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 Gougeon, Philippe Schmitt-Kopplin

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

2

3   Chloé Roullier-Gall[1*], Julie Signoret[1], Christian Coelho[1], Daniel Hemmler[2,3], Mathieu

4   Kajdan[1], Marianna Lucio[3], Bernhard Schäfer[4], Régis D. Gougeon[1] and Philippe Schmitt-

5   Kopplin[2,3*].

6

7   [1] UMR PAM Université de Bourgogne/AgroSup Dijon, Institut Universitaire de la Vigne et du Vin, Jules Guyot,

8   Dijon, France

9   [2]Comprehensive Foodomics Platform, Chair of Analytical Food Chemistry, Technische Universität München,

10   Freising, Germany

11   [3] Research Unit Analytical BioGeoChemistry, Department of Environmental Sciences, Helmholtz Zentrum

12   München, Neuherberg, Germany

13   [4]Bernhard Schäfer : Independent Spirit consultant, Nuremberg, Germany

14

15

16   *Correspondence:

17   Chloé Roullier-Gall: chloe.roullier-gall@u-bourgogne.fr

18   Philippe Schmitt-Kopplin: schmitt-kopplin@helmholtz-muenchen.de

19

22

23   **Abstract**

24   Understanding the chemical composition of whisky and the impact of each step in the

25   manufacturing process provides a basis for responding to the challenges of producing high

26   quality spirits. In this study, the objective was to discriminate whiskies according to their

27   geographical origin and authenticate the maturation time in cask based on the non-volatile

28   profiles. The combination of FT-ICR-MS and chemometrics allowed the distinction of

29   whiskies from four geographical origins in Scotland (Highlands, Lowlands, Speyside and

30   Islay). Statistical modeling was also used to discriminate whiskies according to the

31   maturation time in cask and reveal chemical markers associated with the ageing regardless of

32   the origin or the production process. Interestingly, the flow of transfer of compounds from

33   wood barrels to distillates is not constant and homogeneous over the maturation time. The

34   largest transfer of compounds from the barrel to the whisky was observed around twelve

35   years of maturation

36

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). Infrared 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.

102

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

134 The repeatability of the measurements was evaluated on QCs samples injected in triplicate at

135 the beginning and at the end of the sequence analysis. The repeatability over time of FT-ICR-

136 MS analyses during the entire runtime was also evaluated , by monitoring the precision of the

137 recorded peak intensities from ten analyzed QC samples and manual comparison of the raw

138 mass spectra (Roullier-Gall et al., 2018). In the QC samples, 92.8% of all detected

139 monoisotopic signals showed a relative standard deviation lower than 20% (Roullier-Gall et

140 al., 2018).

141

142 **Excitation Emission Matrix Fluorescence**

143 Whisky fluorescence was acquired on a spectrofluorometer (Horiba Aqualog). Excitation

144 ranged from 225 to 600 nm and emission was recorded from 230 to 600 nm (C. Coelho et al.,

145 2015). The spectrofluorometer used a conventional right-angle optical setup. Excitation

146 Emission Matrices (EEM) were corrected daily for their Rayleigh and Raman scattering, as

147 well as inner filtering effects by using the functions provided within the FluorEssence

148 software of the Horiba Aqualog system. The fluorescence intensity was normalized to a

149 Starna 1 ppm quinine sulfate reference cell, measured every day.

150

151 **UPLC targeted analysis of polyphenols**

152 An Acquity Waters ultra performance liquid chromatography (UPLC) with a diode array

153 fluorescence detector (DAD) was used to separate and quantify individual polyphenolic

154 compounds in whisky samples. The column was a BEH C18, 1.7 µm, 2.1 mm × 150 mm.

155 Column temperature was kept constant at 30 °C, and samples were held at 8 °C. An elution

156 was applied, starting isocratic from 100% A (ultrapure water, 0.1% formic acid) from 0 to 6

157 min and then the gradient was increased linearly over 56 min to 100% B (methanol, 0,1%

158 formic acid), where it was held again until reaching 60 min, with a rate of 0.3 mL·min^{-1}.

159 Injection volume was 5 µL. Detection limits (LODs), quantification limits (LOQs) and intra-

160 and inter-day variation of the assay were determined by injecting a series of dilute solutions

161 with known concentrations.

162

163 **Total phenolic content (TPC) of whiskies**

164 Total phenolic content (TPC) in whisky was determined using the Folin-Ciocalteu assay: 0.1

165 mL of undiluted whisky, 0.5 mL of Folin-Ciocalteu reagent, 2 mL of a 20 (*w/v*) sodium

166 carbonate solution and 7.4 mL of ultrapure water were mixed and reacted for 30 min at room

167   temperature. Then, the absorbance at 750 nm was measured and calculated as mg.L^{-1} of gallic

168   acid equivalent using a calibration curve between 0 and 500 mg.L^{-1} of gallic acid. All

169   samples were analyzed in triplicates.

170

171   **Data analysis**

172   The statistical analysis was performed with Perseus, Simca 12.0 (Umetrics, Umeå, Sweden)

173   and R (Version 1.0.136).

