Summary

Scottish malt whisky is renowned for its diversity and complexity of flavor and odor profiles, as well as distinctive colors. From the deep, rich, smoky whiskies of Islay to the gentle, heather-tinted whiskies of the Spey River, the rich complexity of this fundamentally simple distilled drink has resulted in a world-wide mystique. The resulting worldwide market (expected to reach $8BN by 2027[1]) derives substantially from the fact that the distilled drink is matured in casks for at least 3 years before they can be called single malts. Although much of the flavors can be attributed to this wooden cask barreling, there are clear identifiers that can be associated with the individual distillery, as well as the barrel and age of the whisky. However, no single compound is likely to be attributable to one distillery or barrel type, and therefore the complexity of the "fingerprint" obtained in a gas chromatograph is analogous to systems as diverse as odorants in cancer patients (see Guest et al. 2021[2]) or compounds detectable in a complex Extractables and Leachables study[3]. In the work reported here, we use Gas Chromatography-Mass Spectrometry (GC-MS) to profile a small selection of malt whisky samples from various regions of Scotland and aged over a range of years in casks (or multiple casks).

Introduction

Whisky is the most widely consumed distilled beverage in the world. Generally, it is produced by distilling a beer mash made from grains, i.e. corn, barley, rye, wheat, etc., and fermented by yeast, to concentrate the alcohol. In malt whiskies, barley is initially steeped in water and allowed to germinate[4]. This process converts some of the starch into sugar. The "green malt" is then dried to stop the germination and, in some cases, add flavor through the choice of drying fuel (peat for the "smokey" whiskies). This dried mash is then ground and "mashed" to release the sugars into solution. Once the solids have been removed this "wort" is then fermented for up to two days to convert the sugars to alcohol and additional trace alcohol biproducts "congeners". Prior to distillation the soup of fermented liquids therefore contains multiple trace compounds arising from the water, drying and fermentation process.

Following fermentation, the wash is distilled in pot stills. Depending on the whisky, the wash may be distilled from 1-3 times, but for Scottish whisky it is generally distilled twice. Distilleries all possess uniquely shaped stills, many of them hundreds of years old, and the shape of the still is believed to make a difference in the final flavor of the whisky. Only the middle portion of the distillate is

[1] https://www.prnewswire.com/news-releases/scottish-whisky-market-to-reach-us-7-89-billion-globally-by-the-end-of-2027--coherent-market-insights-301026802.html
[2] https://www.campoly.com/educational-resources/publications/
[3] https://www.campoly.com/cpg-services/analytical-testing/extractablesleachables/
[4] https://www.scotchwhiskyexperience.co.uk/about-whisky/making

captured for further processing, the most and least volatile components are recycled in a subsequent distillation (foreshots and feints respectively). This "center cut" is generally approximately 68% alcohol. The resulting spirit is diluted slightly to maturing strength (63%) using local water and aged.

The colorless distillate or spirit (sometimes called moonshine in some parts of the US) is usually aged in wooden barrels for minimum period of 3 years (for Scottish malt whisky), and often much longer. The wood barrels are traditionally oak, often previously used for bourbon manufacture, but increasingly barrels from other industries (sherry or rum, for example) are used singly, or in conjunction with, the traditional barrels to impart unique properties. The ageing process is what gives the whisky the straw to amber to brown coloring and the bulk of the distinctive unique flavors and aromas, but it is not solely due to the barrel. The true source of the complexity of whisky is still debated. Even the mash and distillation processes have a role to play because the composition of the wash is not just alcohol, but also other secondary fermentation products, as well as components arising from the preparation steps (such as peat smoke used in drying), and while the distillation process is "purifying" the liquid, it is also selecting for compounds other than the alcohol that provide flavor and aroma.

Partly because of the economic impact of this industry, but probably also because of the complexity and mystique of the process, the analysis of whisky has been a growing scientific endeavor for many years. Numerous Ph.D. theses have been written on the subject. One of the main questions of this research has been what are the chemical constituents that produce the distinctive flavors and aromas of the different whisky products, and what is their origin? More prosaically, analysis of whisky is also performed to identify counterfeit product or adulteration. Chromatography, both gas and liquid, is a primary analytical tool for the analysis of whisky. Many different organic compounds are found in whisky. These include short and long chain alcohols and alcohol congeners, volatile fatty acids and fatty alcohols, phenols, terpenes, complex sugars to name a few. These compounds may be introduced during the fermentation and subsequent distillation of the grain mash or from the wood cask in which the whisky is aged (or indeed may be selectively "lost" during the aging process as part of the "angel's share").

