best of knowledge, there are no studies that have been published specifically on the elucidation of the chemical composition of NS or gins using LC-MS. This highlights the substantial work that remains to be done with these spirits.

In the realm of food authentication, it is becoming more evident that the MVA nature of chemical fingerprints demands a rigorous approach (Riedl, Esslinger et al. 2015). Unlike traditional TA which evaluates results compound-by-compound, fingerprinting approaches evaluate data on the level of chemical patterns, making the optimization and validation processes essential. Especially when dealing with complex matrices like alcoholic beverages, these steps are not only procedural formalities but are foundational to ensure the reproducibility and the credibility of the study results (Esslinger, Riedl et al. 2014). Proper method validation establishes the reliability of the findings and the potential replicability of the method for future studies. As the field of analytical chemistry in the context of NTA food authentication evolves, the importance of being rigorous cannot be overstated. ESI and APCI are two prominent ionization techniques pivotal for NTA in food authentication. While APCI excels for thermally stable, less polar

compounds, minimizing ion suppression effects (Ismaiel, Halquist et al. 2008), ESI is adept for polar compounds and large biomolecules (King, Bonfiglio et al. 2000, Kostiainen and Kauppila 2009). However, each technique has its limitations, such as matrix effects (Thurman, Ferrer et al. 2001) and fragmentation challenges (Li, Gan et al. 2015, Commisso, Anesi et al. 2017), with a lack of clarity on their comparative efficiencies for characterizing NS and gins. The effectiveness of these ionization techniques in detecting molecular features part of the chemical profiles of NS and gins remains unexplored. MVA employing software tools like with Profinder and MPP plays a crucial role in extracting, filtering, and analyzing primary LC-MS amenable compounds. While previous studies have demonstrated the utility of MVA in analyzing complex matrices of other alcoholic beverages (Collins, Zweigenbaum et al. 2014, Einfalt 2020, Uttl, Bechynska et al. 2023), the application and efficiency of these tools in studying NS and gins are not well established. Employing Venn diagrams for the delineation of unique entities in NS and gins is a novel approach, with its effectiveness in identifying unique chemical fingerprints in alcoholic beverages remaining scarce (Cao, Shu et al. 2023), despite being used with other matrices (Ueda, Iwamoto et al. 2019, Zhang, Cui et al. 2023). Other clustering methods like PCA have shown promise in distinguishing between different types of alcoholic beverages (Phan and Tomasino 2021), different botanical origins (Tzachristas, Dasenaki et al. 2021), and different geographical origins (Pan, Gu et al. 2022), yet its comprehensive application to NS and gins is lacking. Utilizing software like SIRIUS for the elucidation of chemical structure and formula of previously identified compounds within NS and gins is yet another critical step in understanding the complex chemical composition of NS and gins. Although SIRIUS has proven its accuracy across various fields (Dührkop, Nothias et al. 2021), its application with alcoholic beverage matrices is still an uncommon sight (Mallmann, O. Rios et al. 2023, Uttl, Bechynska et al. 2023).

The primary aim of this study is to characterize the chemical fingerprints of NS and gins for authentication purposes. The first step will be the validation of the proposed NTA LC-MS method followed by a rigorous assessment of the effectiveness of the different ionization modes. Following this, the use of MVA will allow for further characterization of the chemical composition of NS and gins by isolating entities among the comprehensive datasets obtained that are unique to specific spirit categories. Simultaneously, the applicability of PCA in distinguishing between NS and gins based on their respective chemical fingerprints will be evaluated to identify whether it represents a significant aspect of the variance in the dataset. Finally, utilizing SIRIUS software

and reference standards, the tentative identification of some compounds of interest will be pursued to potentially reveal a candidate marker for authentication.

3.3 Materials & Methods

3.3.1 Chemicals and reagents

HPLC grade methanol and acetonitrile were purchased from Agilent (Santa Clara, CA, USA). Deionized water was purified using a Milli-Q^{TM} water purification system (Millipore, Bedford, MA, USA). HPLC grade ammonium acetate was purchased from Sigma-Aldrich. Mobile phases were prepared immediately before LC-MS analysis. The volume of solvents and additives were carefully measured with clean and autoclaved graduated glassware. The solvents and additives were then sonicated for 5 minutes to ensure proper mixing and reduce the presence of bubbles in a Branson 3510 ultrasonic device by Emerson (Markham, Ontario, Canada). The labelled internal standards (caffeine-D_3, carbamazepine-D_{10}, DEHP-D_{38}, terephthalic acid-D_4, Bisphenol S $^{13}C_{12}$, and triclosan-D_3) were purchased from Sigma-Aldrich (St Louis, MO, USA), Toronto Research Chemicals (North York, ON, CA), and CDN Isotopes (Pointe-Claire, QC, CA). These were selected to have early eluting, mid-eluting, and late-eluting analytes to cover the complete analysis run. The reference standards were purchased from Sigma-Aldrich for α-pinene and methyl cinnamate (St. Louis, MO, USA), and from Cayman Chemicals for β-elemene (Ann Arbor, MI, USA).

3.3.2 Samples

Forty neutral spirits/vodkas and 23 gins were obtained from local producers in Quebec or bought at the provincial liquor board (Société des Alcools du Québec) in the Greater Montréal region. Samples were made of various raw materials that were grown in various geographical regions. Samples were catalogued by botanical origin, geographical origin, production type, product name, lot number, distillation date, sampling date, and delivery date. After opening the samples to transfer some into 2 mL amber vials, the original containers were resealed using parafilm and stored in a cool, dark, and dry location. The amber vials were then stored in the freezer, at -20°C for the duration of the experiment.

Sample volumes of 990 µL were transferred from sealed glass bottles to plastic syringes using glass microsyringes. PTFE filters (25mm × 0.2 µm) were then installed on the plastic syringe to

filtrate the samples into 2 mL HPLC amber glass screw vials. 10 µL of the labelled internal standard mixture was added to each vial and the vials were mixed with a vortex.

Pooled QC samples served as a critical quality control measure, ensuring the consistency and reliability of the analytical process (Evans, O'Donovan et al. 2020). They were made by mixing in equal parts all samples and were then spiked with the labelled internal standards mixture appropriate for the ionization mode. Solvent blanks consisting of 1:1 methanol in water was determined as the most suitable available blank since all samples had an ethanol concentration in the 40-60% range and no HPLC grade ethanol could be purchased in Canada.

3.3.3 UHPLC-QTOF-HRMS Analysis

UHPLC 1290 Infinity II (Agilent Technologies, Santa Clara, CA, USA) coupled with the G6546 QTOF-MS (Agilent) was used, along with a G7120 binary pump, G7167 multisampler, and G7116 multicolumn thermosampler, all provided by Agilent. The method was initially adapted from a study focused on contaminant identification in alcoholic beverage matrices (He and Bayen 2020). However, some parameters were optimized to better suit the specific requirements of NS and gins analysis and to improve the quality of chromatograms. After conducting preliminary tests, it was determined that the best combination was an injection volume of 10 µL and a flow rate of 0.3 mL/min. This injection volume provided a sufficient amount of sample for accurate analysis while minimizing band broadening and tailing, while the flow provided a reasonable timeframe with excellent resolution. The mobile phase for ESI consisted of 5 mM ammonium acetate in HPLC grade water (A) and 1:1 acetonitrile/methanol (B) and the elution gradient was as follows: 0.50 min (95% A; 0.300ml/min^{-1}), 4.00 min (0% A; 0.300ml/min^{-1}), 8.00 min (0% A; 0.300ml/min^{-1}), 8.01 min (95% A; 0.300ml/min^{-1}), and 9.00 min (95% A; 0.300ml/min^{-1}). The mobile phase for APCI consisted of 5 mM ammonium acetate in HPLC grade water (A) and 100% methanol (B) and the elution gradient was as follows: 0.50 min (95% A; 0.300ml/min^{-1}), 4.00 min (0% A; 0.300ml/min^{-1}), 9.50 min (0% A; 0.300ml/min^{-1}), 9.51 min (95% A; 0.300ml/min^{-1}), and 10.00 min (95% A; 0.300ml/min^{-1}). These conditions were chosen for their ability to provide high-resolution separations within a reasonable analysis time of 10 min. The APCI method was adapted from the ESI parameters with adjustments made to improve the suitability for this specific ionization mode (Fischer and Dunca 2007). The chromatographic system was equipped with a Poroshell 120 C18 column (3.0 mm x 100 mm x 2.7 µm) from Agilent Technologies. A pre-guard

column (3.0 mm x 5 mm x 2.7 µm) by Agilent was used to enhance performance, protect the column, and reduce the occurrence of contaminants. The UHPLC-QTOF-HRMS was equipped with two non-simultaneous ionization techniques for analysis: electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI) from Agilent. ESI was operated in both positive and negative ion modes, while APCI was only applied APCI+ mode. This decision was based on empirical observations during preliminary tests, which showed that APCI- did not offer the suitability required for the analysis due to mostly absent signals (data not shown). To the best of current knowledge, this specific limitation of APCI- has not been reported in existing literature, suggesting a potential area for future research.

The parameters for the mass spectrometer set with the ESI+ ion source were as follows: sheath gas flow of 12 L/min, sheath gas temperature of 375 ℃, gas flow of 11 L/min, gas temperature of 150℃, nebulizer pressure of 30 PSIG, capillary voltage of 4,000V, nozzle voltage of 1,000V, and skimmer voltage of 45 V. The detection conditions of the instrumental methods were identical for both ESI+ and ESI-, except for gas temperature (175℃ for ESI-) and nozzle voltage (2,000V for ESI-). The difference in temperature and voltage settings were selected based on optimization tests that indicated these were the most effective parameters in their respective mode. For the APCI ion source, the parameters were as follows: gas temperature of 300℃, vaporizer temperature of 350℃, gas flow of 5 L/min, nebulizer pressure of 30 PSIG, capillary voltage of 4,000 V, corona voltage of 4 V, skimmer voltage of 45 V. The instrument was operated in full scan mode in the *m/z* range 50-1,700Da. For the MSMS data, the same parameters were applied with a range of ±0.2min of target RT and 10 V, 20 V, and 40 V were used to fragment compounds and obtain comprehensive MSMS spectra.

To limit potential compound loss or discrimination, a direct-injection strategy that overcomes the need for any extraction or dilution of samples was employed, thus ensuring that the integrity of the original samples is maintained. The method of QC employed in this study was designed with the best contemporary practices in mind and surpasses the standard methodologies outlined in the literature in key areas. The use of randomized injections, solvent and system blanks, and internal standards for normalization post-acquisition was adopted. The methodology also incorporated the injection of pooled samples after every 10 samples to ensure reproducibility. Furthermore, system suitability was evaluated by injecting a blank of chemical standards and subsequently reviewing

the chromatographic peaks' width. Moreover, the method takes a step further in assessing the accuracy and sensitivity of the chemical standards' mass and retention times, leading to more robust and reliable results.

3.3.4 Compound Identification

A five-tiered confidence level system for molecular structure identification via mass spectrometry was used (Schymanski, Jeon et al. 2014). Level 1 offers definitive molecular identification through a reference standard and utilizes MS, MS/MS, and retention time matching, ideally corroborated by an orthogonal method. Level 2 proposes an exact structure based on different forms of evidence, either unambiguous literature or library spectrum data, or MS/MS fragments and ionization behaviour, if no reference standards are available. Level 3 provides a grey zone where evidence suggests possible structures, but none can be confirmed. Level 4 provides an unequivocal molecular formula based on spectral data yet lacks sufficient evidence for structure identification. Finally, Level 5 identifies only the exact mass (m/z) without additional structure or formula identification. The use of this robust classification system aims to standardize the identification confidence across other studies.

3.3.5 Method Validation

A rigorous method validation analysis was conducted to ensure its robustness, specifically targeting the precise characterization of NS and gins. Intra-reproducibility was measured using relative standard deviance (RSD). The RSD for the intra-day precision was calculated based on the analysis of spiked pooled QC samples at 20 µg/L injected every 10 samples (n=10). Mean mass measurement error (MME) was determined using the mass of internal standards to assess the accuracy of mass measurements. It was calculated using the difference between the observed mass and the true mass of internal standards in parts per million (ppm) with the following formula: $MME_{ppm} = \left(\frac{Observed\ mass - Theoretical\ mass}{Theoretical\ mass}\right) \times 10^6$. Suitable mass measurement accuracy was considered to be below 2 ppm (Villar-Pulido, Gilbert-López et al. 2011). LOD and LOI were determined with the use of the signal-to-noise ratio (S/N) with thresholds equal to S/N=3 and S/N =6, respectively, as suggested in the literature (Vogelgesang and Hädrich 1998, Lappas and Lappas 2016). The stability of these metrics was assessed across the duration of the analyses to ensure they provided consistent and reliable detection of MFs, thereby enhancing the robustness of the

analytical method. Matrix effects were evaluated by using the same spiked pooled QC as previously described for intra-reproducibility. The concentration of the analytes in the post-spike samples was then compared to those in solvent blanks to determine the impact of the matrix on the analytes. The matrix effect (ME) was calculated using the formula:$ME = \left(\frac{I_{matrix}}{I_{blank}}\right) \times 100$. ME within the range of 85-115% was considered acceptable (Viswanathan, Bansal et al. 2007).

3.3.6 Data Analysis

Assessing peak quality was performed using Agilent MassHunter Qualitative. Data processing, such as peak alignment, normalization, and gap filling was carried out before any data analysis using Profinder. The molecular feature extraction (MFE) procedure was made using the "Batch Recursive Small Molecules Feature Extraction" wizard. The first step set a feature count of 2000, focusing on ion species that included protonated (H^+) and deprotonated (H^-) species, as well as common organic molecules. This was followed by the second step, which utilized the default settings of the software for analysis. In the third step, the RT shift was set to 0% ± 0.1 based on the RT shift of internal standards observed in pooled QC runs, and a mass tolerance of 20 ppm ± 0 mDa was set. The fourth step expanded the feature count from step one to 4000 and forced a minimum matching score of 70, with the condition that the features must be present in at least one file in one sample group. In step 5, settings were adjusted to include a symmetric tolerance of ± 0.35 ppm, while raising the minimum RT matching score to 80. Steps six and seven maintained default settings. Finally, the eighth step stipulated a minimum absolute height of 2500 and required a minimum matching score of 80 in at least one file in one sample group. This configuration based on preliminary data ensured optimal feature extraction from the data, leading to the identification of significant compounds. The statistical analysis undertaken in this research aimed to compare data among various sets of samples to identify chemical patterns or unique features with the potential to serve as authentication markers.

MPP software by Agilent was utilized to filter the MFs based on the pooled QC data, optimizing the data analysis through a series of filtering parameters. Initially, a frequency cut-off percentage was set, retaining only features that occurred in more than 40.0% of all conditions to ensure consistency across samples. This was followed by setting a sample variability threshold at less than 40.0%, thereby filtering out features with high variability and potentially less reliability.

Further refinement was carried out by implementing a minimum abundance criterion, where only features displaying an abundance of over 100,000 in all samples were retained. This stringent filtration process ensured that the remaining MFs were of high interest for the study, based on their consistent and frequent presence across the samples, their low variability and high reliability, and their high abundance (Dudzik, Barbas-Bernardos et al. 2018, Gravert, Vuaille et al. 2021). Accounting for variations in sample size between different categories, a normalization step was performed to ensure a fair comparison of the analyte counts across different ionization modes. As the number of samples for NS was about twice the number of gin samples, normalizing the data based on sample size helps in mitigating this bias when comparing the ionization efficiency between these spirits. The normalization was performed by calculating the average analyte count per gin samples for each ionization mode, which was then projected to the sample size of NS to estimate the equivalent analyte counts.

The research extends this practice by conducting comprehensive multivariate analysis including PCA, Venn diagrams, and Volcano plots. While PCA is commonly used to assess sample group clustering, the analysis benefits from the further incorporation of Venn diagrams to identify unique molecular features among sample groups. This inclusion improves upon the conventional methods by providing a more detailed comparison of the sample groups. Additionally, the research exploits the use of Volcano plots to compare two sample groups and find statistically significant ($p<0.05$) molecular features with a significant fold change difference (>2.0). This use of Volcano plots provides a detailed comparison that enhances the detection of statistically significant differences.

3.3.7 Compound Identification

Utilizing the Venn diagrams generated from prior analyses, molecular features of interest (MFOIs) unique to each spirit type were isolated. Concurrently, a suspect screening list was compiled based on literature regarding the chemical composition of gins to provide an extensive list of candidate markers for gins. The procedure entailed evaluating the list of unique MFs satisfying one of these two criteria: 1) either presence or absence in over 90% of all samples within a specified category (e.g., gins or NS) and correspondingly absence or presence in the opposite category, and 2) a significantly higher or lower intensity of presence or absence compared to the other category. Alongside, a suspect list of potential key compounds identified from an extensive

literature review was examined. SIRIUS 5.8, a software tool developed by Böcker group that uses algorithms to deduce the molecular structures from MS^2 data through molecular ion detection and preprocessing, molecular formula annotation, fragmentation tree computation, tree scoring and ranking, and molecular structure prediction (CANOPUS), was used with imported MS^2 data for compounds reconstitution which helped in the process of compound identification and characterization (Böcker and Dührkop 2016, Dührkop, Fleischauer et al. 2019).

3.4 Results and discussion

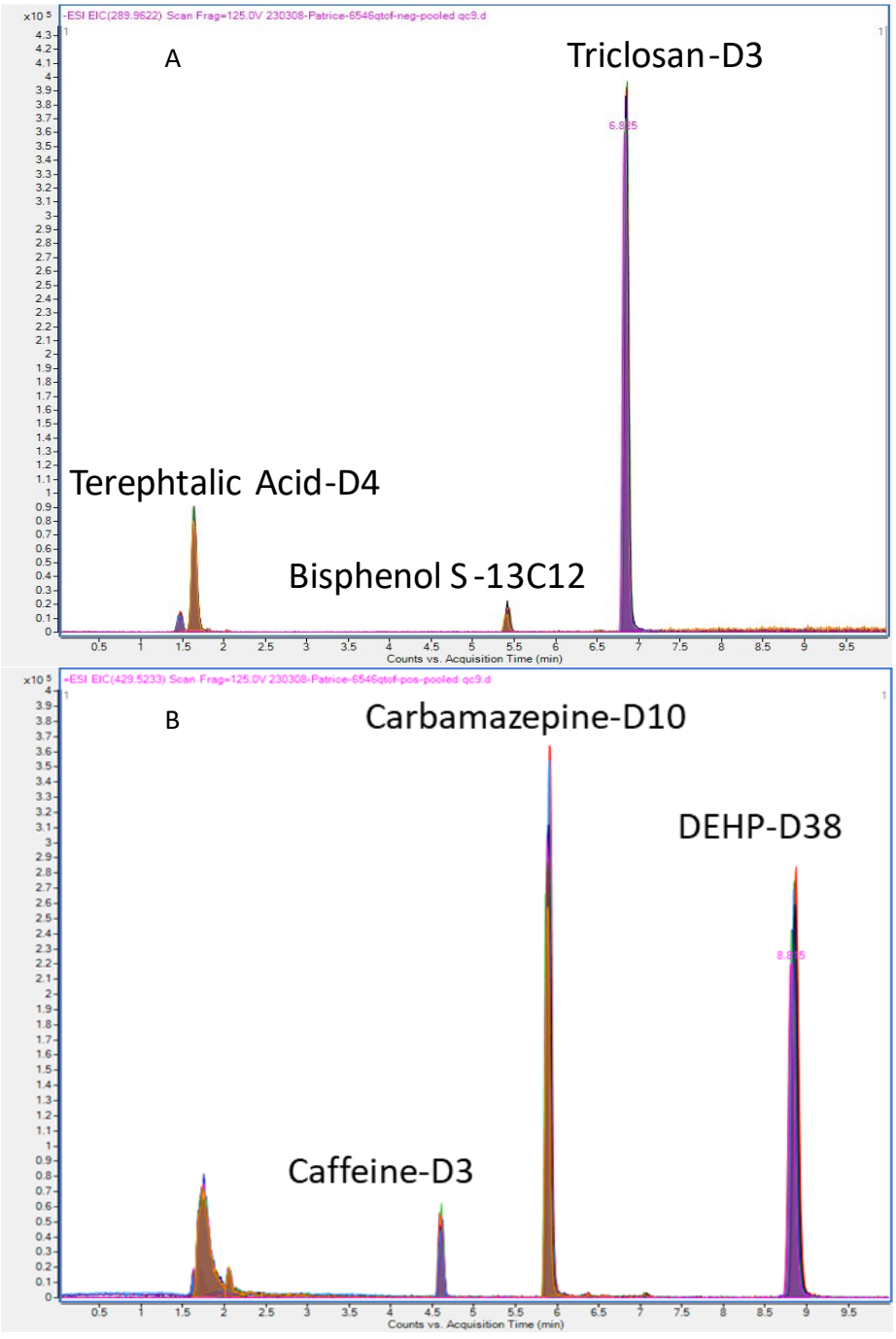
3.4.1 Method Validation

To ensure that the proposed method, adapted from He and Bayen (2020), was suitable for distilled spirits analysis, method validation parameters were measured, therefore ensuring efficiency and robustness. Multiple analyses were conducted, manipulating analytical parameters such as mobile phases, flow rates, and injection volumes. An APCI method was also developed based on the previously described ESI, making minor parameter changes to adapt to this ionization process. For the APCI, the mobile phase was modified to 100% methanol due to its suitability for maximum analyte coverage in APCI. Acetonitrile's use in APCI was avoided since it can prevent the formation of ions in both the positive and the negative ion polarities (Fischer and Dunca 2007). Due to the absence of internal standards' signals in the negative APCI polarity, only the positive polarity data was kept for further analysis.

MME assessment ensures the mass error is within the established parameters for the data processing. The values were obtained from the comparison of observed mass in internal standards (see Figure 3.1) to theoretical mass of the same compounds. The MME ranged from 0.23 to 2.52 ppm in ESI+ and 0.41 ppm to 2.07 ppm in ESI- mode. Despite the recommendation of having MME below 2 ppm (Villar-Pulido, Gilbert-López et al. 2011), these results fall below 5 ppm, which is often considered to be acceptable for most studies (Nácher-Mestre, Ibáñez et al. 2013).

ME were assessed through measurements of the intensities of the labelled internal standards in a single run, as shown in Table 3.1. The resulting ME ranged from 0.52% to -10.73% for ESI+, 0.86% to -37.84% for APCI+, and 2.49% to 52.22% for ESI-. It is important to acknowledge that with NTA methods, it is challenging to account for every analyte due to the variability in their properties and behaviour. Normalization or compensation techniques, like matrix matching

(Giacinti, Raynaud et al. 2016) or dilution (Yang, Chang et al. 2015), would be impossible to work well for all analytes and could have unintended effects on analytes that are within the acceptable range.



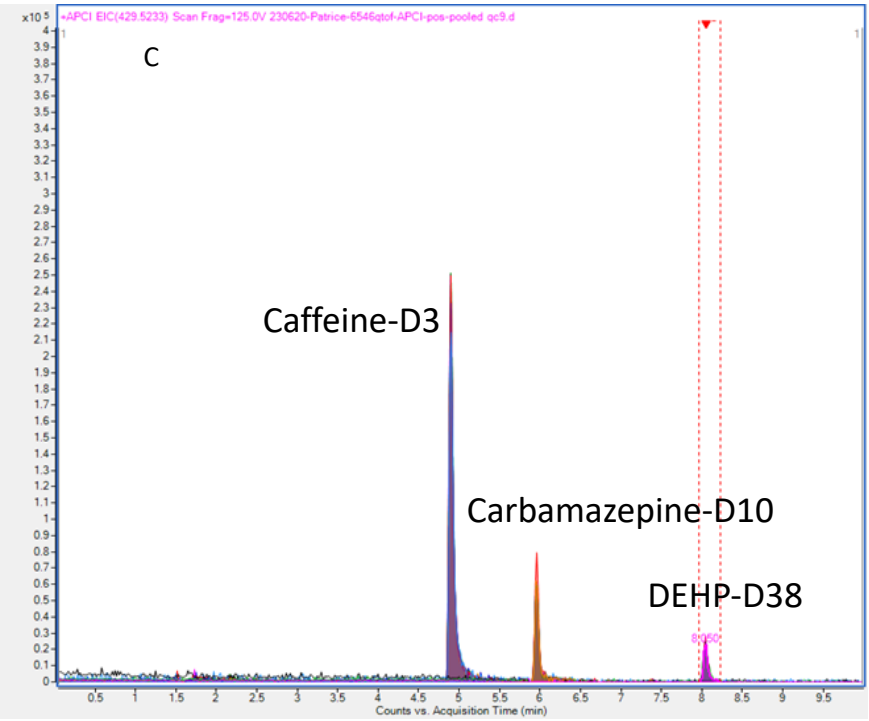


Figure 3.1 Internal standards reproducibility testing for ESI- (A), ESI+ (B), and APCI+ (C).

The LOD and LOI values were utilized to establish the minimum level at which compounds could be detected and identified, respectively. The signal values were calculated as the average of the signals in each spiked blank sample, while the noise value was determined by averaging the noise around each of the internal standard peaks. The final S/N values were obtained by dividing the averaged internal standards' absolute intensities by the averaged noise values. The S/N ratios ranged from 11 to 850, which were deemed more than enough to achieve a good detection of signals and the accurate identification of these compounds, suggesting the robustness of the methodology.

The intra-day intensity relative standard deviation (RSD%) varies within acceptable ranges for the methods tested: 5.67-11.98% (ESI+), 3.31-10.47% (ESI-), and 6.16-11.03% (APCI+). Furthermore, the intra-day retention time shift (RTS RSD%) exhibited minimal fluctuations for all ionization modes: 0.19-0.29 min (ESI+), 0.15-0.68 min (ESI-), and 0.05-0.10 (APCI+), thus confirming the stability and reproducibility of the experimental conditions. Since all RSD% values were below 15-20% (see Table 3.1), it demonstrates that the experiment had a commendable degree of precision and repeatability in its measurements.

	Internal standards					
Ionization mode	**Caffeine-D3**	**Carbamazepine-D10**	**DEHP-D38**	**Terephthalic Acid-D4**	**Bisphenol S-13C12**	**Triclosan-D3**
	Intra-day (Intensity) RSD (%)					
ESI+	11	12	6.9			
ESI-				7.9	10	3.3
APCI+	6.2	11	10			
	Intra-day RTS RSD (%)					
ESI+	0.2	0.2	0.3			
ESI-				0.7	0.2	0.2
APCI+	0.1	0.1	0.1			
	ME (%)					
ESI+	4.7	0.5	-11			
ESI-				52	-6.3	2.5
APCI+	3.2	0.9	-38			

Table 3.1: Matrix effects (ME%) and relative standard deviation (RSD%) for intra-day intensities and retention times across three ionization modes (ESI+, ESI-, AND APCI+) using pooled qc samples spiked with labelled internal standards.

With the method validation parameters assessed, the results demonstrate high levels of efficiency, robustness, and reproducibility. The ME and RSD% data corroborate the method's precision, meeting the overarching aim of achieving a rigorous approach to NTA NS and gins characterization. Our data show acceptable ranges for intra-day intensity and RT RSD%, affirming the stability and repeatability of the method. The high S/N ratios further establish the efficacy of the approach for capturing a broad spectrum of analytes with high accuracy. These metrics align well with the stated research objectives, particularly in enhancing data quality and enabling robust identification of compounds of interest.

These findings are in accordance with prior studies emphasizing the necessity for stringent method validation in NTA (Esslinger, Riedl et al. 2014, Beger, Dunn et al. 2019). Notably, comprehensive method validation reporting, including parameters like MME, ME and RSD%, is not as widespread in NTA conducted on alcoholic beverages as it is in other NTA fields. Most studies in this area do not provide an exhaustive validation, lacking key metrics such as those previously mentioned. In this context, the present study provides a much-needed contribution to fill this gap in the literature. The observed ME and RSD% values of this study compare favourably

with the rare but pertinent studies that do report these measures (Tzachristas, Dasenaki et al. 2021). By situating the method's performances within the objectives and against the backdrop of scarce, albeit relevant, existing literature, its robustness and applicability for the complex task of characterizing NS and gins is affirmed. The detailed validation approach adopted in this study sets a precedent for future research, aiming to standardize validation procedures in NTA.

3.4.2 Comparison of the different ionization methods (ESI-, ESI+, APCI+) for the characterization of neutral spirits and gins

In the endeavour to characterize the chemical fingerprints of NS and gins, the developed LC-MS method was used for the untargeted examination of a broad range of analytes. Figure 3.2 and S1.1 provide the resulting total ion chromatograms (TIC) obtained from the samples in various ionization modes, showcasing the potential richness of the data to be analyzed. Given the need for a nuanced understanding of the chemical composition of NS and gins, multiple ionization methods were used, namely ESI and APCI, each offering distinct advantage. To mitigate risk of compound loss or discrimination, a direct-injection strategy was chosen. This method bypasses the need for any sample extraction or dilution, thereby maintaining the integrity of the original samples.

The methods provided distinct insights into NS and gins' chemical composition. ESI's larger detection capacity made it the preferred method for comprehensive characterization. However, APCI+ was not considered redundant as it helped in detecting different types of molecular features enriching the overall characterization, while the joint use of these two ionization methods helped obtain more detailed chemical fingerprints. For instance, some MFs were only detectable with ESI (ex., m/z 205.1956 at RT 6.958min), while others, only with APCI (ex., m/z 199.1487 at RT 7.05 min). This was expected since both ionization methods work best with different classes of compounds due to their intrinsic difference in ionizing mechanisms (Thurman, Ferrer et al. 2001, Commisso, Anesi et al. 2017). This suggests that both methods can simultaneously offer a different perspective and validate data seen in the other method.

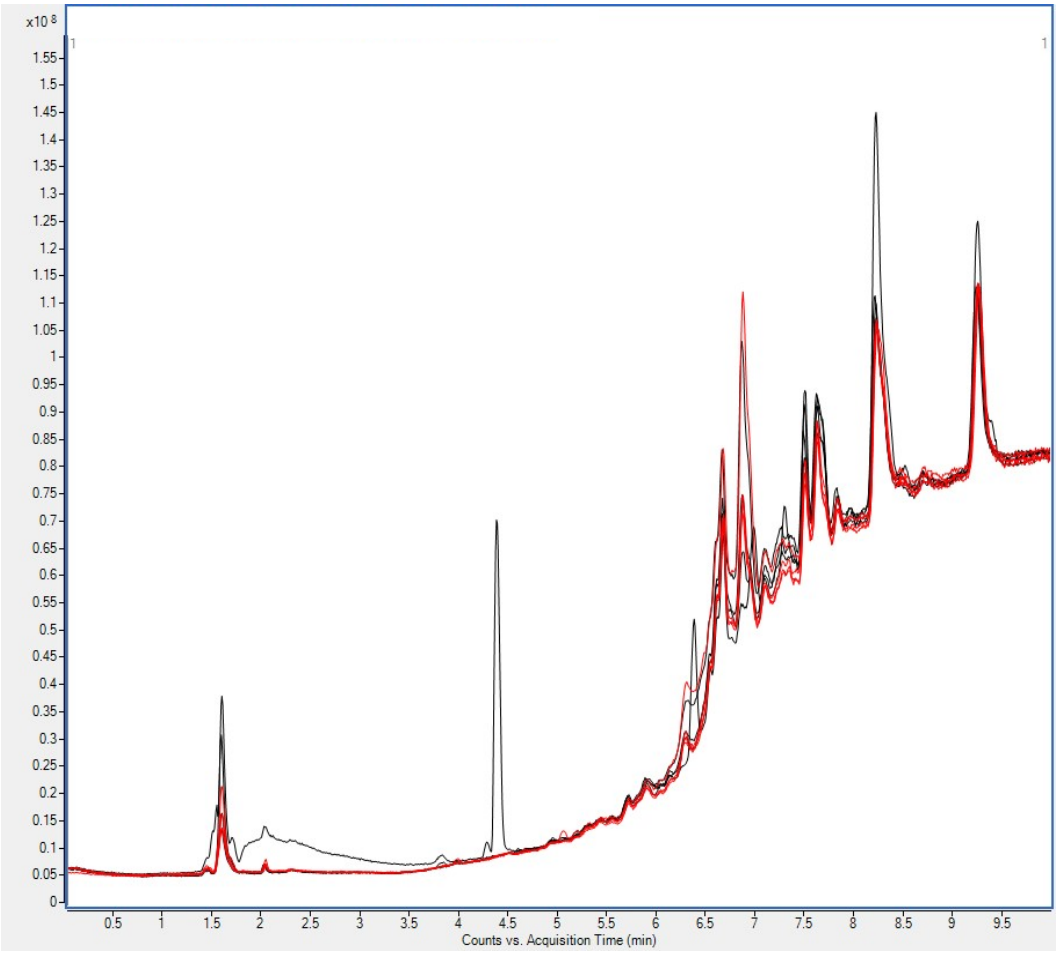


Figure 3.2 Exemplary overlapped TICs of NS (red) and gins (black) for ESI-. See supplementary materials for ESI+ and APCI+ (S1.1).

