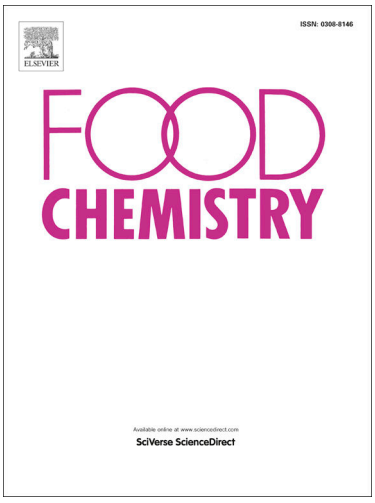
Journal Pre-proofs



Influence of regionality and maturation time on the chemical fingerprint of whisky

Chloé Roullier-Gall, Julie Signoret, Christian Coelho, Daniel Hemmler, Mathieu Kajdan, Marianna Lucio, Bernhard Schäfer, Régis D. Gougeon, Philippe Schmitt-Kopplin

PII: S0308-8146(20)30610-5
DOI: https://doi.org/10.1016/j.foodchem.2020.126748
Reference: FOCH 126748

To appear in: *Food Chemistry*

Received Date: 23 September 2019
Revised Date: 23 March 2020
Accepted Date: 3 April 2020

Please cite this article as: Roullier-Gall, C., Signoret, J., Coelho, C., Hemmler, D., Kajdan, M., Lucio, M., Schäfer, B., Gougeon, R.D., Schmitt-Kopplin, P., Influence of regionality and maturation time on the chemical fingerprint of whisky, *Food Chemistry* (2020), doi: https://doi.org/10.1016/j.foodchem.2020.126748

Influence of regionality and maturation time on the chemical fingerprint of whisky

Chloé Roullier-Gall[1*], Julie Signoret[1], Christian Coelho[1], Daniel Hemmler[2,3], Mathieu Kajdan[1], Marianna Lucio[3], Bernhard Schäfer[4], Régis D. Gougeon[1] and Philippe Schmitt-Kopplin[2,3*].

[1] UMR PAM Université de Bourgogne/AgroSup Dijon, Institut Universitaire de la Vigne et du Vin, Jules Guyot, Dijon, France

[2]Comprehensive Foodomics Platform, Chair of Analytical Food Chemistry, Technische Universität München, Freising, Germany

[3] Research Unit Analytical BioGeoChemistry, Department of Environmental Sciences, Helmholtz Zentrum München, Neuherberg, Germany

[4]Bernhard Schäfer : Independent Spirit consultant, Nuremberg, Germany

*Correspondence:
Chloé Roullier-Gall: chloe.roullier-gall@u-bourgogne.fr
Philippe Schmitt-Kopplin: schmitt-kopplin@helmholtz-muenchen.de

Keywords: Whisky, metabolomics, chemometrics, authentication, geographical origin and maturation time

Abstract

Understanding the chemical composition of whisky and the impact of each step in the manufacturing process provides a basis for responding to the challenges of producing high quality spirits. In this study, the objective was to discriminate whiskies according to their geographical origin and authenticate the maturation time in cask based on the non-volatile profiles. The combination of FT-ICR-MS and chemometrics allowed the distinction of whiskies from four geographical origins in Scotland (Highlands, Lowlands, Speyside and Islay). Statistical modeling was also used to discriminate whiskies according to the maturation time in cask and reveal chemical markers associated with the ageing regardless of the origin or the production process. Interestingly, the flow of transfer of compounds from wood barrels to distillates is not constant and homogeneous over the maturation time. The largest transfer of compounds from the barrel to the whisky was observed around twelve years of maturation

Introduction

Across Scotland there are up to 100 operating distilleries producing whisky distributed in five main regions commonly recognized as Lowlands, Highlands, Speyside, Islands (including whisky from Orkney, Skye, Mull, Jura, Arran and the Campbeltown peninsula) and Islay(Barnard, 2013).

The composition of whisky is complex and mainly impacted by the water and the cereals used, the fermentation, the distillation, the maturation and the blending process(Garcia et al., 2013; John R. Piggott, Sharp, Duncan, & others, 1989; Reid, Swan, & Gutteridge, 1993). Traditionally, whiskies are matured in oak wood barrels for a minimum of three years before consumption, but many whiskies are matured for 12 or more years(JOHN R. Piggott, Conner, Paterson, & Clyne, 1993). As for wine, sensory properties of whiskies are known to improve during the period of maturation in the oak barrel, where the composition changes through a complex array of chemical reactions(MacNamara, Dabrowska, Baden, & Helle, 2011; Macnamara, van Wyk, Brunerie, Augustyn, & Rapp, 2001). Previous studies have attempted to describe the chemical changes during maturation, with particular emphasis on the color, the aroma, the volume, the strength and the flavor compounds(Lee, Paterson, Piggott, & Richardson, 2001a; Liebmann & Scherl, 1949; MacNamara et al., 2011; Macnamara et al., 2001).

Due to its large commercialization and relatively high prices, whisky counterfeiting and adulteration is quite common worldwide(Garcia et al., 2013; Wiśniewska, Dymerski, Wardencki, & Namieśnik, 2015). In case of fraudulent Whisky, maturation time and history of the oak casks can be mislabeled(Stupak, Goodall, Tomaniova, Pulkrabova, & Hajslova, 2018). Authenticity is one of the major concerns for the distillers, dealers and consumers of whiskies around the world(R. I. Aylott & MacKenzie, 2010; Ross I. Aylott, Clyne, Fox, & Walker, 1994; Heller, Vitali, Oliveira, Costa, & Micke, 2011; M. MacKenzie & I. Aylott, 2004; Møller, Catharino, & Eberlin, 2005; Parker, Kelly, Sharman, Dennis, & Howie, 1998; Stupak et al., 2018). Both, maturation period and history of the casks in which maturation occurred are important for the final composition(Kew, Goodall, Clarke, & Uhrín, 2016; Roullier-Gall et al., 2018). Recently, Whisky authentication has been the subject of several studies employing various analytical techniques, such as: GC and LC with different detectors (FID, AED, UV-Vis), electronic nose, atomic absorption spectroscopy and mass spectrometry(Wiśniewska et al., 2015). The aroma and flavor profiles are classically used to

