Spring 5-8-2021

Initial investigation of wildflower honey using headspace solid-phase microextraction coupled with gas chromatography-mass spectrometry for geographical information

Tyra Countiss
Coastal Carolina University

Initial investigation of wildflower honey using headspace solid-phase microextraction coupled with gas chromatography-mass spectrometry for geographical information

By

Tyra Countiss

Biology

Submitted in Partial Fulfillment of the
Requirements for the Degree of Bachelor of Science
In the HTC Honors College at
Coastal Carolina University

Spring 2021

Louis E. Keiner
Director of Honors
HTC Honors College

Drew Budner
Associate Professor
Chemistry
Gupta College of Science

Table of contents

Abstract

Honey has been used as a food, sugar substitute, and flavor enhancer forever. The uses for honey are extremely varied from food to medicine. It is widely touted that you can address seasonal allergies, especially those following a move, by eating local honey. For this to be true the composition of the local honey, including trapped pollen, would allow allergy symptoms to be eliminated. In this project, the volatile and semi-volatile aroma compounds in wild flower honey from several different locations were analyzed. Headspace solid phase microextraction (HS-SPME) coupled with gas chromatography-mass spectrometry (GC-MS) was used to help develop the characteristic flavor and aroma profiles of each honey analyzed. Then in combination with statistical analysis, such as principle component and cluster analysis, the data allowed for the characterizing of these honeys based on location. The goal of this project is to detect regional variations in honey to begin to establish the idea of terroir for honey.

Introduction

Honey is a natural sweetener produced by honey bees using the nectar of plants that bees pollinate. The nectar that the bees collect is passed from bee to bee and is mixed with the enzyme invertase that turns the nectar into honey. Invertase activity differs in different types of honey causing different chemical composition of these honeys. [1]

While honey is mostly comprised of sugar and water, there are also components which include minerals, phenolic compounds, organic acids, proteins, vitamins and volatile compounds, usually referred to as volatile organic compounds (VOCs). [2] These compounds can include esters, ethers, alcohols, carboxylic acids, aldehydes, ketones, terpenes, nonisoprenoids, carotenoid derivatives, furan and pyran derivatives, and phenolic volatiles. [2] These components are specific to geographical location.

Each honey has a unique aroma profile. Some studies have been done to isolate the different aroma profiles in honey samples. [2-5] The aroma profiles come from the unique chemical properties listed above. There is a connection between the types and relative concentrations of these aroma compounds and the floral source the bees' sample. Monofloral honey comes from bees visiting a single floral type, i.e. clover honey. Bees can also collect from a variety of floral sources and is termed as polyfloral honey and often is marketed as wildflower honey. The presence of signature compounds has been used to verify the floral origin of some honeys.

Because of the differences in chemical makeup of each honey, it is thought that consuming honey made locally will help aide in treating seasonal allergies. This idea comes from the thought that the composition of the honey will build up the tolerance of local allergens like

pollen. The aroma of each honey also makes it more “attractive” giving it a distinct flavor. These distinctive qualities come from the floral source of the honey, otherwise known as the plant where the nectar originated. [4]

Solid phase microextraction, or SPME, is a way to eliminate the need to use toxic organics in the extraction of aroma compounds. [5] Head space gas chromatography, or HS GC-MS, is able to detect the VOCs from each sample and isolate them from one another. Instead of using liquid-liquid extraction and having honey and an organic solvent to remove compounds from honey for analysis, the fiber used in HS GC-MS traps VOCs and then transfers them into the GC-MS. GC-olfactometry can also be utilized in isolating odor-active compounds within the sample. [5]

Using SPME GCMS analysis of local honey samples, the aroma profiles can be determined. Once composition of each sample is known, cluster analysis can be used to separate each honey into their different regional geography.

Materials and Methods

Honey

Seventeen honeys were analyzed comparing aroma profiles to determine the composition of each honey and separate each honey into their regional areas as shown in Table 1. These samples were purchased locally from areas around South Carolina, North Carolina, as well as various other locations shown in Table 1. There were kept sealed at room temperature until analysis was conducted. The results were obtained using solid-phase microextraction and gas chromatography-mass spectrometry methods.