This study focuses on a small subset of the world of Scotch Whisky. There are 131 distilleries located in 6 regions of Scotland, the Lowlands, Speyside, the Highlands, Campbeltown, Islay and the Islands (not considered a true region by all). In this study, seventeen samples from four of the regions, namely the Highlands, Speyside, Islay and the Islands, were analyzed. They ranged in age from 3 to 17 years. All samples, except three, were taken from previously opened bottles. The three unopened bottles were taken from the same distillery with different aging times. The samples are listed in Table 1 and the specimen colors are shown in Figure 1.

Figure 1: Color shades of whisky described in Table 1. Samples are ordered left to right 1 to 14 (15, 16 and 17 not shown).

Table 1: Samples discussed with distillery, region, age, wood (if known) and color (in the form of Hue, Saturation and Luminance).

Sample	Distillery	Region	Age (yr)	wood	H	S	L
14518-1	Fettercarin	Highland	12	sherry	42.4	53.2%	59.8%
14518-2	Dalmore	Highland	12	Bourbon/Sherry	36.6	49.0%	49.2%
14518-3	Glenmorangie	Highland	10	white oak	46.3	35.4%	68.4%
14518-4	Dalwhinnie	Speyside	15	oak	43.7	48.4%	62.7%
14518-5	Glenmorangie “Nectar D’or”	Highland	12	Bourbon/Sautern	42.3	56.0%	60.8%
14518-6	Laphroaig	Islay	10	oak	43.5	63.1%	64.9%
14518-7	Ardbeg (single cask)	Islay	7	?	46.5	56.0%	68.8%
14518-8	Bunnahabhain Cruach-Mhona	Islay	3	Oak	48.7	64.9%	74.3%
14518-9	Lagavulin	Islay	16	Oak	42.6	69.7%	61.2%
14518-10	Bruichladdich (private single cask)	Islay	11	Bourbon	44.7	74.6%	75.3%
14518-11	Balvenie (Carribbean cask)	Speyside	14	Rum	42.5	82.7%	68.2%
14518-12	Ledaig (sherry casks)	Island	17	Sherry	43.5	79.0%	72.0%
14518-13	Caol Ila	Islay	11	Oak	44.1	77.3%	82.7%
14518-14	Caol Ila (single cask)	Islay	16	Bourbon?	43.7	86.4%	74.1%
14518-15	Arran	Island	10	Oak			
14518-16	Arran	Island	12	Oak			
14518-17	Arran	Island	14	Oak			

Experimental

The samples were analyzed with two different Gas Chromatography Techniques. One technique analyzed the samples directly by headspace GC-MS (HS-GC-MS). This method gently heats the sample and then analyzes the gases in the “headspace” above the liquid sample. It is therefore most likely to correlate with the compounds associated with the distinct smell of the whisky. The second technique extracted the samples with ethyl acetate using a process known as liquid-liquid extraction and the extract was analyzed by GC-MS using direct injection. The ethyl acetate is not miscible with the water in the spirit but will extract organic species and then be separated from the water to allow injection onto the instrument.

For analysis by HS-GC-MS, the samples were prepared by mixing 1 mL of samples with 4 mL of ASTM Type II water in a 10 mL head space vial. The headspace vial was placed in an Agilent G1888 automated headspace analyzer connected to an Agilent 6890/5973 GC-MSD. The headspace analyzer was programed to equilibrate the sample to 80 °C for 30 minutes with agitation, then sample 100 µL of headspace gas which was subsequently injected into the GC-MSD. The GC was programed to hold for 2 minutes at an initial temperature of 35°C, then ramp at a rate of 20°C/min to 60°C, hold for 2 minutes, then ramp at 10°C/min to 300°C and hold for 10 minutes. The GC column was a HP-5MS UI, 30 m x 0.25 mm with a phase coating of 0.25 µm. The carrier gas was helium, and the flow rate was set at 1mL/min. The MS was programed to collect scan data from 50 amu to 450 amu.

For the extraction experiment, 1 mL of sample was mixed with 4 mL of ASTM Type II water in 10 mL glass screw cap vial. The sample mixture was extracted once by shaking vigorously with 2 mL of ethyl acetate for 2 minutes. The extraction process created a thick emulsion which was dissipated by adding approximately 1.5 g of sodium chloride and shaking gently. The clear solvent layer was drawn off with a Pasteur pipette and placed in a 4 mL vial. The solvent volume was reduced to dryness under a gentle stream of nitrogen. The dry samples were reconstituted in 250 µL of ethyl acetate, vialed 2 mL auto injector vials and injected by auto sampler into the GC-MS. The column and oven program used for the direct injection were the same as that used for the HS-GC-MS analysis.