Interestingly, it was found that ESI+ detected the most features, followed by ESI-, and then APCI+. This finding corroborates previous research conducted other types of alcoholic beverages that also explored the differences between ESI modes (Vaclavik, Lacina et al. 2011, Ehlers, Uttl et al. 2023) and between ESI and APCI modes (Mazerolles, Preys et al. 2010). However, such findings have not been previously reported specifically for NS or gins. Existing literature suggests that certain analytes ionize more in the positive mode, while others produce stronger signals in the negative mode possibly explaining the discrepancies between ESI polarities (Liigand, Kaupmees et al. 2017). To ensure a comprehensive detection of analytes, data from all ionization modes was utilized. Regarding the peak intensities observed, most gin samples displayed maximum intensity signals within a consistent range, varying from 1.45×10^8 (ESI-) and 2×10^8 (ESI+ and APCI+). A similar pattern was observed for NS samples, with maximum intensity signals falling between 1.75×10^8 (ESI-) and 2.5×10^8 (ESI+ and APCI+), indicating the general robustness of the method.

It is worth noting that these ranges fall within reasonable limits for the analysis, making it unlikely that the observed variability arises from methodological limitations. However, there were some notable outliers. Specifically, two samples, Quebec corn gin at a RT of 4.75 min (3.2×10^8) and Quebec rye gin at a RT of 7.61 min (2.5×10^8), presented very high intensity peaks when analyzed with APCI+. Similarly, in the NS samples, two outliers were observed, both being Quebec corn-barley NS at a RT of 7.62 min (5.0×10^8) in APCI+. Importantly, some instances of ion saturation were observed in the chromatograms, and not only in these outliers. Ion saturation could potentially distort quantification and should be considered when interpreting these specific results.

Delving into the specifics of the findings, NS revealed the detection of 48,345 (range of 916-2,126 per sample) MFs in ESI+, 26,138 (range of 530-1,185 per sample) MFs in ESI-, and 8,695 (range of 110-670 per sample) MFs in ACPI+. In contrast, for gins, these figures stood at 28,102 (range of 1,213-2,896 per sample) MFs in ESI+, 16,068 (range of 675-2,715 per sample) MFs in ESI-, and 5,777 (range of 261-545 per sample) MFs in APCI+. For a clearer comparison, the data was normalized based on sample size, as described in the materials and methods section, to account for the different number of samples in NS and gins categories. It is suggested that gins would have had an estimated 48,873 MFs in ESI+ (1.1% higher than NS), 27,944 MFs in ESI- (6.9% higher than NS), and 10,047 MFs in APCI+ (15.5% higher than NS) after normalization. This data indicates that the difference in analyte count between NS and gins ranges from 1% to 15%, a smaller disparity than might be expected given the addition of botanicals in the production process of gins. Existing literature on their respective chemical compositions would suggest a much wider difference in chemical complexity between NS and gins, which are further elaborated with botanicals (Riu Aumatell 2012, Aylott 2016, Einfalt 2020). However, the narrower than expected disparity observed may reflect the specific sensitivities and characteristics of the LC-MS technique, highlighting the potential for uncovering nuanced differences between NS and gins when employing this methodology.

A notable observation is the increased disparity in the normalized number of MFs between NS and gins when utilizing APCI as opposed to ESI. This insight could be attributed to either a diminished detection rate of MFs in NS, or conversely, an enhanced detection rate in gins. The augmentation in detection within gins can possibly be traced back to the inclusion of botanicals, which introduces compounds of non-polar to moderately polar nature, such as terpenes. The

literature asserts that APCI is well adapted to the ionization of such compounds due to its weaker matrix effects and its superior handling of the high volatility characteristic of some gin compounds during the nebulization process (Ismaiel, Halquist et al. 2008, Trufelli, Palma et al. 2011, Kevin J. McHale 2018, Brecht, Uteschil et al. 2020). Despite this narrower difference between the count of MFs in NS and gins when using APCI, the overall detection rate of APCI was significantly lower compared to ESI, making the latter more appropriate for the overall characterization due to this higher range of detected compounds per sample. Therefore, despite the interesting findings with APCI+, the decision was made to prioritize ESI as the method of choice for this study. The number of detected MFs in both NS and gins in the present study constitutes a novel contribution to the existing body of literature on alcoholic beverages, especially for those specific spirits. Overall, these findings constitute a novel contribution to the existing body of literature on alcoholic beverages, especially for these specific spirits. This study represents the first comprehensive comparative analysis of TICs and the number of detected MFs in NS and gins using LC-Ms, providing a baseline for future investigations and a deeper understanding of the chemical complexities of these spirits. The discrepancies observed in comparison to general expectations based on known chemical compositions highlight the potential for LC-MS to provide new insights into the molecular profiles of NS and gins.

3.4.3 Selection and filtering of key features in neutral spirits and gins

The combination of MFE, FbI, and filtration led to significant reduction in the total number of MFs.This stringent protocol effectively filtered out unreliable data, including noise, false positives, and false negatives. Thus, 1,350 MFs in ESI- and 2,259 in ESI+ remained, all of which were amenable to LC-MS. Furthermore, the expected reduction of MFs, a crucial step for this study, aligns well with existing studies that also employ this approach for cleaner and more statistically significant data (Ehlers, Uttl et al. 2023, Uttl, Bechynska et al. 2023). This reduction was expected yet remains a crucial step for this study. Given that this is the first application of such methodology to NS and gins, this reduction is instrumental in providing a cleaner dataset specific to these beverages for subsequent MVA. The remaining MFs were presumed to be substantial contributors to the chemical composition of NS and gins due to their high abundance and statistical significance, thereby strengthening the validity of this study. This rigorous selection process enables the focus on MFs that are most representative of the unique chemical composition of NS and gins. It also

forms the basis for actual compound identification and further statistical analysis fulfilling the specific objective of this study. These results align well with existing studies employing a reduction of MFs to produce cleaner, more statistically significant data. However, the focus on NS and gins marks this research as the first of its kind to apply such methods to these beverages. The following sections will detail the results of further investigations, showcasing the unique and shared chemical components that define the characteristic of NS and gins.

3.4.4 Isolation of unique entities in neutral spirits and gins

Figure 3.3 demonstrates the efficiency of the filtration steps in reducing the number of MFs, ensuring the retention of only the most reliable and abundant MFs for further analysis. The resultant dataset post-filtration provides a clear representation of the core MFs common to both spirits, which are crucial for comparative analysis. Delving into the comparative aspect, it was observed that a large proportion of MFs was shared between NS and gins. This commonality was expected given that the base spirit of gins is oftentimes a neutral spirit. Moreover, the higher number of MFs observed in gins is similarly explained by the additional botanicals used in gins that contribute to its broader chemical profile (Riu Aumatell 2012). The identification of these unique entities is pivotal for authentication purposes and provides a deeper understanding of the impact of botanicals on the chemical complexity of gins. These comparative results of MF count are novel in the context of NS and gins when analyzed with LC-MS. In total, the combined MF count for ESI- and ESI+ revealed close 50,000 MFs pre-filtration and 7,000 MFs post-filtration. In comparison, a study performed on 16 gin samples with FT-ICR-MS reported close to 3,000 MFs with ESI- and APPI+ combined (Dou, Mäkinen et al. 2023). Another study conducted on whiskeys using UHPLC-QTOF-MS reported 7,000 MFs pre-filtration and 3,100 MFs post-filtration in about 60 samples (Collins, Zweigenbaum et al. 2014). Thus, the analyses revealed that the total MF count in both NS and gins was significantly higher than what has been reported in other alcoholic beverages studies. This reflects the efficiency and sensitivity of the developed NTA LC-MS method employed in this study and underscores its suitability in covering a broad chemical range for NS and gins. When comparing pre-filtration (A and B) to post-filtration (C and D), it becomes evident that the filtration process favoured only MFs characterized by high abundance, high reliability, and low variability, resulting in significant reduction of total MFs, approximating a total reduction of 87% in ESI- and 85% in ESI+ respectively. For NS, the filtration efficiency

decreased the MF count by 99.58% and 97.77% for ESI- and ESI+, respectively. Similarly, for gins, the filtration steps were able to reduce the count of MFs from 1,710 to 105 MFs (filtration rate of 93.86%) in ESI- and from 1,297 to 64 (filtration rate of 95.06%) in ESI+.

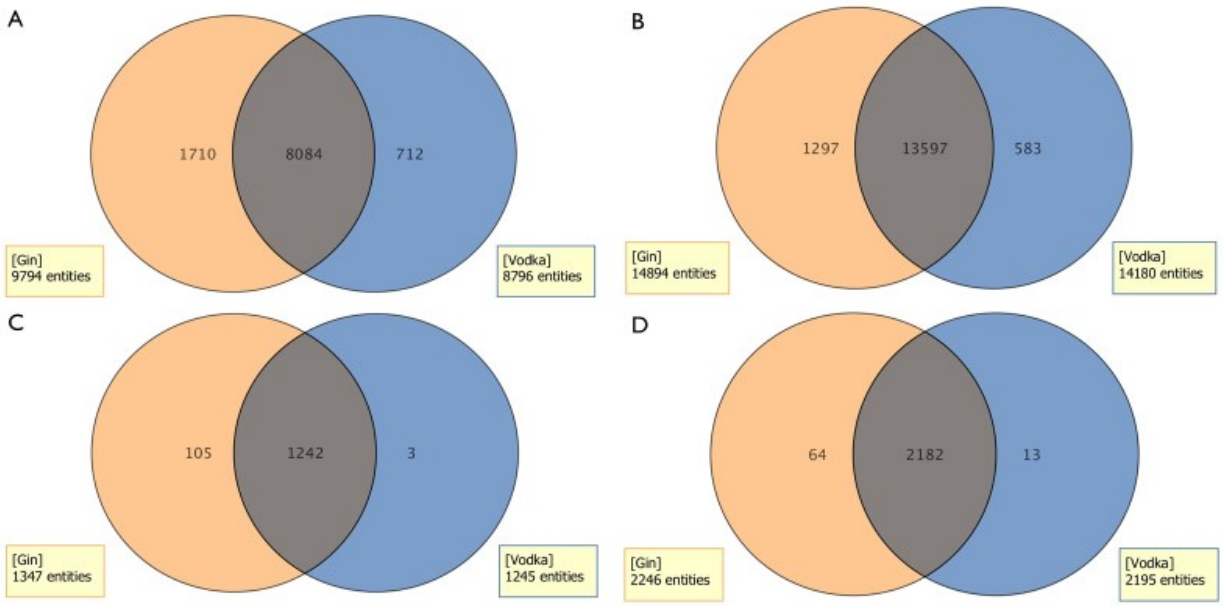


Figure 3.3 Venn diagrams of unique features in NS (blue) and gins (orange) in ESI- (A: before filtration, C: after filtration) and ESI+ (B: before filtration, D: after filtration).

Upon closer examination of the shared compounds between NS and gins, it is observed that before molecular feature filtration, 76.9% of entities were shared by NS and gins in ESI-, compared to 92% after filtration, while the overlap was of 87.9% for ESI+ before filtration, and of 96.6% after filtration. The increase in shared MFs affirms the inherent chemical similarities between NS and gins. This can be rationalized by understanding the inherent attributes of the filtration process, which is designed to retain MFs that exhibit high reliability, low variability, and high abundance, filtering out less reliable or less abundant MFs. Consequently, the entities remaining post-filtration represent a core set of MFs shared between both spirits, an indication of their chemical similarities. The amplification in overlap is not merely a theoretical expectation based on the filtration process's design, but also is empirically observed in the data, thereby serving as a practical validation of the filtration method's effectiveness. The developed method showcased its aptitude for this study by efficiently isolating meaningful MFs in both NS and gins. The substantial overlap of MFs between spirits elucidates their chemical similarities, while the unique chemical fingerprint of each spirit was preserved. The detection of unique entities for NS and gins is a novel contribution as this type

of comparative study has never been performed before using an LC-MS approach. These findings highlight the robustness of the method and offer valuable insights for future chemical characterization studies in the field.

3.4.5 Clustering analysis using PCA as a means to differentiate the chemical fingerprints of NS and gins

Following preliminary analyses and MF selection, PCAs in MPP for both ESI- and ESI+ modes compared NS and gin chemical fingerprints. Rigorously selected MFs ensured the PCA's statistical significance, ensuring a reliable comparison. Tight clustering of methanol and ethanol blanks, along with pooled QC samples in PCA (Figure 3.4), confirmed LC-MS analysis reliability (Beger, Dunn et al. 2019). However, pooled QC samples diverged from central clustering due to unique honey-based spirit samples, which introduced bias (Evans, O'Donovan et al. 2020). This outlier, attributed to the novel use of honey, was excluded from further analyses to avoid bias. ESI- PCA showed clear separation between NS and gins, reflecting differences in their chemical composition. This result was expected due to different production methods and the addition of botanicals in gins (Riu Aumatell 2012, Aylott 2016). This result supports the hypothesis that NS and gins can be delineated based on their chemical fingerprints. The covariance for ESI- was mainly explained by PC1 (33.66%), PC2 (7.02%), and PC3 (5.14%), highlighting group differentiation. In ESI+ mode PCA, separation was less defined due to more MFs creating more variance, with less of it being solely attributed to the spirit categories. However, distinct planes for NS and gins still indicated different chemical compositions. The explained covariance for ESI+ was PC1 (23.74%), PC2 (10.6%), and PC3 (6.04%). These results, consistent with other studies, shows significant compositional differences between spirits due to varied production methods and ingredients. Yet, intergroup dispersion was moderate compared to tighter clustering in other studies, possibly due to different instrumental methods. The explained covariance in the PCA was slightly lower than in other studies (Wiśniewska, Śliwińska et al. 2016, Fleming, Chen et al. 2020). Nevertheless, this was the first PCA performed on NS and gins using LC-MS to the best of knowledge, thus contributing significantly to the literature on these beverages in the context of the characterization of their chemical fingerprints. These findings not only validate the unique chemical fingerprints of NS and gins both visually and statistically but also serve as a foundational step for subsequent phases of the present study and for future research on these spirits.

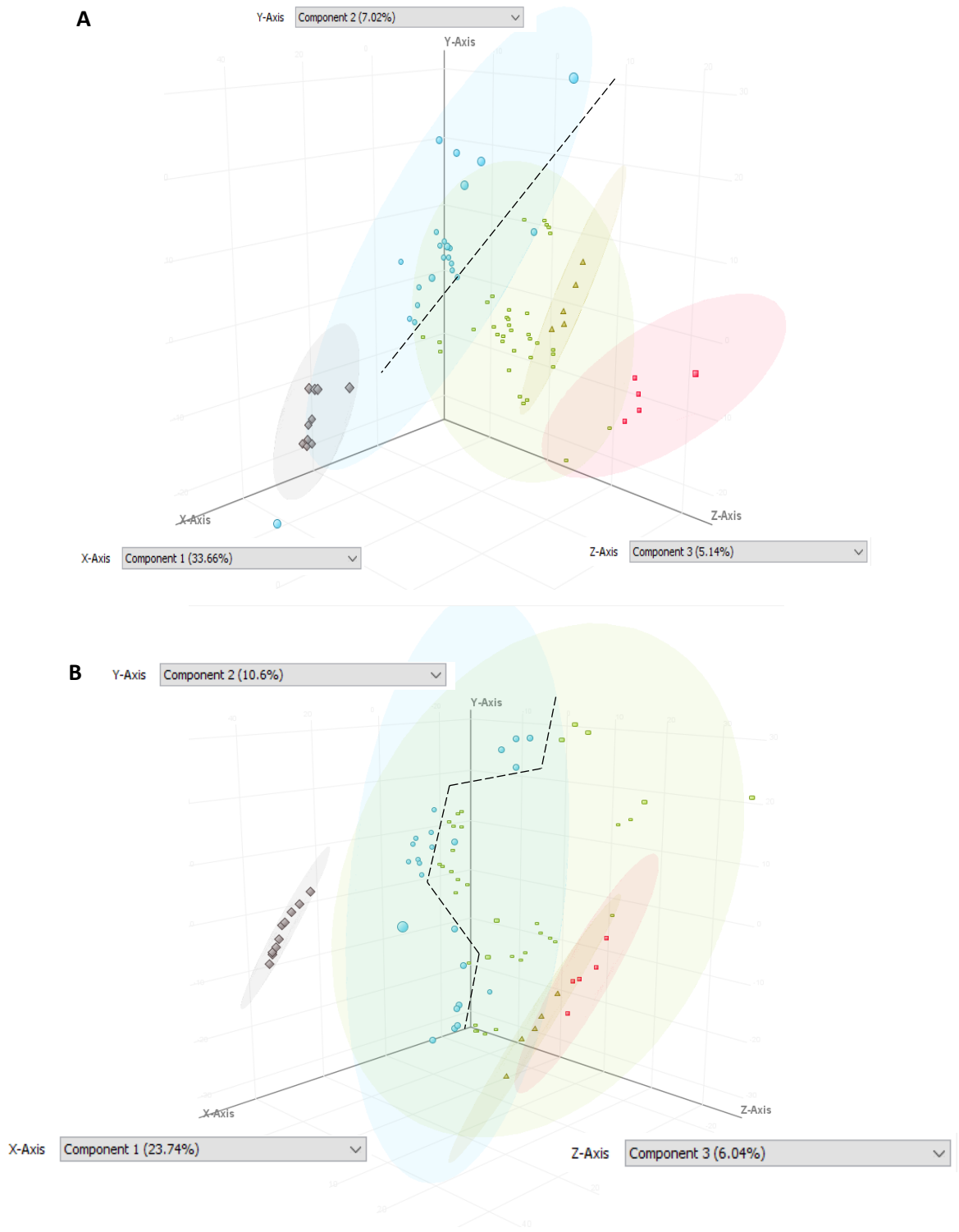


Figure 3.4 3D PCAs of NS and gins for ESI- (A) and ESI+ (B). Pooled QC samples (grey diamonds) and blanks (MeOH: dark-yellow triangles and EtOH: red squares) clustered separately and tightly. Gins (blue circles) were located nearer the plane formed by the y- and x-axes, while NS (green ovals) were located nearer the one formed by the y- and z-axes. Black dotted lines indicate the separation line between NS and gins.

3.4.6 Compound identification

From the initial suspect screening list and the MFOIs from previous analyses, after evaluation for data quality, the list of candidate markers was narrowed down to three potential MFs for gins: α-pinene, β-elemene, and methyl cinnamate. Unfortunately, no MFs of interest (MFOIs) were considered relevant for further data analysis concerning unique entities for NS, primarily due to false positives, lack of consistent presence in over 90% of NS samples, or inadvertent presence in some gin samples. Following MSMS analysis, SIRIUS software was utilized for further structure elucidation. However, SIRIUS analysis proposed an incorrect structure of benzoylacetone with 66.94% score for the MF and 71.88% for the reference standard of methyl cinnamate, albeit correctly identifying the chemical formula ($C_{10}H_{10}O_2$)(Figure 3.6). To ascertain the identity of these MFs, a confirmation step using reference standards was performed. The comparison for methyl cinnamate was promising as the retention time, m/z, and MSMS spectra matched with the reference standard (Figure 3.5), despite slight differences in fragment intensity. This substantiated the potential presence of methyl cinnamate in gin samples. For the terpenes, both the reference standard comparison and the SIRIUS analysis yielded poor results. Since both α-pinene and β-elemene are terpenes, this result was expected as they are known to poorly ionize in ESI due to the lack of regions favourable for protonation (Banerjee and Mazumdar 2012). The absence of MFs that passed the rigorous selection evaluation underscores the complexity of the spirits' composition and the challenges inherent in isolating unique entities, despite rigorous selection processes. This outcome, while seen as unfavourable for the objectives of this study, also highlights the complexities involved in isolating distinct entities even with the application of rigorous selection processes. The scant literature in the detection or identification of unique compounds in NS using LC-MS makes this result a foundational stone which underscores the need for further exploration into this uncharted territory. As for the MFOIs found in gins, the decision to challenge the known limitations regarding α-pinene and β-elemene came from experimental testing conducted on cannabis where these specific terpenes, along with others, could still be detected when present in sufficient concentration (not published). Additionally, both terpenes are commonly found in high concentration in juniper berries and multiple other botanicals that are frequently added to gins (Barjaktarović, Sovilj et al. 2005, Buck, Goblirsch et al. 2020).

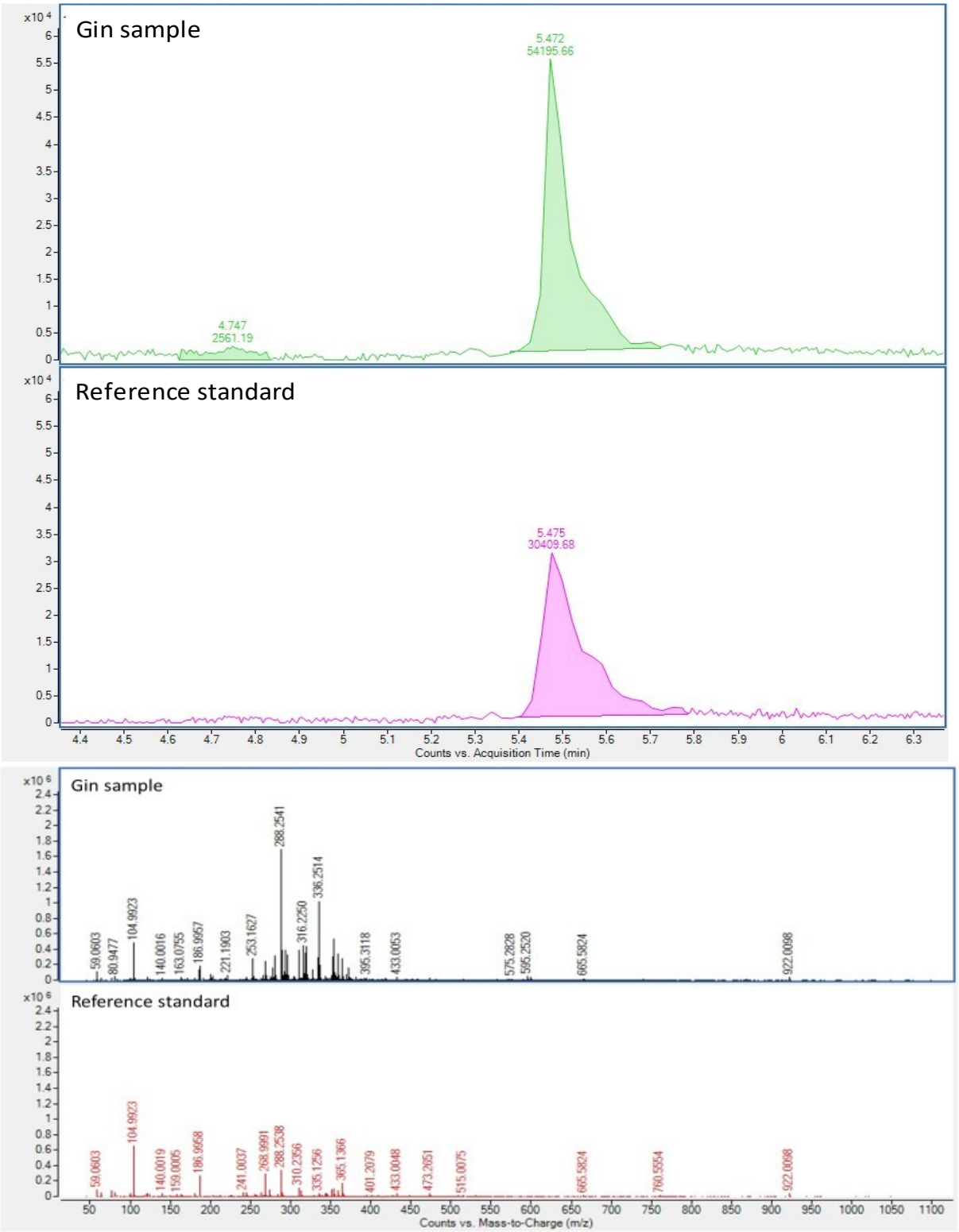


Figure 3.5 Comparisons of MS chromatograms (above) and MSMS spectra (below) for MF (m/z: 163.0759, RT: 5.494 min) and for methyl cinnamate.

Methyl cinnamate, an ester found in cinnamon, strawberries, basil, and vanilla (Lunkenbein, Bellido et al. 2006, Giachino, Sönmez et al. 2014, Mogoşanu, Grumezescu et al. 2017, Karatoprak 2022), shows potential as an authentication marker for gins as it could easily be ionized in ESI+ (Chen, Green et al. 2015, Lee, Kochhar et al. 2015). Although the MSMS spectra were not perfect match, the confidence level of identification is considered to be 2 (Schymanski, Jeon et al. 2014), since the MS, the MSMS, the RT, and the comparison with the reference standard all point to this compound or a potential isomer. The lack of coherence provided by SIRIUS shows that even if advanced software tools are allowing for much faster characterization, the use of reference standards for confirmation of identity remains essential. Concurrently, methyl cinnamate has never been reported in gins using LC-ESI-MS specifically, neither in the context of authentication by comparing them against other spirit categories, such as NS. While methyl cinnamate's promising role as an authentication marker for gins has been identified using LC-ESI-MS, exploring its validation through GC-MS could offer additional insights. Given that methyl cinnamate is a volatile compound, GC-MS, which is particularly adept at analyzing volatile and semi-volatile compounds, might provide enhanced discriminatory capabilities. This approach could complement the current findings, offering a comprehensive understanding of methyl cinnamate's presence in gins. It may also aid in confirming its identity with greater accuracy, validating the role of methyl cinnamate in the authentication of gins and broadening the methodological capabilities for spirit characterization.

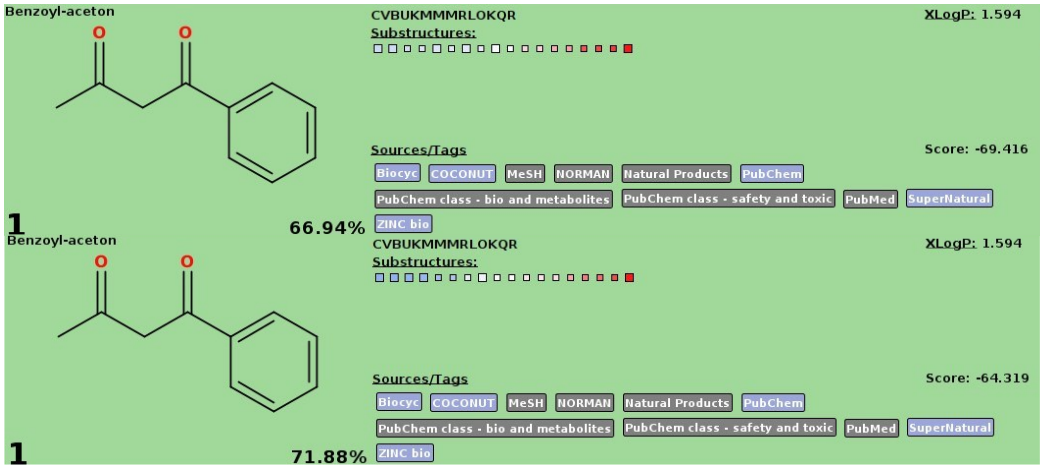


Figure 3.6 Proposed chemical structures for MF (m/z: 163.0759, RT: 5.494 min)(above) and for methyl cinnamate (reference standard)(below) by SIRIUS software.

3.5 Conclusion

This research pioneers an advanced analytical method, NTA RP-LC-QTOF-HRMS, tailored for characterizing the complex chemical fingerprints of NS and gins, marking a significant stride in alcoholic spirits' literature. Through meticulous execution and integration of ESI and APCI ionization techniques, this method refined a vast dataset to reveal distinct molecular entities between NS and gins. The method, also showed a distinct clustering of spirit categories via PCA, therefore presenting a promising tool for authenticating spirits. The end result of the analysis notably identified a strong candidate marker specific to gins, methyl cinnamate, which was never reported before in this context. This addresses crucial concerns in both scientific and beverage industry fields. The findings not only substantiate the distinct chemical compositions of these spirits, but also lay a robust foundation for future studies, expanding the horizon of analytical chemistry in alcoholic beverage characterization.

CONNECTING TEXT

Chapter 3 marked a significant milestone in the understanding of the chemical fingerprints of neutral spirits and gins, employing the NTA RP-LC-QTOF-HRMS method along with MVA. This chapter deepened the scientific comprehension of these spirits as well as pioneered the identification of unique markers, like methyl cinnamate in gins. The results obtained have been instrumental on characterizing the distinct chemical compositions of neutral spirits and gins, setting a robust foundation for further analytical exploration inf the field.

While Chapter 3 focused on characterization, Chapter 4 aims to revolutionize the process of authenticating the botanical and geographical origins of neutral spirits and gins. This novel exploration considered multiple botanical origins such as corn, rye, wheat, and potato, and grouped geographical origins into Quebec, Ontario, and an 'Others' category, encompassing a diverse range of countries. Utilizing the same method, this chapter delves into developing a more focused approached for spirit authentication by detecting thousands of molecular features related to these origins. Although faced with challenges, such as inconclusive PCA results and disparities in ionization modes, the research offers new insights and underscores the complexity of authenticating spirits based on their origin. The chapter also showcased the application of PLS-DA models that achieved high accuracy rates in classifying spirits by origins, demonstrating the method's effectiveness in the novel context of authenticating neutral spirits and gins. It highlights the mixed success of current methods and the pressing need for further refinement. This progression aligns with global efforts to combat food fraud and ensures the integrity and safety of alcoholic beverages. The advancements in Chapter 4 place us at the forefront of technological innovations in the alcoholic beverage industry, signaling a promising future for more robust and precise authentication techniques.