Table 1: List of honey purchase either directly from the producer or local merchant. The name of the honey, the laboratory ID, and the approximate production location are all listed.

Sample	ID	Identity	Locale
A	SS	Silver Spoon	Wilmington, NC
B	Beach	Beach Road	Southport, NC
C	Lowe	Lowe Honey	Southport, NC
D	GALL	Hive-Gallberry	Calabash, NC
E	SPALM	Hive-Gallberry	Southern, GA
F	BG	Bee Gee	Calabash, NC
G	SER	Kirkland	Mix
H	UNC	Uncle Jim's	Latta, SC
I	GRIS	David Grissett	Ocean Isle, NC
J	LOUG	Louisiana Gold	New Orleans, LA
K	Craic	Craic Honey Co.	Naches, WA
L	WFLWR	Wildflower Honey	Roseville, MN
M	OBLOSM	Orange Blossom	Hamptonville, NC
N	ASUE	Aunt Sue's	Sioux City, IA
O	BRAZIL	Wildflower Brazil	Brazil
P	UNC2	Uncle Jim's #2	Latta, SC
Q	Mnt2	Mountain Man #2	Conway, SC

Chemicals

Sodium chloride (NaCl) was obtained from EMD Chemicals Inc. (Darmstadt, Germany), 2-heptanol, guaiacol, 2-methyl-butanal, and n-octanal from TCI (Tokyo, Japan), furfural from Acros Organics (Fairfield, NJ) and benzaldehyde from Alfa Aesar (Haverhill, MA). All chemicals were used as supplied without additional purification. The internal standard for the GC-MS analysis was prepared using 200 mg/L of 2-heptanol and 100 mg/L of guaiacol in ethanol and used throughout the study.

Sample Preparation

Samples were made using 5g of each honey sample along with 1g of NaCl, 5mL of water, and 50 µL of standard solution of 2-heptanol and guiacol in a 20 mL headspace vial. This mixture was heated with stirring at 45ºC for 15 minutes. A divinylbenzene-carboxen-polydimethylsiloxane 50/30 µm (DVB-CAR-PDMS) SPME fiber was then injected into the honey headspace for 40 minutes while the sample continues stirring at 45ºC. Finally, the fiber is removed from the headspace and injected into the gas chromatography injector for two minutes. Each sample was analyzed three times.

GC-MS

Gas chromatography–mass spectroscopy (GC-MS) was carried out using a Shimadzu GC-2010 coupled to a QP2010 SE quadrupole mass spectrometer. A Rxi-5Sil MS column (30 m X 0.25 µm I.D.) with a film thickness of 0.25 µm was used. The GC was equipped with a split-splitless injector which was held at 250 °C. The analysis was performed with a splitless injection over the two-minute desorption time. The GC oven was initially set to 30 °C with a two minutes hold and then was raised in three steps: 30-70 °C at 10 °C/min and held for one minute; 70-220

°C at 4 °C/min and 220-270 °C at 20 °C/min and finally held at 270 °C for 6 minutes. The response of the mass spectrometer was monitored in TIC mode from 35-280 m/z. Compounds were identified via match to the NIST Mass spectra library. The area of each identified peak was reported relative to the area of the internal standard. These relative responses were then subjected to statistical analysis.

Statistical Analysis

A principle component analysis (PCA) was conducted on the scaled data using *prcomp* and *sparcepca* in R. In addition, cluster analysis was performed using a variety of methods to determine distance between the groups using traditional approaches (Euclidian, Manhattan, Minowski) and correlation based (Pearson, Kendall, and Spearman). The elbow method in K-means clustering determines the optimal number of clusters by comparing the within cluster sum of squares against the number of clusters. The majority of the statistical analysis was completed by Dr. Lindsey Bell from the Department of Mathematics and Statistics at Coastal Carolina University.