Results

The headspace analysis yielded 106 compounds above a threshold level of 0.1% of a reference compound. The majority of these peaks were only detected intermittently or rarely across many samples. The most common compounds (>5% total area) comprised only 16 compounds. A similar story is observed in the direct injection samples, with 408 compounds detected, 51 of which were present at above 5% total area. There is no doubt from these data that whisky is not a "simple" system, and indeed the significance of these compounds to the drinker's experience is likely not limited to simple percentage concentration.

The challenge for analytically separating these whiskies based on their composition is that although the sample size of 17 whiskies may appear large, the selection across multiple distilleries, ages and barrels makes patterns difficult to discern, and, more importantly, the huge number of compounds present in these samples makes it difficult to reduce the data set in a meaningful way. However, because we are forearmed with some information about what drives taste and aroma in whiskies, we can use this information to reduce the data set considerably. Examining the list of common "expected" compounds (plus a more expansive list of esters which one would expect to be involved in the taste and aroma), the direct injection GC analysis identifies 25 compounds of interest, and the headspace only detects six of these, all of which are esters (a subset of the GC compounds). These are shown in Table 2.

Table 2: Detected compounds in either GC mode from the shortlist, and whether they were detected in HS (headspace) mode. Note the very limited number of compounds from the shortlist detected in HS and the prevalence of esters. Shortlist of compounds from [5]).

Compound	Characteristic	HS?
2(3H)-Furanone, 5-butyldihydro-4-methyl-, cis-	Whisky Lactones (from cask)	
trans-3-methyl-4-octanolide	Whisky Lactones (from cask)	
Phenol, 2-methoxy- (Guaiacol)	Phenolic compounds (smoke)	
Eugenol	Phenolic compounds	
Phenol, 2-methyl- (o-Cresol)	Phenolic compounds (smoke)	
Phenol, 3-methyl-	Phenolic compounds	
p-Cresol	Phenolic compounds	
Benzaldehyde, 4-hydroxy-3,5-dimethoxy- (syringaldehyde)	Aldehydes	
Vanillin	Aldehydes (Vanilla note)	
Furfural	Aldehydes	
2-Buten-1-one, 1-(2,6,6-trimethyl-1,3-cyclohexadien-1-yl)-, (E)-	Other compounds	
Phenylethyl alcohol	Other compounds (floral note)	
Esters		
Ethyl acetate	nail polish remover, model paint, model airplane glue	Yes
Benzoic acid, ethyl ester	sweet, wintergreen, fruity, medicinal, cherry, grape	
Butanoic acid, ethyl ester	banana, pineapple, strawberry	Yes
Hexanoic acid, ethyl ester	pineapple, waxy-green banana	Yes
Heptanoic acid, ethyl ester	apricot, cherry, grape, raspberry	
Butanoic acid, 3-methyl-, ethyl ester	apple	Yes
Nonanoic acid, ethyl ester	grape	
Pentanoic acid, ethyl ester	apple	
Isobutyl acetate	cherry, raspberry, strawberry	Yes
1-Butanol, 3-methyl-, acetate	pear, banana (flavoring in Pear drops)	Yes
1,6-Octadien-3-ol, 3,7-dimethyl-, formate	apple, peach	
Propanoic acid, 2-methyl-, propyl ester	rum	

[5] https://i1.wp.com/www.compoundchem.com/wp-content/uploads/2015/03/Chemistry-of-Whisky.png?ssl=1

When the GC data (not shown) are examined, there are few discernable patterns, although some qualitative trends are observed. Furfural, Hexanoic acid and 1-Butanol are largely present in all samples. The p-Cresol appears more associated with the Spey whiskies, and the o-Cresol appears more associated with the Islay whiskies. In terms of the three different aged whiskies (-15, -16 and -17, from Arran), there is no compound that appears to rank with age. Reviewing the HS data, the compounds that one would expect to provide the closest match to the drinker's initial perception of the whisky yields a similar result, with no clear pattern emerging.