4. DEVELOPMENT OF A NTA RP-LC-QTOF-HRMS METHOD TO AUTHENTICATE THE GEOGRAPHICAL AND BOTANICAL ORIGIN OF QUEBEC NEUTRAL SPIRITS AND GINS

4.1 Abstract

The discrimination of alcoholic beverages based on parameters such as geographical origin, botanical origin, ageing time, and the presence of adulterants has been well documented in the literature. Previous studies have successfully discriminated spirits based on the geographical origin of the raw materials, and analytical strategies have been reviewed to support protected designations of origins for spirits. However, the use of LC-MS for these purposes has not been extensively explored, especially concerning neutral spirits and gins. This study primarily focused on the investigation of MFs for botanical origin authentication using MVA. From the detected MFs, a list of MFOIs was made through a selection process based on peak quality, abundance, and consistency across samples. Following this, the tentative structural and chemical formula elucidation of these MFOIs using SIRIUS software was pursued, aiming to identify new authentication markers. Finally, the classification of the samples in their respective categories using PLS-DA was achieved with overall accuracy above 88% with all categories and ionization modes. By focusing on these objectives, this research sought to employ NTA LC-QTOF-HRMS as a novel approach for the discrimination of neutral spirits and gins based on botanical origins, thereby filling a significant gap in the existing literature. Additionally, authentication based on geographical origin was also explored with preliminary results, as a way to broaden the scope of the study.

4.2 Introduction

The authentication of alcoholic beverages based on botanical and geographical origins is an old topic, but the recent introduction of powerful analytical instruments has allowed to provide crucial insights into combatting alcoholic beverage fraud. However, despite notable advancements in this domain through studies on wine (Tzachristas, Dasenaki et al. 2021, Pan, Gu et al. 2022) and various spirits (Contreras, Barbosa-García et al. 2010, Collins, Zweigenbaum et al. 2014, Roullier-Gall, Signoret et al. 2020), a significant gap persists concerning the authentication of NS and gins.

Given the outlined challenge, a deeper exploration into the chemical compositions of NS and gins is required to uncover novel insights into their authentication. These spirits encompass a multitude of volatile compounds, and their volatile composition has only been explored by a minority of studies, usually with GC-MS (Buck, Goblirsch et al. 2020, Einfalt 2020, Dou, Mäkinen et al. 2023). Despite these few studies, the exploration, characterization, and identification of authentication markers related to origins for these spirits remain largely unexplored. Furthermore, the scant investigations into the volatile composition of NS and gins have only scratched the surface of the complexity inherent to the chemical compositions of NS and gins. Therefore, the current study ambitiously aims to delve into the completely unexplored territory of non-volatile composition with the goal of discovering unique authentication markers related to origin, that can significantly contribute to the literature on spirit authentication. The potential to address this challenge relies on a previously developed NTA LC-QTOF-HRMS method, which has been used to characterize the chemical composition of NS and gins (see Chapter 3).

To address the overarching aim, three main objectives are delineated. Firstly, the aim was to use the NTA RP-LC-QTOF-HRMS to study the chemical fingerprints of NS and gins. Coupling this with MVA to compare samples based on the geographical and botanical origins of their raw materials, MFs specific to certain origins were isolated. Secondly, from these MFs, a rigorous evaluation of peak quality, abundance level, and consistency across samples was conducted to create a list of MFOIs. All of them underwent MSMS to provide detailed mass spectra allowing for the tentative elucidation of their chemical structure and formula with SIRIUS software. This objective served to potentially identify and characterize key compounds as candidate markers of authenticity. Lastly, a PLS-DA classification model was built with the objective of determining the accuracy with which the samples could be classified based solely on their chemical composition. Centred around the hypothesis, this research proposes that through MVA, especially PCA and PLS-DA, it would be possible to discriminate and classify neutral spirits and gins based on the botanical and geographical origins of their raw materials. It was anticipated that statistical tools like volcano plots, and Venn diagrams of unique entities, would enable the isolation of key molecular features that are highly correlated with specific origin categories. The aim was then to characterize these key features and identify them using the SIRIUS machine-learning software. With the foundational success of the NTA LC-QTOF-HRMS method in characterizing the chemical fingerprint of neutral spirits and gins (see Chapter 3), this research is poised to shed light

on the nuances of spirit authentication. This will bridge a significant knowledge gap in the authentication of these spirits and lay the groundwork for an analytical framework in combating alcoholic beverage fraud.

4.3 Materials & Methods

The method, instrument as well as the chemicals for this experiment have been described in Chapter 3. Please refer to it for more details.

4.3.1 Samples

Samples used in this study were consistent with those described in Chapter 3, thus encompassing 21 gins and 40 NS (see supplementary materials Table S1 for complete details). This alignment maintains the rigour of the research and allows for a better understanding of results across different sections of the research. To ensure a robust and comprehensive characterization of botanical and geographical origins in NS and gins, four primary botanical origins were considered: corn (n=12), potato (n=11), rye (n=11), and wheat (n=13). The inclusion of other botanical origins, such as barley and honey, as well as combinations of botanical sources was deliberate to enhance variability and comprehensiveness of the characterization process. While geographical selection was somewhat influenced by availability at the provincial liquor board (SAQ), two key geographical origins, Quebec (n=48) and Ontario (n=3), were highlighted. All other geographical origins, such as Sweden, France, Netherlands, and the United States, were thoughtfully grouped into one category labelled "Others" to maintain focus on the primary regions while acknowledging the broader diversity in spirit origins. The study's overarching principle was to align sample selection within the research objective of obtaining spirits from as diverse a range of botanical and geographical origins as possible, even within the constraints of availability. By setting a criterion that only botanical and geographical origins with three or more samples were considered as individual categories, the research ensured that the results would be statistically significant and representative of each category.

4.3.2 Compound Identification

Please refer to Chapter 3 for more details on the molecular feature extraction process by Profinder, the filtration process by MPP, the use of SIRIUS to elucidate chemical formula and

structure of molecular features. The confidence levels used in this study are also discussed in detail in Chapter 3.

4.3.3 Data Analysis

The same data processing, normalization, and statistical analyses as those described in Chapter 3 were used in the present study with the addition of the supervised statistical PLS-DA classification model to evaluate the accuracy of classification of the samples based on their botanical or geographical characteristics. In this study, PLS-DA models were developed using MPP to classify samples based on botanical and geographical origins, using both ESI+ and ESI- ionization modes. The PLS-DA model was constructed using four components, capturing the essential variance and correlation structure in the data, while reducing dimensionality. This approach allowed for the isolation of the most relevant features in the analysis. Each variable was automatically scaled, normalizing them to have a mean of zero and a standard deviation of one, thus allowing variables of different magnitudes to contribute equally to the model. The dataset was partitioned into three equal-sized folds for cross-validation, ensuring that each segment of the data was used once as a validation set while the remaining two segments served as training sets. This 3-fold cross-validation was repeated ten times, thereby enhancing the robustness of the validation and providing more stable results of the model's performance (Rubingh, Bijlsma et al. 2006). The performance of the model in classifying the samples accurately was evaluated using accuracy, sensitivity, and specificity as metrics. Accuracy is the proportion of correctly classified samples out of the total number of samples. It is calculated as $\frac{(TP + TN)}{(TP+TN+FP+FN)}$, where TP is the number of true positive, TN is the number of true negative, FP is the number of false positive, and FN is the number of false negative. Sensitivity, also known as the true positive rate, measures the proportion of actual positives that are correctly identified as such and is calculated as $\frac{TP}{(TP + FN)}$. Specificity, on the other hand, measures the proportion of actual negatives that are correctly identified as such and is calculated as $\frac{(TN)}{(TN+FP)}$. These metrics provide a comprehensive understanding of the model's performance across different classes and are crucial for evaluating the robustness and reliability of the PLS-DA model in distinguishing between different botanical and geographical origins of the samples (Lee, Liong et al. 2018).

4.4 Results and discussion

4.4.1 Comparison of the different ionization methods for the characterization of botanical origins

Utilizing various ionization modes, this section, building on the methodology from Chapter 3, delves into the chemical fingerprinting of NS and gins to detect and isolate MFs indicative of their botanical origins, showcasing the inherent complexities of these spirits. Due to discrepancies among the number of samples per categories, this study chose to normalize the number of MFs based on the number of samples to account for this variability (as detailed in Chapter 3), ensuring a more accurate and fair comparison of MFs across different botanical origins. This method, although novel, provides a robust basis for comparing MF distribution and abundance, crucial for the characterization and differentiation of spirits based on their origins.

Using the ESI- ionization mode, the analysis detected 11,362 MFs for corn, 7,923 for rye, 7,438 for wheat, and 6,288 for potato in the context of botanical origins (Figure 4.1). Transitioning to the ESI+ ionization mode, corn samples showed 21,438 MFs, rye had 15,396 MFs, wheat samples revealed 14,611 MFs, and potato samples comprised 12,040 MFs (S1.2). The normalization method adopted in this study revealed intriguing patterns post-normalization. The percentages represent the range of the ratio of MFs per samples within a category, indicating the relative abundance and distribution of MFs across different botanical origins. For instance, in ESI- ionization mode, the botanical dataset showed MF differences across categories ranging from -5% to +7% post-normalization, contrasting with the pre-normalized +7% to +81% (see Table 4.1). For ESI+, the differences in MFs ranged from -7% to +5% post-normalization as opposed to the initial +5% to +78%. Delving into the APCI+ ionization mode, corn samples had 4,589 MFs, rye had 3,151 MFs, wheat presented 2,983 MFs, and potato revealed 2,325 MFs. The APCI+ ionization mode presented a trend similar to the ESI modes when normalizing the data. For botanical origins, the difference in MFs shifted from +6% to +97% pre-normalization to -3% to +8% post-normalization. The percentage calculations aim to present a normalized comparison to elucidate the variations in MF distribution, facilitating a more balanced comparison across different categories.

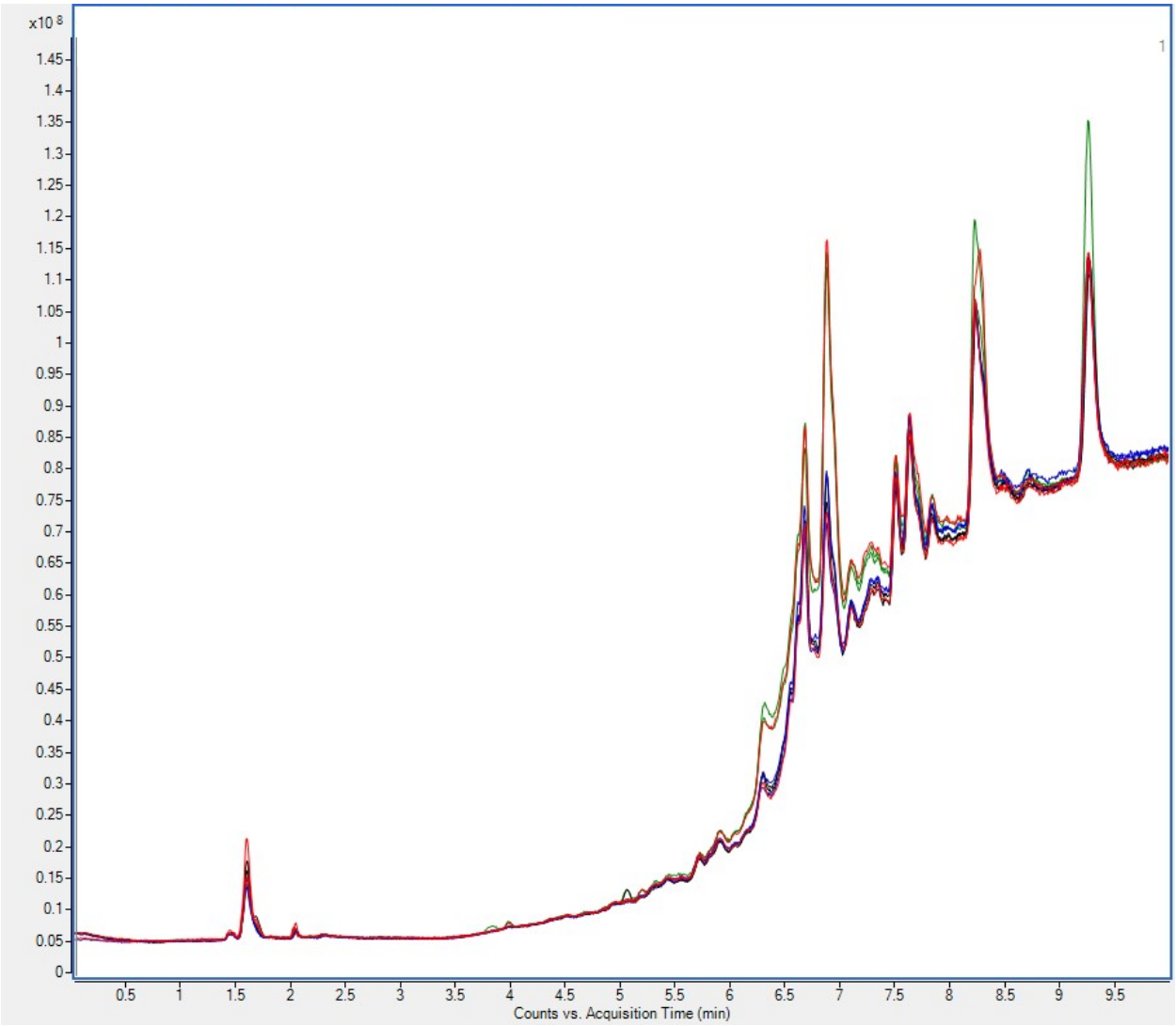


Figure 4.1 Exemplary overlapped TICs of NS and gins for ESI-, colorized based on botanical origins. Represented categories are corn (black), potato (blue), wheat (green), and rye (red). See supplementary materials for ESI+ and APCI+ (S1.2).

This study introduced a novel approach in characterizing the number of MFs associated with botanical origins in NS and gins, a realm not explored in the existing literature. The endeavour to comprehensively catalogue and compare MFs across diverse origins unveils a significant contribution to the understanding of their chemical compositions. This characterization enriches the existing body of knowledge and sets a precedent for subsequent investigations aiming at the authentication of the botanical origins of these spirits. Interestingly, the botanical differences in MFs, when normalized by sample size, remained relatively negligible across both ESI and APCI modes. Yet, a critical observation was the APCI+ mode's reduced capability in detecting MFs when compared to the ESI modes. This echoed the findings from Chapter 3. Given this reduced

detection rate, there are concerns about APCI+'s adequacy in capturing comprehensively the chemical differences for origin differentiation. As a result, and prioritizing the modes with more potential for authentication, APCI+ was excluded from further analyses. Nevertheless, to the best of knowledge, this study represents the first comparative characterization of the number of MFs associated with botanical origins in NS and gins employing ESI+, ESI-, and APCI+ ionization modes. This novel approach, highlighting the varying efficacy of these ionization modes in capturing the diversity of MFs, has not been reported in the literature on the authentication of alcoholic beverages before. These results showcase that the usage of multiple ionization modes in NTA LC-MS has the potential to characterize the chemical fingerprints of these beverages more comprehensively, thus setting a precedent in the field.

In conclusion, the analysis of overlapped TICs has yielded intriguing insights into the characterization of the botanical origins of NS and gins. The findings highlight the disparities in the effectiveness of the various ionization modes in distinguishing the spirits according to their origins' chemical compositions. The normalization procedure provided a more refined perspective by significantly reducing discrepancies in the data, revealing nuances that warrant further exploration. These variations point towards future studies that could delve into the underlying factors contributing to this distinction.

BOTANICAL ORIGIN DATA								
ESI-								
	Before Filtration				After Filtration			
	Corn	Rye	Wheat	Potato	Corn	Rye	Wheat	Potato
Total	11362	7923	7438	6288	7533	6758	6261	6201
Normalized	11362	11885	11157	11790	7533	10137	9392	11627
	Difference Before Normalization				Difference Before Normalization			
		Corn	Wheat			Corn	Wheat	
	Rye	143%	107%		Rye	111%	108%	
	Wheat	153%	126%		Wheat	120%	109%	
	Potato	181%	118%		Potato	121%	101%	
	Difference After Normalization				Difference After Normalization			
		Corn	Wheat			Corn	Wheat	
	Rye	96%	107%		Rye	74%	108%	
	Wheat	102%	101%		Wheat	80%	87%	

	Potato	96%	95%		Potato	65%	81%	

ESI+								
	Before Filtration				After Filtration			
	Corn	Rye	Wheat	Potato	Corn	Rye	Wheat	Potato
Total	21438	15396	14611	12040	13332	12425	11822	11688
Normalized	21438	23094	21917	22575	13332	18638	17733	21915

Difference Before Normalization		
	Corn	Wheat
Rye	139%	105%
Wheat	147%	128%
Potato	178%	121%

Difference After Normalization		
	Corn	Wheat
Rye	93%	105%
Wheat	98%	102%
Potato	95%	97%

Difference Before Normalization		
	Corn	Wheat
Rye	107%	105%
Wheat	113%	106%
Potato	114%	101%

Difference After Normalization		
	Corn	Wheat
Rye	72%	105%
Wheat	75%	85%
Potato	61%	81%

APCI+				
	Corn	Rye	Wheat	Potato
Total	4589	3151	2983	2325
Normalized	4589	4727	4475	4359

Difference Before Normalization		
	Corn	Wheat
Rye	146%	106%
Wheat	154%	136%
Potato	197%	128%

Difference After Normalization		
	Corn	Wheat
Rye	97%	106%
Wheat	103%	108%
Potato	105%	103%

Table 4.1: Number of botanical origin MFs during various steps of the normalization and filtration procedures.

4.4.2 Isolation of unique entities in botanical origins

To comprehensively analyze the unique chemical constituents, present within different botanical origins of NS and gins, this study employed a Venn diagram approach. Figure 4.2 displays the distribution of shared and unique MFs across various categories both before and after

the filtration process (see Chapter 3), showcasing distinctions in botanical origins. This visual representation, never before utilized in conjunction with NS and gins, reveals a realm of origin specificities in terms of chemical composition.

In ESI- ionization mode, the reduction of unique entities in botanical origins ranged from 91.7% to 97.9%, while the total reduction of MFs was 83.6%. With ESI+ ionization mode, the range of reduction in unique entities spanned from 92.3% to 98.8%, with a total reduction of 83.4%. Notably, only about 3% of all MFs in ESI- and 1% in ESI+ were considered unique to specific botanical categories. Conversely, only 21% and 23% of MFs were common to all categories in ESI- and ESI+ modes, respectively. While ESI- mode typically yielded fewer MFs compared to ESI+ mode for both NS and gins, consistent with findings in Chapter 3 and corroborated by literature (Cech and Enke 2001, Kebarle and Verkerk 2010), this distinction becomes particularly pronounced in the context of gins. Given the chemical composition of this spirit, which contains multiple compounds amenable to positive ionization (such as aldehydes, ketones, esters, ethers, terpenoids, etc.) (Riu Aumatell 2012), the superiority of ESI+ in detecting a higher number of MFs aligns well with established chemical principles. This underscores not only the consistency of these findings with existing literature but also emphasizes the relevance of method choice for specific beverage types. The data reflected notable overlaps within certain botanical categories, hinting at common chemical profiles. This commonality necessitates further exploration into authentication markers shared among multiple botanical sources, as it suggests the presence of exclusion markers which are compounds absent in a category but present in all others. More interestingly, the analysis also revealed certain unique MFs within specific categories, which indicate the potential for compounds serving as authentication markers of specific origins. The normalized data, presented in Table 4.1, provided assistance to better compare these unique and shared entities across different categories, enhancing the evaluation of MFs for botanical origins. However, it is important to acknowledge that this relationship is not linear and more complex than it might initially seem. Intriguingly, post-normalization revealed that sample categories originally perceived as richer in MFs were in fact less abundant compared to other categories.

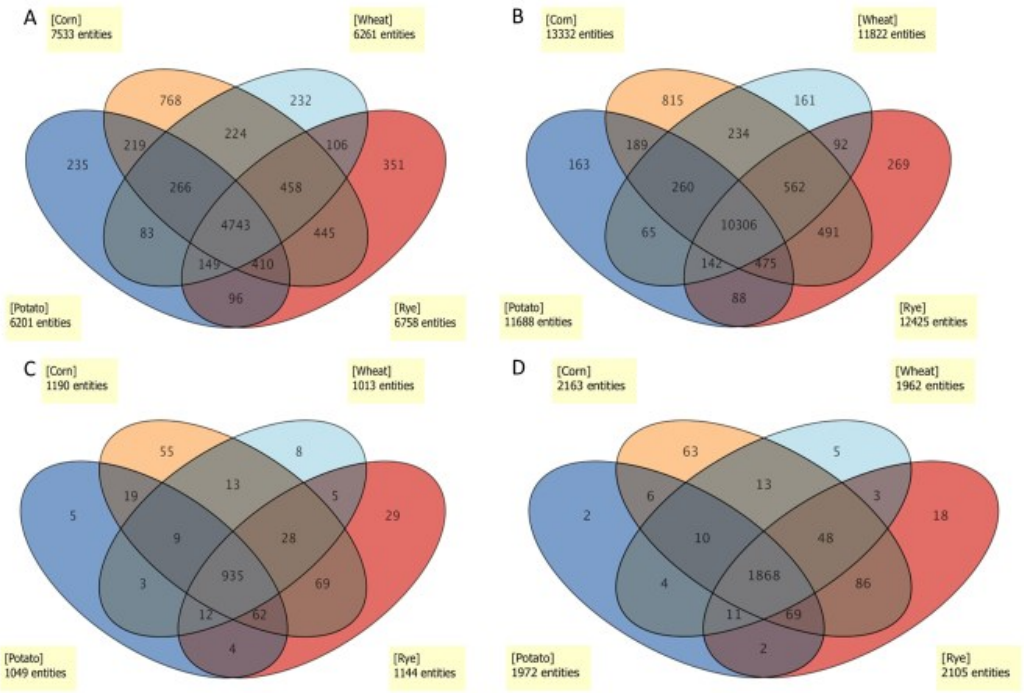


Figure 4.2 Venn diagrams of botanical origins (potato (blue), corn (orange), wheat (light blue), and rye (red)) for ESI- before filtration (A) and after filtration (C) and for ESI+ before filtration (B) and after filtration (D).

The literature on alcoholic beverage authentication has primarily employed methods such as MF filtration (Collins, Zweigenbaum et al. 2014, Uttl, Bechynska et al. 2023), selection of MFs based on loading plots in PCA (Yang, Somogyi et al. 2020), or simply proceeded with MSMS on all MFs after a rigorous filtration process. These conventional approaches have their merits but often emphasize the singular pathway of meticulous filtration followed by analysis using MSMS. This study introduced the Venn diagram approach to augment the understanding of origin-specific chemical entities, a methodology not previously explored in alcoholic beverages studies, especially in the context of NS and gins. The inclusion of this approach aimed to provide a robust tool for isolating unique entities within categories. Utilizing this analysis, a critical evaluation of the data was conducted to highlight the commonalities and uniqueness of different origins. The results suggest that this pipeline might be more stringent than traditionally employed techniques, as evidenced by the unique insights derived from the analyses. This stringency allows for more efficient data treatment by isolating potential unique entities within a category. The use of Venn diagrams to depict the relative proportions of unique and shared entities across different botanicals has not been reported in the literature to the best of knowledge. Thus, it not only provides a novel

method but addresses an uncharted area in the literature, underscoring the Venn diagram's capability to elucidate intricate relationships.

4.4.3 Statistical analyses to evaluate the discriminative nature of the chemical fingerprints of the various botanical origins

PCA is a valuable statistical tool used to discern patterns within data by capturing its underlying variance structure. In the present study, PCAs were conducted to investigate potential clustering relating to botanical origins in both ESI- and ESI+ modes. However, the results did not reveal the anticipated clustering for botanical origins (Figure 4.3). Instead, the analyses displayed considerable overlap of all categories along with notable dispersion. Specifically, the covariance explained was 31.68% in ESI- and 30.59% in ESI+ for botanical origins. The absence of clear clustering within the first three principal components of the PCA didn't imply an absence of differences based on botanical origins. Instead, it suggests that the main sources of variability in this dataset might not be primarily driven by these parameters. This observation highlights that while PCA reveals major variations in a dataset, it might not always detect smaller differences within the dominant principal components. These results highlight the challenge in authentication based on chemical fingerprints alone, especially given the intrinsic complexity of botanical factors. These categories may represent a vast spectrum of chemical compositions, making significant differentiation elusive. Similar challenges in clustering based on origins with PCA were observed in studies on wine (Rubert, Lacina et al. 2014, Ehlers, Uttl et al. 2023) and whisky (Roullier-Gall, Signoret et al. 2020). While PCA didn't clearly discriminate different categories, other statistical tools could capture subtle variations, implying that a multifaceted analytical approach is beneficial in such contexts. Notably, the Venn diagrams, as well as the Volcano plots and the PLS-DA models (discussed below) did reveal significant findings in relation to origins, further underscoring the multifaceted nature of the data, requiring multiple analytical perspectives.

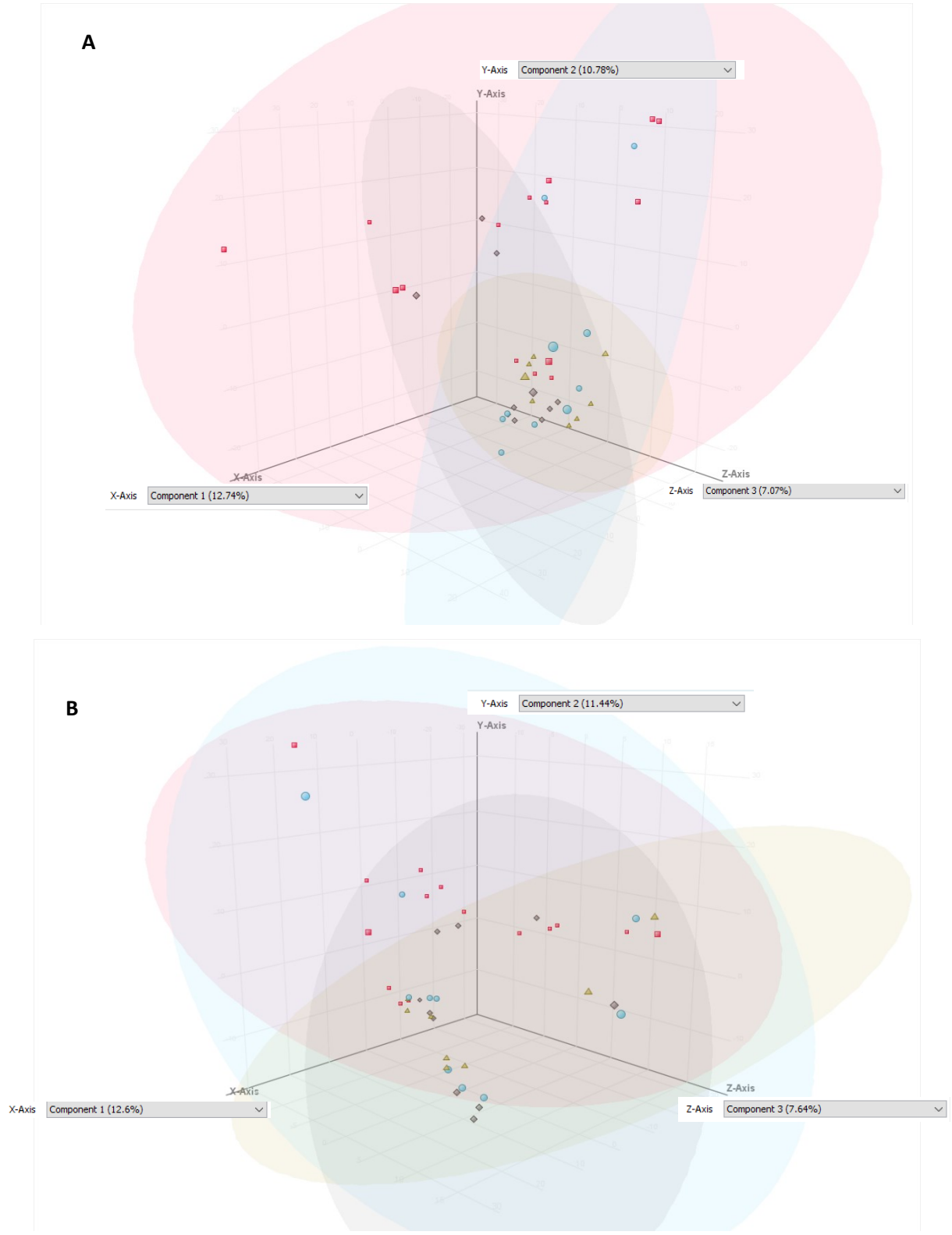


Figure 4.3 PCAs of botanical origins for both ESI ionization modes (ESI- (A), ESI+ (B)). Botanical categories are represented by colours: corn (red), wheat (blue), rye (grey), and potato (dark yellow).

The volcano plots were employed to uncover significant MFs that differentiates the samples based on their botanical origins. These plots allowed for the comparison of specific categories, providing insights into the distinct chemical fingerprints associated with each origin. All combinations of one botanical origin against another were explored in both ESI- and ESI+ modes (1 vs. 1). These pairwise comparisons allow for a more nuanced understanding and identification of significant MFs specific to each category. By doing so, certain distinctions that would be otherwise obscured in broader one vs. rest comparison can be seen more easily detected. Among these, only the corn-potato combination revealed significant differences, while the other combinations failed to exhibit any MFs satisfying the criteria for significance (fold change > 2, *p*-value < 0.05). The ESI+ mode captured more MFs of significant importance in distinguishing corn samples from potato samples. Interestingly, more MFs were present in significantly lower intensity in potato (see Figure 4.4), aligning with the previous Venn diagram results that highlighted more unique entities in corn. The prominence of corn-related MFs accentuated the potential to authenticate spirits based on their primary botanical ingredients. This observation was also made in the Venn diagrams for both ionization modes, where corn-based spirits had a higher abundance of unique entities (Figure 4.2). Many other studies have succeeded in differentiating between various botanical sources in spirits using mostly GC-MS, IRMS, or rarely LC-MS (Collins, Zweigenbaum et al. 2014). However, this distinction between NS and gins made from corn as raw material and those made from potato is a novel contribution to the literature on these categories of spirits and on alcoholic beverage authentication.

This comprehensive analysis delved into the complex differentiation of NS and gins based on their botanical origins using PCA and Volcano plots as statistical analyses. While the PCAs indicated that these origins were not the main sources of variance in the dataset, the Volcano plots helped in pinpointing potential markers, especially when considering corn and potato as botanical origins. If a TA is to be derived from the present study, it is essential that the candidate markers can be detected with high reproducibility. Therefore, MFOIs extracted from the previously described statistical analyses including the unique entities in Venn diagrams underwent subsequent meticulous investigations of peak quality leading to the shortlisting of molecular features of interest (MFOIs) that demonstrated consistency and sufficient abundance across samples (see Table 4.2).