Results

The fifteen different honey were sampled in triplicate and the relative response for each peak was normalized against the response from the 2-heptanol internal standard. An example of a typical honey chromatogram is shown in Figure 1. The total number of peaks identified from the 45 individual samples was over 2000 compounds. Compounds likely resulting from either the SPME fiber or column bleed were removed from the results. The resulting peak information was averaged for each honey type. The vast majority, approximately 80%) of relative responses were approximately zero. This is, these compounds were in very few of the samples. Upon closer inspection, there were over 1200 compounds that were present within only one honey type. The remaining analysis, specifically the statistical analysis, was limited to those compounds present in at least half the honey types. This limitation resulted in 116 unique compounds of interest. All remaining statistical analysis were performed with this subset of 116 compounds.

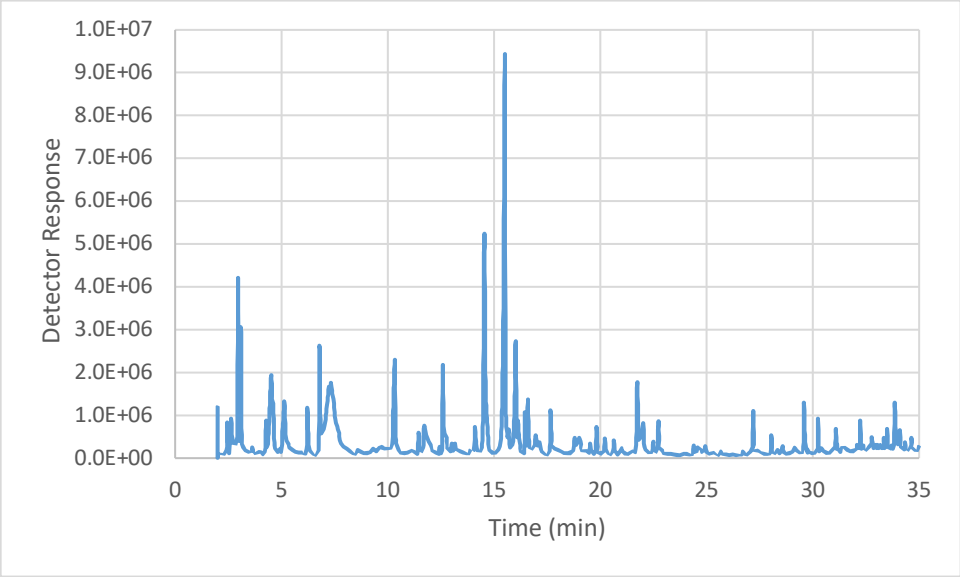


Figure 2. Example chromatogram of the Mountain Man honey.

A subset of compound which were present in all honey types was chosen for closer inspection and quantification. These compounds include 1-methyl-butanal, furfural, benzaldehyde, and octanal. These compounds the calculated response factor and average concentration (mg/L) is shown in Table 2.

Table 2: The calculated response factor and average concentration (mg/L) for 2-methyl-butanal, furfural, benzaldehyde, and octanal in the 15 honey samples.

	2-methyl-butanal	Furfural	Benzaldehyde	Octanal
	mg/L	mg/L	mg/L	mg/L
Response Factor	3.69E+01	3.98E+01	3.72E+03	7.46E+03
BRd	2.62E-05	7.82E-04	1.08E-04	2.78E-06
Gall	5.86E-05	4.12E-03	5.38E-05	1.67E-06
Gris	4.11E-05	1.04E-02	1.26E-05	4.98E-07
Kirk	2.94E-04	8.83E-03	1.51E-04	3.33E-06
LOUG	1.06E-04	2.17E-03	3.40E-05	5.34E-07
Lowe	2.05E-05	1.43E-04	1.46E-04	2.22E-06
SPALM	3.15E-05	3.07E-03	5.05E-05	9.66E-07
SS	2.15E-05	5.78E-04	3.13E-04	1.73E-06
Unc	3.78E-05	5.75E-03	4.02E-05	2.76E-06
ASUE	8.13E-05	1.24E-03	2.01E-04	1.29E-06
Brazil	1.09E-04	5.06E-04	7.83E-05	1.40E-06
Craic	3.73E-05	4.88E-03	3.56E-05	5.17E-07
WFLWR	3.05E-04	9.13E-03	8.76E-05	3.18E-06
OBLSM	2.03E-05	1.70E-03	1.08E-05	5.42E-08
UNC2	8.42E-05	1.16E-03	1.52E-04	1.27E-06
Mnt2	1.87E-03	1.30E-02	3.14E-04	3.02E-06

A principle component analysis (PCA) was conducted on the scaled data using prcomp() in R. The compounds present in most abundance were used in this analysis by using average gas chromatogram peak height. This plot produced clustering as show in the Figure 3. 34.89% of variation can be explained by the first two principle components and 8 principle components are

needed to explain at least 80% of the variation. The variation associated with each principle component is listed in Table 3.