distinguish whiskies by distillery or even to reflect their places of origin(Jack & Steele, 2002; Jackson, 2015; Lee, Paterson, Piggott, & Richardson, 2001b). Gas chromatography coupled to tandem mass spectrometry (GC-Q-ToF) was employed to distinguish malt whiskies according to the type of cask in which they were matured (bourbon versus bourbon and wine)(Stupak et al., 2018). lnfrared spectroscopy with statistical analysis allowed to distinguish Scottish, Irish, and American whiskies and 2 and 3 years old beverages from 6 and 12 years old whiskies(Sujka & Koczoń, 2018). Recently, two different, hypothesis-free, sensor arrays based upon three fluorophores successfully discriminated whisky samples with respect to origin (American, Irish and Scotch Whisky) and taste (rich vs. light)(Han et al., 2017). Mass spectrometry is more and more used for quality control and proof of authenticity of whisky samples. Direct infusion electrospray ionization mass spectrometry (ESI-MS) was used in order to provide a direct, rapid and sensitive method for the characterization of distilleries and authenticity of whisky samples(Møller et al., 2005). Garcia et al., highlighted the use of Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) as an approach to screen for ageing and counterfeiting(Garcia et al., 2013). FT-ICR-MS was used by Kew et al. (2016) for the discrimination of whiskies according to their blending process and type of casks used in maturation(Kew et al., 2016). More recently, non-targeted FT-ICR-MS combined with LC-MS/MS showed the impact of the wood and in particular the history of the barrel on the distillate composition during ageing and revealed the importance of the initial composition of the distillate and the distillery process(Roullier-Gall et al., 2018).
The objective of this work was to go beyond our previous work and implement strategies by using statistical modeling to discriminate whiskies according to their geographical origin and authenticate the maturation time in the barrel based on the non-volatile profiles. Here, we show that the combination of FT-ICR-MS and statistical analysis of whiskies from 3 to 43 years of cask maturation allowed the authentication of the geographical origin and maturation time. The combination of metabolomics and chemometrics was able to reveal chemical markers associated with the ageing, regardless of the origin or the production process.

Materials and Methods

Whisky samples

106 whisky samples from 32 different distilleries in Scotland and from 3 to 43 years of maturation were analyzed (Supplemental table 1). All samples were collected directly from the bottle and stored in 10 mL amber vials at room temperature until analysis.

Direct infusion FT-ICR-MS

Ultrahigh-resolution FT-ICR mass spectra were acquired with a 12 T Bruker SolariX mass spectrometer (Bruker Daltonics, Bremen, Germany) equipped with an APOLLO II electrospray source in negative ionization mode. For MS analysis, whiskies were diluted 5:100 (*v/v*) with methanol (LC-MS grade, Fluka, Germany). The diluted samples were infused into the electrospray ion source with a flow rate of 120 µL h^{-1}. Operating parameters of direct infusion FT-ICR-MS were carried out according to Roullier-Gall *et al*., 2018 previous study. Settings for the ion source were: drying gas temperature 180 °C, drying gas flow 4.0 L min^{-1}, capillary voltage 3,600 V. Spectra were first externally calibrated by ion clusters of arginine (10 mg mL^{-1} in methanol). Internal calibration of each spectrum was conducted with a reference list including selected whisky markers and ubiquitous fatty acids. The spectra were acquired with a time-domain of 4 megawords and 400 scans were accumulated within a mass range of *m/z* 92 to 1000 corresponding to a run time of 18 min. A resolving power of 400,000 at *m/z* 300 was achieved. Quality control (QC) samples were prepared by pooling equal amounts of all samples. QC samples were analyzed at the beginning and after every tenth sample to monitor the reproducibility of the measurements (Roullier-Gall et al., 2018).

Processing of FT-ICR-MS data

Raw spectra were post-processed by Compass DataAnalysis 4.2 (Bruker Daltonics, Bremen, Germany) and peaks with a signal-to-noise ratio (S/N) of at least 6 were exported to mass lists. All exported *m/z* features were aligned into a matrix containing averaged *m/z* values (peak alignment window width: ±1 ppm) and corresponding peak intensities of all analyzed samples. Molecular formulae were assigned to the exact *m/z* values by mass difference network analysis using an in-house developed software tool. In total, 5979 detected features could be assigned to distinct and unique molecular formulae. More than 90% of all assignments were found within an error range lower than 0.2 ppm. All further calculations and filtering were done in Perseus 1.5.1.6 (Max Planck Institute of Biochemistry, Germany) and R Statistical Language (version 3.1.1).

Repeatability of FT-ICR-MS whisky measurements

The repeatability of the measurements was evaluated on QCs samples injected in triplicate at the beginning and at the end of the sequence analysis. The repeatability over time of FT-ICR-MS analyses during the entire runtime was also evaluated , by monitoring the precision of the recorded peak intensities from ten analyzed QC samples and manual comparison of the raw mass spectra (Roullier-Gall et al., 2018). In the QC samples, 92.8% of all detected monoisotopic signals showed a relative standard deviation lower than 20% (Roullier-Gall et al., 2018).

Excitation Emission Matrix Fluorescence

Whisky fluorescence was acquired on a spectrofluorometer (Horiba Aqualog). Excitation ranged from 225 to 600 nm and emission was recorded from 230 to 600 nm (C. Coelho et al., 2015). The spectrofluorometer used a conventional right-angle optical setup. Excitation Emission Matrices (EEM) were corrected daily for their Rayleigh and Raman scattering, as well as inner filtering effects by using the functions provided within the FluorEssence software of the Horiba Aqualog system. The fluorescence intensity was normalized to a Starna 1 ppm quinine sulfate reference cell, measured every day.

UPLC targeted analysis of polyphenols

An Acquity Waters ultra performance liquid chromatography (UPLC) with a diode array fluorescence detector (DAD) was used to separate and quantify individual polyphenolic compounds in whisky samples. The column was a BEH C18, 1.7 µm, 2.1 mm × 150 mm. Column temperature was kept constant at 30 °C, and samples were held at 8 °C. An elution was applied, starting isocratic from 100% A (ultrapure water, 0.1% formic acid) from 0 to 6 min and then the gradient was increased linearly over 56 min to 100% B (methanol, 0,1% formic acid), where it was held again until reaching 60 min, with a rate of 0.3 $mL \cdot min^{-1}$. Injection volume was 5 µL. Detection limits (LODs), quantification limits (LOQs) and intra- and inter-day variation of the assay were determined by injecting a series of dilute solutions with known concentrations.