174   Van Krevelen and elemental diagrams were used to visualize FT-ICR MS data (Gonsior et

175   al., 2009; Gougeon et al., 2011; Hertkorn et al., 2008). Multivariate statistical methods (.

176   Principal component analysis and hierarchical cluster analysis) were used to explore

177   similarities and hidden patterns among samples(Barker & Rayens, 2003; Granato, Santos,

178   Escher, Ferreira, & Maggio, 2018). As within-groups variability dominates the among-groups

179   variability for the geographical origin and maturation time impact, partial least square

180   discriminant analysis were used(Barker & Rayens, 2003). We set up diverse classification

181   models in order to reveal the different geographical origin and maturation time for barrel

182   specific metabolites. In order to improve the efficiency of such classification and the possible

183   presence of overfitting and noise, we preprocessed the entire dataset applying the ReliefF

184   algorithm(Witten, Frank, Hall, & Pal, 2016). The algorithm identified a subset of variables

185   that was used for the next classification models maximizing the model performance in term

186   of accuracy. The features' selection was based on the highest rank value attributed to each

187   variable by the algorithm. Consequently, with the reduced datasets we built several partial

188   least square discriminant analysis (PLS-DA) models. The goodness of the fit was evaluated

189   by the coefficient of determination (R²X), the proportion of the variance of the response that

190   is explained by the model (R²Y), the predictive ability (Q²Y) and validation was carried out

191   using permutation tests. For the ageing time model, whisky samples were randomly divided

192   into two groups: 101 samples selected to build the statistical model and the 5 remaining

193   samples employed as external validation. A Bland–Altman plot was used to analyze the

194   agreement between the known ageing time and the model prediction.

195

196   **Results and discussion**

197   **Geographical origin impact of Scotch whisky**

198

199   A visual comparison of FT-ICR-MS spectra of Scotch whiskies from Highland, Lowland,

200   Speyside and Islay distilleries and from 3 to 43 years of cask maturation is presented in

201   Figure 1. Several thousands of signals were found between m/z 100 to 600. The enlargements

202   at m/z 261 exemplarily show the differences in chemical composition between the four

203   geographical areas (Figure 1). Between m/z 261.0 and 261.2, whisky spectra only differed in

204   the intensities of the eight detected peaks. For example, signals at m/z 261.04044 and at m/z

205   261.17075 correspond to the molecularformulas $[C_{13}H_9O_6]^-$ and $[C_{13}H_{25}O_5]^-$, respectively,

206   and show comparable intensities in all four whisky samples (Figure 1). The Venn and van

207   Krevelen diagrams (Figure 1) show that 2200 formulas (out of 5979) were found in all

208   samples of the four groups of whisky samples confirming a great consistency in the chemical

209   composition regardless of the origin(Kew et al., 2016).

210   Despite the high similarities between whiskies (36.8% of formulas were found in all samples

211   of the four groups of whisky samples) as shown in the Venn diagrams and the direct

212   comparison of spectra (Figure 1), 533 (8%), 610 (10%), 401 (6.7%) and 464 (7.7%) formulas

213   were uniquely found in at least 50% of Highland, Islay, Speyside and Lowland whisky

214   samples, respectively. In addition to the unique formulas, differences in signal intensities

215   between whiskies can be found which could result from particular distilleries, maturation

216   times, geographical features or environmental impacts. As example, the signal at m/z

217   261.06158 corresponding to the molecular formula $[C_{10}H_{13}O_8]^-$ is less intense in the whisky

218   from Speyside ($2 \cdot 10^7$) compared to the other three whiskies ($5\text{-}6 \cdot 10^7$) (Figure 1).

219

220   The impact of the type of whisky (blend or malt), the maturation wood type and the wood

221   history on the Whisky composition has recently been studied(Kew et al., 2016; Roullier-Gall

222   et al., 2018). However, barley, peat and water origin, malting, brewing, fermentation and

223   distillation process, nature of the cellars, climates as well as environmental factors could also

224   impact the complex composition of whisky. Figure 2 shows the first two principal

225   components obtained from principal component analysis computed on the 106 analyzed

226   whisky samples from Scotland. Whiskies showed a rough separation by their maturation time

227   in the barrel on principal component 2. By comparison, the impact of the geographical origin

228   - Highlands, Lowlands, Speyside and Islay - on the whisky metabolome appears less

229   significant in the PCA compared to the maturation time (Figure 2). The samples from the

230   Highlands and Lowlands appeared in a more discrete cluster than the Speyside and Islay

231   samples, which are more distributed all around the plot. Interestingly, the four samples in the

232   top left corner of the PCA were from the same distillery (but with different maturation times)

233   and the eight samples (in red) in the top right corner were from different distilleries but all in

234   Islay and all with six years of barrel maturation (Figure 2). The great diversity in the

235   composition of the whisky samples demonstrates the large number of environmental

236   parameters (including production, geographical origin, maturation time…), which eventually

237   impacts the final product.