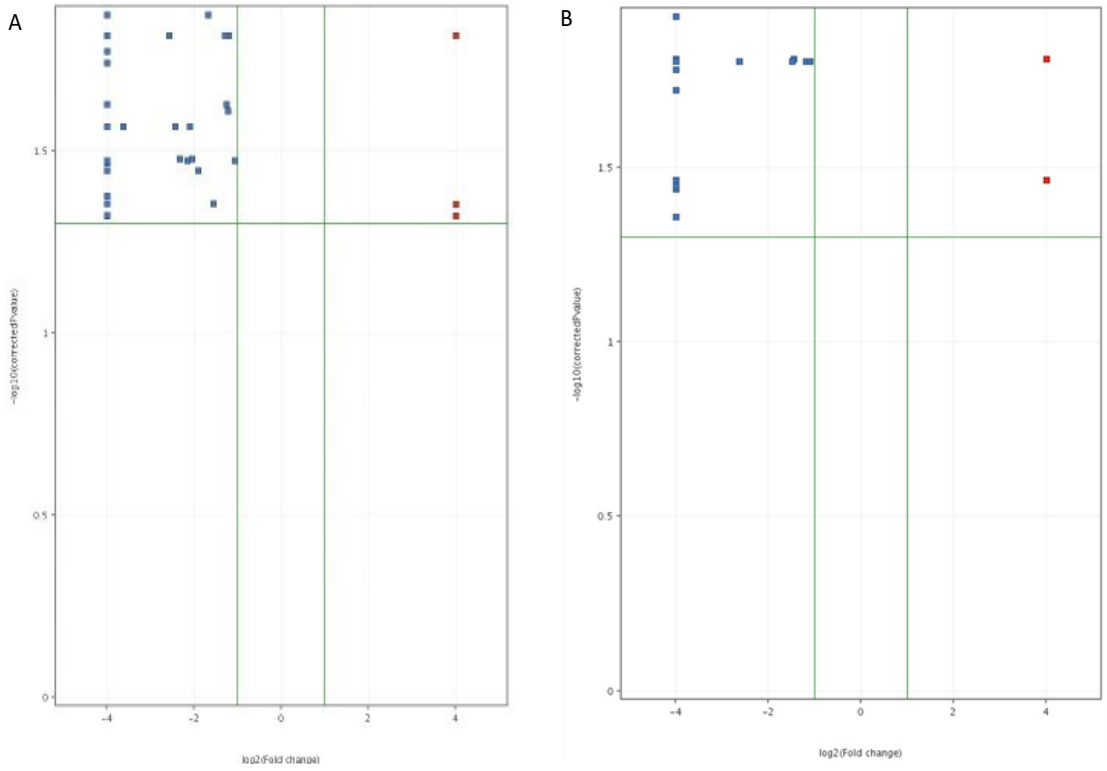


Figure 4.4. Volcano plots of potato vs. corn categories in ESI+ (A) and ESI- (B). Blue squares are MFs present in significant downfold abundance in potato samples, while red squares are MFs significantly more abundant in potato samples.

4.4.4 Identification of candidate markers of authenticity for botanical origins

The primary aim here was to determine if the proposed analytical methods and statistical analyses could generate sufficient information from the dataset to identify candidate markers of authenticity. While several studies have explored the identification of markers for various alcoholic beverage matrices (Collins, Zweigenbaum et al. 2014, Roullier-Gall, Signoret et al. 2018, Ciepielowski, Pacholczyk-Sienicka et al. 2019, Kew, Goodall et al. 2019), the specific focus on NS and gins using the methods described remains underrepresented in the literature. The objective was to identify and characterize key features that correlate strongly with specific botanical origins. The MS^2 spectra of these MFOIs were imported into SIRIUS to aid in the elucidation and characterization of these MFs' chemical formulas and structures. Following the meticulous evaluation of the MFOIs described previously, a selection of 20 MFOIs with good peak quality, high intensity, and strong consistency across representative categories was made for subsequent MS^2 analysis. Extracted ion chromatograms of those MFOIs are represented in Figure 4.5. This approach is consistent with other studies where relevant MFs are evaluated then shortlisted to

improve the focus of the study (Stupak, Goodall et al. 2018, Phan and Tomasino 2021). This selection is presented in Table 4.2 detailing the mass, m/z, retention time, ion species, and labels for all MFOIs related to botanical origins.

Mass	m/z	RT	Formula	Possible ID (%SIRIUS Structure Score)	Label
568.0129	567.007	1.603	C22H13F3N4O5S3	2-[4-nitro-2-[(Z)-[4-oxo-2-sulfanylidene-3-[3-(trifluoromethyl)phenyl]-1,3-thiazolidin-5-ylidene]methyl]phenoxy]-N-(1,3-thiazol-2-yl)acetamide (46.90%)	CORN
536.3824	537.39	7.401	-	-	POTATO
485.2052	486.212	6.910	C18H35N3O10S	[[(2R)-1-(6-aminopurin-9-yl)propan-2-yl]oxymethyl-(2,2-dimethylpropanoyloxy)phosphoryl]oxymethyl 2,2-dimethylpropanoate (48.97%)	POTATO
465.2365	466.244	6.953	-	-	POTATO
448.2933	449.303	7.338	C25H36B2N2O4	1,3,5,6-Tetramethyl-2,7-bis(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)indole-4-carbonitrile (24.58%)	POTATO
358.2724	359.28	6.731	C20H38O5	7-Hydroxyicosanedioic acid (66.01%)	POTATO
803.4094	804.417	7.020	C40H61N5O10S	(2S,4R)-4-[[2-[(1R,3R)-1-acetyloxy-3-[2-hydroxyethoxymethyl-[(2S)-3-methyl-2-[[(2R)-1-methylpiperidine-2-carbonyl]amino]pentanoyl]amino]-4-methylpentyl]-1,3-thiazole-4-carbonyl]amino]-5-(4-hydroxyphenyl)-2-methylpentanoic acid (49.07%)	POTATO
710.4701	711.476	7.106	C38H66N2O10	[(2S,3R,4S,6R)-4-(dimethylamino)-2-[[(3R,5R,6R,7R,8R,9R,12R,13R,14S,19R)-12-ethyl-5,8-dihydroxy-3,5,7,9,13,19-hexamethyl-17-methylidene-10-oxo-13-prop-2-enoxy-11,15-dioxa-1-azabicyclo[12.4.1]nonadecan-6-yl]oxy]-6-methyloxan-3-yl] acetate (61.59%)	POTATO
699.3255	700.333	7.434	C36H42FN9O5	H-His-Arg-Gly-Thr-Thr-Glu-OH (61.97%)	POTATO
201.1728	202.18	3.616	C11H23NO2	4-Hydroxy-1-(2-hydroxyethyl)-2,2,6,6-tetramethylpiperidine (93.17%)	POTATO
431.9973	217.006	1.602	C5H4F4N2OS	-	WHEAT
295.2146	296.223	6.459	C17H29NO3	2-(2-methyl-3-oxocyclopentyl)-N-[(2-propan-2-yloxan-3-yl)methyl]acetamide (61.18%)	WHEAT

550.362	551.371	7.515	C25H46BN7O6	(2S)-5-[[amino(nitramido)methylidene]amino]-2-[[2-(dimethylamino)acetyl]amino]-N-[(1R)-3-methyl-1-[(1S,2S,6R,8S)-2,9,9-trimethyl-3,5-dioxa-4-boratricyclo[6.1.1.02,6]decan-4-yl]butyl]pentanamide (48.56%)	WHEAT
738.5389	739.547	7.668	C37H7ON8O7	Unk-Leu-Ile-Gly-Arg-NH^2 (55.67%)	WHEAT
241.9564	242.962	1.603	C6H5ClF2N2S2	5-chloro-3-(4,4-difluorobut-3-enylsulfanyl)-1,2,4-thiadiazole (41.67%)	WHEAT
307.9269	308.934	1.609	C10H7BrClFO3	Ethyl 2-(5-bromo-2-chloro-3-fluorophenyl)-2-oxoacetate (57.87%)	WHEAT
283.9656	284.973	1.608	C7H8O8S2	Prop-2-ynyl 2,2,4,4-tetraoxo-1,5,2,4-dioxadithiepane-6-carboxylate (53.70%)	WHEAT
201.9624	202.969	1.610	C4H2N4O2S2	[2,2"-Bi-1,3,4-oxadiazole]-5,5"(4H,4"H)-dithione (42.66%)	WHEAT

Table 4.2: List of the 18 selected molecular features of interest considered as potential marker candidates for botanical origin and their tentative identification.

None of the MFs and unique entities found in rye passed the subsequent evaluation based on peak quality, intensity, and sample consistency. This outcome contrast with a study conducted on American whiskeys, where non-volatile compounds were used to discriminate rye whiskeys from other types (Collins, Zweigenbaum et al. 2014). All the MFOIs that passed the quality, intensity, and consistency investigation belonged to the potato and wheat categories. As mentioned before, this is the first study conducted on the authenticity of NS and gins from these botanical origins, which makes these results novel contributions to the field. The analysis aimed to identify unique compounds or those present at higher concentrations in specific origins. While several MFs were deemed potentially significant, the SIRIUS software evaluation did not conclusively identify any compound. However, most chemical formulas were identified with over 99% confidence, suggesting identification confidence level of at least 3. This is the first contribution of the sort for NS and gins, potentially shedding some light on novel compounds never before described in the literature on those spirits.

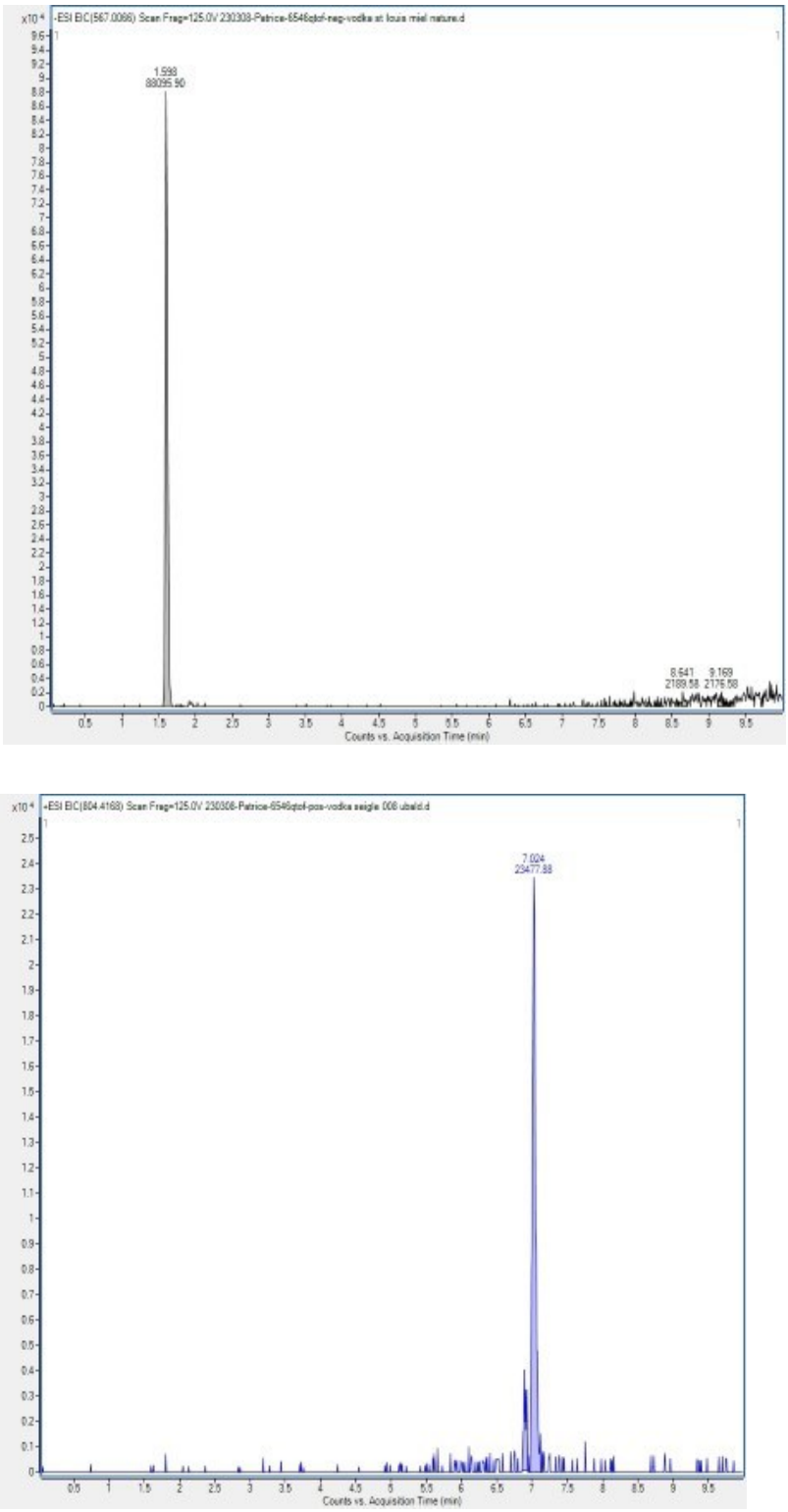


Figure 4.5 Exemplary chromatograms of MFOIs for wheat (m/z 567.007, left) and potato (m/z 804.417, right).

To further these findings, future research could benefit from the integration of two-dimensional LC-MS (2D LC-MS). This advanced technique allows for a more comprehensive separation and analysis of complex mixtures, offering superior resolution and detection capabilities compared to traditional methods. The implementation of 2D LC-MS could lead to a more detailed and accurate identification of specific MFOIs, potentially elevating the authenticity assessment of spirits to a new level of precision (Stoll and Carr 2017).

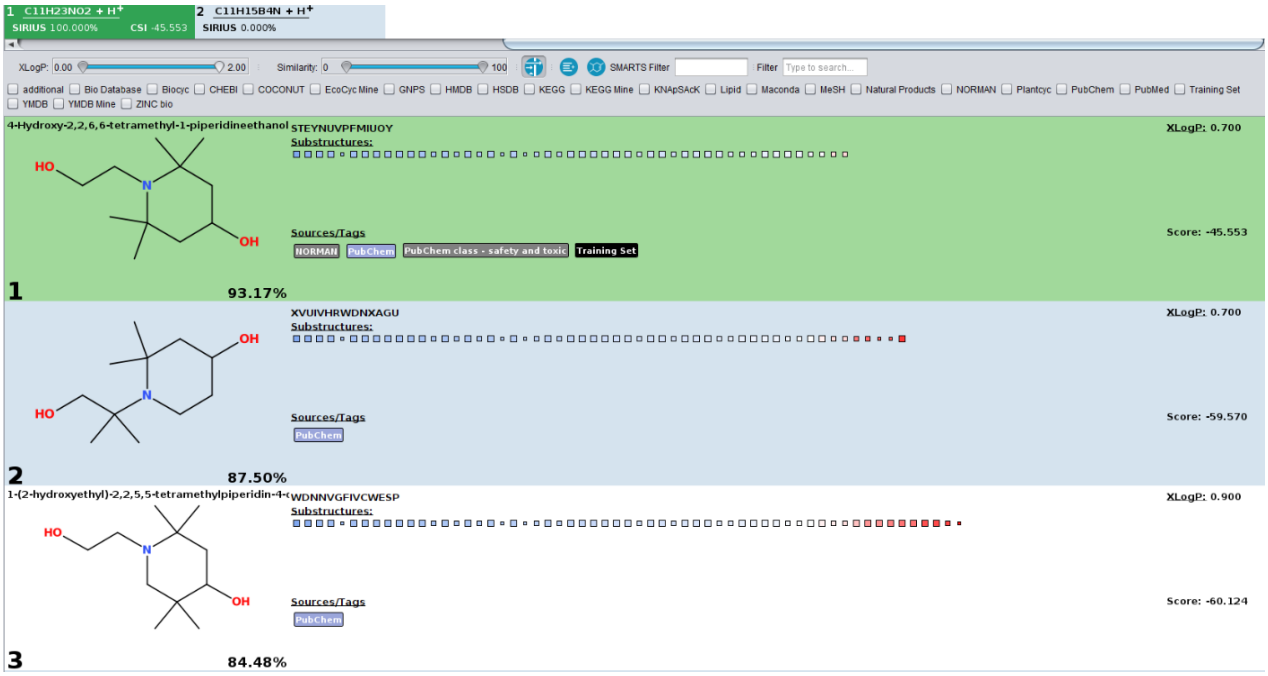


Figure 4.6. SIRIUS results for the identification of MFOI 202.1802 (*m/z*) at retention time 3.62 minutes.

Among the MFOIs, only one compound, 4-hydroxy-2,2,6,6-tetramethyl-1-piperidineethanol (shown in Figure 4.6), had a structural confidence above 90%. However, literature cross-referencing revealed it as a known soil and water contaminant, unrelated to the chemical composition of NS and gins (Hale, Neumann et al. 2022). This unexpected result highlights the complexity of tracing botanical origins and the potential interferences from external sources. While no definitive authentication markers for botanical origins were identified, the methodologies and insights can contribute to future studies. The challenges encountered underscores the novelty and complexity of authenticating origin in NS and gins, where established markers are yet to be found. The presence of botanicals in gins, such as flowers, herbs, spices, and barks, can significantly influence chemical composition and thus marker identification (Riu

Aumatell 2012). The robustness of the analysis, as evidenced by the confidence levels from SIRIUS, emphasizes the need for ongoing research, possibly with refined methods or broader data sets, to explore the potential of chemical fingerprinting in these areas.

4.4.5 PLS-DA classification model

PLS-DA is a powerful statistical technique that combines the properties of both PCA and multiple linear regression. It has been used in this study to classify samples based on botanical origins by leveraging their respective chemical composition. It has become an essential tool in the field of chemometrics to handle complex and high-dimensional data. Both ESI ionization modes yielded identical results for all metrics and classifications, achieving an overall accuracy of 93.02%, with an overall sensitivity of 95%, and an overall specificity of 97.83% (see Table 4.3). These metrics reflect a robust performance of the PLS-DA model in distinguishing between different botanical origins, aligning well with other studies which built classification models on alcoholic beverages, showcasing excellent accuracy, sensitivity, and specificity metrics (Wiśniewska, Boqué et al. 2017, Costa, Llobodanin et al. 2018, Urvieta, Jones et al. 2021). The specificity values, ranging from 94.3% to 100% across different classes, further substantiate the model's ability to correctly identify negative cases, thus minimizing false positive rates. The potato, rye, and wheat samples were correctly classified in 100% of cases, reflecting a sensitivity of 100% as well for these classes. Conversely, the corn samples experienced three misclassifications, with two samples being mistaken for potato and one for wheat, resulting in a sensitivity of 80%. This asymmetry, where corn samples were misclassified as potato or wheat but not the reverse, may point to underlying complexities in the chemical fingerprints shared between these categories that allow nonreciprocal misclassifications. These complexities could stem from similarities in production processes, metabolic analytes, or other uncontrolled variables, making one-way misclassifications quite informative as they can indicate chemical similarities, model bias, feature importance, or other factors that could be explored further.

The PLS-DA model has demonstrated strong potential in classifying samples based on botanical origins, with the accuracy, sensitivity, and specificity metrics reflecting its effectiveness in distinguishing between different botanical origins. However, the results do uncover areas that require careful considerations. First, the misclassification of corn samples sheds light on the

necessity for a more in-depth understanding of the underlying chemical or process-driven commonalities between the corn, potato, and wheat samples.

	Corn (P)	Potato (P)	Rye (P)	Wheat (P)	Sensitivity	Specificity	Accuracy
Corn (T)	**12**	2	0	1	80%	100%	80%
Potato (T)	0	**8**	0	0	100%	94.3%	100%
Rye (T)	0	0	**10**	0	100%	100%	100%
Wheat (T)	0	0	0	**10**	100%	97%	100%
Overall Score					95%	97.83%	93.02%

Table 4.3: Confusion matrix (True (T) against Predicted (P)) of PLS-DA model for the classification of NS and gins based on botanical origins.

This hints at a potential for future investigation that could further refine the classification model by identifying and accounting for these commonalities. The high accuracy in classifying potato, rye, and wheat samples, with a sensitivity and specificity of 100% and above 94.3%, respectively, implies a distinctiveness in the chemical profiles of these botanical origins as captured by the model, allowing for precise classification. This contrast sharply with the lower accuracy in classifying corn, which stands at 80%, suggesting that its chemical profile might share similarities with the other categories, or that the model requires further refinement to accurately capture the unique chemical signature of corn-based NS and gins. The PLS-DA model employed in this study offers robust insights into the classification of botanical origins, offering promising accuracies. The misclassifications of corn samples and require a deeper exploration, indicating that the complexity of the relationships between chemical fingerprints and origins demands larger sample sizes and more nuanced categorizations to allow for a better classification accuracy. This research lays a solid foundation for the ongoing endeavour to authenticate botanical origins in distilled spirits, contributing valuable insights to the field and paving the way for more comprehensive future investigations.

4.4.6 Comparison of the different ionization methods for the characterization of the botanical and geographical origins

The exact same methodologies and statistical analyses were performed to detect and isolate MFs indicative of their geographical origins with the goal of identifying candidate authentication markers. Due to the small number of samples for some categories, these results are presented as preliminary results for the geographical authentication of NS and gins.

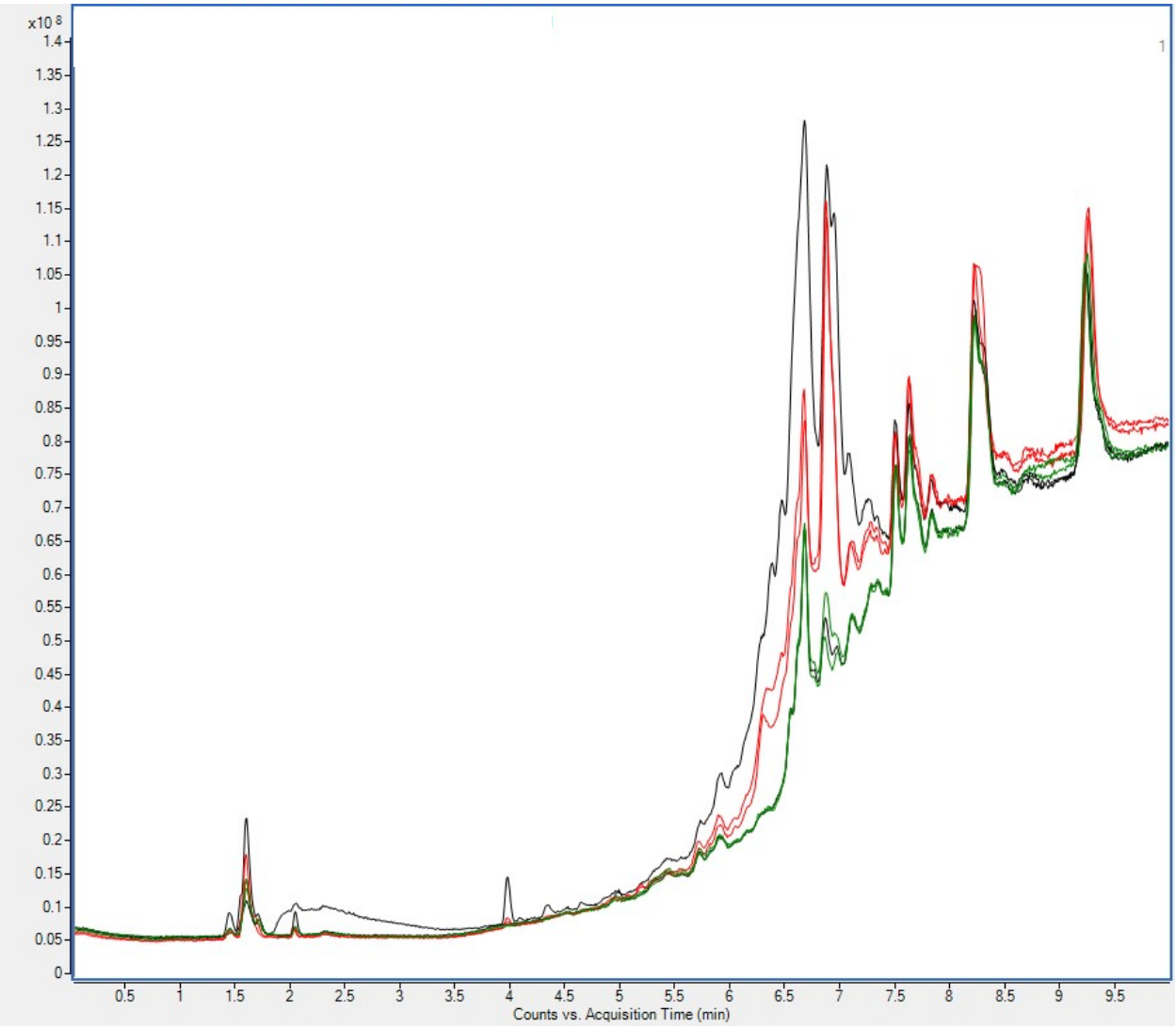


Figure 4.7 Exemplary overlapped TICs of NS and gins for ESI-, colorized based on geographical origins. Represented categories are Quebec (red), Ontario (black), and "Others" (green). See supplementary materials for ESI+ and APCI+ (S1.3).

Using the ESI- ionization mode, Quebec samples revealed 33,444 MFs, Ontario samples had 2,171 MFs, and the category labelled "Others" comprised 3,198 MFs (Figure 4.7). After normalization, the MF difference between categories shifted from -32% to +1,440% to +13% to +34% (Table 4.4). Transitioning to the ESI+ ionization mode, Quebec samples had 61,126 MFs, Ontario samples had 4,291 MFs, and the "Others" category had 6,122 MFs (S1.3). The geographical normalization range changed from -30% to +1,325% pre-normalization to +10% to +28% post-normalization. Delving into the APCI+ ionization mode, Quebec samples showed 15,097 MFs, Ontario had 857 MFs, and the "Others" category consisted of 975 MFs. The APCI+ ionization mode also presented a trend similar to the ESI modes when normalizing the data for

geographical origins. The range changed from an initial -12% to +1,662%, to +36% to +99% after normalization. The same novel approach described with botanical origins was applied to geographical origins, and provided results not reported in the existing literature. This paves the way for subsequent in-depth studies aiming at the authentication of geographical origins for these spirits. Interestingly, where the botanical origins showed negligible differences across ionization modes, the geographical origins presented more pronounced variations. More specifically, Quebec spirits demonstrated a higher number of MFs across both modes. Similarly, with APCI+, the geographical origins exhibited a broader post -normalization range of 136%-199%, continuing to highlight the more distinct variation compared with the findings in ESI modes. The same observation concerning APCI+ mode's reduced capability in detecting MFs when compared to the ESI modes was made leading to the decision to exclude it from further analyses. Nevertheless, this study represented the first comparative characterization of the number of MFs associated with geographical origins in NS and gins employing these ionization modes. The results showcased that the usage of multiple ionization modes in NTA LC-MS has the potential to characterize the chemical fingerprints of these beverages more comprehensively, thus setting a precedent in the field. In conclusion, the analysis of overlapped MF count provided significant insights into the characterization of the geographical origins of NS and gins. These preliminary results highlight the potential of distinguishing these spirits according to their origins' chemical compositions. Interestingly, the close similarities in MFs related to botanical origins contrasted with more pronounced differences in geographical origins. These variations point towards future studies that could delve into the underlying factors contributing to this distinction.

GEOGRAPHICAL ORIGIN DATA						
ESI-						
	Before Filtration			After Filtration		
	QC	ON	Others	QC	ON	Others
Total	33444	2171	3198	8517	4589	5429
Normalized	33444	28223	24944	8517	59657	42346
	Difference Before Normalization			Difference Before Normalization		
		QC	ON		QC	ON
	ON	1540%		ON	186%	
	Others	1046%	68%	Others	157%	85%
	Difference After Normalization			Difference After Normalization		
		QC	ON		QC	ON

	ON	118%		ON	14%	
	Others	134%	113%	Others	20%	141%
ESI+						
	Before Filtration			After Filtration		
	QC	ON	Others	QC	ON	Others
Total	61126	4291	6122	14076	9933	10593
Normalized	61126	55783	47752	14076	129129	82625
	Difference Before Normalization			Difference Before Normalization		
		QC	ON		QC	ON
	ON	1425%		ON	142%	
	Others	998%	70%	Others	133%	94%
	Difference After Normalization			Difference After Normalization		
		QC	ON		QC	ON
	ON	110%		ON	11%	
	Others	128%	117%	Others	17%	156%

APCI+				
		QC	ON	Others
	Total	15097	857	975
	Normalized	15097	11141	7605
		Difference Before Normalization		
			QC	ON
		ON	1762%	
		Others	1548%	88%
		Difference After Normalization		
			QC	ON
		ON	136%	
		Others	199%	146%

Table 4.4: Number of geographical origin MFs during various steps of the normalization and filtration procedures.

4.4.7 Isolation of unique entities in botanical and geographical origins

To isolate unique MF among different geographical origins for NS and gins, the same Venn diagram approach was employed. Figure 4.8 describes the same analysis as performed with botanical origins, but now for geographical origins. The reduction in unique entities ranged from 90.9% to 97.1% with ESI- ionization mode, with a total reduction of MFs of 83.1%. ESI+ showed a reduction range from 92.0% to 95.7%, and the total reduction in MFs was 82.9%. Surprisingly, about 8% of MFs in ESI- and 2.5% in ESI+ were unique, with a large majority attributed to Quebec samples. Only 27% (ESI-) and 29% (ESI+) of MFs were common to all geographical categories, indicating significant overlap between certain categorical combinations. Post-normalization, the

range for ESI- shifted from 85%-186% to 14%-141%, and for ESI+, it changed from 94%-142% to 11%-156% (see Table 4.4). The normalization procedure for geographical origins revealed that the "Others" category showed an abundance of MFs five times higher when normalized. This may be explained due to this category encompassing spirits from diverse regions, reflecting the richness of different botanical, geographical, climatic, and production factors. Another interesting aspect is that this apparent higher number of MFs in "Others" upon normalization doesn't result in a higher number of unique entities for this category mainly because these compounds might only be representative of a particular country, within the "Others" category, making them vulnerable to exclusion during the filtration process. The data also reflected notable overlaps within geographical categories, suggesting commonalities across categories. Once again, this suggests the presence of exclusion markers which are compound absent in a category but present in all others. More interestingly, the analysis revealed unique MFs within specific categories, especially for the Quebec category.

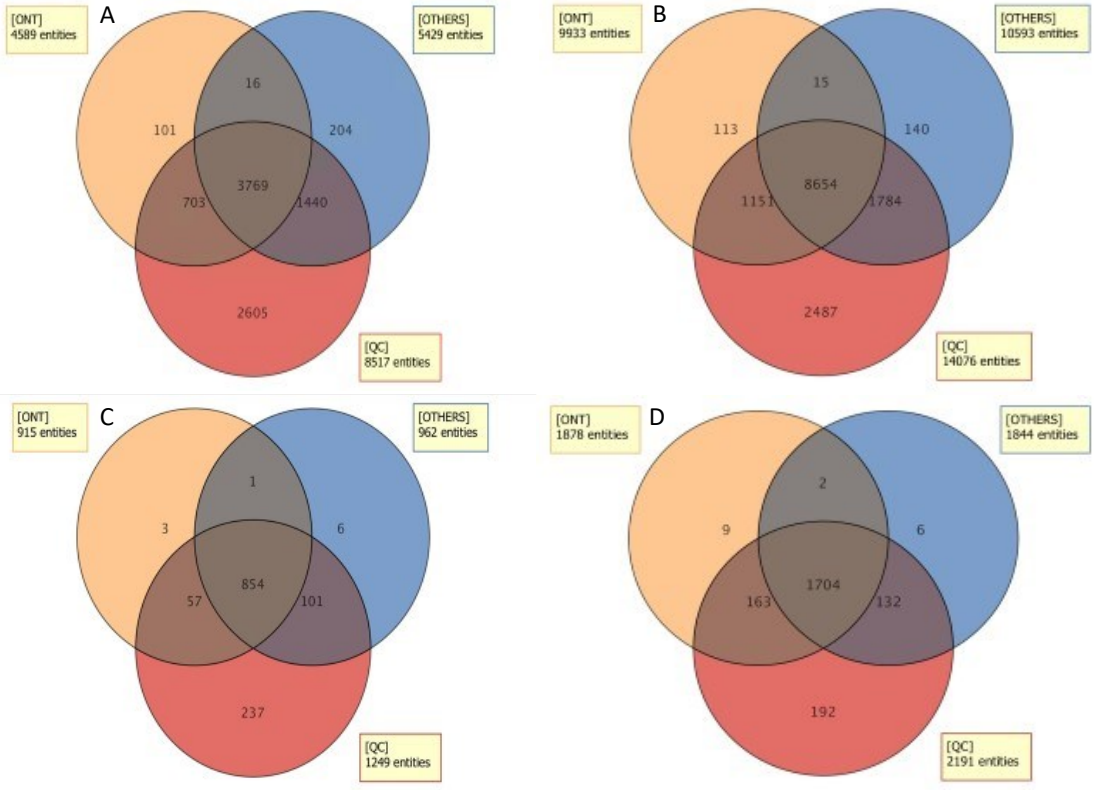


Figure 4.8 Venn diagrams of geographical origins (Others (blue), Ontario (orange), and Quebec (red)) for ESI- (A) and ESI+ before filtration and after filtration (C and D, respectively).