Total phenolic content (TPC) of whiskies

Total phenolic content (TPC) in whisky was determined using the Folin-Ciocalteu assay: 0.1 mL of undiluted whisky, 0.5 mL of Folin-Ciocalteu reagent, 2 mL of a 20 (*w/v*) sodium carbonate solution and 7.4 mL of ultrapure water were mixed and reacted for 30 min at room

temperature. Then, the absorbance at 750 nm was measured and calculated as $mg.L^{-1}$ of gallic acid equivalent using a calibration curve between 0 and 500 $mg.L^{-1}$ of gallic acid. All samples were analyzed in triplicates.

Data analysis

The statistical analysis was performed with Perseus, Simca 12.0 (Umetrics, Umeå, Sweden) and R (Version 1.0.136).

Van Krevelen and elemental diagrams were used to visualize FT-ICR MS data (Gonsior et al., 2009; Gougeon et al., 2011; Hertkorn et al., 2008). Multivariate statistical methods (. Principal component analysis and hierarchical cluster analysis) were used to explore similarities and hidden patterns among samples(Barker & Rayens, 2003; Granato, Santos, Escher, Ferreira, & Maggio, 2018). As within-groups variability dominates the among-groups variability for the geographical origin and maturation time impact, partial least square discriminant analysis were used(Barker & Rayens, 2003). We set up diverse classification models in order to reveal the different geographical origin and maturation time for barrel specific metabolites. In order to improve the efficiency of such classification and the possible presence of overfitting and noise, we preprocessed the entire dataset applying the ReliefF algorithm(Witten, Frank, Hall, & Pal, 2016). The algorithm identified a subset of variables that was used for the next classification models maximizing the model performance in term of accuracy. The features' selection was based on the highest rank value attributed to each variable by the algorithm. Consequently, with the reduced datasets we built several partial least square discriminant analysis (PLS-DA) models. The goodness of the fit was evaluated by the coefficient of determination (R^2X), the proportion of the variance of the response that is explained by the model (R^2Y), the predictive ability (Q^2Y) and validation was carried out using permutation tests. For the ageing time model, whisky samples were randomly divided into two groups: 101 samples selected to build the statistical model and the 5 remaining samples employed as external validation. A Bland–Altman plot was used to analyze the agreement between the known ageing time and the model prediction.

Results and discussion

Geographical origin impact of Scotch whisky

A visual comparison of FT-ICR-MS spectra of Scotch whiskies from Highland, Lowland, Speyside and Islay distilleries and from 3 to 43 years of cask maturation is presented in Figure 1. Several thousands of signals were found between *m/z* 100 to 600. The enlargements at *m/z* 261 exemplarily show the differences in chemical composition between the four geographical areas (Figure 1). Between *m/z* 261.0 and 261.2, whisky spectra only differed in the intensities of the eight detected peaks. For example, signals at *m/z* 261.04044 and at *m/z* 261.17075 correspond to the molecularformulas $[C_{13}H_9O_6]^-$ and $[C_{13}H_{25}O_5]^-$, respectively, and show comparable intensities in all four whisky samples (Figure 1). The Venn and van Krevelen diagrams (Figure 1) show that 2200 formulas (out of 5979) were found in all samples of the four groups of whisky samples confirming a great consistency in the chemical composition regardless of the origin(Kew et al., 2016).

Despite the high similarities between whiskies (36.8% of formulas were found in all samples of the four groups of whisky samples) as shown in the Venn diagrams and the direct comparison of spectra (Figure 1), 533 (8%), 610 (10%), 401 (6.7%) and 464 (7.7%) formulas were uniquely found in at least 50% of Highland, Islay, Speyside and Lowland whisky samples, respectively. In addition to the unique formulas, differences in signal intensities between whiskies can be found which could result from particular distilleries, maturation times, geographical features or environmental impacts. As example, the signal at *m/z* 261.06158 corresponding to the molecular formula $[C_{10}H_{13}O_8]^-$ is less intense in the whisky from Speyside ($2 \cdot 10^7$) compared to the other three whiskies ($5\text{-}6 \cdot 10^7$) (Figure 1).

The impact of the type of whisky (blend or malt), the maturation wood type and the wood history on the Whisky composition has recently been studied(Kew et al., 2016; Roullier-Gall et al., 2018). However, barley, peat and water origin, malting, brewing, fermentation and distillation process, nature of the cellars, climates as well as environmental factors could also impact the complex composition of whisky. Figure 2 shows the first two principal components obtained from principal component analysis computed on the 106 analyzed whisky samples from Scotland. Whiskies showed a rough separation by their maturation time in the barrel on principal component 2. By comparison, the impact of the geographical origin - Highlands, Lowlands, Speyside and Islay - on the whisky metabolome appears less significant in the PCA compared to the maturation time (Figure 2). The samples from the Highlands and Lowlands appeared in a more discrete cluster than the Speyside and Islay samples, which are more distributed all around the plot. Interestingly, the four samples in the

top left corner of the PCA were from the same distillery (but with different maturation times) and the eight samples (in red) in the top right corner were from different distilleries but all in Islay and all with six years of barrel maturation (Figure 2). The great diversity in the composition of the whisky samples demonstrates the large number of environmental parameters (including production, geographical origin, maturation time…), which eventually impacts the final product.

A group separation is not easily observable in PCA models because of the high number of sample characteristics (maturation from 6 to 43 years, 4 geographical origins and 32 manufacturers) and the high number of variables considered (5979 molecular formulas). A PLS regression model was used to better distinguish the geographical origins of the whisky samples. Figure 3 shows results from a PLS regression computed from FT-ICR-MS data of 106 whisky samples from the four areas in Scotland, 32 different distilleries and from 3 to 43 years of maturation. The first two components allowed to separate whiskies into four distinct groups according to the geographical origin (Highlands, Lowlands, Speyside and Islay) and independent of the distillery or the maturation time (Figure 3). Whiskies from Lowlands and Islay distilleries appear to be chemically more different from the samples from Highlands and Speyside distilleries (separation on the 1st component). The second component allowed to distinguish whiskies from the Highlands and Lowlands from those produced in Speyside and Islay distilleries. Based on the PLS, specific masses for each location were extracted (VIP masses) and projected into van Krevelen diagrams (Figure 3). For example, samples from Speyside clearly showed an enrichment of polyphenols or polyphenol-type compounds (Figure 3). By comparison, Highland samples seem to be characterized by higher alcohols with a profile reminiscent to the distillate(Roullier-Gall et al., 2018). Lowland and Islay appear to be characterized by carbohydrates. The schematic classification of single malts according to their origin could be based, in part, on a geological reality, of which water would be the main factor. Meir-Augenstein showed that the source water used for and during whisky production correlates with the authenticity of whisky(Meier-Augenstein, Kemp, & Hardie, 2012). Water is involved in several stages of processing: added to ground barley to produce the mash, when diluting the distilled alcohol to cask strength for maturation and when reducing cask strength whisky for bottling. Meir-Augenstein's study suggests that whisky keeps a signature of the geographic provenance of the used water(Meier-Augenstein et al., 2012). However, from all elements necessary for whisky production, water is the one

whose impact on the final composition is probably the most difficult to evaluate(Witten et al., 2016).