238

239   A group separation is not easily observable in PCA models because of the high number of

240   sample characteristics (maturation from 6 to 43 years, 4 geographical origins and 32

241   manufacturers) and the high number of variables considered (5979 molecular formulas). A

242   PLS regression model was used to better distinguish the geographical origins of the whisky

243   samples. Figure 3 shows results from a PLS regression computed from FT-ICR-MS data of

244   106 whisky samples from the four areas in Scotland, 32 different distilleries and from 3 to 43

245   years of maturation. The first two components allowed to separate whiskies into four distinct

246   groups according to the geographical origin (Highlands, Lowlands, Speyside and Islay) and

247   independent of the distillery or the maturation time (Figure 3). Whiskies from Lowlands and

248   Islay distilleries appear to be chemically more different from the samples from Highlands and

249   Speyside distilleries (separation on the 1st component). The second component allowed to

250   distinguish whiskies from the Highlands and Lowlands from those produced in Speyside and

251   Islay distilleries. Based on the PLS, specific masses for each location were extracted (VIP

252   masses) and projected into van Krevelen diagrams (Figure 3). For example, samples from

253   Speyside clearly showed an enrichment of polyphenols or polyphenol-type compounds

254   (Figure 3). By comparison, Highland samples seem to be characterized by higher alcohols

255   with a profile reminiscent to the distillate(Roullier-Gall et al., 2018). Lowland and Islay

256   appear to be characterized by carbohydrates. The schematic classification of single malts

257   according to their origin could be based, in part, on a geological reality, of which water

258   would be the main factor. Meir-Augenstein showed that the source water used for and during

259   whisky production correlates with the authenticity of whisky(Meier-Augenstein, Kemp, &

260   Hardie, 2012). Water is involved in several stages of processing: added to ground barley to

261   produce the mash, when diluting the distilled alcohol to cask strength for maturation and

262   when reducing cask strength whisky for bottling. Meir-Augenstein's study suggests that

263   whisky keeps a signature of the geographic provenance of the used water(Meier-Augenstein

264   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

297 Based on the huge impact of maturation in wood barrels on the chemical diversity of whisky,

298 as shown in our previous study(Roullier-Gall et al., 2018), we here focused on the impact of

299 the maturation time on the final composition. We first studied three samples from one

300 distillery (W3) in Islay after 6, 8 and 10 years of cask maturation (Supplementary Figure 2).

301 The samples revealed 376 mass peaks whose intensity significantly increased with maturation

302 time, and 81 mass peaks with decreasing intensity (Supplementary Figure 2a-b). The higher

303 number of features which showed an increase in peak intensity confirms that chemicals

304 transfer from wood to the distillate as a function of ageing.

305

306 In order to extract age-related metabolite signatures, a PLS discriminant analysis was

307 designed using 7 groups with different maturation times: whiskies that had been aged in casks

308 for 6 years, 8 years, 10 years, 12 years, 14 years, 16 years and 18 years (Supplementary Table

309 3). We selected 76 from the 106 analyzed whisky samples which ranged from 6 to 18 years.

310 A PLS-DA model was built based on 71 of these whisky samples (Figure 4c). Five of the 76

311 samples were used for model validation (Figure 4e). The separation of whiskies according to

312 the ageing time (Figure 4c) showed a good predictive power with Q^2 of 0.49 and R^2Y of 0.77.

313 The scatter plot of the model group versus the predicted group for each sample confirmed the

314 robustness of the prediction model (Figure 4 D), with a prediction average close to the

315 maturation time and an excellent standard deviation (between 0.69 to 1.57 years,

316 Supplementary Table 3). However, it was not possible to really differentiate groups of 12 and

317 14 years of maturation. These samples were predicted as 12.3 years for both, the group of 12

318 years and 14 years old whiskies (maximum standard deviation was 1 year).

319 Finally, the five whisky samples, which were not used for the model construction (maturation

320 of 6, 10, 12, 12 and 16 years in barrel) were used to validate the model. The good predictions,

321 achieved for these five whisky samples, confirmed the predictive ability of the maturation

322 time model (Figure 4 C and E). After model validation, 256 discriminating and characteristic

323 masses (VIP variables higher than 1), whose peak intensities increased with maturation, were

324 extracted (Figure 4 F and H). As example, compounds from wine such as syringic acid and

325 caftaric acid and from wood such as lyoniresinol, patuletin and digallic acid were found as

326 increasing with maturation time (Figure 4 and supplementary Figure 2). In contrast, 213

327 masses with decreasing intensity led to 54 annotations from databases including stearyl

328 acetate and syringaledehyde (Figure 4).

329

330   ANOVA tests were applied to extract subsets of features that discriminate between whiskies

331   with a maturation time of 6 and 8 years (supplementary figure 4A), 8 and 10 years

332   (supplementary figure 4B), 10 and 12 years (supplementary figure 4C), 12 and 14 years

333   (supplementary figure 4D), 14 and 16 years (supplementary figure 4E), and 16 and 18 years

334   (supplementary figure 4F), respectively. Statistically extracted masses can be used to

335   highlight the impact of maturation time on the whisky composition between the different

336   ageing groups. According to the statistics, 27 molecular formulas were significantly enriched

337   in the youngest whiskies (6 years of maturation) while 36 formulas were specific to the 66

338   whiskies older than 6 years of maturation. In the same way, 74 molecular formulas were

339   statistically identified as markers to differentiate between whiskies younger than 8 years of

340   maturation (6 and 8 years of maturation) and 132 formulas were identified as markers for

341   whiskies older than 8 years of maturation. The number of significant masses statistically

342   extracted increased until the distinction of whiskies between 10 (121 molecular formulas) and

343   12 (1182 molecular formulas) years of casks maturation. Then the number of significant

344   masses starts to decrease.