As mentioned with botanical origins, the use of Venn diagrams to isolate unique and shared entities across different geographical origins has not been reported in the literature, thus providing a novel method to elucidate intricate chemical relationships across categories of origins.

4.4.8 Statistical analyses to evaluate the discriminative nature of the chemical fingerprints of the various botanical and geographical origins

PCAs were also conducted with the objective to detect clustering related to geographical origins with both ESI- and ESI+ modes analyses. As with botanical origins, the results did not reveal the anticipated clustering for geographical origins, showing notable dispersion (Figure 4.9). Specifically, the covariance explained was 31.39% in ESI- and 30.42% in ESI+. These results also highlight the challenge in authentication based on origins.

Volcano plots were then also employed, although only with Quebec against Ontario since the "Others" category would not have yielded meaningful insights in a one vs. one context given its diverse sample types. For geographical differentiation, Ontario samples revealed nuanced chemical profiles with the ESI+ dataset. The distribution of MFs in both lower and higher abundance regions suggests a complex chemical composition potentially stemming from various factors intrinsic to their geographical origin (see Figure 4.10). This discrepancy of MFs when comparing Quebec and Ontario samples could serve as a cornerstone for future in-depth authentication studies. Similarly to botanical origins, multiple studies have been able to show differences among spirits of various geographical origins using various instruments and statistical tools (Ciepielowski, Pacholczyk-Sienicka et al. 2019, Roullier-Gall, Signoret et al. 2020). However, such discernment for spirits produced from raw material of Quebec or Ontario origin is a novel insight contributing significantly to the authentication literature on these spirits and these geographical origins.

Despite being only preliminary results, the MFs extracted from the previously described statistical analyses were also evaluated for their potential to be shortlisted for MFOIs, with a specific focus on the Quebec and Ontario apparent compositional differences.

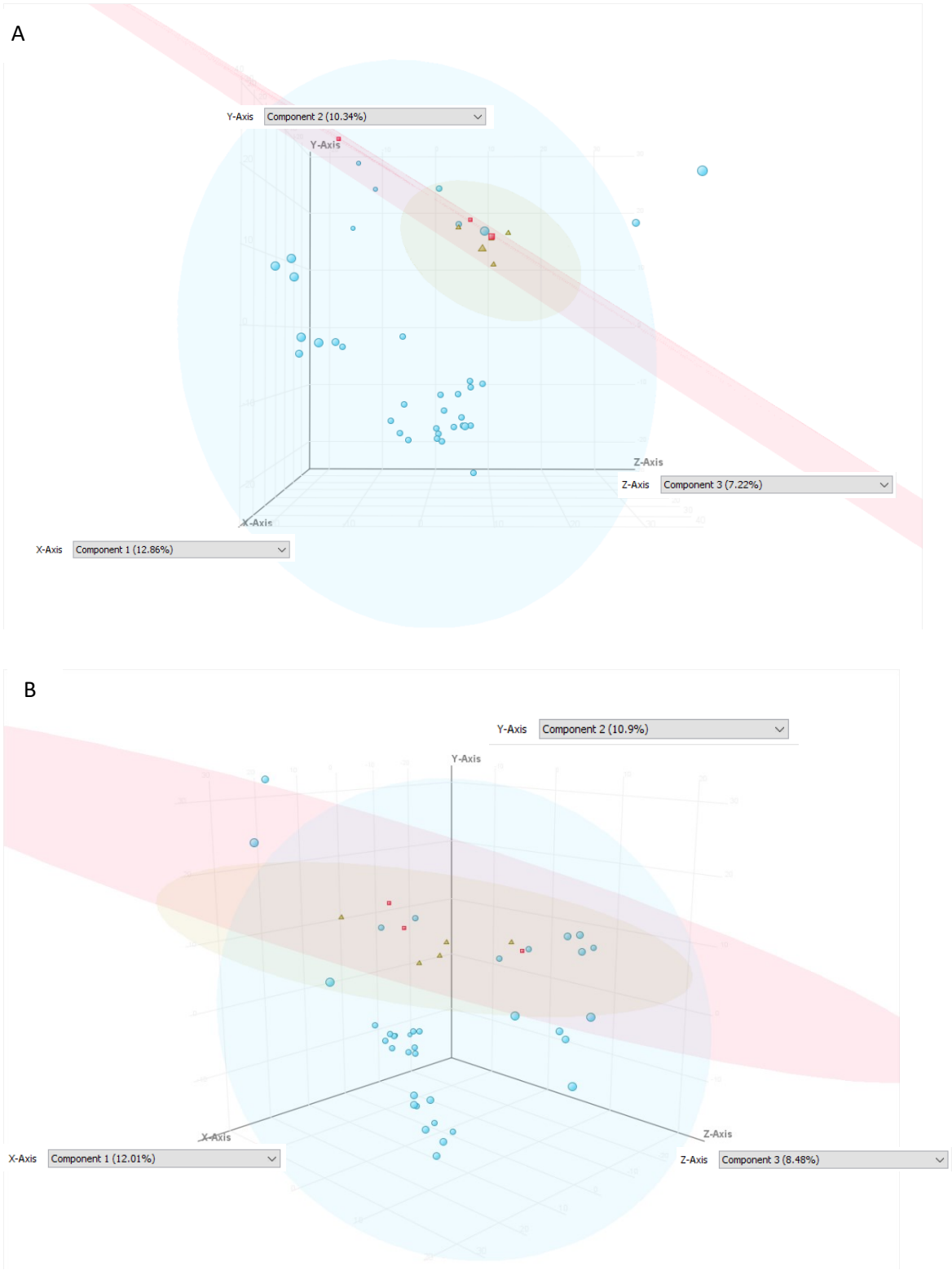


Figure 4.9 PCAs of geographical origins for both ESI ionization modes (ESI- (A), ESI+ (B)). Geographical origins are represented by colours: Quebec (blue), Ontario (red), Others (dark yellow.

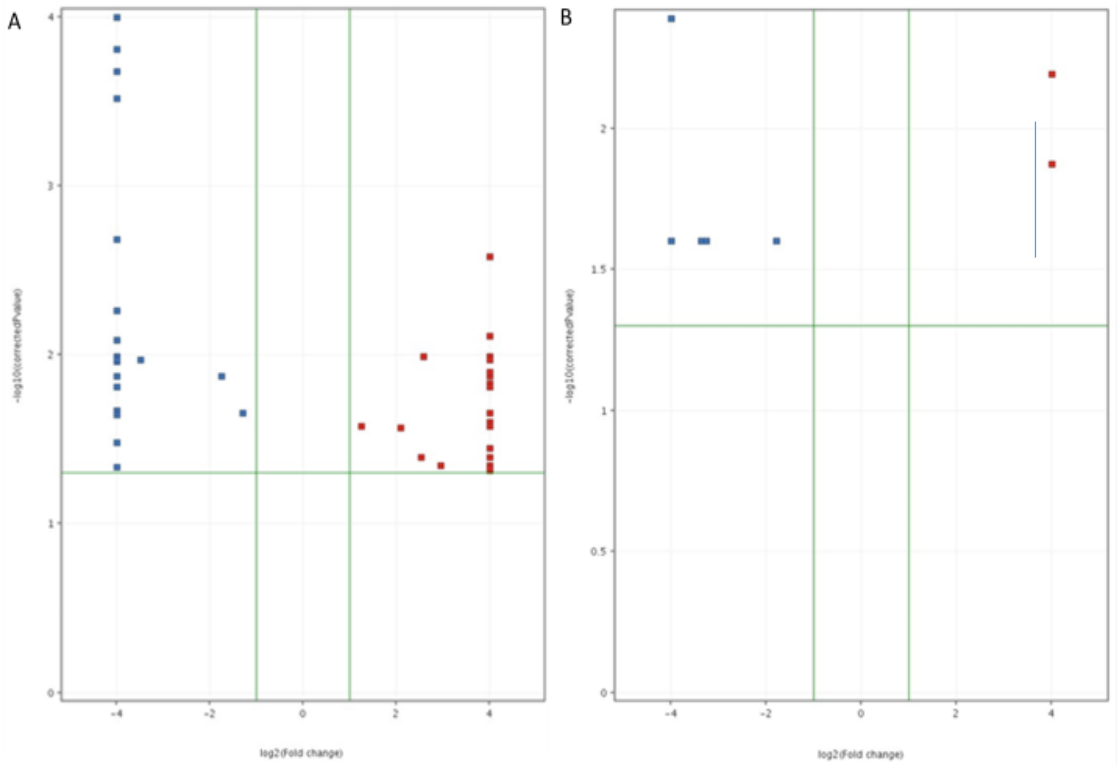


Figure 4.10 Volcano plots of Ontario vs. Quebec categories in ESI+ (A) and ESI- (B). Blue squares are MFs present in significant downfold abundance in Ontario samples, while red squares are MFs significantly more abundant in Ontario samples.

4.4.9 Identification of candidate markers of authenticity for geographical origins

The selection of the shortlisted MFOIs, extracted from previous statistical analyses is presented in Table 4.5 detailing the mass, m/z, retention time, ion species, and labels for the different geographical origins. Both peaks related to these MFOI are presented through their extracted ion chromatograms in Figure 4.11.

Mass	m/z	RT	Formula	Possible ID (%SIRIUS Structure Score)	Label
103.0998	104.107	1.727	C5H13NO	-	QC
99.1047	100.112	4.117	C6H13N	-	QC

Table 4.5: List of the two selected molecular features of interest considered as potential marker candidates for geographical origins and their tentative identification.

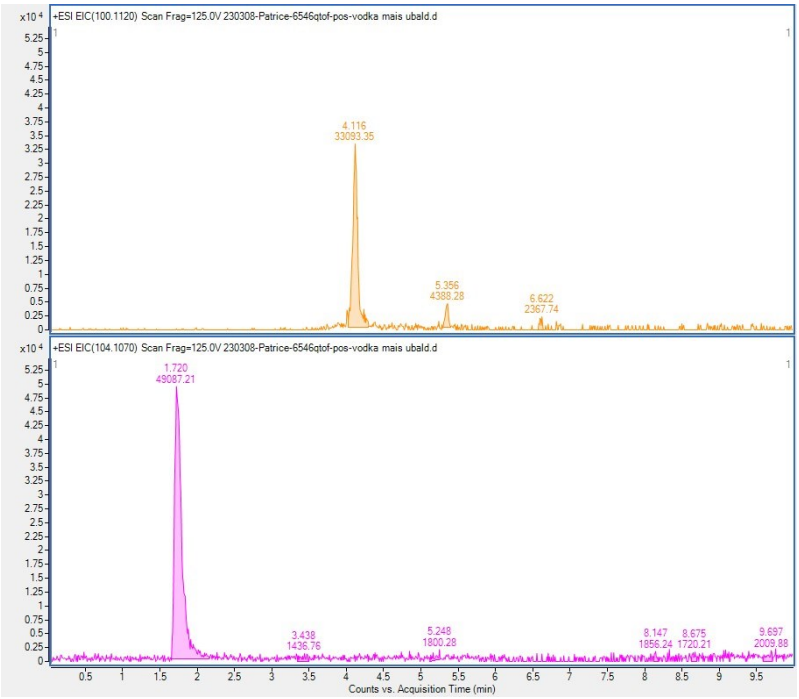


Figure 4.11 Exemplary chromatograms of MFOIs for Quebec (m/z 100.112 (top) and 104.107 (bottom)).

None of the MFOIs from the Ontario or "Others" categories passed the evaluation, and only two Quebec MFOIs passed the evaluation for geographical origin. These results have never been reported in the literature before for any kind of spirits related to these specific origins, as well as for geographical authentication of NS and gins. The subsequent SIRIUS analysis did not conclusively identify the compounds related to Quebec origin. However, the chemical formulas were identified with over 99% confidence, providing an identification confidence level of 3. This is the first contribution of the sort for NS and gins, potentially shedding some light on novel compounds never before detected or described in the literature on those spirits.

4.4.10 PLS-DA classification model

For the geographical origins, the model performed with slightly varying accuracies: 88.89% in ESI- and 91.11% in ESI+, with overall sensitivity scores of 95.5% and 96.4%, and specificity scores of 96.2% and 96.9%, respectively. While both models classified all Ontario and "Others" samples correctly, achieving a sensitivity of 100% for these categories, the ESI+ model misclassified two samples as Ontario and two as "Others". The ESI- model added one more misclassification of a Quebec sample as Ontario (see Table 4.6). These misclassifications were consistent across the two models for most of the samples, except for an additional one in ESI-.

Interestingly, the categories of misclassification were not constant between the two ionization modes. For example, two samples misclassified as "Others" in ESI+ were identified as Ontario in ESI-. Similarly, of the two samples misclassified as Ontario in ESI+, one was also misidentified as Ontario in ESI-, while the other was misclassified as "Others". This variation in misclassification suggests a nuanced relationship between the chemical fingerprints and the geographical origins. The fact that the same samples were misclassified in different ways in the two modes might reflect complexities in the data or differences in how the ionization modes impact these relationships.

ESI-						
	ONT (P)	Others (P)	QC (P)	Sensitivity	Specificity	Accuracy
ONT (T)	3	0	0	100%	93.3%	100.00%
Others (T)	0	5	0	100%	95.2%	100.00%
QC (T)	3	2	32	86.5%	100%	86.49%
Overall Score				95.5%	96.2%	88.89%
ESI+						
	ONT (P)	Others (P)	QC (P)	Sensitivity	Specificity	Accuracy
ONT (T)	3	0	0	100%	95.5%	100.00%
Others (T)	0	5	0	100%	95.2%	100.00%
QC (T)	2	2	33	89.2%	100%	89.19%
Overall Score				96.4%	96.9%	91.11%

Table 4.6 Confusion matrix (True (T) against Predicted (P)) of PLS-DA model for the classification of NS and gins based on geographical origins.

The PLS-DA model has demonstrated strong potential in classifying samples based on both botanical and geographical origins. However, the results do uncover areas that require careful considerations. The disparity in the number of samples between Quebec, Ontario, and the "Others" categories might introduce bias in the statistical models and future work should seek to apply the classification model to larger and more varied sample sets to enhance the evaluation of the model's accuracy. It's important to note that these are preliminary results due to the unbalanced and small sample size. Concurrently, the broader categorization of regions not identified as Quebec or Ontario into the "Others" category may mask significant differences in chemical fingerprints for these regions. These differences could be attributed to various factors such as botanical varieties, seasonal variations, production methods, soil composition, and other factors with potential impact

on chemical composition. Nevertheless, the preliminary PLS-DA model employed in this study offers promising insights into the classification of geographical origins.

4.5 Conclusion

This study, utilizing a UHPLC-QTOF-HRMS method, revealed distinct chemical signatures within different origins of NS and gins, advancing a method for spirit authentication. While Venn diagrams, Volcano plots, and PLS-DA models demonstrated the method's potential, the inconclusive PCA results and other identified challenges highlight the need for further refinement. Disparities in ionization modes' effectiveness provide insights for future analytical method development. The results partially support the hypothesis of grouping spirits based on their origins and identifying key features. Despite certain limitations, this study lays a solid foundation for future research, emphasizing the importance of expanding sample sets and advanced statistical or machine learning techniques for more precise spirit authentication. Overall, the mixed success underscores the method's promise and the complexities in authenticating gins based on origin, pushing the field towards more robust analytical solutions for NS and gin authentication.

5. SUMMARY AND CONCLUSIONS

Alcoholic beverages, with a market value of USD1,369.4 billion in 2020, account for over 10% of the global food and beverage sector (Lu 2020). Their growing demand over the years has not only expanded their share in the food industry but has also created more adulteration opportunities. Historically, alcoholic beverages have been prime targets for adulteration, primarily for economic benefits, ranking them among the top adulterated food commodities (Goodall, Harrison et al. 2018, Europol - OPSON IX 2020). Food authenticity ensures that a product aligns with all the regulations and standards pertaining to its truthful labelling and representation. Contrarily, food fraud, aims to bypass these regulations for monetary gains (GFSI 2018). Alcoholic beverages, have both intrinsic adulterations, which relate to agricultural and processing practices, and extrinsic adulteration, which concern the finished product, such as mislabelling and counterfeiting (GFSI 2018). The challenge intensifies when specific spirits are concerned such as neutral spirits and gins, where there happens to be a scarcity of literature on their authentication and chemical composition. While there have been efforts to detect adulterations in other spirits, the unique characteristics of NS and gins have left a significant gap in research, especially in

identifying reliable authentication markers. The recent popularity surge of gins, combined with the absence of established authentication methods, makes them particularly vulnerable to adulteration. This vulnerability is not just a threat to the economy of the industry but poses significant risks to consumer trust and safety (Ellis, Muhamadali et al. 2019, Manning and Kowalska 2021). In light of these challenges, there's an urgent need to develop robust methods to authenticate NS and gins. Interest in developing protected geographical indication for spirits, e.g., for Quebec, underscores this need, emphasizing the importance of ensuring the authenticity of these spirits in the face of potential fraud.

This research project aimed to address this critical gap. By employing advanced techniques like NTA LC-QTOF-HRMS, the goal was to discern the unique chemical fingerprint of NS and gins, in order to identify candidate markers for the botanical and geographical origins of these spirits. In doing so, this research hoped to fortify the defences against fraudulent activities in the alcoholic beverage sector, ensuring the integrity of NS and gins in the market. The core objective of this research was to delve deep into the chemical complexities of NS and gins, aiming to characterize their unique chemical fingerprints based on their respective spirit categories, and both botanical and geographical origins. To achieve this overarching goal, advanced analytical techniques, specifically the NTA LC-QTOF-HRMS method, were harnessed to elucidate the distinct chemical profiles of these spirits. The objectives were achieved for the characterization of NS and gins when considering them as different spirit categories. However, the objective was only partially achieved for their characterization in the context of botanical or geographical origins. Some unique entities could be extracted from these complex matrices but advanced analyses using MSMS and SIRIUS did not allow for the unequivocal elucidation of any chemical structure, only for chemical formulas. This presence of a grey zone regarding the identification of compounds related to origin means the MFOIs detected are considered tentative candidates (Schymanski, Jeon et al. 2014).

Chapter 3 pioneered the analytical characterization of NS and gins as two distinct spirit categories. Utilizing the NTA LC-QTOF-HRMS method and integrating both ESI and APCI ionization techniques, distinct chemical fingerprints were obtained allowing to compare and contrast them. ESI expectedly performed better with these matrices due to the chemical family present in gins (Riu Aumatell 2012) and APCI's performance was deemed unsuitable for the

chemical characterization. Following rigorous data extraction and filtration using Profinder and Mass Profiler Professional software, distinct molecular profiles for each spirit were identified. From thousands of MFs for each spirit and each ionization mode, only a few were considered significant for further statistical analyses. Using Venn diagrams, these MFs were then classified depending on their uniqueness to a specific spirit category or whether it was common in both. The PCA analyses distinctly showcased the chemical fingerprints of NS and gins, with ESI- revealing a clearer separation between the two spirits than ESI+. Despite some unexpected clustering patterns, the rigorous methods employed, and the consistency of controls validated the approach, emphasizing the unique chemical compositions of NS and gins. The resulting unique molecular features of NS and gins were then tentatively identified using advanced tools like SIRIUS software and reference standards. Perhaps one of the most striking findings was the identification of methyl cinnamate as a potential distinguishing compound in gins. Its consistent presence and the validation against the reference standard highlights its potential as a robust authentication marker. Such a discovery is a testament to both the efficacy of the analytical methods employed and the layered complexity of alcoholic beverages. This identification showcases the challenges intrinsic to authentication studies and paves the way for future research in understanding the chemical fingerprints of various spirits. Despite analyzing various compounds and testing them against reference standards, only one compound could be identified to a level 1 confidence level, highlighting a significant challenge in such authentication studies. This groundbreaking work not only validates the method's effectiveness but also sets the stage for future investigations into the rich chemical profiles of alcoholic beverages.

Chapter 4 delved deeper into the characterization of NS and gins by targeting their specific botanical and geographical origins as areas of authentication. Multiple botanical origins were considered, such as corn, rye, wheat, and potato, while three groups were considered for geographical origins, Quebec, Ontario, and "Others", a group containing multiple countries represented by fewer than 3 samples in the study's sample set. Using the previously developed NTA LC-QTOF-HRMS method, promising insights and challenges were revealed in the thousands of MFs detected, in relation to the botanical and geographical origins of NS and gins. ESI and APCI were once again compared and similar results were observed, where ESI performed significantly better than APCI. Pursuing with the ESI datasets, a filtration and extraction step were performed to remove low quality or redundant MFs. Then from the remaining MFs, certain

analytical tools like Venn diagrams and PCA were used with the same purpose as in Chapter 3. The Venn diagram revealed unique entities belonging to specific origin categories with the goal of isolating these to proceed with further statistical analyses to discover candidate markers. The PCA analyses, while effective at capturing large-scale variability, did not reveal expected clustering based on botanical or geographical origins. This lack of distinction suggests the main variability may not be primarily driven by these parameters. Following this, an additional statistical analysis, Volcano plots, was used to further extract and isolate MFs representative of specific categories. From these, twenty candidate markers were selected on which MSMS analysis was performed. The resulting data was used in conjunction with SIRIUS to elucidate the chemical structure and formula of these MFOIs, with the purpose of identifying compounds with the potential of serving as an authentication marker. PLS-DA, a combination of PCA and multiple linear regression, was used to classify samples based on botanical and geographical origins. Achieving 93.02% accuracy for botanical origins and close to 90% for geographical origins, the models were able to classify samples with high overall accuracy. Despite identifying multiple compounds with their chemical formula, the unequivocal identification of their chemical structure was a significant challenge that could not be overcome with the twenty MFOIs selected. Additionally, despite, the rigorous extraction and filtration methods used, the detection of unique entities with high peak quality, high abundance and consistency across representative categories proved to be an unexpected challenge. A potential solution would be to use a broader sample set to encompass more samples per categories and more categories overall. This proposed solution would lead to even more rigorous filtration procedures and statistical analyses simply due to a higher number of samples. The research aimed to group NS and gins based on their origins and subsequently identify authentication markers for these origins, and while some results aligned with expectations, others underscored the complexity in food authentication in general. This comprehensive analysis not only broadens the understanding of spirit authentication but also emphasizes the need for further refinement and exploration in the field.

This study fills a notable gap in comprehensive NTA for NS and gin chemical fingerprinting. Through detailed metrics like ME, MME, and RSD%, the method was deemed robust, validating data quality and identification accuracy. The analysis of NS and gin's chemical fingerprints using LC-MS, marked a significant comparative analysis of detected MFs in these spirits. Through a refined LC-MS method and multiple ionization modes, a narrower disparity in

analyte counts between NS and gins than expected from the literature was revealed for the first time (Riu Aumatell 2012, Aylott 2016), aiding in understanding their relative chemical compositions. Similarly, a higher MF count was detected compared with other studies on spirits, emphasizing the efficacy of the developed NTA LC-MS method (Collins, Zweigenbaum et al. 2014, Dou, Mäkinen et al. 2023). Venn diagrams highlighted the commonalities and uniqueness among NS and gins, shedding light upon their chemical fingerprints, while also focusing the scope of the study by isolating potential authentication markers, a statistical technique applied for the first time in the characterization of NS and gins. PCA were used on NS and gins using LC-MS for the first time, showcasing a clear distinction in their chemical composition. The most notable finding, however, was the identification of methyl cinnamate as a candidate marker of authentication for gins using LC-ESI-MS. This significantly enriched gin characterization and authentication literature. The continued analysis of NS and gin's chemical fingerprints provided a more in-depth perspective by comparatively analyzing four different botanical origins for the first time as well as three geographical origins in a preliminary way. The number of MFs were reported and compared for the first time in various botanical and geographical origins of NS and gins, when analyzed using LC-MS and various ionization modes, providing a novel framework for working with the authentication of these spirits. Venn diagrams were once more applied for the first time, this time in the context of authentication of origins in NS and gins. The isolation of many unique entities provided significant novel insights into the chemical composition associated with various origins in gins. PCAs were also a innovative application in this research context and although it was showcased that the main variability in the dataset was not associated with origins, this observation was paved the way for a better understanding of the chemical composition of these spirits. Volcano plots were a novel statistical analysis that was employed in origin authentication for NS and gins and gave unprecedented results regarding the notable disparities in chemical fingerprints between corn and potato spirits. From the preliminary results on geographical origins, it was also observed that there might be a difference in the chemical signatures of spirits of Quebec origin compared to Ontario origin. The use of SIRIUS with the resulting MSMS spectra didn't provide an unequivocal identification of these compounds, but provided crucial metrics such as m/z, RT, chemical formulas, and tentative structure, for twenty compounds. Since LC-MS analysis performed on NS and gins was a novel approach, there is a high probability that one of these MFOI represents a previously undiscovered compound, indicating a substantial contribution towards

potential authentication markers. Finally, high accuracy, sensitivity, and specificity provided by a PLS-DA classification model was reported for the first time for botanical and geographical origins of NS and gins. These results indicate a strong potential for authentication based on botanical and geographical origins, reinforcing the discrimination potential established in the present study.

This research explored the complex chemical composition of NS and gins and highlighted the challenge of characterizing their unique chemical fingerprints and identifying authentication markers related to their botanical and geographical origins. Using advanced techniques, notable the NTA LC-QTOF-HRMS, promising molecular features were uncovered. However, certain complexities arose, especially in unequivocally identifying chemical structures from the selected MFOIs. Nevertheless, this study has paved a significant path in spirit authentication. The ability to discern chemical composition and potential markers is promising, even as the goal to classify based on origins remains challenging. Increasing the sample set size could significantly improve the results of the statistical analyses, potentially leading to more accurate results and better characterization and identification of candidate markers of authentication. This difficulty, however, doesn't mean that no compounds exist in these origin categories, as has been proven multiple times with other matrices. Future studies should focus on exploring different areas of the chemical composition of these spirits by making slight alteration to the present method. One strategic area of exploration would involve the modification of the LC-QTOF-HRMS parameters. Adjustments in the MS settings, like tuning the ionization modes, could enhance the detection of certain molecular features. Fine-tuning these parameters may allow for a more comprehensive exploration of these spirits' chemical landscape, especially for compounds that are present in lower concentrations or have unique ionization properties. Similarly, adjustments on the mobile phase type, concentration, or gradient, or experimenting with different column types, could favor the detection of MFs that were not detected with the original settings. Furthermore, advanced characterization of the twenty identified MFOIs should be pursued. This would involve in-depth MS/MS analysis and the application of tools like SIRIUS for proper identification of these chemical compounds. Such detailed analysis could provide vital clues in identifying potential authentication markers and contribute significantly to the field of spirit authentication. Considering the complexities observed, integrating 2D LC-MS into this analysis could refine the identification process, isolating specific MFOIs for clearer analysis and aiding in chemical structure elucidation.

6. REFERENCES

27-CFR-5.22 (2008). Title 27—Alcohol, Tobacco Products and Firearms. Subpart C— Standards of Identity for Distilled Spirits. Section 5.22— The standards of identity. 27. Washington, DC., U.S. Government Printing Office.

Abiodun, O. I., A. Jantan, A. E. Omolara, K. V. Dada, A. M. Umar, O. U. Linus, H. Arshad, A. A. Kazaure, U. Gana and M. U. Kiru (2019). "Comprehensive Review of Artificial Neural Network Applications to Pattern Recognition." IEEE Access **7**: 158820-158846.

Abraben, L. A., K. A. Grogan and Z. Gao (2017). "Organic price premium or penalty? A comparative market analysis of organic wines from Tuscany." Food Policy **69**: 154-165.

Adam, T., E. Duthie and J. Feldmann (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.

Alonso-Salces, R. M., S. Guyot, C. Herrero, L. A. Berrueta, J.-F. Drilleau, B. Gallo and F. Vicente (2005). "Chemometric classification of Basque and French ciders based on their total polyphenol contents and CIELab parameters." Food Chemistry **91**(1): 91-98.

Alonso-Salces, R. M., K. Héberger, M. V. Holland, J. M. Moreno-Rojas, C. Mariani, G. Bellan, F. Reniero and C. Guillou (2010). "Multivariate analysis of NMR fingerprint of the unsaponifiable fraction of virgin olive oils for authentication purposes." Food Chemistry **118**(4): 956-965.

Anca, P. O., D. Adriana BĂ, M. D. Anamaria, I. S. Dorin and F. P. Carmen (2021). "Comparison of some DNA extraction methods from monovarietal must and wines." Notulae Botanicae Horti Agrobotanici Cluj-Napoca **49**(2).

ANSA, R. (2022). Falso Made in Italy, sequestrati liquori per 1,5 mln. ANSA. Valle d'Aosta.

Arslan, M., H. E. Tahir, M. Zareef, J. Shi, A. Rakha, M. Bilal, H. Xiaowei, L. Zhihua and Z. Xiaobo (2021). "Recent trends in quality control, discrimination and authentication of alcoholic beverages using nondestructive instrumental techniques." Trends in Food Science & Technology **107**: 80-113.

Arvanitoyannis, I. S. (2010). 8 - Wine authenticity, traceability and safety monitoring. Managing Wine Quality. A. G. Reynolds, Woodhead Publishing**:** 218-270.

Asenstorfer, R. E., G. P. Jones, G. Laurence and U. Zoppi (2011). Authentication of Red Wine Vintage Using Bomb-Pulse ¹⁴C. Progress in Authentication of Food and Wine, American Chemical Society. **1081:** 89-99.

Associated local press (2020). Mexico: at least 35 people die in mass bad-alcohol poisoning. The Guardian. Mexico City.

Associated Press (2020). Mexico: at least 35 people die in mass bad-alcohol poisoning. The Guardian. Mexico.

Ávila-Fernández, Á., X. Rendón-Poujol, C. Olvera, F. González, S. Capella, A. Peña-Álvarez and A. López-Munguía (2009). "Enzymatic Hydrolysis of Fructans in the Tequila Production Process." Journal of Agricultural and Food Chemistry **57**(12): 5578-5585.

Aylott, R. (2013). Chapter 16 - Analytical Strategies Supporting Protected Designations of Origin for Alcoholic Beverages. Comprehensive Analytical Chemistry. M. de la Guardia and A. Gonzálvez, Elsevier. **60:** 409-438.

Aylott, R. I. (2016). Vodka. Encyclopedia of Food and Health. B. Caballero, P. M. Finglas and F. Toldrá. Oxford, Academic Press**:** 442-445.

Azeroual, O. (2020). "Data Wrangling in Database Systems: Purging of Dirty Data." Data **5**(2): 50.