Another source of the variability in the chemical composition is the distillery brand. Compositional specificities from manufacturers can contribute to the geographical region model. Studying samples from Islay and from three distilleries including various maturation times highlights the ability of FT-ICR-MS to differentiate whiskies according to the distilleries and independently of the maturation time. Non-supervised statistics allowed to visualize the impact of the process, composition and factory on the final whisky independently of the maturation (supplementary Figure 1). According to the hierarchical cluster analysis (supplementary Figure 1), whiskies from the manufacturers W2 and W1 manufacturers seem close and distinct from W3 although W2 and W3 are geographically very close to one another (less than 3 km). According to the van Krevelen diagrams in supplementary Figure S1, the whisky from distillery W1 (including whiskies from 12 to 18 years in barrel) seems to be characterized by a higher concentration in polyphenolic compounds whereas whiskies from distillery W2 (12 to 21 years in barrel) are mainly characterized by carbohydrate-type compounds as well as sulfur and nitrogen containing compounds. The whisky from distillery W3 (6 to 12 years in barrel) seems to be characterized by higher alcohols and shows a profile reminiscent to distillate (Roullier-Gall et al., 2018).

The samples differing in brand and geographical origin can be properly grouped and characterized by FT-ICR-MS and chemometric analysis. These combined techniques enable the fast evaluation and authentication of Scotch whiskies independently of the maturation time or the process used. They allowed an excellent determination of whisky composition as well as distinguishing between different brands of whiskies produced in areas only 20 km apart from one another.

Impact of maturation time on the chemical composition of Scotch whiskies

Based on the huge impact of maturation in wood barrels on the chemical diversity of whisky, as shown in our previous study(Roullier-Gall et al., 2018), we here focused on the impact of the maturation time on the final composition. We first studied three samples from one distillery (W3) in Islay after 6, 8 and 10 years of cask maturation (Supplementary Figure 2). The samples revealed 376 mass peaks whose intensity significantly increased with maturation time, and 81 mass peaks with decreasing intensity (Supplementary Figure 2a-b). The higher number of features which showed an increase in peak intensity confirms that chemicals transfer from wood to the distillate as a function of ageing.

In order to extract age-related metabolite signatures, a PLS discriminant analysis was designed using 7 groups with different maturation times: whiskies that had been aged in casks for 6 years, 8 years, 10 years, 12 years, 14 years, 16 years and 18 years (Supplementary Table 3). We selected 76 from the 106 analyzed whisky samples which ranged from 6 to 18 years. A PLS-DA model was built based on 71 of these whisky samples (Figure 4c). Five of the 76 samples were used for model validation (Figure 4e). The separation of whiskies according to the ageing time (Figure 4c) showed a good predictive power with Q^2 of 0.49 and R^2Y of 0.77. The scatter plot of the model group versus the predicted group for each sample confirmed the robustness of the prediction model (Figure 4 D), with a prediction average close to the maturation time and an excellent standard deviation (between 0.69 to 1.57 years, Supplementary Table 3). However, it was not possible to really differentiate groups of 12 and 14 years of maturation. These samples were predicted as 12.3 years for both, the group of 12 years and 14 years old whiskies (maximum standard deviation was 1 year).
Finally, the five whisky samples, which were not used for the model construction (maturation of 6, 10, 12, 12 and 16 years in barrel) were used to validate the model. The good predictions, achieved for these five whisky samples, confirmed the predictive ability of the maturation time model (Figure 4 C and E). After model validation, 256 discriminating and characteristic masses (VIP variables higher than 1), whose peak intensities increased with maturation, were extracted (Figure 4 F and H). As example, compounds from wine such as syringic acid and caftaric acid and from wood such as lyoniresinol, patuletin and digallic acid were found as increasing with maturation time (Figure 4 and supplementary Figure 2). In contrast, 213 masses with decreasing intensity led to 54 annotations from databases including stearyl acetate and syringaledehyde (Figure 4).

ANOVA tests were applied to extract subsets of features that discriminate between whiskies with a maturation time of 6 and 8 years (supplementary figure 4A), 8 and 10 years (supplementary figure 4B), 10 and 12 years (supplementary figure 4C), 12 and 14 years (supplementary figure 4D), 14 and 16 years (supplementary figure 4E), and 16 and 18 years (supplementary figure 4F), respectively. Statistically extracted masses can be used to highlight the impact of maturation time on the whisky composition between the different ageing groups. According to the statistics, 27 molecular formulas were significantly enriched in the youngest whiskies (6 years of maturation) while 36 formulas were specific to the 66 whiskies older than 6 years of maturation. In the same way, 74 molecular formulas were statistically identified as markers to differentiate between whiskies younger than 8 years of maturation (6 and 8 years of maturation) and 132 formulas were identified as markers for whiskies older than 8 years of maturation. The number of significant masses statistically extracted increased until the distinction of whiskies between 10 (121 molecular formulas) and 12 (1182 molecular formulas) years of casks maturation. Then the number of significant masses starts to decrease.

Interestingly, the higher differences were found between whiskies after 10 and 12 years of maturation directly followed by whiskies after 14 and 16 years of maturation (supplementary Figure 4). The similarities between samples characterized by 12 and 14 years of maturation, which probably can be explained by the low ability of the model prediction to distinguish those two groups (Figure 4). The difference between young (less than 12 years of maturation, group 1) and old whiskies (more than 12 years of maturation, group 2) was further studied by Excitation Emission Matrix (EEM) fluorescence spectroscopy and analysis of polyphenols (supplementary figure 5 and 6). As illustrated in Supplementary figure 6A, PARAFAC results enabled to easily discriminate whiskies from both groups on the first component. Whiskies of group 1 (younger than 12 years of maturation) were driven by wood barrel ageing processes, conferring a relative homogeneity to this group. Whiskies from the group 2 (older than 12 years of maturation) showed a higher within-group-variation (confirming results from FT-ICR-MS), which may be because of different cask finishing practices. This may include re-casking for a secondary maturation into other wood casks that were used for ageing of wines, spirits or even beers (Scotch Whisky Association, n.d.). Such practice tends to diversify Scottish whisky profiles by masking the original distillate. The fluorescence positions of the PARAFAC components suggest differences in polyphenolic compounds (Figure 6B) which could be confirmed in a TPC assay and quantification of syringic acid,

gallic acid and scopoletin by UPLC (Figure 6C). Scopoletin was also identified as wood marker increasing in intensity with maturation time. No significant differentiations were found for tyrosol between our two groups, surely due to a moderated retention of this compound to wood surfaces (Christian Coelho et al., 2019).