345   Interestingly, the higher differences were found between whiskies after 10 and 12 years of

346   maturation directly followed by whiskies after 14 and 16 years of maturation (supplementary

347   Figure 4). The similarities between samples characterized by 12 and 14 years of maturation,

348   which probably can be explained by the low ability of the model prediction to distinguish

349   those two groups (Figure 4). The difference between young (less than 12 years of maturation,

350   group 1) and old whiskies (more than 12 years of maturation, group 2) was further studied by

351   Excitation Emission Matrix (EEM) fluorescence spectroscopy and analysis of polyphenols

352   (supplementary figure 5 and 6). As illustrated in Supplementary figure 6A, PARAFAC

353   results enabled to easily discriminate whiskies from both groups on the first component.

354   Whiskies of group 1 (younger than 12 years of maturation) were driven by wood barrel

355   ageing processes, conferring a relative homogeneity to this group. Whiskies from the group 2

356   (older than 12 years of maturation) showed a higher within-group-variation (confirming

357   results from FT-ICR-MS), which may be because of different cask finishing practices. This

358   may include re-casking for a secondary maturation into other wood casks that were used for

359   ageing of wines, spirits or even beers (Scotch Whisky Association, n.d.). Such practice tends

360   to diversify Scottish whisky profiles by masking the original distillate. The fluorescence

361   positions of the PARAFAC components suggest differences in polyphenolic compounds

362   (Figure 6B) which could be confirmed in a TPC assay and quantification of syringic acid,

363   gallic acid and scopoletin by UPLC (Figure 6C). Scopoletin was also identified as wood

364   marker increasing in intensity with maturation time. No significant differentiations were

365   found for tyrosol between our two groups, surely due to a moderated retention of this

366   compound to wood surfaces (Christian Coelho et al., 2019).

367   Out of the 1182 molecular formulas found significantly increased for whiskies with at least

368   12 years of maturation, only 222 annotations were found in databases: 167 were found in our

369   home-built plant database, 154 in Metlin, 130 in KEGG and 25 in HMDB (Supplementary

370   Figure 7 and supplementary Table 4). Interestingly, 467 of the 1182 markers could be

371   previously identified as wood specific markers (Roullier-Gall et al., 2018) (Supplementary

372   Figure 8). Moreover, 675 of the 1182 features could be detected in wine samples including

373   337 compounds already identified as wood markers (Supplemental Figure 8). The higher

374   presence of wine-based compounds in whiskies older than 12 years, could correspond to

375   finishing practices that could take place after 10 years of ageing independently of the region

376   of production.

377

378   **Conclusion**

379   In conclusion, FT-ICR-MS based metabolomics was used as a powerful tool to discriminate

380   whiskies according to maturation time and geographical origin, regardless of the

381   manufacturer, the barrel used and the production process. Untargeted metabolomics allowed

382   to study a broad range of metabolites and the extraction of markers that were significantly

383   different depending on the geographical origin and/or the maturation time. A prediction

384   model was further developed for the comprehensive evaluation of maturation time

385   authenticity. Interestingly, the flow of transfer of compounds from wood barrels to distillates

386   is not constant and homogeneous over the maturation time. The largest transfer of compounds

387   from the barrel to the whisky was observed between ten and twelve years of maturation with

388   up to 1182 molecular formulas that showed a significant increase in their peak intensities as a

389   function of the maturation time. 467 of these were identified as wood-specific compounds

390   and 338 are typically also found in wine. The explanatory hypothesis for the higher presence

391   of wine-related compounds in whiskies older than 12 years, compared to whiskies younger

392   than 12 years could be the finishing practice that takes place after 12 years of ageing

393   independently of the region of production.

394

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

403

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

408

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

412

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

426

427

428 **Literature**

429

430 Aylott, R. I., & MacKenzie, W. M. (2010). Analytical strategies to confirm the generic

431   authenticity of Scotch whisky. *Journal of the Institute of Brewing*, *116*(3), 215–229.

432 Aylott, Ross I., Clyne, A. H., Fox, A. P., & Walker, D. A. (1994). Analytical strategies to

433   confirm Scotch whisky authenticity. *The Analyst*, *119*(8), 1741.