Baldwin, E. A., J. Bai, A. Plotto and S. Dea (2011). "Electronic Noses and Tongues: Applications for the Food and Pharmaceutical Industries." Sensors **11**(5): 4744-4766.
Ballabio, D. and R. Todeschini (2009). Chapter 4 - Multivariate Classification for Qualitative Analysis. Infrared Spectroscopy for Food Quality Analysis and Control. D.-W. Sun. San Diego, Academic Press**:** 83-104.
Balmforth, T. (2022). Death toll from Russian alcohol poisoning incident rises to 34. Reuters.
Banerjee, S. and S. Mazumdar (2012). "Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte." International Journal of Analytical Chemistry **2012**: 282574.
Barjaktarović, B., M. Sovilj and Ž. Knez (2005). "Chemical Composition of Juniperus communis L. Fruits Supercritical CO2 Extracts: Dependence on Pressure and Extraction Time." Journal of Agricultural and Food Chemistry **53**(7): 2630-2636.
Barker, M. and W. Rayens (2003). "Partial least squares for discrimination." Journal of Chemometrics **17**(3): 166-173.
Basalekou, M., M. Kyraleou and S. Kallithraka (2022). Chapter 38 - Authentication of wine and other alcohol-based beverages—Future global scenario. Future Foods. R. Bhat, Academic Press**:** 669-695.
Beger, R. D., W. B. Dunn, A. Bandukwala, B. Bethan, D. Broadhurst, C. B. Clish, S. Dasari, L. Derr, A. Evans, S. Fischer, T. Flynn, T. Hartung, D. Herrington, R. Higashi, P.-C. Hsu, C. Jones, M. Kachman, H. Karuso, G. Kruppa, K. Lippa, P. Maruvada, J. Mosley, I. Ntai, C. O'Donovan, M. Playdon, D. Raftery, D. Shaughnessy, A. Souza, T. Spaeder, B. Spalholz, F. Tayyari, B. Ubhi, M. Verma, T. Walk, I. Wilson, K. Witkin, D. W. Bearden and K. A. Zanetti (2019). "Towards quality assurance and quality control in untargeted metabolomics studies." Metabolomics **15**(1): 4.
Bellido-Milla, D., J. M. Moreno-Perez and M. a. P. Hernández-Artiga (2000). "Differentiation and classification of beers with flame atomic spectrometry and molecular absorption spectrometry and sample preparation assisted by microwaves." Spectrochimica Acta Part B: Atomic Spectroscopy **55**(7): 855-864.
Belmonte-Sánchez, J. R., R. Romero-González, J. L. Martínez Vidal, F. J. Arrebola and A. Garrido Frenich (2020). "1H NMR and multi-technique data fusion as metabolomic tool for the classification of golden rums by multivariate statistical analysis." Food Chemistry **317**: 126363.
Bennett, M. (2017). World's most expensive dram of Scotch was a fake. BBC News. Scotland.
Boccacci, P., W. Chitarra, A. Schneider, L. Rolle and G. Gambino (2020). "Single-nucleotide polymorphism (SNP) genotyping assays for the varietal authentication of 'Nebbiolo' musts and wines." Food Chemistry **312**: 126100.
Böcker, S. and K. Dührkop (2016). "Fragmentation trees reloaded." Journal of Cheminformatics **8**(1): 5.
Böhme, K., P. Calo-Mata, J. Barros-Velázquez and I. Ortea (2019). "Review of Recent DNA-Based Methods for Main Food-Authentication Topics." Journal of Agricultural and Food Chemistry **67**(14): 3854-3864.
Brecht, D., F. Uteschil and O. J. Schmitz (2020). "Development of a fast-switching dual (ESI/APCI) ionization source for liquid chromatography/mass spectrometry." Rapid Communications in Mass Spectrometry **34**(17): e8845.
Breiman, L. (1996). "Bagging Predictors." Machine Learning **24**(2): 123-140.
Breiman, L. (2001). "Random Forests." Machine Learning **45**(1): 5-32.

Broadhurst, D., R. Goodacre, S. N. Reinke, J. Kuligowski, I. D. Wilson, M. R. Lewis and W. B. Dunn (2018). "Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies." Metabolomics **14**(6): 72.
Bronzi, B., C. Brilli, G. M. Beone, M. C. Fontanella, D. Ballabio, R. Todeschini, V. Consonni, F. Grisoni, F. Parri and M. Buscema (2020). "Geographical identification of Chianti red wine based on ICP-MS element composition." Food Chemistry **315**: 126248.
Buck, N., T. Goblirsch, J. Beauchamp and E. Ortner (2020). "Key Aroma Compounds in Two Bavarian Gins." Applied Sciences **10**(20).
Bujak, R., W. Struck-Lewicka, M. J. Markuszewski and R. Kaliszan (2015). "Metabolomics for laboratory diagnostics." Journal of Pharmaceutical and Biomedical Analysis **113**: 108-120.
Bureau National Interprofessionnel du Cognac (BNIC) (2018). CAHIER DES CHARGES DE L'APPELLATION D'ORIGINE CONTRÔLÉE

« Cognac » ou « Eau-de-vie de Cognac » ou « Eau-de-vie des Charentes ». Partie I Fiche technique, D. - Description de la méthode d'obtention. Décret N°2015-10. O. d. D. e. d. G. (ODG). France. **Partie I - D.9**.
Burns, R. L., R. Alexander, L. Snaychuk, J. C. Edwards, N. Fitzgerald, P. Gao, D. Quan, C. Douvris, T. Vaughan and D. D. Bussan (2021). "A Fast, Straightforward and Inexpensive Method for the Authentication of Baijiu Spirit Samples by Fluorescence Spectroscopy." Beverages **7**(3): 65.
Cajka, T. and O. Fiehn (2016). "Toward Merging Untargeted and Targeted Methods in Mass Spectrometry-Based Metabolomics and Lipidomics." Analytical Chemistry **88**(1): 524-545.
Cao, W., N. Shu, J. Wen, Y. Yang, Y. Wang and W. Lu (2023) "Widely Targeted Metabolomics Was Used to Reveal the Differences between Non-Volatile Compounds in Different Wines and Their Associations with Sensory Properties." Foods **12** DOI: 10.3390/foods12020290.
Capron, X., J. Smeyers-Verbeke and D. L. Massart (2007). "Multivariate determination of the geographical origin of wines from four different countries." Food Chemistry **101**(4): 1585-1597.
Cardoso, D. R., L. G. Andrade-Sobrinho, A. F. Leite-Neto, R. V. Reche, W. D. Isique, M. M. C. Ferreira, B. S. Lima-Neto and D. W. Franco (2004). "Comparison between Cachaça and Rum Using Pattern Recognition Methods." Journal of Agricultural and Food Chemistry **52**(11): 3429-3433.
Cassago, A. L. L., M. M. Artêncio, J. De Moura Engracia Giraldi and F. B. Da Costa (2021). "Metabolomics as a marketing tool for geographical indication products: a literature review." European Food Research and Technology **247**(9): 2143-2159.
Cavanna, D., L. Righetti, C. Elliott and M. Suman (2018). "The scientific challenges in moving from targeted to non-targeted mass spectrometric methods for food fraud analysis: A proposed validation workflow to bring about a harmonized approach." Trends in Food Science & Technology **80**: 223-241.
Cavicchi, A. and C. Santini (2011). "Brunellopoli: a wine scandal under the tuscan sun." Tourism Review International **15**(3): 253-267.
Ceballos-Magaña, S. G., J. M. Jurado, M. a. J. s. Martín and F. Pablos (2009). "Quantitation of Twelve Metals in Tequila and Mezcal Spirits as Authenticity Parameters." Journal of Agricultural and Food Chemistry **57**(4): 1372-1376.
Cech, N. B. and C. G. Enke (2001). "Practical implications of some recent studies in electrospray ionization fundamentals." Mass Spectrometry Reviews **20**(6): 362-387.

Cellier, R., S. Berail, J. Barre, E. Epova, F. Claverie, A.-L. Ronzani, S. Milcent, P. Ors and O. F. X. Donard (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**: 122433.
Česlová, L., K. Pravcová, M. Juričová and J. Fischer (2022). "Rapid HPLC/MS/MS analysis of phenolic content and profile for mead quality assessment." Food Control **134**: 108737.
Cetó, X., M. Llobet, J. Marco and M. d. Valle (2013). "Application of an electronic tongue towards the analysis of brandies." Analytical Methods **5**(5): 1120-1129.
Cevallos-Cevallos, J. M., J. I. Reyes-De-Corcuera, E. Etxeberria, M. D. Danyluk and G. E. Rodrick (2009). "Metabolomic analysis in food science: a review." Trends in Food Science & Technology **20**(11): 557-566.
Chao, E. and D. Krewski (2008). "A risk-based classification scheme for genetically modified foods II: Graded testing." Regulatory Toxicology and Pharmacology **52**(3): 223-234.
Charlebois, S., M. Juhasz, L. Foti and S. Chamberlain (2017). "Food Fraud and Risk Perception: Awareness in Canada and Projected Trust on Risk-Mitigating Agents." Journal of International Food & Agribusiness Marketing **29**(3): 260-277.
Charlebois, S., A. Schwab, R. Henn and C. W. Huck (2016). "Food fraud: An exploratory study for measuring consumer perception towards mislabeled food products and influence on self-authentication intentions." Trends in Food Science & Technology **50**: 211-218.
Chaudhary, V., P. Kajla, A. Dewan, R. Pandiselvam, C. T. Socol and C. M. Maerescu (2022). "Spectroscopic techniques for authentication of animal origin foods." Frontiers in Nutrition **9**.
Chávez-Márquez, A., A. A. Gardea, H. González-Rios and L. Vazquez-Moreno (2022). "Characterization of Cabernet Sauvignon Wines by Untargeted HS-SPME GC-QTOF-MS." Molecules **27**(5): 1726.
Chen, J., K. B. Green and K. K. Nichols (2015). "Characterization of Wax Esters by Electrospray Ionization Tandem Mass Spectrometry: Double Bond Effect and Unusual Product Ions." Lipids **50**(8): 821-836.
Chen, L., F. Zhong and J. Zhu (2020). "Bridging Targeted and Untargeted Mass Spectrometry-Based Metabolomics via Hybrid Approaches." Metabolites **10**(9): 348.
Ciepielowski, G., B. Pacholczyk-Sienicka, T. Frączek, K. Klajman, P. Paneth and Ł. Albrecht (2019). "Comparison of quantitative NMR and IRMS for the authentication of 'Polish Vodka'." Journal of the Science of Food and Agriculture **99**(1): 263-268.
Ciosek, P. and W. Wróblewski (2007). "Sensor arrays for liquid sensing – electronic tongue systems." Analyst **132**(10): 963-978.
Cisneros, E. R. (2001). The Protection of Geographical Indications in Mexico. Symposium in the International Protection of Geographical Indication. Montevideo, Uruguay, World Intellectual Property Organization.
Collins, T. S., J. Zweigenbaum and S. E. Ebeler (2014). "Profiling of nonvolatiles in whiskeys using ultra high pressure liquid chromatography quadrupole time-of-flight mass spectrometry (UHPLC–QTOF MS)." Food Chemistry **163**: 186-196.
Comité Consultivo Nacional de Normalización de Seguridad al Usuario, I. C. y. P. d. C. C. (2012). NORMA OFICIAL MEXICANA NOM-006-SCFI-2012, BEBIDAS ALCOHOLICAS-TEQUILA-

ESPECIFICACIONES. NOM-006-SCFI-2012. CCNNSUICPC. Mexico. **5**.
Commisso, M., A. Anesi, S. Dal Santo and F. Guzzo (2017). "Performance comparison of electrospray ionization and atmospheric pressure chemical ionization in untargeted and targeted

liquid chromatography/mass spectrometry based metabolomics analysis of grapeberry metabolites." Rapid Communications in Mass Spectrometry **31**(3): 292-300.
Contreras, U., O. Barbosa-García, J. L. Pichardo-Molina, G. Ramos-Ortíz, J. L. Maldonado, M. A. Meneses-Nava, N. E. Ornelas-Soto and P. L. López-de-Alba (2010). "Screening method for identification of adulterate and fake tequilas by using UV–VIS spectroscopy and chemometrics." Food Research International **43**(10): 2356-2362.
Cook, G. T., E. Dunbar, B. G. Tripney and D. Fabel (2020). "Using Carbon Isotopes to Fight the Rise in Fraudulent Whisky." Radiocarbon **62**(1): 51-62.
Cosme, F., J. Milheiro, J. Pires, F. I. Guerra-Gomes, L. Filipe-Ribeiro and F. M. Nunes (2021). "Authentication of Douro DO monovarietal red wines based on anthocyanin profile: Comparison of partial least squares – discriminant analysis, decision trees and artificial neural networks." Food Control **125**: 107979.
Costa, N. L., L. A. G. Llobodanin, I. A. Castro and R. Barbosa (2018). "Geographical Classification of Tannat Wines Based on Support Vector Machines and Feature Selection." Beverages **4**(4): 97.
Cozzolino, D., M. Holdstock, R. G. Dambergs, W. U. Cynkar and P. A. Smith (2009). "Mid infrared spectroscopy and multivariate analysis: A tool to discriminate between organic and non-organic wines grown in Australia." Food Chemistry **116**(3): 761-765.
Cremers, D. A. R., L.J. (2013). Introduction. Handbook of Laser‐Induced Breakdown Spectroscopy**:** 1-27.
Cubero-Leon, E., R. Peñalver and A. Maquet (2014). "Review on metabolomics for food authentication." Food Research International **60**: 95-107.
Cusa, M., K. St John Glew, C. Trueman, S. Mariani, L. Buckley, F. Neat and C. Longo (2022). "A future for seafood point-of-origin testing using DNA and stable isotope signatures." Reviews in Fish Biology and Fisheries **32**(2): 597-621.
David W. Koppenaal, C. J. B., M. Bonner Denton, Roger P. Sperline, Gary M. Hieftje, Gregory D. Schilling, Francisco J. Andrade, James H. Barnes (2005). "MS Detectors." Analytical Chemistry **77**(21): 418 A-427 A.
de Aquino, F. W. B., S. Rodrigues, R. F. do Nascimento and A. R. S. Casimiro (2006). "Simultaneous determination of aging markers in sugar cane spirits." Food Chemistry **98**(3): 569-574.
De León-Rodríguez, A., L. González-Hernández, A. P. Barba de la Rosa, P. Escalante-Minakata and M. G. López (2006). "Characterization of Volatile Compounds of Mezcal, an Ethnic Alcoholic Beverage Obtained from Agave salmiana." Journal of Agricultural and Food Chemistry **54**(4): 1337-1341.
Dou, Y., M. Mäkinen and J. Jänis (2023). "Analysis of Volatile and Nonvolatile Constituents in Gin by Direct-Infusion Ultrahigh-Resolution ESI/APPI FT-ICR Mass Spectrometry." Journal of Agricultural and Food Chemistry **71**(18): 7082-7089.
Dudzik, D., C. Barbas-Bernardos, A. García and C. Barbas (2018). "Quality assurance procedures for mass spectrometry untargeted metabolomics. a review." Journal of Pharmaceutical and Biomedical Analysis **147**: 149-173.
Dührkop, K., M. Fleischauer, M. Ludwig, A. A. Aksenov, A. V. Melnik, M. Meusel, P. C. Dorrestein, J. Rousu and S. Böcker (2019). "SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information." Nature Methods **16**(4): 299-302.
Dührkop, K., L.-F. Nothias, M. Fleischauer, R. Reher, M. Ludwig, M. A. Hoffmann, D. Petras, W. H. Gerwick, J. Rousu, P. C. Dorrestein and S. Böcker (2021). "Systematic classification of

unknown metabolites using high-resolution fragmentation mass spectra." Nature Biotechnology **39**(4): 462-471.
Đurđić, S., M. Pantelić, J. Trifković, V. Vukojević, M. Natić, Ž. Tešić and J. Mutić (2017). "Elemental composition as a tool for the assessment of type, seasonal variability, and geographical origin of wine and its contribution to daily elemental intake." RSC Advances **7**(4): 2151-2162.
Ehlers, M., B. Horn, J. Raeke, C. Fauhl-Hassek, A. Hermann, J. Brockmeyer and J. Riedl (2022). "Towards harmonization of non-targeted 1H NMR spectroscopy-based wine authentication: Instrument comparison." Food Control **132**: 108508.
Ehlers, M., L. Uttl, J. Riedl, J. Raeke, I. Westkamp, J. Hajslova, J. Brockmeyer and C. Fauhl-Hassek (2023). "Instrument comparability of non-targeted UHPLC-HRMS for wine authentication." Food Control **144**: 109360.
Einfalt, D. (2020). "Characterization of Volatile Compounds in Quality-Ranked Gins." Mitteilungen Klosterneuberg **70**: 14.
El Naqa, I. and M. J. Murphy (2015). What Is Machine Learning? Machine Learning in Radiation Oncology: Theory and Applications. I. El Naqa, R. Li and M. J. Murphy. Cham, Springer International Publishing**:** 3-11.
Elliott, C. (2014). Review into the integrity and assurance of food supply networks. London, HM Government**:** 432.
Ellis, D. I., H. Muhamadali, Y. Xu, R. Eccles, I. Goodall and R. Goodacre (2019). "Rapid through-container detection of fake spirits and methanol quantification with handheld Raman spectroscopy." The Analyst **144**(1): 324-330.
Espinosa-Vega, L. I., A. Belio-Manzano, C. A. Mercado-Ornelas, I. E. Cortes-Mestizo and V. H. Mendez-Garcia (2019). "Aging spectral markers of tequila observed by Raman spectroscopy." European Food Research and Technology **245**(5): 1031-1036.
Esslinger, S., J. Riedl and C. Fauhl-Hassek (2014). "Potential and limitations of non-targeted fingerprinting for authentication of food in official control." Food Research International **60**: 189-204.
Euromonitor International (2015). Illicit Alcohol in Russia.
European Commission. (2003). "TRACES." from https://ec.europa.eu/food/animals/traces_en#:~:text=TRACES%20allows%20for%20the%20quick,cooperation%20and%20collaboration%20between%20authorities.
European Commission (2018). The EU Food Fraud Network and the System for Administrative Assistance - Food Fraud - Annual Report 2018**:** 16.
European Commission (2022). Monthly Summary of Articles on Food Fraud and Adulteration - September 2016 - September 2022.
European Commission (EC) (2008). Council Regulation (EC) No 479/2008 of 29 April 2008 on the common organisation of the market in wine, amending Regulations (EC) No 1493/1999, (EC) No 1782/2003, (EC) No 1290/2005, (EC) No 3/2008 and repealing Regulations (EEC) No 2392/86 and (EC) No 1493/1999. E. C. (EC). **ANNEX VI - Restrictions - A. 1**.
European Commission (EC) (2008). Regulation (EC) No 110/2008 of the European Parliament and of the Council of 15 January 2008 on the definition, description, presentation, labelling and the protection of geographical indications of spirit drinks and repealing Council Regulation (EEC) No 1576/89. **Article 3**.
Europol - OPSON IX (2020). 320 tonnes of potentially dangerous dairy products taken off the market in operation OPSON IX targeting food fraud.

Europol - OPSON V (2016). Largest-ever seizures of fake food and drink in INTERPOL-Europol operation.
Europol - OPSON X (2021). 15 000 tonnes of illegal food and beverages off the market.
Evans, A. M., C. O'Donovan, M. Playdon, C. Beecher, R. D. Beger, J. A. Bowden, D. Broadhurst, C. B. Clish, S. Dasari, W. B. Dunn, J. L. Griffin, T. Hartung, P.-C. Hsu, T. Huan, J. Jans, C. M. Jones, M. Kachman, A. Kleensang, M. R. Lewis, M. E. Monge, J. D. Mosley, E. Taylor, F. Tayyari, G. Theodoridis, F. Torta, B. K. Ubhi, D. Vuckovic and Q. C. C. on behalf of the Metabolomics Quality Assurance (2020). "Dissemination and analysis of the quality assurance (QA) and quality control (QC) practices of LC–MS based untargeted metabolomics practitioners." Metabolomics **16**(10): 113.
Evert, R. F., S. E. Eichhorn and P. H. Raven (2013). Raven Biology of plants. Chapter 7: Photosynthesis, Light, and Life. New York, W.H. Freeman and Company Publishers.
Eyring, M. B. (2003). Spectroscopy in Forensic Science. Encyclopedia of Physical Science and Technology (Third Edition). R. A. Meyers. New York, Academic Press**:** 637-643.
Fahrni, S. M., B. T. Fuller and J. R. Southon (2015). "Angel's Share Combats Wine Fraud: 14C Dating of Wine without Opening the Bottle." Analytical Chemistry **87**(17): 8646-8650.
Fan, C., M. Chen, X. Wang, J. Wang and B. Huang (2021). "A Review on Data Preprocessing Techniques Toward Efficient and Reliable Knowledge Discovery From Building Operational Data." Frontiers in Energy Research **9**.
Feider, C. L., A. Krieger, R. J. DeHoog and L. S. Eberlin (2019). "Ambient Ionization Mass Spectrometry: Recent Developments and Applications." Analytical Chemistry **91**(7): 4266-4290.
Feldmann, J., A. Raab and E. M. Krupp (2018). "Importance of ICPMS for speciation analysis is changing: future trends for targeted and non-targeted element speciation analysis." Analytical and Bioanalytical Chemistry **410**(3): 661-667.
Fernández-Sánchez, M. L. (2019). Atomic Emission Spectrometry | Inductively Coupled Plasma☆. Encyclopedia of Analytical Science (Third Edition). P. Worsfold, C. Poole, A. Townshend and M. Miró. Oxford, Academic Press**:** 169-176.
Fernández Pierna, J. A., L. Duponchel, C. Ruckebusch, D. Bertrand, V. Baeten and P. Dardenne (2012). "Trappist beer identification by vibrational spectroscopy: A chemometric challenge posed at the 'Chimiométrie 2010' congress." Chemometrics and Intelligent Laboratory Systems **113**: 2-9.
Fischer, S. and W. Dunca (2007). Optimizing the Agilent Multimode Source. Santa Clara, California, U.S.A.**:** 4.
Fisher, C. M., T. R. Croley and A. M. Knolhoff (2021). "Data processing strategies for non-targeted analysis of foods using liquid chromatography/high-resolution mass spectrometry." TrAC Trends in Analytical Chemistry **136**: 116188.
Fleming, H., M. Chen, G. D. Bruce and K. Dholakia (2020). "Through-bottle whisky sensing and classification using Raman spectroscopy in an axicon-based backscattering configuration." Analytical Methods **12**(37): 4572-4578.
Food Safety Authority of Ireland (FSAI) (2013). FSAI Survey Finds Horse DNA in Some Beef Burger Products. Ireland, Food Safety Authority of Ireland.
Food Safety Authority of Ireland (FSAI) (2019). Labelling of Irish Whiskey. F. S. A. o. I. (FSAI). Ireland**:** 20.
Forina, M., P. Oliveri, H. Jäger, U. Römisch and J. Smeyers-Verbeke (2009). "Class modeling techniques in the control of the geographical origin of wines." Chemometrics and Intelligent Laboratory Systems **99**(2): 127-137.

Fougere, E., E. K. Kaplan and C. A. Collins (2020). "Pricing uncertainty in wine markets following the Rudy Kurniawan scandal." Journal of Wine Research **31**(1): 1-5.
Freund, M. S. and N. S. Lewis (1995). "A chemically diverse conducting polymer-based "electronic nose"." Proceedings of the National Academy of Sciences **92**(7): 2652-2656.
Fung, F., H.-S. Wang and S. Menon (2018). "Food safety in the 21st century." Biomedical Journal **41**(2): 88-95.
Gabrieli, G., M. Muszynski and P. W. Ruch (2022). A reconfigurable integrated electronic tongue and its use in accelerated analysis of juices and wines. 2022 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN).
Gajek, M., A. Pawlaczyk, P. Wysocki and M. I. Szynkowska-Jozwik (2021). "Elemental Characterization of Ciders and Other Low-Percentage Alcoholic Beverages Available on the Polish Market." Molecules **26**(8): 2186.
García-Ruiz, S., M. Moldovan, G. Fortunato, S. Wunderli and J. I. García Alonso (2007). "Evaluation of strontium isotope abundance ratios in combination with multi-elemental analysis as a possible tool to study the geographical origin of ciders." Analytica Chimica Acta **590**(1): 55-66.
Geană, E.-I., C. T. Ciucure and C. Apetrei (2020). "Electrochemical Sensors Coupled with Multivariate Statistical Analysis as Screening Tools for Wine Authentication Issues: A Review." Chemosensors **8**(3): 59.
Geană, E.-I., C. T. Ciucure, C. Apetrei and V. Artem (2019). "Application of Spectroscopic UV-Vis and FT-IR Screening Techniques Coupled with Multivariate Statistical Analysis for Red Wine Authentication: Varietal and Vintage Year Discrimination." Molecules **24**(22): 4166.
Gertsman, I. and B. A. Barshop (2018). "Promises and pitfalls of untargeted metabolomics." Journal of Inherited Metabolic Disease **41**(3): 355-366.
GFSI. (2014). "GFSI position on mitigating the public health risk of food fraud." Retrieved March 19th, 2022.
GFSI. (2018). "Tackling Food Fraud Through Food Safety Management Systems." Retrieved March 19th, 2022, from https://mygfsi.com/wp-content/uploads/2019/09/Food-Fraud-GFSI-Technical-Document.pdf.
Ghasemi-Varnamkhasti, M., S. S. Mohtasebi, M. Siadat, J. Lozano, H. Ahmadi, S. H. Razavi and A. Dicko (2011). "Aging fingerprint characterization of beer using electronic nose." Sensors and Actuators B: Chemical **159**(1): 51-59.
Giachino, R. R. A., Ç. Sönmez, F. A. Tonk, E. Bayram, S. Yüce, I. Telci and M. A. Furan (2014). "RAPD and essential oil characterization of Turkish basil (Ocimum basilicum L.)." Plant Systematics and Evolution **300**(8): 1779-1791.
Giacinti, G., C. Raynaud, S. Capblancq and V. Simon (2016). "Matrix-Matching as an Improvement Strategy for the Detection of Pesticide Residues." Journal of Food Science **81**(5): T1342-T1350.
Giannetti, V., M. B. Mariani, F. Marini, P. Torrelli and A. Biancolillo (2020). "Grappa and Italian spirits: Multi-platform investigation based on GC–MS, MIR and NIR spectroscopies for the authentication of the Geographical Indication." Microchemical Journal **157**: 104896.
Godelmann, R., F. Fang, E. Humpfer, B. Schütz, M. Bansbach, H. Schäfer and M. Spraul (2013). "Targeted and Nontargeted Wine Analysis by 1H NMR Spectroscopy Combined with Multivariate Statistical Analysis. Differentiation of Important Parameters: Grape Variety, Geographical Origin, Year of Vintage." Journal of Agricultural and Food Chemistry **61**(23): 5610-5619.

Goethem, B. V. and C. Elliott (2018). Handbook of Food Authenticity: A Guide to Food Authenticity Issues and Analytical Solutions, Elsevier.
González-Arjona, D., V. González-Gallero, F. Pablos and A. Gustavo González (1999). "Authentication and differentiation of irish whiskeys by higher-alcohol congener analysis." Analytica Chimica Acta **381**(2): 257-264.
González-Pereira, A., P. Otero, M. Fraga-Corral, P. Garcia-Oliveira, M. Carpena, M. A. Prieto and J. Simal-Gandara (2021). "State-of-the-Art of Analytical Techniques to Determine Food Fraud in Olive Oils." Foods **10**(3): 484.
Goodall, I., S. Harrison, R. Eccles and P. Cockburn (2018). 14. Spirit Drinks. Food Integrity Handbook. M. L. Jean-François Morin. France, Eurofins Analytics France**:** 229-250.
Gordon, R., J. Chapman, A. Power, S. Chandra, J. Roberts and D. Cozzolino (2018). "Unfrazzled by Fizziness: Identification of Beers Using Attenuated Total Reflectance Mid-infrared Spectroscopy and Multivariate Analysis." Food Analytical Methods **11**(9): 2360-2367.
Goyal, K., P. Kumar and K. Verma (2022). "Food Adulteration Detection using Artificial Intelligence: A Systematic Review." Archives of Computational Methods in Engineering **29**(1): 397-426.
Granados, J. Q., M. V. Mir, H. L. G. Serrana and M. C. L. Martinez (1996). "The influence of added caramel on furanic aldehyde content of matured brandies." Food Chemistry **56**(4): 415-419.
Gravert, T. K. O., J. Vuaille, J. Magid and M. Hansen (2021). "Non-target analysis of organic waste amended agricultural soils: Characterization of added organic pollution." Chemosphere **280**: 130582.
Greer, D., L. Pfahl, J. Rieck, T. Daniels and O. Garza (2008). "Comparison of a Novel Distillation Method versus a Traditional Distillation Method in a Model Gin System Using Liquid/Liquid Extraction." Journal of Agricultural and Food Chemistry **56**(19): 9030-9036.
Haag, A. M. (2016). Mass Analyzers and Mass Spectrometers. Modern Proteomics – Sample Preparation, Analysis and Practical Applications. H. Mirzaei and M. Carrasco. Cham, Springer International Publishing**:** 157-169.
Hajeb, P., L. Zhu, R. Bossi and K. Vorkamp (2022). "Sample preparation techniques for suspect and non-target screening of emerging contaminants." Chemosphere **287**: 132306.
Hale, S. E., M. Neumann, I. Schliebner, J. Schulze, F. S. Averbeck, C. Castell-Exner, M. Collard, D. Drmač, J. Hartmann, R. Hofman-Caris, J. Hollender, M. de Jonge, T. Kullick, A. Lennquist, T. Letzel, K. Nödler, S. Pawlowski, N. Reineke, E. Rorije, M. Scheurer, G. Sigmund, H. Timmer, X. Trier, E. Verbruggen and H. P. H. Arp (2022). "Getting in control of persistent, mobile and toxic (PMT) and very persistent and very mobile (vPvM) substances to protect water resources: strategies from diverse perspectives." Environmental Sciences Europe **34**(1): 22.
Harper, W. J. (2001). The Strengths and Weaknesses of the Electronic Nose. Headspace Analysis of Foods and Flavors: Theory and Practice. R. L. Rouseff and K. R. Cadwallader. Boston, MA, Springer US**:** 59-71.
Hassoun, A., I. Måge, W. F. Schmidt, H. T. Temiz, L. Li, H.-Y. Kim, H. Nilsen, A. Biancolillo, A. Aït-Kaddour, M. Sikorski, E. Sikorska, S. Grassi and D. Cozzolino (2020). "Fraud in Animal Origin Food Products: Advances in Emerging Spectroscopic Detection Methods over the Past Five Years." Foods **9**(8): 1069.
He, N. X. and S. Bayen (2020). "An overview of chemical contaminants and other undesirable chemicals in alcoholic beverages and strategies for analysis." Compr Rev Food Sci Food Saf **19**(6): 3916-3950.