The complex brew of hundreds of compounds makes it highly unlikely that a simple correlation can be determined using conventional means. This issue can be highlighted by examining three specific samples, 14518-15, -16 and -17. These three were freshly opened and all from the same distillery. The only difference between them was age of maturation (although of course these whiskies have been blended). Both the headspace and GC data look very similar across the three ages, although it is interesting to note that the complexity of the chromatograms (number of compounds) appears to correlate with the age of the whisky. This would appear to be consistent with the idea of barrel age adding complexity and character. One might therefore think that the number of compounds should be related to the age of the whisky but ranking the samples in order of age does not suggest any particular pattern. We are therefore left with, if anything, too much data to handle in a conventional manner. This issue arises fairly frequently in chromatographic analysis, and we have had prior success with some forms of correlation analysis[2], where attempts were made to correlate chromatographic analysis with canine olfaction in the detection of cancer. Unfortunately, in this non-systematic study, the number of permutations does not lend itself to a simple analysis (17 malts, 13 distilleries, 5 whisky "regions", 9 different barrel types and a range of ages from 3 to 17 years).

The challenge in this large data set is how to distill (pun intended) the multiple potential permutations into a manageable analysis. 400 compounds, across 17 samples, where the samples themselves are combinations of properties (barrel, age, distillery etc.) is an almost impossible data set to handle by conventional means. Rather than focus on individual compounds, are there any trends or correlations that can be extracted from the data set as a whole? Multivariate analysis is a means of analyzing large, interconnected, data sets using statistical tools to handle situations where multiple measurements are made on individual specimens. There are a range of tools that fall within this general class, but one popular tool is Principal Component Analysis (PCA). This is a method for collapsing a large number of variables down into fewer "principal components". It can be used to reduce the number of dimensions that are considered and is therefore frequently used in exploratory data analysis. Although the math is complex, many statistical packages allow simple use of PCA to examine and screen data prior to building models and analyzing specific outputs. Initially, Origin 9.1 was used to perform PCA on the 400 compounds identified, but no correlation was generated that appeared to match with the known properties of the whiskies. Clearly, that does not mean that these correlations are invalid, just that the huge number of compounds (400 compounds) correlated across multiple permutations in a limited number of samples (17) does not generate useful correlations. If instead the short-list of compounds discussed above is analyzed, a more useful correlation can be obtained. In this case, PCA suggests that the first four eigenvalues cover almost 90% of the variability in the data. However, examining the coefficients for each component, a large number of all whiskies are in the same general area for PC1 and there is no obvious separation arising in this PC. Most likely this PC is dominated by compounds common to all of the samples, and therefore is not of any particular value here.

However, examining PC2 to PC4 (**Figure 2**), it is immediately clear that for these components, there are three general groups (vectors that are at similar angles are correlated). PC2 and PC3 demonstrate three clear groups, one with almost zero contribution from PC3 and the two more in positive and negative PC3. The group with the correlation with PC2 has only a small contribution in PC3. This relationship can be most clearly seen in the inset of Figure 2, where this group is comprised of samples 14518-1, -10, -15, 16 and -17. These specimens are the Fettercairn, Bruichladdich and the three from Arran. The two other groups, both with low contributions from PC2, are separated in PC3, with one group comprised of 14518-6, -8, -9 and -13. These samples are the Laphroaig, Bunnahabhain, Lagavulin and Caol Ila. The other group is comprised of -2, -3, -4, -5, -11, -12 and -14. These are Dalmore, Glenmorangie, Dalwhinnie, Glenmorangie, Balvenie, Ledaig, and single cask Caol Ila suggesting that this PC is correlated generally with the difference between the Islay and Spey/Highland whiskies. Curiously, one final sample appears to stand alone, sample -7, the Ardbeg sample. Given the smoky nature of the Ardbeg, one would consider that this should appear correlated with the same group as the other Islay samples. PC4 also provides further separation.