434   https://doi.org/10.1039/an9941901741

435 Barker, M., & Rayens, W. (2003). Partial least squares for discrimination. *Journal of*

436   *Chemometrics*, *17*(3), 166–173. https://doi.org/10.1002/cem.785

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

438 Coelho, C., Aron, A., Roullier-Gall, C., Gonsior, M., Schmitt-Kopplin, P., & Gougeon, R. D.

439   (2015). Fluorescence Fingerprinting of Bottled White Wines Can Reveal Memories

440   Related to Sulfur Dioxide Treatments of the Must. *Analytical Chemistry*.

441 Coelho, Christian, Gougeon, R. D., Perepelkine, L., Alexandre, H., Guzzo, J., & Weidmann,

442   S. (2019). Chemical Transfers Occurring Through Oenococcus oeni Biofilm in

443   Different Enological Conditions. *Frontiers in Nutrition*, *6*.

444   https://doi.org/10.3389/fnut.2019.00095

445 Garcia, J. S., Vaz, B. G., Corilo, Y. E., Ramires, C. F., Saraiva, S. A., Sanvido, G. B., …

446   others. (2013). Whisky analysis by electrospray ionization-Fourier transform mass

447   spectrometry. *Food Research International*, *51*(1), 98–106.

448 Gonsior, M., Peake, B. M., Cooper, W. T., Podgorski, D., D'Andrilli, J., & Cooper, W. J.

449   (2009). Photochemically Induced Changes in Dissolved Organic Matter Identified by

450   Ultrahigh Resolution Fourier Transform Ion Cyclotron Resonance Mass

451   Spectrometry. *Environmental Science & Technology*, *43*(3), 698–703.

452   https://doi.org/10.1021/es8022804

453 Gougeon, R. D., Lucio, M., Boutegrabet, L., Peyron, D., Feuillat, F., Chassagne, D., …

454   Schmitt-Kopplin, P. (2011). Authentication Approach of the Chemodiversity of Grape

455   and Wine by FTICR-MS. In S. E. Ebeler, G. R. Takeoka, & P. Winterhalter (Eds.),

456   *Progress in Authentication of Food and Wine* (Vol. 1081, pp. 69–88). Washington,

457   DC: American Chemical Society.

458 Granato, D., Santos, J. S., Escher, G. B., Ferreira, B. L., & Maggio, R. M. (2018). Use of

459   principal component analysis (PCA) and hierarchical cluster analysis (HCA) for

460   multivariate association between bioactive compounds and functional properties in

461   foods: A critical perspective. *Trends in Food Science & Technology*, *72*, 83–90.

462   https://doi.org/10.1016/j.tifs.2017.12.006

463 Han, J., Ma, C., Wang, B., Bender, M., Bojanowski, M., Hergert, M., … Bunz, U. H. F.

464   (2017). A Hypothesis-Free Sensor Array Discriminates Whiskies for Brand, Age, and

465   Taste. *Chem*, *2*(6), 817–824. https://doi.org/10.1016/j.chempr.2017.04.008

466 Heller, M., Vitali, L., Oliveira, M. A. L., Costa, A. C. O., & Micke, G. A. (2011). A rapid

467   sample screening method for authenticity control of whiskey using capillary

468   electrophoresis with online preconcentration. *Journal of Agricultural and Food*

469   *Chemistry*, *59*(13), 6882–6888.

470 Hertkorn, N., Frommberger, M., Witt, M., Koch, B. P., Schmitt-Kopplin, Ph., & Perdue, E.

471   M. (2008). Natural Organic Matter and the Event Horizon of Mass Spectrometry.

472   *Analytical Chemistry*, *80*(23), 8908–8919. https://doi.org/10.1021/ac800464g

473 Jack, F. R., & Steele, G. M. (2002). Modelling the sensory characteristics of Scotch whisky

474   using neural networks—a novel tool for generic protection. *Food Quality and*

475   *Preference*, *13*(3), 163–172. https://doi.org/10.1016/S0950-3293(02)00012-5

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

477   Kew, W., Goodall, I., Clarke, D., & Uhrín, D. (2016). Chemical Diversity and Complexity of

478   Scotch Whisky as Revealed by High-Resolution Mass Spectrometry. *Journal of The*

479   *American Society for Mass Spectrometry*, 1–14.

480   Lee, K.-Y. M., Paterson, A., Piggott, J. R., & Richardson, G. D. (2001a). Origins of flavour

481   in whiskies and a revised flavour wheel: a review. *Journal of the Institute of Brewing*,

482   *107*(5), 287–313.

483   Lee, K.-Y. M., Paterson, A., Piggott, J. R., & Richardson, G. D. (2001b). Sensory

484   discrimination of blended Scotch whiskies of different product categories. *Food*

485   *Quality and Preference*, *12*(2), 109–117.

486   Liebmann, A. J., & Scherl, B. (1949). Changes in whisky while maturing. *Industrial &*

487   *Engineering Chemistry*, *41*(3), 534–543.

488   MacNamara, K., Dabrowska, D., Baden, M., & Helle, N. (2011). *Advances in the Ageing*

489   *Chemistry of Distilled Spirits Matured in Oak Barrels*.

490   Macnamara, K., van Wyk, C. J., Brunerie, P., Augustyn, O. P. H., & Rapp, A. (2001).

491   Flavour components of whiskey. III. Ageing changes in the low-volatility fraction. *S.*

492   *Afr. J. Enol. Vitic*, *22*(2), 82.