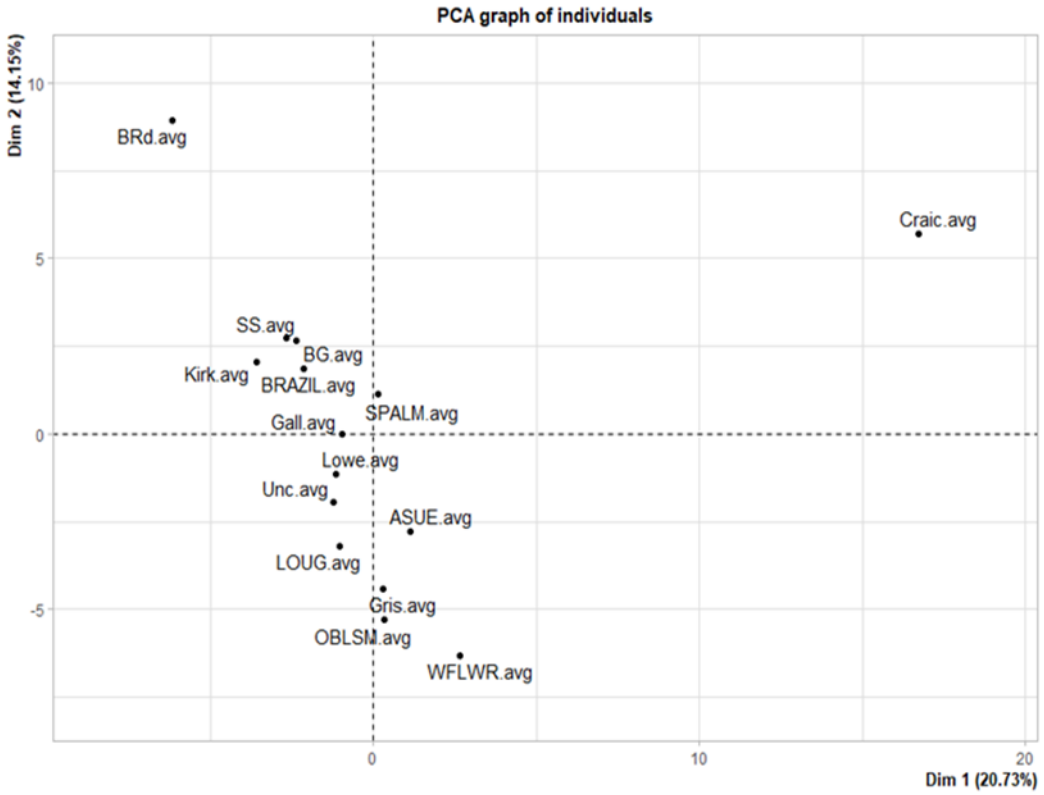


Figure 1. PCA Graph of Individual Honeys.

Table 3: Importance of Components in PCA Graph.

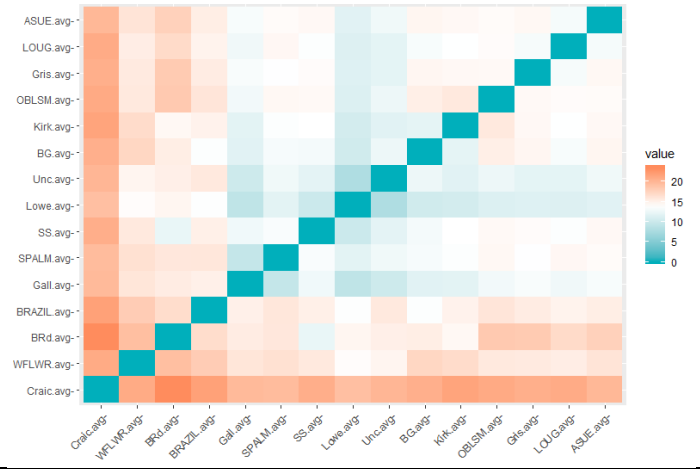
	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11	PC12	PC13	PC14
Standard deviation	4.90	4.05	3.46	3.33	2.91	2.80	2.72	2.60	2.42	2.26	1.92	1.83	1.53	1.24
Proportion of Variance	0.21	0.14	0.10	0.10	0.07	0.07	0.07	0.06	0.05	0.04	0.03	0.03	0.02	0.01
Cumulative Proportion	0.21	0.35	0.45	0.55	0.62	0.69	0.75	0.81	0.86	0.91	0.94	0.97	0.99	1.00