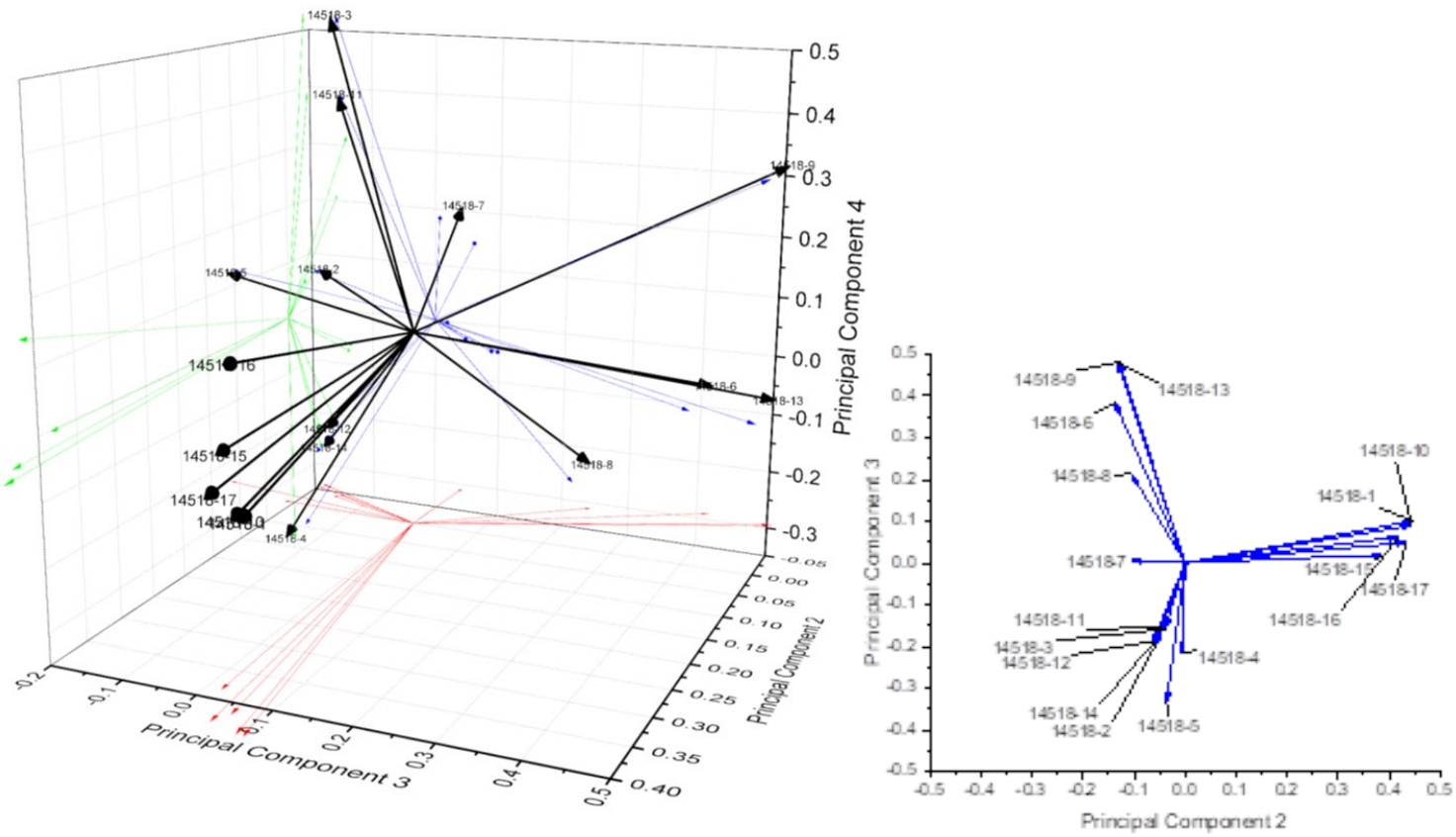

Figure 2: PC2, PC3 and PC4 (main plot) with inset for PC2 versus PC3 for the samples reviewed here.

Conclusions

In conclusion, our initial hope of teasing out simple correlations with known traits of the whiskies was not borne out. At least partially this is due to the intrinsic complexity of these liquids, with combinations of age, processing, environment and ingredients resulting in a complexity that is difficult to handle analytically. In addition, most of the liquids presented here are in fact blended from multiple barrels (even though they are termed "single malts") to create a certain profile, further complicating any analytical process.

This study was an internal project, conducted to demonstrate some of CPG's chromatographic abilities without compromising any confidential client data and as an excuse to hold a whisky tasting with leftover supplies. When CPG analyzes a product on behalf of a client, we have access to the manufacturer's IP (including key information about processing, environment, and materials), which allows us to better isolate compound correlations with known traits.

Complexity is increasingly becoming the norm in analytical chemistry. CPG conducts complex profiles in Extractable and Leachable (E&L) studies in support of biocompatibility determinations, and for new devices and drugs requiring analysis *in vivo* for residual products or change products. However, because we keep our clients' IP confidential, we can't share most of those examples (unless our client has given permission or credited us in publication). CPG is very proud of our part in the cancer study noted earlier, in which we performed a complex direct analysis of physiological fluids for analytes generated by the body, which we can discuss because our participation was published by our client.

Although the hoped-for patterns were not found in this small sampling, the tools (both analytically and processing) have potential value across a range of fields and understanding how to apply and tailor them to specific applications will become more important in the future. Plus, unlike usual laboratory experiments, it was encouraged to taste these samples!