493   Meier-Augenstein, W., Kemp, H. F., & Hardie, S. M. L. (2012). Detection of counterfeit

494   scotch whisky by 2H and 18O stable isotope analysis. *Food Chemistry*, *133*(3), 1070–

495   1074. https://doi.org/10.1016/j.foodchem.2012.01.084

496   M. MacKenzie, W., & I. Aylott, R. (2004). Analytical strategies to confirm Scotch whisky

497   authenticity. Part II: Mobile brand authentication. *Analyst*, *129*(7), 607–612.

498   https://doi.org/10.1039/B403068K

499   Møller, J. K., Catharino, R. R., & Eberlin, M. N. (2005). Electrospray ionization mass

500   spectrometry fingerprinting of whisky: immediate proof of origin and authenticity.

501   *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

527   Wiśniewska, P., Dymerski, T., Wardencki, W., & Namieśnik, J. (2015). Chemical

528   composition analysis and authentication of whisky. *Journal of the Science of Food*

529   *and Agriculture*, *95*(11), 2159–2166.

530   Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical Machine*

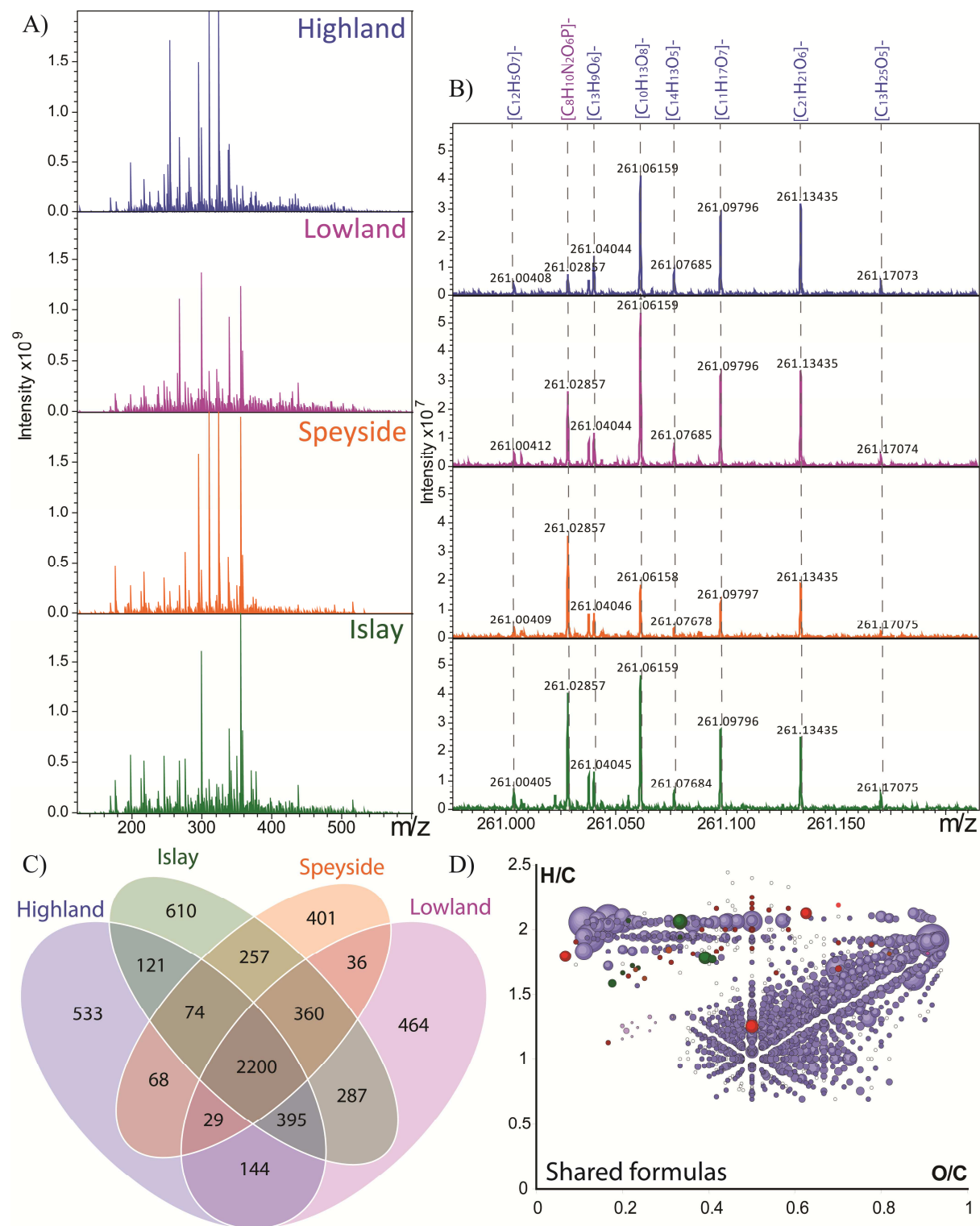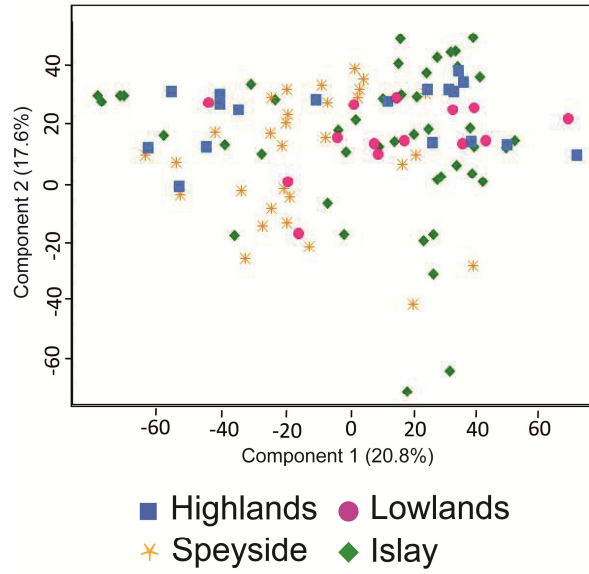531   *Learning Tools and Techniques*. Morgan Kaufmann.