Hermosín-Gutiérrez, I., N. Castillo-Muñoz, S. Gómez-Alonso and E. García-Romero (2011). Flavonol Profiles for Grape and Wine Authentication. Progress in Authentication of Food and Wine, American Chemical Society. **1081:** 113-129.
Hodel, J., M. Pauley, M. C. J. K. Gorseling and A. E. Hill (2019). "Quantitative Comparison of Volatiles in Vapor Infused Gin versus Steep Infused Gin Distillates." Journal of the American Society of Brewing Chemists **77**(3): 149-156.
Holmberg (2010). "Wine fraud." International Journal of Wine Research: 105.
Hopfer, H., G. Gilleland, S. E. Ebeler and J. Nelson (2017). "Elemental Profiles of Whisk(e)y Allow Differentiation by Type and Region." Beverages **3**(1): 8.
Hopke, P. K. (2003). "The evolution of chemometrics." Analytica Chimica Acta **500**(1): 365-377.
Hubert, P., F. Perrot, J. Gaye, B. Médina and M. S. Pravikoff (2009). "Radioactivity measurements applied to the dating and authentication of old wines." Comptes Rendus Physique **10**(7): 622-629.
IARC (1988). "Chemical composition of alcoholic beverages, additives and contaminants." IARC Monogr Eval Carcinog Risks Hum **44**: 71-99.
IFPRI, I. F. P. R. I. "Food Security." Retrieved March 18th, 2022, from https://www.ifpri.org/topic/food-security#:~:text=Food%20security,%20as%20defined%20by,an%20active%20and%20healthy%20life.
Ingrid Peignier, N. D. M.-W. (2017). Food Fraud and Consumer Risk

Percep3on in Quebec (Canada). Food Fraud Global Understanding. Château Frontenac, Québec City, Canada.
International Alliance for Drinking Responsibly (IARD) (2018). Alcohol in the Shadow Economy - Unregulated, Untaxed, and Potentially Toxic.
Ismaiel, O. A., M. S. Halquist, M. Y. Elmamly, A. Shalaby and H. Thomas Karnes (2008). "Monitoring phospholipids for assessment of ion enhancement and ion suppression in ESI and APCI LC/MS/MS for chlorpheniramine in human plasma and the importance of multiple source matrix effect evaluations." Journal of Chromatography B **875**(2): 333-343.
Jain, A. K., M. Jianchang and K. M. Mohiuddin (1996). "Artificial neural networks: a tutorial." Computer **29**(3): 31-44.
Jennings, K. R. and G. G. Dolnikowski (1990). [2] Mass analyzers. Methods in Enzymology, Academic Press. **193:** 37-61.
Ji, Q. (2020). 5 - Computer vision applications. Probabilistic Graphical Models for Computer Vision. Q. Ji. Oxford, Academic Press**:** 191-297.
Johnson, S. R., J. M. Sutter, H. L. Engelhardt, P. C. Jurs, J. White, J. S. Kauer, T. A. Dickinson and D. R. Walt (1997). "Identification of Multiple Analytes Using an Optical Sensor Array and Pattern Recognition Neural Networks." Analytical Chemistry **69**(22): 4641-4648.
JRC (2023). Geel Unit F.4 “Fraud Detection and Prevention”, Food Fraud Summary Report. Europe, European Commission (EC).
Jung, J., T. Jaufmann, U. Hener, A. Münch, M. Kreck, H. Dietrich and A. Mosandl (2006). "Progress in wine authentication: GC–C/P–IRMS measurements of glycerol and GC analysis of 2,3-butanediol stereoisomers." European Food Research and Technology **223**(6): 811-820.
Kakouri, E., P.-K. Revelou, C. Kanakis, D. Daferera, C. S. Pappas and P. A. Tarantilis (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): 1565.

Kang, J., S. He, W. Zhou, Z. Shen, Y. Li, M. Chen, Q. Zhang and Y. Wang (2020). "Single-pass transformation of syngas into ethanol with high selectivity by triple tandem catalysis." Nature Communications **11**(1): 827.
Karabagias, I. K., V. K. Karabagias and A. V. Badeka (2021). "Volatilome of white wines as an indicator of authenticity and adulteration control using statistical analysis." Australian Journal of Grape and Wine Research **27**(3): 269-279.
Karatoprak, G. Ş. (2022). Chapter4.20 - Vanillin. Antioxidants Effects in Health. S. M. Nabavi and A. S. Silva, Elsevier**:** 517-526.
Katajamaa, M. and M. Orešič (2005). "Processing methods for differential analysis of LC/MS profile data." BMC Bioinformatics **6**(1): 179.
Kayacan, E. and M. A. Khanesar (2016). Chapter 2 - Fundamentals of Type-1 Fuzzy Logic Theory. Fuzzy Neural Networks for Real Time Control Applications. E. Kayacan and M. A. Khanesar, Butterworth-Heinemann**:** 13-24.
Kebarle, P. and U. H. Verkerk (2010). On the Mechanism of Electrospray Ionization Mass Spectrometry (ESIMS). Electrospray and MALDI Mass Spectrometry**:** 1-48.
Kelly, S., C. Brodie and A. Hilkert (2018). Chapter 11 - Isotopic-Spectroscopic Technique: Stable Isotope-Ratio Mass Spectrometry (IRMS). Modern Techniques for Food Authentication (Second Edition). D.-W. Sun, Academic Press**:** 349-413.
Kelly, S., K. Heaton and J. Hoogewerff (2005). "Tracing the geographical origin of food: The application of multi-element and multi-isotope analysis." Trends in Food Science & Technology **16**(12): 555-567.
Kendall, H., B. Clark, C. Rhymer, S. Kuznesof, J. Hajslova, M. Tomaniova, P. Brereton and L. Frewer (2019). "A systematic review of consumer perceptions of food fraud and authenticity: A European perspective." Trends in Food Science & Technology **94**: 79-90.
Kevin J. McHale, R. D. (2018). Measurement of Terpenes in Plant Extracts via LC-MS/MS.
Kew, W., I. Goodall and D. Uhrín (2019). "Analysis of Scotch Whisky by 1H NMR and chemometrics yields insight into its complex chemistry." Food Chemistry **298**: 125052.
Khaksar, R., T. Carlson, D. W. Schaffner, M. Ghorashi, D. Best, S. Jandhyala, J. Traverso and S. Amini (2015). "Unmasking seafood mislabeling in U.S. markets: DNA barcoding as a unique technology for food authentication and quality control." Food Control **56**: 71-76.
Khalil, G., R. Doss and M. Chowdhury (2019). "A Comparison Survey Study on RFID Based Anti-Counterfeiting Systems." Journal of Sensor and Actuator Networks **8**(3): 37.
King, R., R. Bonfiglio, C. Fernandez-Metzler, C. Miller-Stein and T. Olah (2000). "Mechanistic investigation of ionization suppression in electrospray ionization." Journal of the American Society for Mass Spectrometry **11**(11): 942-950.
Kizil, R. and J. Irudayaraj (2018). Chapter 6 - Spectroscopic Technique: Fourier Transform Raman (FT-Raman) Spectroscopy. Modern Techniques for Food Authentication (Second Edition). D.-W. Sun, Academic Press**:** 193-217.
Kostiainen, R. and T. J. Kauppila (2009). "Effect of eluent on the ionization process in liquid chromatography–mass spectrometry." Journal of Chromatography A **1216**(4): 685-699.
Kruzlicova, D. and S. Gruberova (2022). "Classification and authentication of Slovak varietal wines by attenuated total reflectance Fourier-transform infrared spectrometry and multidimensional data analysis." Chemical Papers **76**(4): 2385-2396.
Krzywinski, M. and N. Altman (2017). "Classification and regression trees." Nature Methods **14**(8): 757-758.

Kuballa, T., T. Hausler, A. O. Okaru, M. Neufeld, K. O. Abuga, I. O. Kibwage, J. Rehm, B. Luy, S. G. Walch and D. W. Lachenmeier (2018). "Detection of counterfeit brand spirits using 1H NMR fingerprints in comparison to sensory analysis." Food Chemistry **245**: 112-118.
Lachenmeier, D. W. (2016). 21 - Advances in the Detection of the Adulteration of Alcoholic Beverages Including Unrecorded Alcohol. Advances in Food Authenticity Testing. G. Downey, Woodhead Publishing**:** 565-584.
Lachenmeier, D. W., R. Attig, W. Frank and C. Athanasakis (2003). "The use of ion chromatography to detect adulteration of vodka and rum." European Food Research and Technology **218**(1): 105-110.
Lachenmeier, D. W., K. Schoeberl, F. Kanteres, T. Kuballa, E.-M. Sohnius and J. Rehm (2011). "Is contaminated unrecorded alcohol a health problem in the European Union? A review of existing and methodological outline for future studies." Addiction **106**(s1): 20-30.
Lachenmeier, D. W., E.-M. Sohnius, R. Attig and M. G. López (2006). "Quantification of Selected Volatile Constituents and Anions in Mexican Agave Spirits (Tequila, Mezcal, Sotol, Bacanora)." Journal of Agricultural and Food Chemistry **54**(11): 3911-3915.
Lai, Y. W., E. K. Kemsley and R. H. Wilson (1994). "Potential of Fourier Transform Infrared Spectroscopy for the Authentication of Vegetable Oils." Journal of Agricultural and Food Chemistry **42**(5): 1154-1159.
Lappas, N. T. and C. M. Lappas (2016). Chapter 10 - Methods of Detection, Identification, and Quantitation. Forensic Toxicology. N. T. Lappas and C. M. Lappas. San Diego, Academic Press**:** 161-181.
Le Gall, G. and I. J. Colquhoun (2003). NMR spectroscopy in food authentication, Elsevier**:** 131-155.
Lecat, B., J. Brouard and C. Chapuis (2017). "Fraud and counterfeit wines in France: an overview and perspectives." British Food Journal **119**(1): 84-104.
Ledauphin, J., C. Le Milbeau, D. Barillier and D. Hennequin (2010). "Differences in the Volatile Compositions of French Labeled Brandies (Armagnac, Calvados, Cognac, and Mirabelle) Using GC-MS and PLS-DA." Journal of Agricultural and Food Chemistry **58**(13): 7782-7793.
Lee, H.-R., S. Kochhar and S.-M. Shim (2015). "Comparison of Electrospray Ionization and Atmospheric Chemical Ionization Coupled with the Liquid Chromatography-Tandem Mass Spectrometry for the Analysis of Cholesteryl Esters." International Journal of Analytical Chemistry **2015**: 650927.
Lee, K.-Y. M., A. Paterson, J. R. Piggott and G. D. Richardson (2001). "Origins of Flavour in Whiskies and a Revised Flavour Wheel: a Review." Journal of the Institute of Brewing **107**(5): 287-313.
Lee, L. C., C.-Y. Liong and A. A. Jemain (2018). "Partial least squares-discriminant analysis (PLS-DA) for classification of high-dimensional (HD) data: a review of contemporary practice strategies and knowledge gaps." Analyst **143**(15): 3526-3539.
Legin, A., A. Rudnitskaya, L. Lvova, Y. Vlasov, C. Di Natale and A. D'Amico (2003). "Evaluation of Italian wine by the electronic tongue: recognition, quantitative analysis and correlation with human sensory perception." Analytica Chimica Acta **484**(1): 33-44.
Legin, A., A. Rudnitskaya, B. Seleznev and Y. Vlasov (2005). "Electronic tongue for quality assessment of ethanol, vodka and eau-de-vie." Analytica Chimica Acta **534**(1): 129-135.
León-Rodríguez, A. D., P. Escalante-Minakata, M. I. Jiménez-García, L. G. Ordoñez-Acevedo, J. L. F. Flores and A. P. B. d. l. Rosa (2007). "Characterization of Volatile Compounds from Ethnic Agave Alcoholic Beverages by Gas Chromatography-Mass

Spectrometry." Food Technology and Biotechnology **46**(4): 448-455.
Li, D.-X., L. Gan, A. Bronja and O. J. Schmitz (2015). "Gas chromatography coupled to atmospheric pressure ionization mass spectrometry (GC-API-MS): Review." Analytica Chimica Acta **891**: 43-61.
Li, S., Y. Wang, G. M. L. Tacken, Y. Liu and S. J. Sijtsema (2021). "Consumer trust in the dairy value chain in China: The role of trustworthiness, the melamine scandal, and the media." Journal of Dairy Science **104**(8): 8554-8567.
Liigand, P., K. Kaupmees, K. Haav, J. Liigand, I. Leito, M. Girod, R. Antoine and A. Kruve (2017). "Think Negative: Finding the Best Electrospray Ionization/MS Mode for Your Analyte." Analytical Chemistry **89**(11): 5665-5668.
Lin, S. and S. Salcido-Keamo (2021). Chapter 12 - Fraud in wine and other alcoholic beverages. Food Fraud. R. S. Hellberg, K. Everstine and S. A. Sklare, Academic Press**:** 233-247.
Lindley, J. (2022). "Food regulation and policing: innovative technology to close the regulatory gap in Australia." Journal of Consumer Protection and Food Safety.
López-Monroy, A. P. and J. S. García-Salinas (2022). Chapter 9 - Neural networks and deep learning. Biosignal Processing and Classification Using Computational Learning and Intelligence. A. A. Torres-García, C. A. Reyes-García, L. Villaseñor-Pineda and O. Mendoza-Montoya, Academic Press**:** 177-196.
López-Ruiz, R., R. Romero-González and A. Garrido Frenich (2019). "Metabolomics approaches for the determination of multiple contaminants in food." Current Opinion in Food Science **28**: 49-57.
Lopez, M. G., N. A. Mancilla-Margalli and G. Mendoza-Diaz (2003). "Molecular Structures of Fructans from Agave tequilana Weber var. azul." Journal of Agricultural and Food Chemistry **51**(27): 7835-7840.
Lorenc, W., A. Hanć, A. Sajnóg and D. Barałkiewicz (2022). "LC/ICP-MS AND COMPLEMENTARY TECHNIQUES IN BESPOKE AND NONTARGETED SPECIATION ANALYSIS OF ELEMENTS IN FOOD SAMPLES." Mass Spectrometry Reviews **41**(1): 32-50.
Lu, Y. (2020). Alcoholic Drinks Report 2020. Statista Consumer Market Outlook – Market Report**:** 128.
Lu, Y. (2020). Spirits Report 2020. Statista Consumer Market Outlook – Market Report**:** 36.
Lukić, I., M. Banović, Đ. Peršurić, S. Radeka and B. Sladonja (2006). "Determination of volatile compounds in grape distillates by solid-phase extraction and gas chromatography." Journal of Chromatography A **1101**(1): 238-244.
Lunkenbein, S., M. Bellido, A. Aharoni, E. M. J. Salentijn, R. Kaldenhoff, H. A. Coiner, J. Muñoz-Blanco and W. Schwab (2006). "Cinnamate Metabolism in Ripening Fruit. Characterization of a UDP-Glucose:Cinnamate Glucosyltransferase from Strawberry." Plant Physiology **140**(3): 1047-1058.
Luzzani, G., E. Grandis, M. Frey and E. Capri (2021). "Blockchain Technology in Wine Chain for Collecting and Addressing Sustainable Performance: An Exploratory Study." Sustainability **13**(22): 12898.
Magdas, D. A., G. Cristea, A. Pîrnau, I. Feher, A. R. Hategan and A. Dehelean (2021). "Authentication of Transylvanian Spirits Based on Isotope and Elemental Signatures in Conjunction with Statistical Methods." Foods **10**(12): 3000.
Magdas, D. A., M. David and C. Berghian-Grosan (2022). "Fruit spirits fingerprint pointed out through artificial intelligence and FT-Raman spectroscopy." Food Control **133**: 108630.

Malek, S., C. Hui, N. Aziida, S. Cheen, S. Toh and P. Milow (2019). Ecosystem Monitoring Through Predictive Modeling. Encyclopedia of Bioinformatics and Computational Biology. S. Ranganathan, M. Gribskov, K. Nakai and C. Schönbach. Oxford, Academic Press**:** 1-8.
Mallmann, L. P., A. O. Rios and E. Rodrigues (2023). "MS-FINDER and SIRIUS for phenolic compound identification from high-resolution mass spectrometry data." Food Research International **163**: 112315.
Mamani-Huanca, M., A. G. de la Fuente, A. Otero, A. Gradillas, J. Godzien, C. Barbas and Á. López‐Gonzálvez (2021). "Enhancing confidence of metabolite annotation in Capillary Electrophoresis-Mass Spectrometry untargeted metabolomics with relative migration time and in-source fragmentation." Journal of Chromatography A **1635**: 461758.
Mannina, L., F. Marini, R. Antiochia, S. Cesa, A. Magrì, D. Capitani and A. P. Sobolev (2016). "Tracing the origin of beer samples by NMR and chemometrics: Trappist beers as a case study." ELECTROPHORESIS **37**(20): 2710-2719.
Manning, L. (2017). "Guest editorial." British Food Journal **119**(1): 2-6.
Manning, L. and A. Kowalska (2021). "Illicit Alcohol: Public Health Risk of Methanol Poisoning and Policy Mitigation Strategies." Foods **10**(7): 1625.
Mardia, K. V., J. T. Kent and J. M. Bibby (1979). Multivariate analysis. Chapter 8: Principal Component Analysis. London ;, Academic Press.
Marenco, L. F. L., L. P. de Oliveira, D. L. Vale and M. O. Salles (2021). "Predicting Vodka Adulteration: A Combination of Electronic Tongue and Artificial Neural Networks." Journal of The Electrochemical Society **168**(11): 117513.
Marriott, P. J. (2005). GAS CHROMATOGRAPHY | Principles. Encyclopedia of Analytical Science (Second Edition). P. Worsfold, A. Townshend and C. Poole. Oxford, Elsevier**:** 7-18.
Martí, M. P., O. Busto and J. Guasch (2004). "Application of a headspace mass spectrometry system to the differentiation and classification of wines according to their origin, variety and ageing." Journal of Chromatography A **1057**(1): 211-217.
Martin, A. E., R. J. Watling and G. S. Lee (2012). "The multi-element determination and regional discrimination of Australian wines." Food Chemistry **133**(3): 1081-1089.
Martinelli, L. A., G. B. Nardoto, M. A. Z. Perez, G. A. Junior, F. C. Fracassi, J. G. G. Oliveira, I. S. Ottani, S. H. Lima, E. A. Mazzi, T. F. Gomes, A. Soltangheisi, A. L. A. Filho, E. Mariano, F. J. V. Costa, P. J. Duarte-Neto, M. Z. Moreira and P. B. Camargo (2020). "Carbon and Nitrogen Isotope Ratios of Food and Beverage in Brazil." Molecules **25**(6): 1457.
Matijašević, S., J. Popović-Djordjević, R. Ristić, D. Ćirković, B. Ćirković and T. Popović (2019). "Volatile Aroma Compounds of Brandy ‘Lozovača’ Produced from Muscat Table Grapevine Cultivars (Vitis vinifera L.)." Molecules **24**(13): 2485.
Mattarucchi, E., M. Stocchero, J. M. Moreno-Rojas, G. Giordano, F. Reniero and C. Guillou (2010). "Authentication of Trappist Beers by LC-MS Fingerprints and Multivariate Data Analysis." Journal of Agricultural and Food Chemistry **58**(23): 12089-12095.
Mavani, N. R., J. M. Ali, S. Othman, M. A. Hussain, H. Hashim and N. A. Rahman (2022). "Application of Artificial Intelligence in Food Industry—a Guideline." Food Engineering Reviews **14**(1): 134-175.
Mazerolles, G., S. Preys, C. Bouchut, E. Meudec, H. Fulcrand, J. M. Souquet and V. Cheynier (2010). "Combination of several mass spectrometry ionization modes: A multiblock analysis for a rapid characterization of the red wine polyphenolic composition." Analytica Chimica Acta **678**(2): 195-202.

McCallum, C. S., S. Cerroni, D. Derbyshire, W. G. Hutchinson and R. M. Nayga Jr. (2021). "Consumers' responses to food fraud risks: an economic experiment." European Review of Agricultural Economics.
Médina, B., M. H. Salagoïty, F. Guyon, J. Gaye, P. Hubert and F. Guillaume (2013). 8 - Using new analytical approaches to verify the origin of wine. New Analytical Approaches for Verifying the Origin of Food. P. Brereton, Woodhead Publishing**:** 149-188.
Mendel, J. M. (1995). "Fuzzy logic systems for engineering: a tutorial." Proceedings of the IEEE **83**(3): 345-377.
Merkytė, V., E. Longo, G. Windisch and E. Boselli (2020). "Phenolic Compounds as Markers of Wine Quality and Authenticity." Foods **9**(12): 1785.
Minnens, F., N. Lucas Luijckx and W. Verbeke (2019). "Food Supply Chain Stakeholders' Perspectives on Sharing Information to Detect and Prevent Food Integrity Issues." Foods **8**(6): 225.
Mogoşanu, G. D., A. M. Grumezescu, C. Bejenaru and L. E. Bejenaru (2017). 11 - Natural products used for food preservation. Food Preservation. A. M. Grumezescu, Academic Press**:** 365-411.
Mohsenzadeh, A., A. Zamani and M. J. Taherzadeh (2017). "Bioethylene Production from Ethanol: A Review and Techno-economical Evaluation." ChemBioEng Reviews **4**(2): 75-91.
Monakhova, Y. B., H. Schäfer, E. Humpfer, M. Spraul, T. Kuballa and D. W. Lachenmeier (2011). "Application of automated eightfold suppression of water and ethanol signals in 1H NMR to provide sensitivity for analyzing alcoholic beverages." Magnetic Resonance in Chemistry **49**(11): 734-739.
Morales-Serna, J. A., A. Jiménez, R. Estrada-Reyes, C. Marquez, J. Cárdenas and M. Salmón (2010). "Homoisoflavanones from Agave tequilana Weber." Molecules **15**(5): 3295-3301.
Müller, T. M., Q. Zhong, S. Fan, D. Wang and C. Fauhl-Hassek (2021). "What's in a wine? – A spot check of the integrity of European wine sold in China based on anthocyanin composition, stable isotope and glycerol impurity analysis." Food Additives & Contaminants: Part A **38**(8): 1289-1300.
Muñoz-Muñoz, A. C., A. C. Grenier, H. Gutiérrez-Pulido and J. Cervantes-Martínez (2008). "Development and validation of a High Performance Liquid Chromatography-Diode Array Detection method for the determination of aging markers in tequila." Journal of Chromatography A **1213**(2): 218-223.
Nácher-Mestre, J., M. Ibáñez, R. Serrano, J. Pérez-Sánchez and F. Hernández (2013). "Qualitative Screening of Undesirable Compounds from Feeds to Fish by Liquid Chromatography Coupled to Mass Spectrometry." Journal of Agricultural and Food Chemistry **61**(9): 2077-2087.
Næs, T., T. Isaksson, T. Fearn and T. Davies (2004). A user-friendly guide to multivariate calibration and classification. Chichester, NIR Publications.
Nardi, T., M. Petrozziello, R. Girotto, M. Fugaro, R. A. Mazzei and S. Scuppa (2020). "Wine aging authentication through Near Infrared Spectroscopy: a feasibility study on chips and barrel aged wines." OENO One **54**(1): 165-173.
Nascimento, C. F., P. M. Santos, E. R. Pereira-Filho and F. R. P. Rocha (2017). "Recent advances on determination of milk adulterants." Food Chemistry **221**: 1232-1244.
Neath, R. C. and M. S. Johnson (2010). Discrimination and Classification. International Encyclopedia of Education (Third Edition). P. Peterson, E. Baker and B. McGaw. Oxford, Elsevier**:** 135-141.

Ng, L.-K., M. Hupé, J. Harnois and D. Moccia (1996). "Characterisation of Commercial Vodkas by Solid-Phase Microextraction and Gas Chromatography/Mass Spectrometry Analysis." Journal of the Science of Food and Agriculture **70**(3): 380-388.
Nietner, T., S. A. Haughey, N. Ogle, C. Fauhl-Hassek and C. T. Elliott (2014). "Determination of geographical origin of distillers dried grains and solubles using isotope ratio mass spectrometry." Food Research International **60**: 146-153.
Novakowski, W., M. Bertotti and T. R. L. C. Paixão (2011). "Use of copper and gold electrodes as sensitive elements for fabrication of an electronic tongue: Discrimination of wines and whiskies." Microchemical Journal **99**(1): 145-151.
Nunes, J. C., A. Ordanini and G. Giambastiani (2021). "The Concept of Authenticity: What It Means to Consumers." Journal of Marketing **85**(4): 1-20.
Nuti, G., L. A. Jiménez Rugama and A.-I. Cross (2021). "An Explainable Bayesian Decision Tree Algorithm." Frontiers in Applied Mathematics and Statistics **7**.
Pacholczyk-Sienicka, B., G. Ciepielowski and Ł. Albrecht (2021). "The Application of NMR Spectroscopy and Chemometrics in Authentication of Spices." Molecules **26**(2): 382.
Pál, L., T. Muhollari, O. Bujdosó, E. Baranyai, A. Nagy, E. Árnyas, R. Ádány, J. Sándor, M. McKee and S. Szűcs (2020). "Heavy metal contamination in recorded and unrecorded spirits. Should we worry?" Regulatory Toxicology and Pharmacology **116**: 104723.
Palamareva, M. D. (2005). LIQUID CHROMATOGRAPHY | Principles. Encyclopedia of Analytical Science (Second Edition). P. Worsfold, A. Townshend and C. Poole. Oxford, Elsevier**:** 112-118.
Palmioli, A., D. Alberici, C. Ciaramelli and C. Airoldi (2020). "Metabolomic profiling of beers: Combining 1H NMR spectroscopy and chemometric approaches to discriminate craft and industrial products." Food Chemistry **327**: 127025.
Pan, Y., H.-W. Gu, Y. Lv, X.-L. Yin, Y. Chen, W. Long, H. Fu and Y. She (2022). "Untargeted metabolomic analysis of Chinese red wines for geographical origin traceability by UPLC-QTOF-MS coupled with chemometrics." Food Chemistry **394**: 133473.
Panossian, A., G. Mamikonyan, M. Torosyan, E. Gabrielyan and S. Mkhitaryan (2001). "Analysis of Aromatic Aldehydes in Brandy and Wine by High-Performance Capillary Electrophoresis." Analytical Chemistry **73**(17): 4379-4383.
Parpinello, G. P., A. Ricci, P. Arapitsas, A. Curioni, L. Moio, S. Riosegade, M. Ugliano and A. Versari (2019). "Multivariate characterisation of Italian monovarietal red wines using MIR spectroscopy: This article is published in cooperation with the 11th OenoIVAS International Symposium, June 25–28 2019, Bordeaux, France." OENO One **53**(4).
Parra, V., Á. A. Arrieta, J.-A. Fernández-Escudero, M. L. Rodríguez-Méndez and J. A. De Saja (2006). "Electronic tongue based on chemically modified electrodes and voltammetry for the detection of adulterations in wines." Sensors and Actuators B: Chemical **118**(1): 448-453.
Parra, V., Á. A. Arrieta, J. A. Fernández-Escudero, M. Íñiguez, J. A. d. Saja and M. L. Rodríguez-Méndez (2006). "Monitoring of the ageing of red wines in oak barrels by means of an hybrid electronic tongue." Analytica Chimica Acta **563**(1): 229-237.
Passos, M. L. C., M. C. Sarraguça, M. L. M. F. S. Saraiva, T. Prasada Rao and V. M. Biju (2019). Spectrophotometry | Organic Compounds☆. Encyclopedia of Analytical Science (Third Edition). P. Worsfold, C. Poole, A. Townshend and M. Miró. Oxford, Academic Press**:** 236-243.
Pegan, G., D. Vianelli and P. de Luca (2020). Strategic Entry Modes and Country of Origin Effect. International Marketing Strategy: The Country of Origin Effect on Decision-Making in Practice. G. Pegan, D. Vianelli and P. de Luca. Cham, Springer International Publishing**:** 23-38.

Pei, X., A. Tandon, A. Alldrick, L. Giorgi, W. Huang and R. Yang (2011). "The China melamine milk scandal and its implications for food safety regulation." Food Policy **36**(3): 412-420.
Peña-Alvarez, A., S. Capella, R. Juárez and C. Labastida (2006). "Determination of terpenes in tequila by solid phase microextraction-gas chromatography–mass spectrometry." Journal of Chromatography A **1134**(1): 291-297.
Peña-Alvarez, A., L. Díaz, A. Medina, C. Labastida, S. Capella and L. E. Vera (2004). "Characterization of three Agave species by gas chromatography and solid-phase microextraction–gas chromatography–mass spectrometry." Journal of Chromatography A **1027**(1): 131-136.
Pereira, H. V., V. S. Amador, M. M. Sena, R. Augusti and E. Piccin (2016). "Paper spray mass spectrometry and PLS-DA improved by variable selection for the forensic discrimination of beers." Analytica Chimica Acta **940**: 104-112.
Perini, M. and F. Camin (2013). "δ18O of Ethanol in Wine and Spirits for Authentication Purposes." Journal of Food Science **78**(6): C839-C844.
Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fisher and H. E. Hoekstra (2012). "Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species." PLoS ONE **7**(5): e37135.
Petrescu, D. C., I. Vermeir and R. M. Petrescu-Mag (2019). "Consumer Understanding of Food Quality, Healthiness, and Environmental Impact: A Cross-National Perspective." International Journal of Environmental Research and Public Health **17**(1): 169.
Petropoulos, S., C. S. Karavas, A. T. Balafoutis, I. Paraskevopoulos, S. Kallithraka and Y. Kotseridis (2017). "Fuzzy logic tool for wine quality classification." Computers and Electronics in Agriculture **142**: 552-562.
Pezoulas, V. C., T. P. Exarchos and D. I. Fotiadis (2020). Chapter 7 - Machine learning and data analytics. Medical Data Sharing, Harmonization and Analytics. V. C. Pezoulas, T. P. Exarchos and D. I. Fotiadis, Academic Press**:** 227-309.
Phan, Q. and E. Tomasino (2021). "Untargeted lipidomic approach in studying pinot noir wine lipids and predicting wine origin." Food Chemistry **355**: 129409.
Pliny the Elder (1885). Naturalis Historia.
Pomarici, E., A. Corsi, S. Mazzarino and R. Sardone (2021). "The Italian Wine Sector: Evolution, Structure, Competitiveness and Future Challenges of an Enduring Leader." Italian Economic Journal **7**(2): 259-295.
Popîrdă, A., C. E. Luchian, V. V. Cotea, L. C. Colibaba, E. C. Scutarașu and A. M. Toader (2021). "A Review of Representative Methods Used in Wine Authentication." Agriculture **11**(3).
Power, A. C., C. N. Néill, S. Geoghegan, S. Currivan, M. Deasy and D. Cozzolino (2020). "A Brief History of Whiskey Adulteration and the Role of Spectroscopy Combined with Chemometrics in the Detection of Modern Whiskey Fraud." Beverages **6**(3): 49.
Quesada-Granados, J. J., C. Samaniego-Sánchez and R. M. Blanca-Herrera (2016). Rhum–Ron–Rum: Technology and Tradition. Encyclopedia of Food and Health. B. Caballero, P. M. Finglas and F. Toldrá. Oxford, Academic Press**:** 618-627.
Raičević, D., T. Popović, D. Jančić, D. Šuković and R. Pajović-Šćepanović (2022). "The Impact of Type of Brandy on the Volatile Aroma Compounds and Sensory Properties of Grape Brandy in Montenegro." Molecules **27**(9): 2974.
Ranaweera, R. K. R., A. M. Gilmore, D. L. Capone, S. E. P. Bastian and D. W. Jeffery (2021). "Authentication of the geographical origin of Australian Cabernet Sauvignon wines using spectrofluorometric and multi-element analyses with multivariate statistical modelling." Food Chemistry **335**: 127592.