Out of the 1182 molecular formulas found significantly increased for whiskies with at least 12 years of maturation, only 222 annotations were found in databases: 167 were found in our home-built plant database, 154 in Metlin, 130 in KEGG and 25 in HMDB (Supplementary Figure 7 and supplementary Table 4). Interestingly, 467 of the 1182 markers could be previously identified as wood specific markers (Roullier-Gall et al., 2018) (Supplementary Figure 8). Moreover, 675 of the 1182 features could be detected in wine samples including 337 compounds already identified as wood markers (Supplemental Figure 8). The higher presence of wine-based compounds in whiskies older than 12 years, could correspond to finishing practices that could take place after 10 years of ageing independently of the region of production.

Conclusion

In conclusion, FT-ICR-MS based metabolomics was used as a powerful tool to discriminate whiskies according to maturation time and geographical origin, regardless of the manufacturer, the barrel used and the production process. Untargeted metabolomics allowed to study a broad range of metabolites and the extraction of markers that were significantly different depending on the geographical origin and/or the maturation time. A prediction model was further developed for the comprehensive evaluation of maturation time authenticity. Interestingly, the flow of transfer of compounds from wood barrels to distillates is not constant and homogeneous over the maturation time. The largest transfer of compounds from the barrel to the whisky was observed between ten and twelve years of maturation with up to 1182 molecular formulas that showed a significant increase in their peak intensities as a function of the maturation time. 467 of these were identified as wood-specific compounds and 338 are typically also found in wine. The explanatory hypothesis for the higher presence of wine-related compounds in whiskies older than 12 years, compared to whiskies younger than 12 years could be the finishing practice that takes place after 12 years of ageing independently of the region of production.

Figure 1: Visualization of ESI(-) FT-ICR-MS spectra of four samples from Highlands, Lowlands, Speyside and Islay A) in the mass range *m/z* 100 - 600. B) The enlargement at the nominal mass *m/z* 261 shows eight ions detected and assigned to unique molecular formulas. C) The Venn diagram shows the number of detected features which are common and specific to all samples analyzed for the four regions. D) Van Krevelen diagram (H/C vs. O/C) of the common composition found in all analyzed whisky samples. Points in van Krevelen diagrams are colored according to their elemental composition: CHO in blue, CHOS in green, CHON in orange, and CHONS in red. Scaling is according to the mean peak intensity.

Figure 2: Principal component score plots of 106 whisky samples from Scotland, analyzed by FT-ICR-MS and colored according to A) the geographical origin and B) the maturation time (from 6 up to 18 years). The first two components explained 38.4% of the total variability.

Figure 3: Map of the four whisky regions in Scotland (Highland, Speyside, Lowland and Islay). PLS regression model based on 106 whisky samples and van Krevelen diagrams showing the extracted "VIP" features for each region, respectively.

Figure 4: A) Visualization of ESI (-) FT-ICR-MS spectra (*m/z* 100 to 600) of four samples matured in barrel for 6, 10, 14 and 18 years. B) Enlargement of the nominal mass at *m/z* 261 showing eight masses detected and assigned to unique molecular formulas. C) PLS model colored according to maturation time, based on 71 whisky samples ($R^2Y(cum)=0.57$ and $Q^2(cum)=0.53$). Comparison of prediction results versus the maturation time for D) the samples used to build the PLS model and for E) the 5 predicted samples. Van Krevelen diagram (H/C vs. O/C) of the significant features, which contribute to the discrimination of maturation time including F) features which showed an increase in peak intensity and G) features with decreasing peak intensity as a function of the maturation time, respectively. H) Bar chart of the absolute intensity of some biomarkers that showed a significant change in peak intensities. Points in the van Krevelen diagrams are colored according to their elemental composition, CHO in blue, CHOS in green, CHON in orange, and CHONS in red. Scaling is according to the mean peak intensity.

Literature

Aylott, R. I., & MacKenzie, W. M. (2010). Analytical strategies to confirm the generic authenticity of Scotch whisky. *Journal of the Institute of Brewing*, *116*(3), 215–229.

Aylott, Ross I., Clyne, A. H., Fox, A. P., & Walker, D. A. (1994). Analytical strategies to confirm Scotch whisky authenticity. *The Analyst*, *119*(8), 1741. https://doi.org/10.1039/an9941901741

Barker, M., & Rayens, W. (2003). Partial least squares for discrimination. *Journal of Chemometrics*, *17*(3), 166–173. https://doi.org/10.1002/cem.785

Barnard, A. (2013). *The Whisky Distilleries of the United Kingdom*. Aaron Barker Publishing.

Coelho, C., Aron, A., Roullier-Gall, C., Gonsior, M., Schmitt-Kopplin, P., & Gougeon, R. D. (2015). Fluorescence Fingerprinting of Bottled White Wines Can Reveal Memories Related to Sulfur Dioxide Treatments of the Must. *Analytical Chemistry*.

Coelho, Christian, Gougeon, R. D., Perepelkine, L., Alexandre, H., Guzzo, J., & Weidmann, S. (2019). Chemical Transfers Occurring Through Oenococcus oeni Biofilm in Different Enological Conditions. *Frontiers in Nutrition*, *6*. https://doi.org/10.3389/fnut.2019.00095

Garcia, J. S., Vaz, B. G., Corilo, Y. E., Ramires, C. F., Saraiva, S. A., Sanvido, G. B., … others. (2013). Whisky analysis by electrospray ionization-Fourier transform mass spectrometry. *Food Research International*, *51*(1), 98–106.