532
533
534

Figure 1

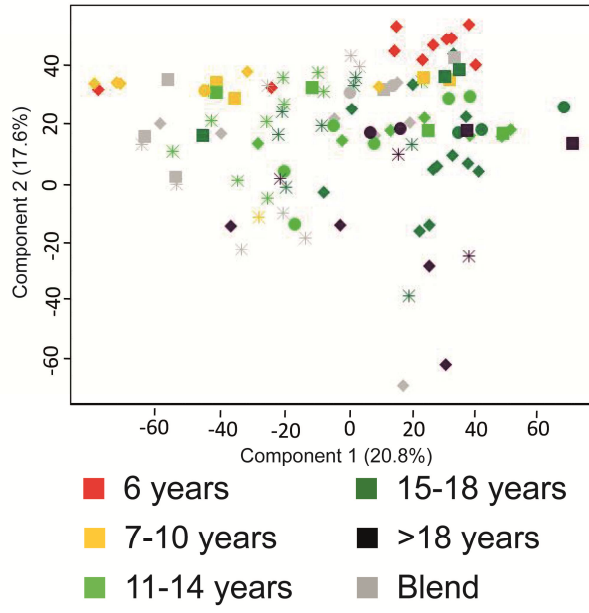A) PCA colored by geographical origin

B) PCA colored by maturation time

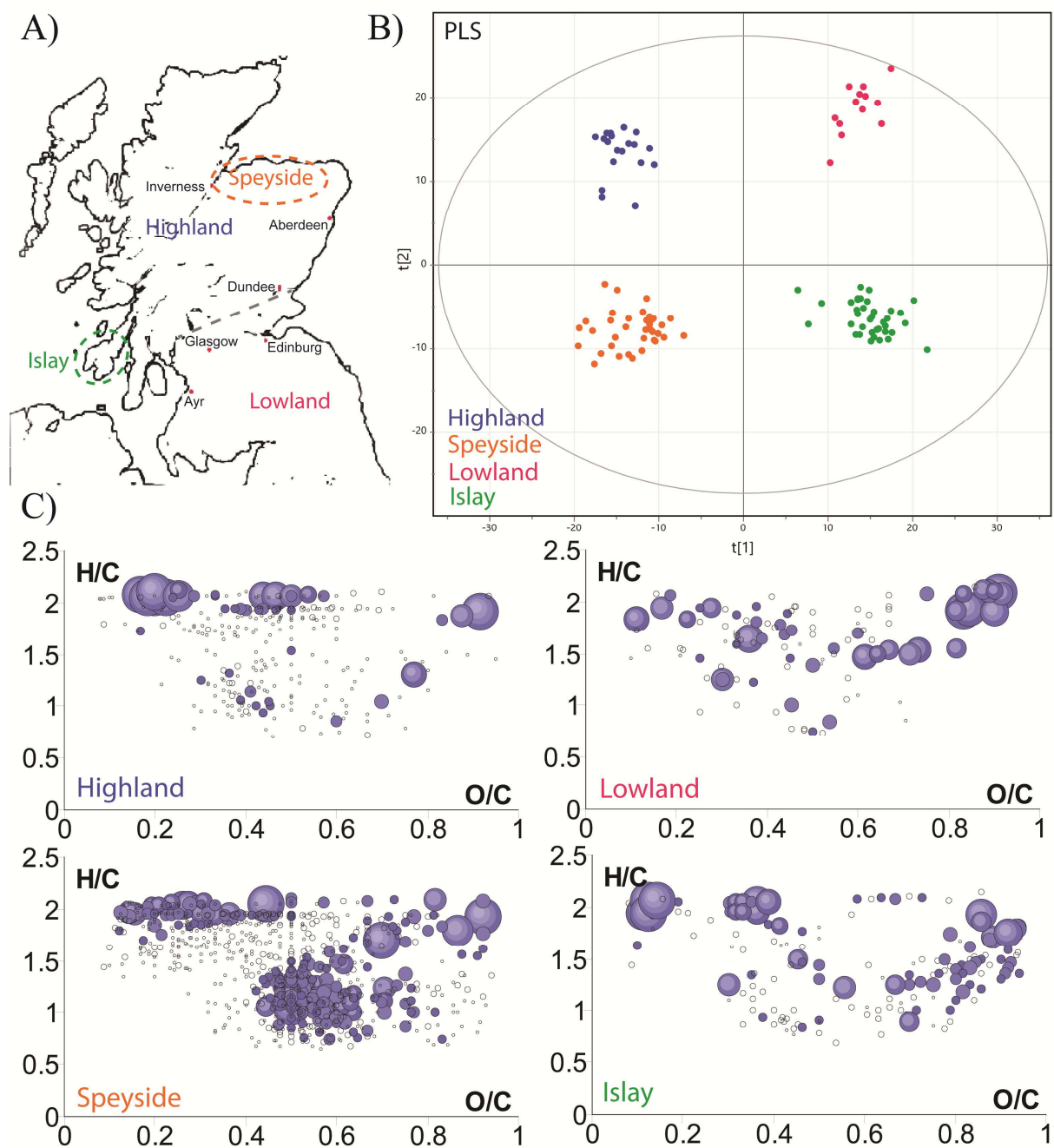Highlands ■ Lowlands ●