Raptis, C. G., C. I. Siettos, C. T. Kiranoudis and G. V. Bafas (2000). "Classification of aged wine distillates using fuzzy and neural network systems." Journal of Food Engineering **46**(4): 267-275.
Reche, R. V., A. F. Leite Neto, A. A. Da Silva, C. A. Galinaro, R. Z. De Osti and D. W. Franco (2007). "Influence of Type of Distillation Apparatus on Chemical Profiles of Brazilian Cachaças." Journal of Agricultural and Food Chemistry **55**(16): 6603-6608.
Regulations, F. a. D. (2022). Division 2 - Alcoholic Beverages - Vodka. C.R.C., c. 870. G. o. Canada. Canada. **B.02.080**.
Rehm, J., S. Kailasapillai, E. Larsen, M. X. Rehm, A. V. Samokhvalov, K. D. Shield, M. Roerecke and D. W. Lachenmeier (2014). "A systematic review of the epidemiology of unrecorded alcohol consumption and the chemical composition of unrecorded alcohol." Addiction **109**(6): 880-893.
Rhodes, C. N., K. Heaton, I. Goodall and P. A. Brereton (2009). "Gas chromatography carbon isotope ratio mass spectrometry applied to the detection of neutral alcohol in Scotch whisky: an internal reference approach." Food Chemistry **114**(2): 697-701.
Riedl, J., S. Esslinger and C. Fauhl-Hassek (2015). "Review of validation and reporting of non-targeted fingerprinting approaches for food authentication." Analytica Chimica Acta **885**: 17-32.
Riu Aumatell, M. (2012). 12 - Gin: production and sensory properties. Alcoholic Beverages. J. Piggott, Woodhead Publishing**:** 267-280.
Room, R. and P. West (1998). "Alcohol and the U.S.-Canada Border: Trade Disputes and Border Traffic Problems." Journal of Public Health Policy **19**(1): 68-87.
Rossmann, A. (2001). "DETERMINATION OF STABLE ISOTOPE RATIOS IN FOOD ANALYSIS." Food Reviews International **17**(3): 347-381.
Roullier-Gall, C., M. Lucio, L. Noret, P. Schmitt-Kopplin and R. D. Gougeon (2014). "How subtle is the "terroir" effect? Chemistry-related signatures of two "climats de Bourgogne"." PLoS One **9**(5): e97615.
Roullier-Gall, C., J. Signoret, C. Coelho, D. Hemmler, M. Kajdan, M. Lucio, B. Schäfer, R. D. Gougeon and P. Schmitt-Kopplin (2020). "Influence of regionality and maturation time on the chemical fingerprint of whisky." Food Chemistry **323**: 126748.
Roullier-Gall, C., J. Signoret, D. Hemmler, M. A. Witting, B. Kanawati, B. Schäfer, R. D. Gougeon and P. Schmitt-Kopplin (2018). "Usage of FT-ICR-MS Metabolomics for Characterizing the Chemical Signatures of Barrel-Aged Whisky." Frontiers in Chemistry **6**.
Rubert, J., O. Lacina, C. Fauhl-Hassek and J. Hajslova (2014). "Metabolic fingerprinting based on high-resolution tandem mass spectrometry: a reliable tool for wine authentication?" Analytical and Bioanalytical Chemistry **406**(27): 6791-6803.
Rubingh, C. M., S. Bijlsma, E. P. P. A. Derks, I. Bobeldijk, E. R. Verheij, S. Kochhar and A. K. Smilde (2006). "Assessing the performance of statistical validation tools for megavariate metabolomics data." Metabolomics **2**(2): 53-61.
Sallans, B., D. Bruckner and G. Russ (2006). Statistical Model-Based Sensor Diagnostics for Automation Systems. Fieldbus Systems and Their Applications 2005. M. L. Chávez. Oxford, Elsevier**:** 239-246.
Sampaio, O. M., R. V. Reche and D. W. Franco (2008). "Chemical Profile of Rums as a Function of their Origin. The Use of Chemometric Techniques for their Identification." Journal of Agricultural and Food Chemistry **56**(5): 1661-1668.
Sarkar, T., M. Salauddin, K. Kirtonia, S. Pati, M. Rebezov, M. Khayrullin, S. Panasenko, L. Tretyak, M. Temerbayeva, N. Kapustina, S. Azimova, L. Gruzdeva, F. Makhmudov, I. Nikitin, A. Kassenov, M. A. Shariati and J. M. Lorenzo (2022). "A Review on the Commonly Used Methods for Analysis of Physical Properties of Food Materials." Applied Sciences **12**(4): 2004.

Schelezki, O. J., A. Deloire and D. W. Jeffery (2020). "Substitution or Dilution? Assessing Pre-Fermentative Water Implementation to Produce Lower Alcohol Shiraz Wines." Molecules **25**(9).
Schermann, J.-P. (2008). 2 - Spectroscopy. Spectroscopy and Modeling of Biomolecular Building Blocks. J.-P. Schermann. Amsterdam, Elsevier**:** 59-128.
Schrimpe-Rutledge, A. C., S. G. Codreanu, S. D. Sherrod and J. A. McLean (2016). "Untargeted Metabolomics Strategies—Challenges and Emerging Directions." Journal of The American Society for Mass Spectrometry **27**(12): 1897-1905.
Schymanski, E. L., J. Jeon, R. Gulde, K. Fenner, M. Ruff, H. P. Singer and J. Hollender (2014). "Identifying Small Molecules via High Resolution Mass Spectrometry: Communicating Confidence." Environmental Science & Technology **48**(4): 2097-2098.
Scott, E. R. and E. E. Crone (2021). "Using the right tool for the job: the difference between unsupervised and supervised analyses of multivariate ecological data." Oecologia **196**(1): 13-25.
Sierra-Padilla, A., J. J. García-Guzmán, D. López-Iglesias, J. M. Palacios-Santander and L. Cubillana-Aguilera (2021). "E-Tongues/Noses Based on Conducting Polymers and Composite Materials: Expanding the Possibilities in Complex Analytical Sensing." Sensors **21**(15): 4976.
Silva, A. J., R. S. Hellberg and R. H. Hanner (2021). Chapter 7 - Seafood fraud. Food Fraud. R. S. Hellberg, K. Everstine and S. A. Sklare, Academic Press**:** 109-137.
Simonnet-Laprade, C., S. Bayen, B. Le Bizec and G. Dervilly (2021). "Data analysis strategies for the characterization of chemical contaminant mixtures. Fish as a case study." Environment International **155**: 106610.
Siuzdak, G. (2004). "An Introduction to Mass Spectrometry Ionization: An Excerpt from The Expanding Role of Mass Spectrometry in Biotechnology, 2nd ed.; MCC Press: San Diego, 2005." JALA: Journal of the Association for Laboratory Automation **9**(2): 50-63.
Śliwińska, M., P. Wiśniewska, T. Dymerski, W. Wardencki and J. Namieśnik (2015). "The flavour of fruit spirits and fruit liqueurs: a review." Flavour and Fragrance Journal **30**(3): 197-207.
Soares, S., J. S. Amaral, M. B. P. P. Oliveira and I. Mafra (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.
Son, H.-S., K. M. Kim, F. van den Berg, G.-S. Hwang, W.-M. Park, C.-H. Lee and Y.-S. Hong (2008). "1H Nuclear Magnetic Resonance-Based Metabolomic Characterization of Wines by Grape Varieties and Production Areas." Journal of Agricultural and Food Chemistry **56**(17): 8007-8016.
Spangenberg, J. E. and V. Zufferey (2018). "Changes in soil water availability in vineyards can be traced by the carbon and nitrogen isotope composition of dried wines." Science of The Total Environment **635**: 178-187.
Spink, J. W. (2019). Food Fraud Prevention Overview (Part 1 of 3): Basics. Food Fraud Prevention: Introduction, Implementation, and Management. J. W. Spink. New York, NY, Springer New York**:** 81-113.
Spink, J. W. (2019). Risk Analysis (Part 1 of 3): Basic Fundamentals. Food Fraud Prevention: Introduction, Implementation, and Management. J. W. Spink. New York, NY, Springer New York**:** 501-528.
Stój, A., I. Kapusta and D. Domagała (2020). "Classification of Red Wines Produced from Zweigelt and Rondo Grape Varieties Based on the Analysis of Phenolic Compounds by UPLC-PDA-MS/MS." Molecules **25**(6): 1342.
Stoll, D.R. and Carr, P.W. (2017). "Two-Dimensional Liquid Chromatography: A State of the Art Tutorial." Analytical Chemistry **89 (1)** 519-531.

Su, W.-H., I. S. Arvanitoyannis and D.-W. Sun (2018). Chapter 18 - Trends in Food Authentication. Modern Techniques for Food Authentication (Second Edition). D.-W. Sun, Academic Press**:** 731-758.
Subasi, A. (2020). Chapter 7 - Clustering examples. Practical Machine Learning for Data Analysis Using Python. A. Subasi, Academic Press**:** 465-511.
Takeoka, G. R. and S. E. Ebeler (2011). Progress in Authentication of Food and Wine. Progress in Authentication of Food and Wine, American Chemical Society. **1081:** 3-11.
Tanatavikorn, H. and Y. Yamashita (2015). Improving Data Reliability for Process Monitoring with Fuzzy Outlier Detection. Computer Aided Chemical Engineering. K. V. Gernaey, J. K. Huusom and R. Gani, Elsevier. **37:** 1595-1600.
Taylor, S. C., G. Laperriere and H. Germain (2017). "Droplet Digital PCR versus qPCR for gene expression analysis with low abundant targets: from variable nonsense to publication quality data." Scientific Reports **7**(1): 2409.
Teodoro, J. A. R., H. V. Pereira, M. M. Sena, E. Piccin, J. J. Zacca and R. Augusti (2017). "Paper spray mass spectrometry and chemometric tools for a fast and reliable identification of counterfeit blended Scottish whiskies." Food Chemistry **237**: 1058-1064.
The Scotch Whisky Regulations (2009). Definition of "Scotch Whisky" and categories of Scotch Whisky. 2009 No. 2890 - Regulation 3.1 (c). U. S. Instruments. United Kingdom. **3.1 (c)**.
Theodoridis, S., A. Pikrakis, K. Koutroumbas and D. Cavouras (2010). CHAPTER 7 - Clustering. Introduction to Pattern Recognition. S. Theodoridis, A. Pikrakis, K. Koutroumbas and D. Cavouras. Boston, Academic Press**:** 159-208.
Thomas, S. N. (2019). Mass spectrometry, Elsevier**:** 171-185.
Thurman, E. M., I. Ferrer and D. Barceló (2001). "Choosing between Atmospheric Pressure Chemical Ionization and Electrospray Ionization Interfaces for the HPLC/MS Analysis of Pesticides." Analytical Chemistry **73**(22): 5441-5449.
Tian, Y., C. Yan, T. Zhang, H. Tang, H. Li, J. Yu, J. Bernard, L. Chen, S. Martin, N. Delepine-Gilon, J. Bocková, P. Veis, Y. Chen and J. Yu (2017). "Classification of wines according to their production regions with the contained trace elements using laser-induced breakdown spectroscopy." Spectrochimica Acta Part B: Atomic Spectroscopy **135**: 91-101.
Trufelli, H., P. Palma, G. Famiglini and A. Cappiello (2011). "An overview of matrix effects in liquid chromatography–mass spectrometry." Mass Spectrometry Reviews **30**(3): 491-509.
Tsakiris, A., S. Kallithraka and Y. Kourkoutas (2014). "Grape brandy production, composition and sensory evaluation." Journal of the Science of Food and Agriculture **94**(3): 404-414.
Tsakiris, A., S. Kallithraka and Y. Kourkoutas (2016). Brandy and Cognac: Manufacture and Chemical Composition. Encyclopedia of Food and Health. B. Caballero, P. M. Finglas and F. Toldrá. Oxford, Academic Press**:** 462-468.
Tzachristas, A., M. Dasenaki, R. Aalizadeh, N. S. Thomaidis and C. Proestos (2021). "LC-MS based metabolomics for the authentication of selected Greek white wines." Microchemical Journal **169**: 106543.
Ueda, S., E. Iwamoto, Y. Kato, M. Shinohara, Y. Shirai and M. Yamanoue (2019). "Comparative metabolomics of Japanese Black cattle beef and other meats using gas chromatography–mass spectrometry." Bioscience, Biotechnology, and Biochemistry **83**(1): 137-147.
Urvieta, R., G. Jones, F. Buscema, R. Bottini and A. Fontana (2021). "Terroir and vintage discrimination of Malbec wines based on phenolic composition across multiple sites in Mendoza, Argentina." Scientific Reports **11**(1): 2863.

Uttl, L., K. Bechynska, M. Ehlers, V. Kadlec, K. Navratilova, Z. Dzuman, C. Fauhl-Hassek and J. Hajslova (2023). "Critical assessment of chemometric models employed for varietal authentication of wine based on UHPLC-HRMS data." Food Control **143**: 109336.
Vaclavik, L., O. Lacina, J. Hajslova and J. Zweigenbaum (2011). "The use of high performance liquid chromatography–quadrupole time-of-flight mass spectrometry coupled to advanced data mining and chemometric tools for discrimination and classification of red wines according to their variety." Analytica Chimica Acta **685**(1): 45-51.
Vaibhav Vikas, S. P. K., Meshram Raviprakash, Bhute Ashish R., Varun Abishek, Durgapal Prashant (2022). "Methanol Poisoning: An Autopsy-Based Study at the Tertiay Care Center of Uttarakhand, India." Cureus.
van Ruth, S. M., W. Huisman and P. A. Luning (2017). "Food fraud vulnerability and its key factors." Trends in Food Science & Technology **67**: 70-75.
Vichi, S., M. Riu-Aumatell, M. Mora-Pons, J. M. Guadayol, S. Buxaderas and E. López-Tamames (2007). "HS-SPME coupled to GC/MS for quality control of Juniperus communis L. berries used for gin aromatization." Food Chemistry **105**(4): 1748-1754.
Vigar, V., S. Myers, C. Oliver, J. Arellano, S. Robinson and C. Leifert (2020). "A Systematic Review of Organic Versus Conventional Food Consumption: Is There a Measurable Benefit on Human Health?" Nutrients **12**(1): 7.
Villanueva-Rodríguez, S. J., B. Rodríguez-Garay, R. Prado-Ramírez and A. Gschaedler (2016). Tequila: Raw Material, Classification, Process, and Quality Parameters. Encyclopedia of Food and Health. B. Caballero, P. M. Finglas and F. Toldrá. Oxford, Academic Press**:** 283-289.
Villar-Pulido, M., B. Gilbert-López, J. F. García-Reyes, N. R. Martos and A. Molina-Díaz (2011). "Multiclass detection and quantitation of antibiotics and veterinary drugs in shrimps by fast liquid chromatography time-of-flight mass spectrometry." Talanta **85**(3): 1419-1427.
Viswanathan, C. T., S. Bansal, B. Booth, A. J. DeStefano, M. J. Rose, J. Sailstad, V. P. Shah, J. P. Skelly, P. G. Swann and R. Weiner (2007). "Quantitative Bioanalytical Methods Validation and Implementation: Best Practices for Chromatographic and Ligand Binding Assays." Pharmaceutical Research **24**(10): 1962-1973.
Vogelgesang, J. and J. Hädrich (1998). "Limits of detection, identification and determination: a statistical approach for practitioners." Accreditation and Quality Assurance **3**(6): 242-255.
Waleckx, E., A. Gschaedler, B. Colonna-Ceccaldi and P. Monsan (2008). "Hydrolysis of fructans from Agave tequilana Weber var. azul during the cooking step in a traditional tequila elaboration process." Food Chemistry **108**(1): 40-48.
Walker, M. and H. Gowland (2015). "Deadly fraud – Food allergen substitution in the food chain." Clinical and Translational Allergy **5**(S3): P137.
Wang, F., Q. Wang, F. Nie, Z. Li, W. Yu and R. Wang (2021). "Unsupervised Linear Discriminant Analysis for Jointly Clustering and Subspace Learning." IEEE Transactions on Knowledge and Data Engineering **33**(3): 1276-1290.
Wilson, M. D. (2008). Support Vector Machines. Encyclopedia of Ecology. S. E. Jørgensen and B. D. Fath. Oxford, Academic Press**:** 3431-3437.
Wiśniewska, P., R. Boqué, E. Borràs, O. Busto, W. Wardencki, J. Namieśnik and T. Dymerski (2017). "Authentication of whisky due to its botanical origin and way of production by instrumental analysis and multivariate classification methods." Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy **173**: 849-853.

Wiśniewska, P., T. Dymerski, W. Wardencki and J. Namieśnik (2015). "Chemical composition analysis and authentication of whisky." Journal of the Science of Food and Agriculture **95**(11): 2159-2166.
Wiśniewska, P., M. Śliwińska, T. Dymerski, W. Wardencki and J. Namieśnik (2016). "Differentiation Between Spirits According to Their Botanical Origin." Food Analytical Methods **9**(4): 1029-1035.
Wiśniewska, P., M. Śliwińska, J. Namieśnik, W. Wardencki and T. Dymerski (2016). "The Verification of the Usefulness of Electronic Nose Based on Ultra-Fast Gas Chromatography and Four Different Chemometric Methods for Rapid Analysis of Spirit Beverages." Journal of Analytical Methods in Chemistry **2016**: 8763436.
Wu, T.-Z. (1999). "A piezoelectric biosensor as an olfactory receptor for odour detection: electronic nose." Biosensors and Bioelectronics **14**(1): 9-18.
Xagoraris, M., P.-K. Revelou, N. Arvanitis, M. Basalekou, C. S. Pappas and P. A. Tarantilis (2021). "The application of right-angle fluorescence spectroscopy as a tool to distinguish five autochthonous commercial Greek white wines." Current Research in Food Science **4**: 815-820.
Yang, K., Á. Somogyi, C. A. Thomas, H. Zhang, Z. Cheng, S. Xu, C. Miller, D. Spivey, C. C. F. Blake, C. Smith, D. Dafoe, N. D. Danielson and M. W. Crowder (2020). "Analysis of Barrel-Aged Kentucky Bourbon Whiskey by Ultrahigh Resolution Mass Spectrometry." Food Analytical Methods **13**: 2301 - 2311.
Yang, P., J. S. Chang, J. W. Wong, K. Zhang, A. J. Krynitsky, M. Bromirski and J. Wang (2015). "Effect of Sample Dilution on Matrix Effects in Pesticide Analysis of Several Matrices by Liquid Chromatography–High-Resolution Mass Spectrometry." Journal of Agricultural and Food Chemistry **63**(21): 5169-5177.
Zaukuu, J.-L. Z., J. Soós, Z. Bodor, J. Felföldi, I. Magyar and Z. Kovacs (2019). "Authentication of Tokaj Wine (Hungaricum) with the Electronic Tongue and Near Infrared Spectroscopy." Journal of Food Science **84**(12): 3437-3444.
Zhang, L., D. Cui, X. Ma, B. Han and L. Han (2023). "Comparative analysis of rice reveals insights into the mechanism of colored rice via widely targeted metabolomics." Food Chemistry **399**: 133926.
Zou, Y., H. Wan, X. Zhang, D. Ha and P. Wang (2015). Electronic Nose and Electronic Tongue. Bioinspired Smell and Taste Sensors. P. Wang, Q. Liu, C. Wu and K. J. Hsia. Dordrecht, Springer Netherlands**:** 19-44.

7. SUPPLEMENTARY MATERIALS

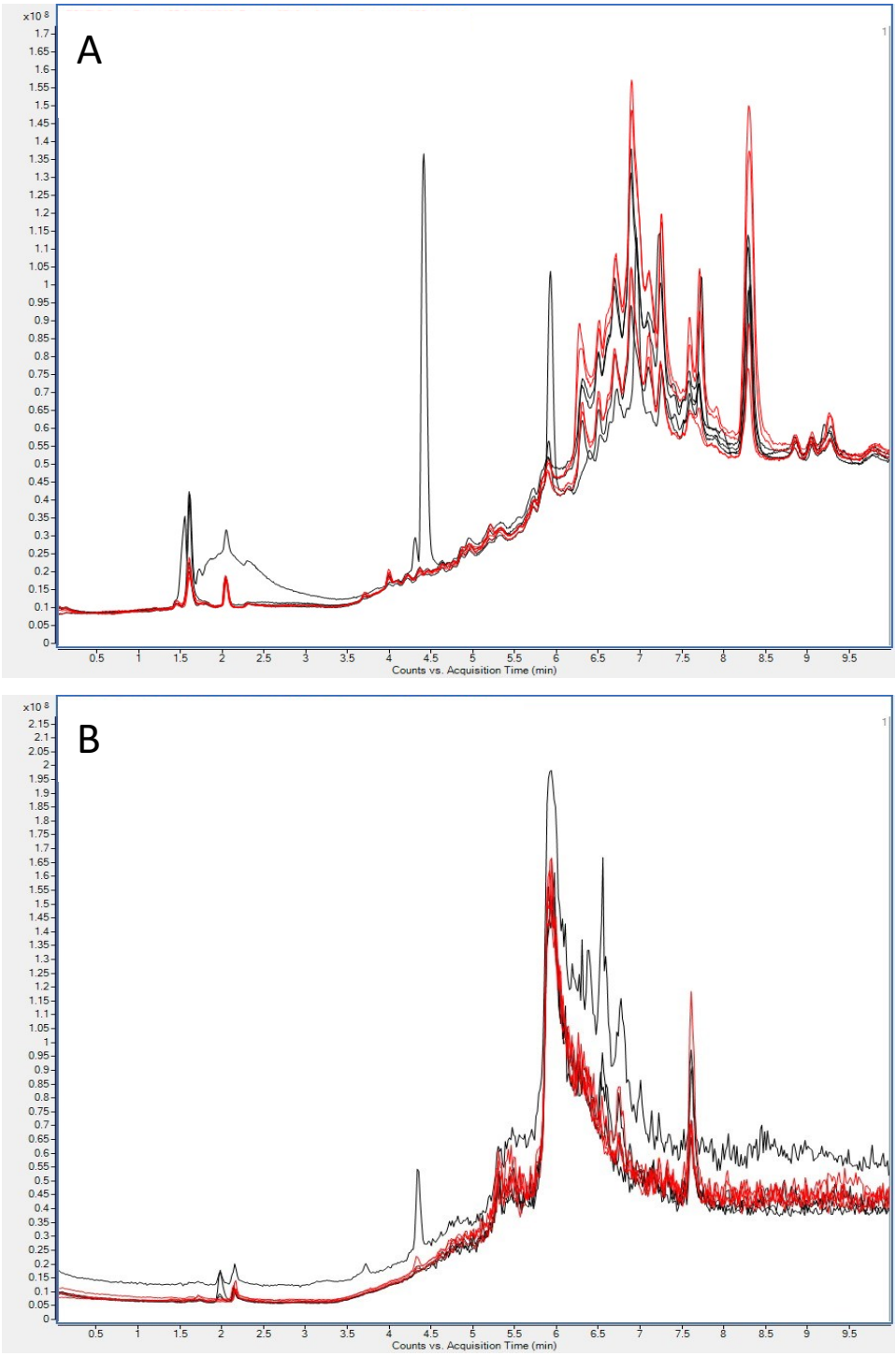


Figure S1.1 TIC of ESI+ (A) and APCI+ (B) for NS (red) vs. gins (black).

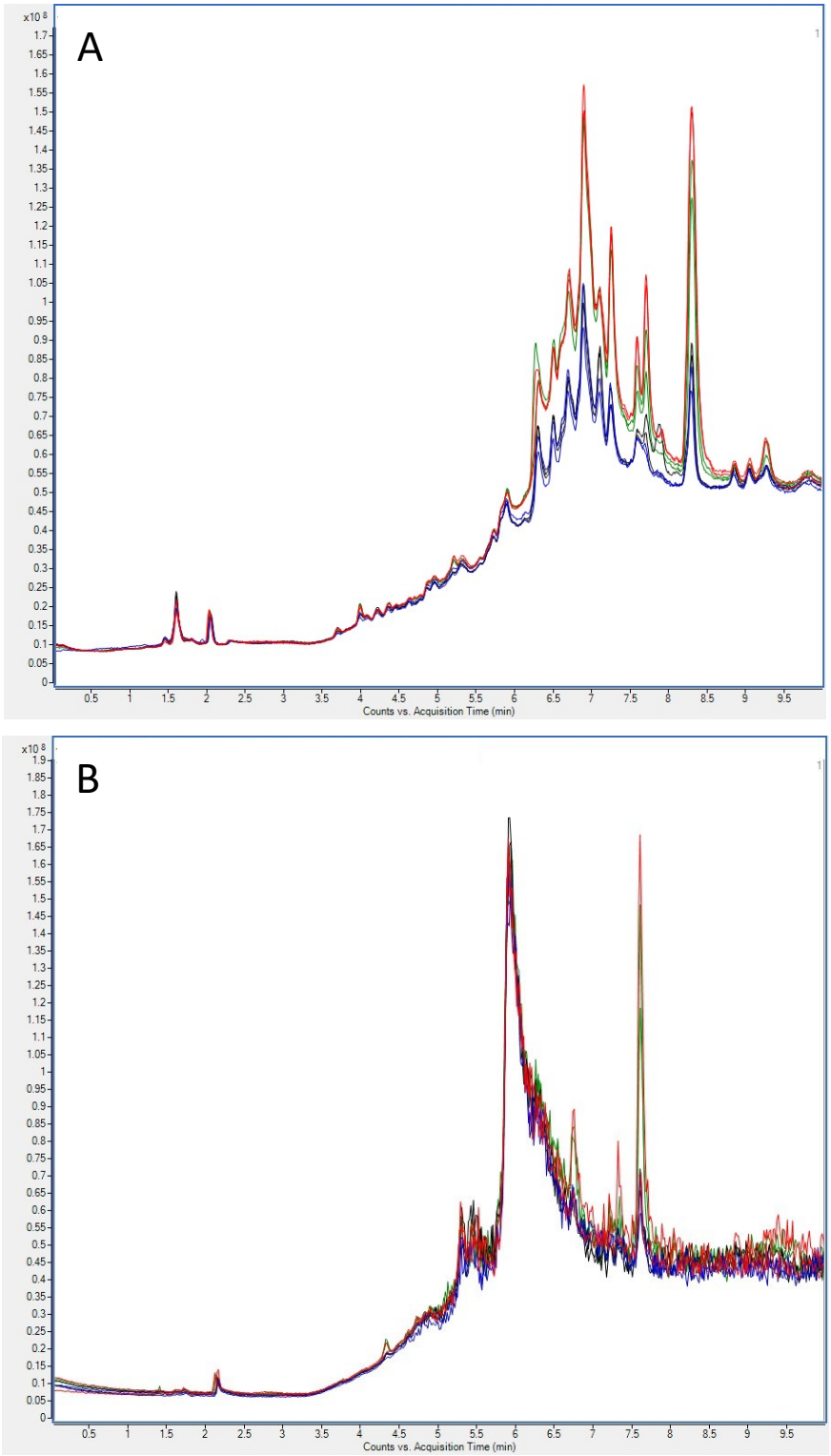


Figure S1.2 TIC of ESI+ (A) and APCI+ (B) for wheat (green), rye (red), corn (Black), and potato (blue).

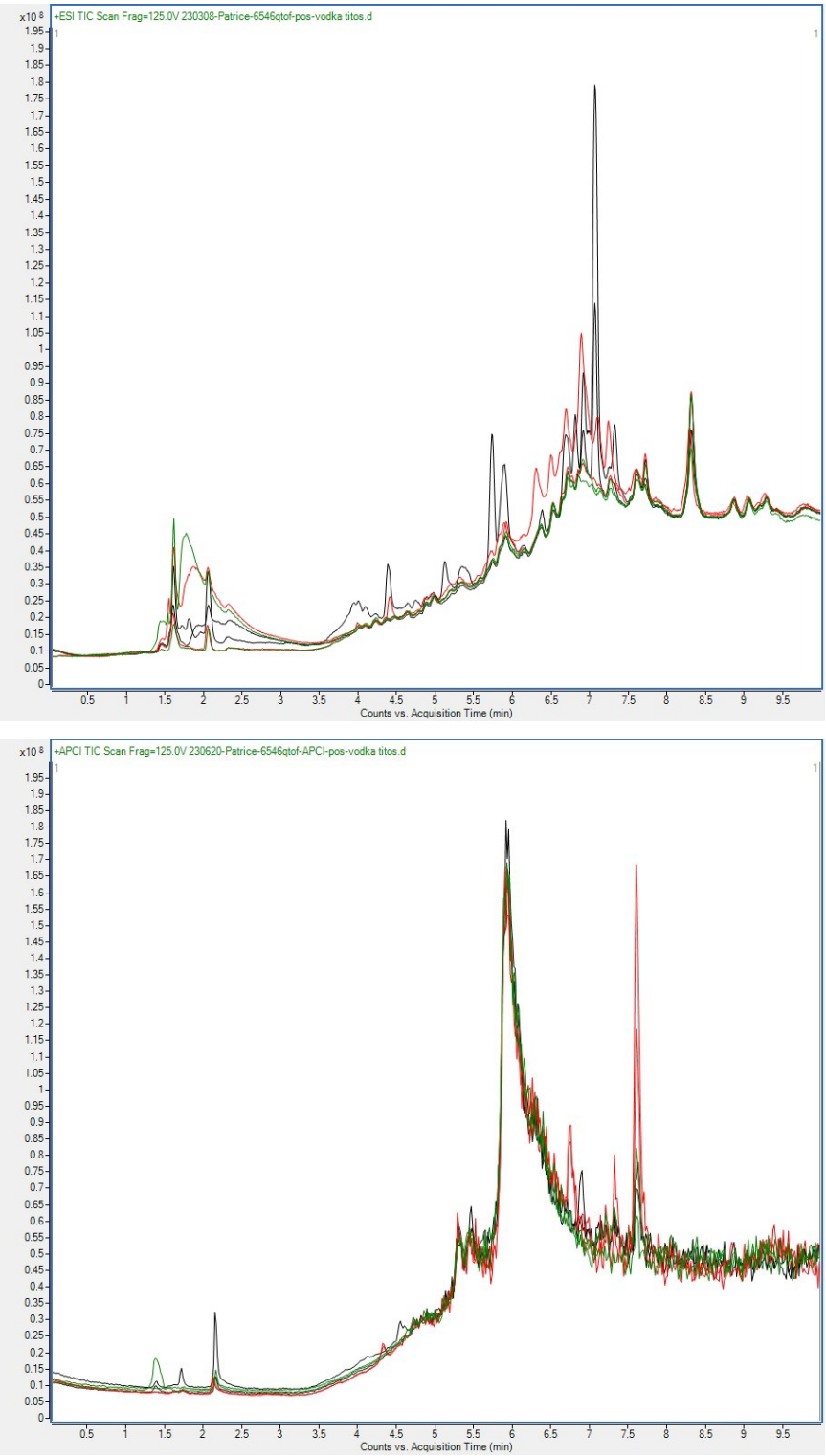


Figure S1.3 TIC of ESI+ (A) and APCI+ (B) for Others (green), QC (red), and ONT (Black).

CATEGORY	BOTANICAL ORIGIN	GEOGRAPHICAL ORIGIN	COUNT
NS	Corn	Quebec	5
NS	Corn	Canada (Excl. Quebec)	3
NS	Corn	USA	2
NS	Mixed (Corn & Malted Barley)	Quebec	4
NS	Wheat	Southern Sweden	1
NS	Wheat	France	1
NS	Wheat	Netherlands	1
NS	Mixed (Wheat & Rye)	Quebec	1
NS	Wheat	Quebec	6
NS	Rye	Quebec	7
NS	Honey	Quebec	1
NS	Potato	Quebec	8
Gin	Corn	Quebec	2
Gin	Mixed (Wheat & Rye)	Quebec	1
Gin	Barley	Quebec	1
Gin	Corn	Canada	4
Gin	Wheat	France	1
Gin	Potato	Quebec	4
Gin	Wheat	Quebec	4
Gin	Rye	Quebec	4

Table S1 Summary of NS and gin samples per botanical and geographical origins.