Gonsior, M., Peake, B. M., Cooper, W. T., Podgorski, D., D'Andrilli, J., & Cooper, W. J. (2009). Photochemically Induced Changes in Dissolved Organic Matter Identified by Ultrahigh Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Environmental Science & Technology*, *43*(3), 698–703. https://doi.org/10.1021/es8022804

Gougeon, R. D., Lucio, M., Boutegrabet, L., Peyron, D., Feuillat, F., Chassagne, D., … Schmitt-Kopplin, P. (2011). Authentication Approach of the Chemodiversity of Grape and Wine by FTICR-MS. In S. E. Ebeler, G. R. Takeoka, & P. Winterhalter (Eds.), *Progress in Authentication of Food and Wine* (Vol. 1081, pp. 69–88). Washington, DC: American Chemical Society.

Granato, D., Santos, J. S., Escher, G. B., Ferreira, B. L., & Maggio, R. M. (2018). Use of principal component analysis (PCA) and hierarchical cluster analysis (HCA) for multivariate association between bioactive compounds and functional properties in foods: A critical perspective. *Trends in Food Science & Technology*, *72*, 83–90. https://doi.org/10.1016/j.tifs.2017.12.006

Han, J., Ma, C., Wang, B., Bender, M., Bojanowski, M., Hergert, M., … Bunz, U. H. F. (2017). A Hypothesis-Free Sensor Array Discriminates Whiskies for Brand, Age, and Taste. *Chem*, *2*(6), 817–824. https://doi.org/10.1016/j.chempr.2017.04.008

Heller, M., Vitali, L., Oliveira, M. A. L., Costa, A. C. O., & Micke, G. A. (2011). A rapid sample screening method for authenticity control of whiskey using capillary electrophoresis with online preconcentration. *Journal of Agricultural and Food Chemistry*, *59*(13), 6882–6888.

Hertkorn, N., Frommberger, M., Witt, M., Koch, B. P., Schmitt-Kopplin, Ph., & Perdue, E. M. (2008). Natural Organic Matter and the Event Horizon of Mass Spectrometry. *Analytical Chemistry*, *80*(23), 8908–8919. https://doi.org/10.1021/ac800464g

Jack, F. R., & Steele, G. M. (2002). Modelling the sensory characteristics of Scotch whisky using neural networks—a novel tool for generic protection. *Food Quality and Preference*, *13*(3), 163–172. https://doi.org/10.1016/S0950-3293(02)00012-5

Jackson, M. (2015). *Malt Whisky Companion*. Dorling Kindersley Ltd.

Kew, W., Goodall, I., Clarke, D., & Uhrín, D. (2016). Chemical Diversity and Complexity of Scotch Whisky as Revealed by High-Resolution Mass Spectrometry. *Journal of The American Society for Mass Spectrometry*, 1–14.

Lee, K.-Y. M., Paterson, A., Piggott, J. R., & Richardson, G. D. (2001a). Origins of flavour in whiskies and a revised flavour wheel: a review. *Journal of the Institute of Brewing*, *107*(5), 287–313.

Lee, K.-Y. M., Paterson, A., Piggott, J. R., & Richardson, G. D. (2001b). Sensory discrimination of blended Scotch whiskies of different product categories. *Food Quality and Preference*, *12*(2), 109–117.

Liebmann, A. J., & Scherl, B. (1949). Changes in whisky while maturing. *Industrial & Engineering Chemistry*, *41*(3), 534–543.

MacNamara, K., Dabrowska, D., Baden, M., & Helle, N. (2011). *Advances in the Ageing Chemistry of Distilled Spirits Matured in Oak Barrels*.

Macnamara, K., van Wyk, C. J., Brunerie, P., Augustyn, O. P. H., & Rapp, A. (2001). Flavour components of whiskey. III. Ageing changes in the low-volatility fraction. *S. Afr. J. Enol. Vitic*, *22*(2), 82.

Meier-Augenstein, W., Kemp, H. F., & Hardie, S. M. L. (2012). Detection of counterfeit scotch whisky by 2H and 18O stable isotope analysis. *Food Chemistry*, *133*(3), 1070–1074. https://doi.org/10.1016/j.foodchem.2012.01.084

M. MacKenzie, W., & I. Aylott, R. (2004). Analytical strategies to confirm Scotch whisky authenticity. Part II: Mobile brand authentication. *Analyst*, *129*(7), 607–612. https://doi.org/10.1039/B403068K

Møller, J. K., Catharino, R. R., & Eberlin, M. N. (2005). Electrospray ionization mass spectrometry fingerprinting of whisky: immediate proof of origin and authenticity. *Analyst*, *130*(6), 890–897.

Parker, I. G., Kelly, S. D., Sharman, M., Dennis, M. J., & Howie, D. (1998). Investigation into the use of carbon isotope ratios (13C/12C) of Scotch whisky congeners to establish brand authenticity using gas chromatography-combustion-isotope ratio mass spectrometry. *Food Chemistry*, *63*(3), 423–428.

Piggott, JOHN R., Conner, J. M., Paterson, A., & Clyne, J. (1993). Effects on Scotch whisky composition and flavour of maturation in oak casks with varying histories. *International Journal of Food Science & Technology*, *28*(3), 303–318.

Piggott, John R., Sharp, R., Duncan, R. E. B., & others. (1989). *Science and technology of whiskies*. Retrieved from http://agris.fao.org/agris-search/search.do?recordID=US201300643960

Reid, K. J. G., Swan, J. S., & Gutteridge, C. S. (1993). Assessment of Scotch whisky quality by pyrolysis—mass spectrometry and the subsequent correlation of quality with the oak wood cask. *Journal of Analytical and Applied Pyrolysis*, *25*, 49–62.

Roullier-Gall, C., Signoret, J., Hemmler, D., Witting, M. A., Kanawati, B., Schäfer, B., … Schmitt-Kopplin, P. (2018). Usage of FT-ICR-MS Metabolomics for Characterizing the Chemical Signatures of Barrel-Aged Whisky. *Frontiers in Chemistry*, *6*. https://doi.org/10.3389/fchem.2018.00029

Scotch Whisky Association. (n.d.). Q&A: Scotch Whisky Technical File.

Stupak, M., Goodall, I., Tomaniova, M., Pulkrabova, J., & Hajslova, J. (2018). A novel approach to assess the quality and authenticity of Scotch Whisky based on gas chromatography coupled to high resolution mass spectrometry. *Analytica Chimica Acta*, *1042*, 60–70. https://doi.org/10.1016/j.aca.2018.09.017

Sujka, K., & Koczoń, P. (2018). The application of FT-IR spectroscopy in discrimination of differently originated and aged whisky. *European Food Research and Technology*, *244*(11), 2019–2025. https://doi.org/10.1007/s00217-018-3113-5

Wiśniewska, P., Dymerski, T., Wardencki, W., & Namieśnik, J. (2015). Chemical composition analysis and authentication of whisky. *Journal of the Science of Food and Agriculture*, *95*(11), 2159–2166.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann.