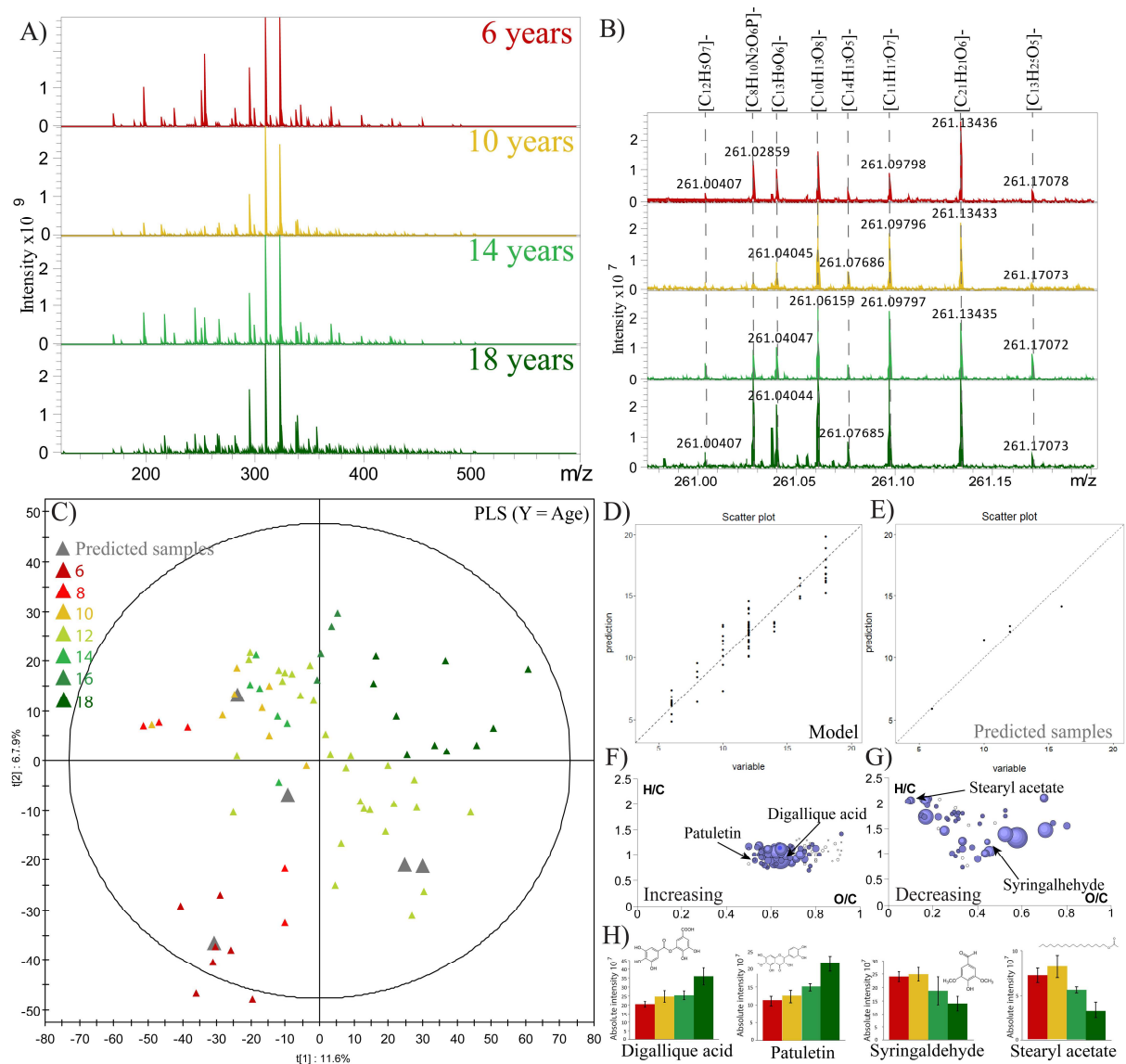Speyside ✳ Islay ◆

6 years ■ 15-18 years ■
7-10 years ■ >18 years ■
11-14 years ■ Blend ■

538

539 **Figure 2**

540

20

541
542 **Figure 3**
543

544

Figure 4

545

546

547

548