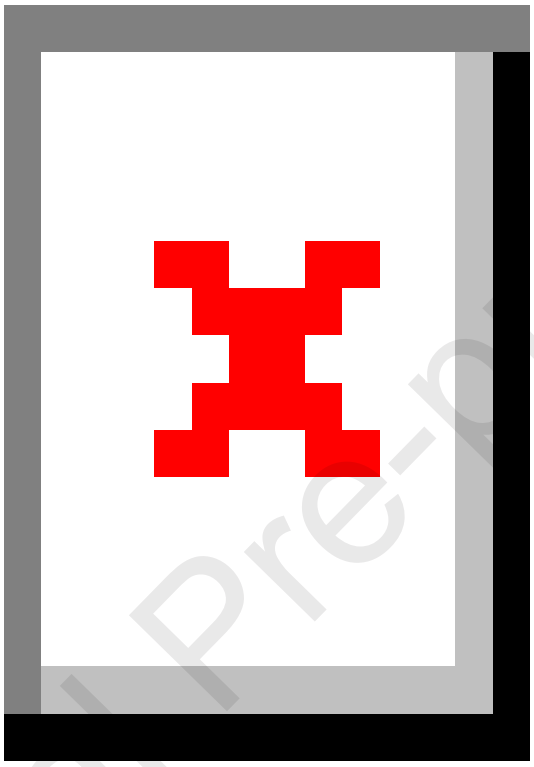


Figure 1

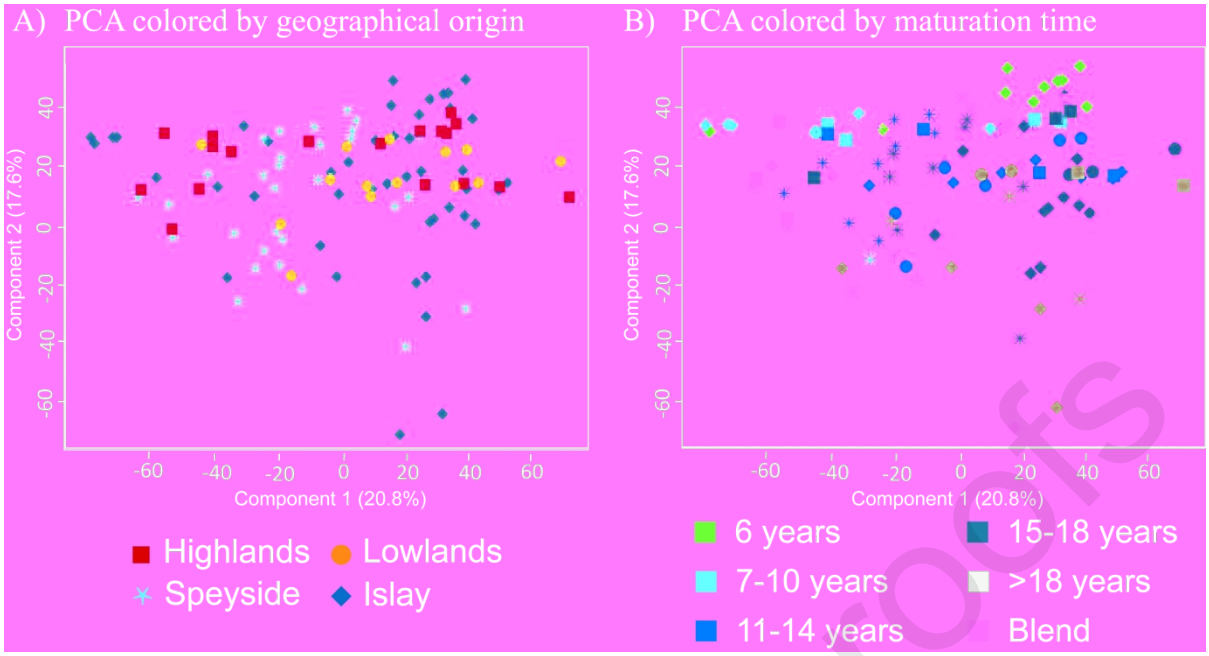


Figure 2

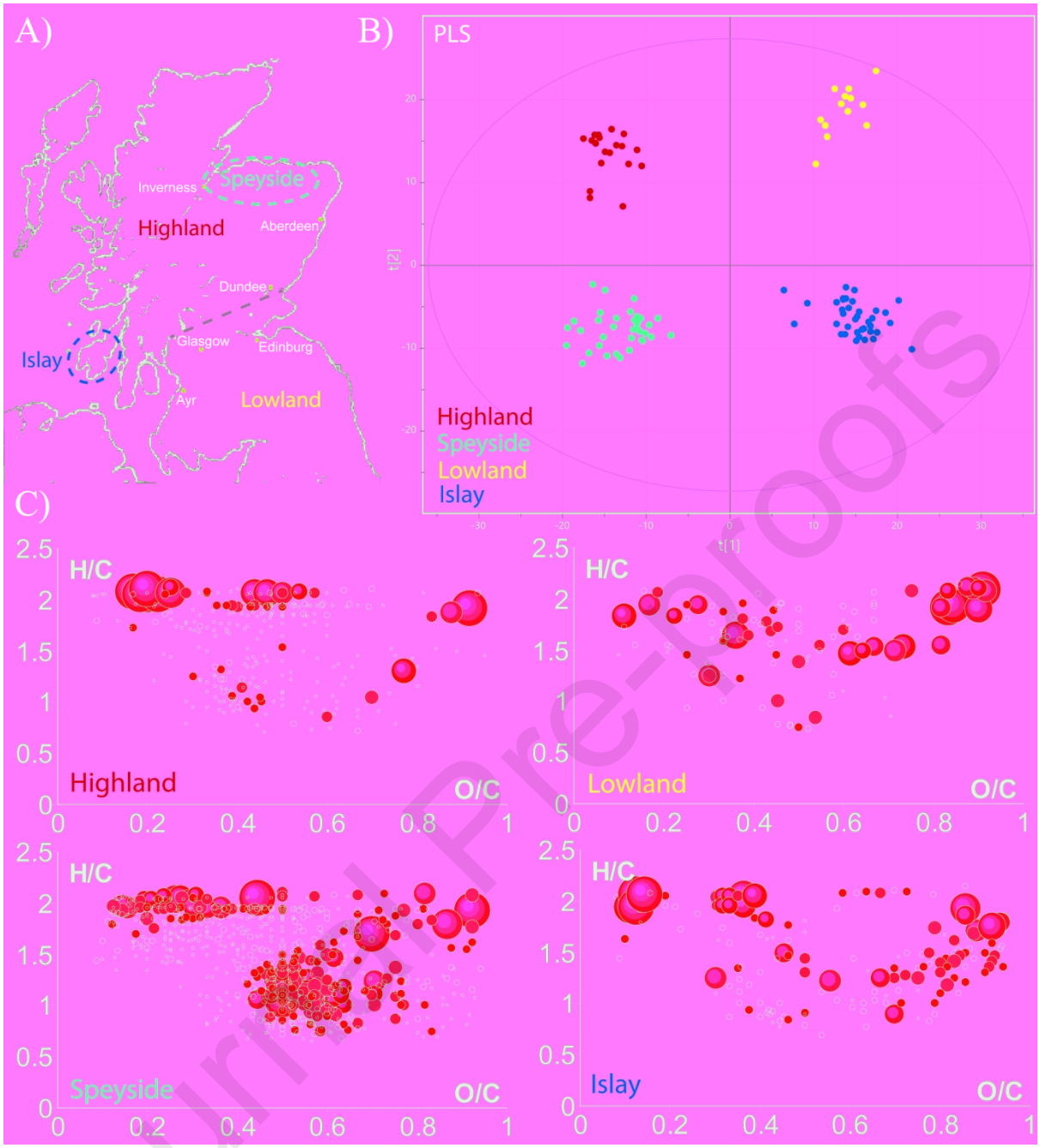


Figure 3

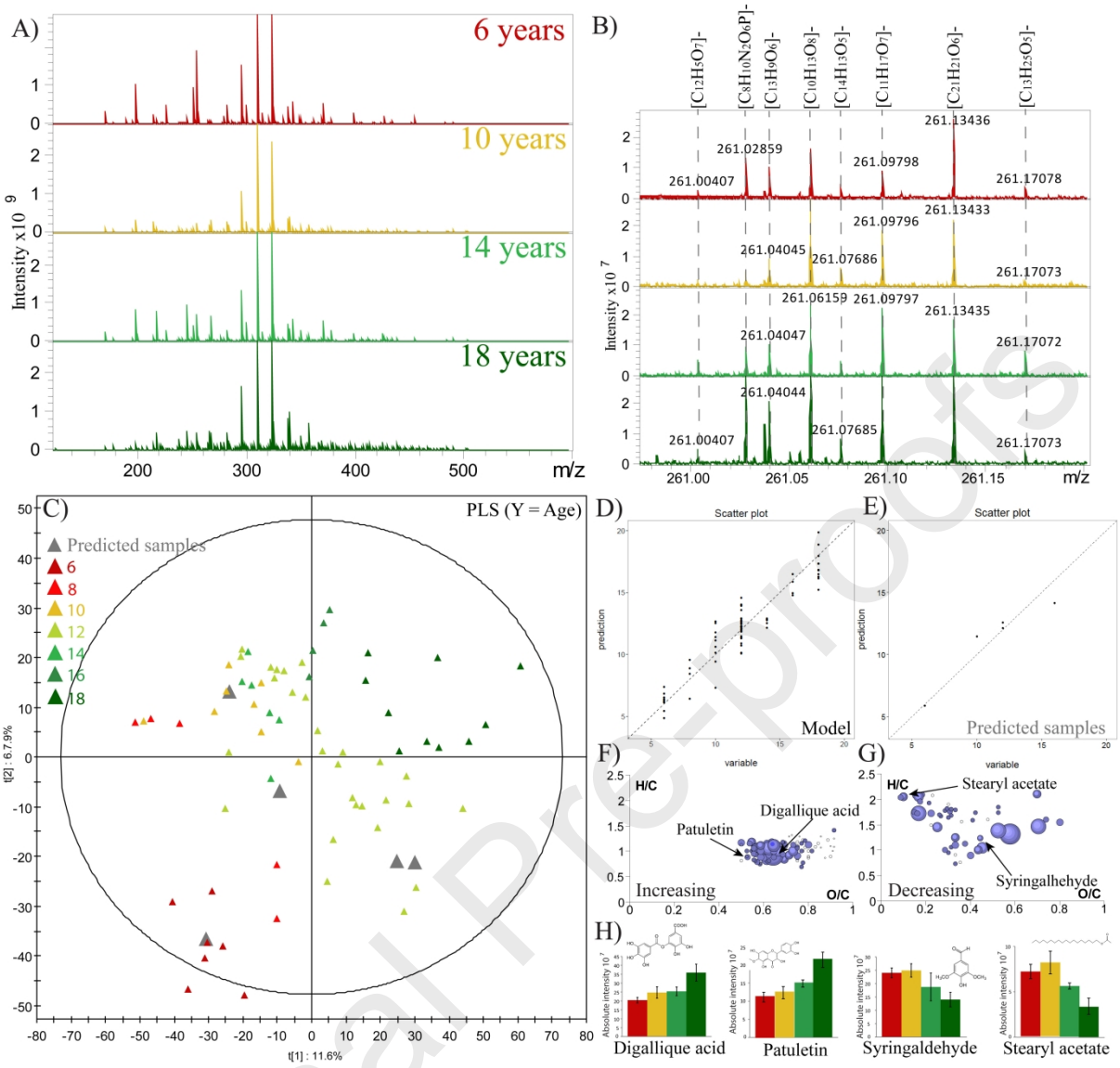


Figure 4

Author contributions:

C.R.-G., R.D.G, and P.S.-K. designed the research.

C.R.-G., J.S., D.H., M.K., C.C., M.L. performed the experiments and analyzed the data.

C.R.-G., D.H., C.C., M.L. B.S., R.D.G. and P.S.-K. wrote the manuscript.

All authors discussed the results and commented on the manuscript at all stages.

Declaration of interests

☒The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

- FT-ICR-MS analysis of 106 whisky
- Four areas in Scotland, 32 distilleries and 3 to 43 years of maturation
- Brand and geographical origin authentication
- Maturation time authenticity
- Flow transfer of compounds from wood barrels to distillates not constant and homogeneous over the maturation time