In addition to the PCA, cluster analysis was also completed. The measure of dissimilarity or distance between groups is of upmost importance in cluster analysis. In order to visualize this distance a series of plots were created using traditional and correlation based measures. These

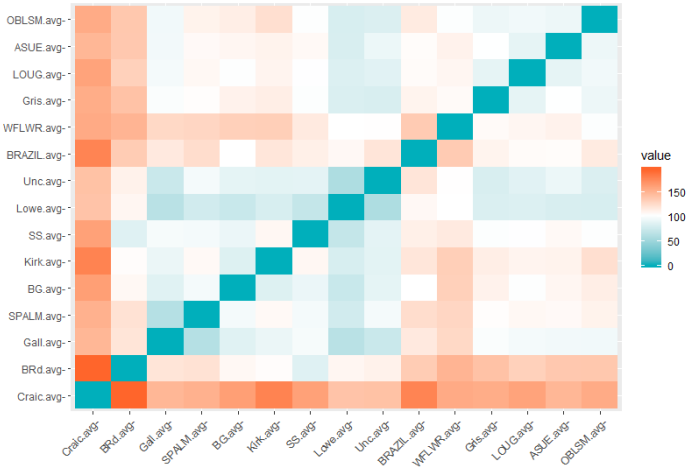
plots are shown in Figure 3. In each plot, the darker the teal, the more similar the honey profile are according to the 116 compounds used. Conversely, the darker the orange, the more dissimilar the honeys are according to the 116 compounds. It is clear that the difference between the more traditional distance and the correlation based measures. The elbow method in K-means clustering determines the optimal number of clusters by comparing the within cluster sum of squares against the number of clusters. A reduction in sum of squares suggests a desirable number of clusters. Using this method along with some different dissimilarity measures, no clear number of clusters was suggested. Clustering results for the Euclidean distance for different numbers of clusters (k) were determined and shown in Figure 4. It should be mentioned that given previous results, these groupings may not be representing strong differences/groupings in the data.

Figure 3: The measure of dissimilarity or distance between groups is of upmost importance in cluster analysis. The following plots explore six different distance measures, the last three being correlation based. In each plot, the darker the teal, the more similar the honey profiles are according to the 116 compounds. The darker the orange, the more dissimilar the honey profiles are according to the 116 compounds.

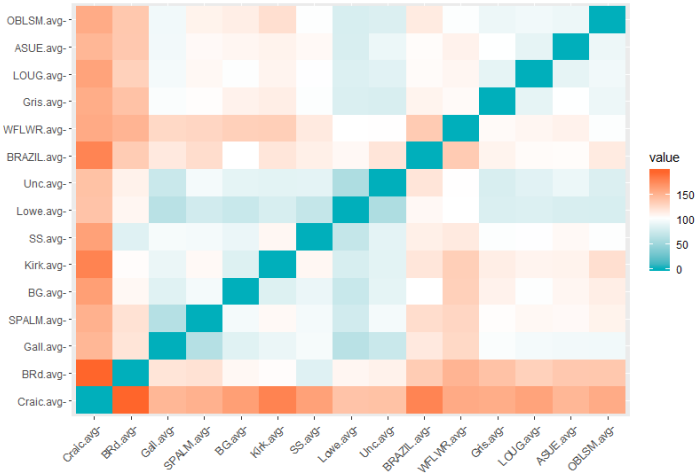
A) Distance Measure: Euclidean



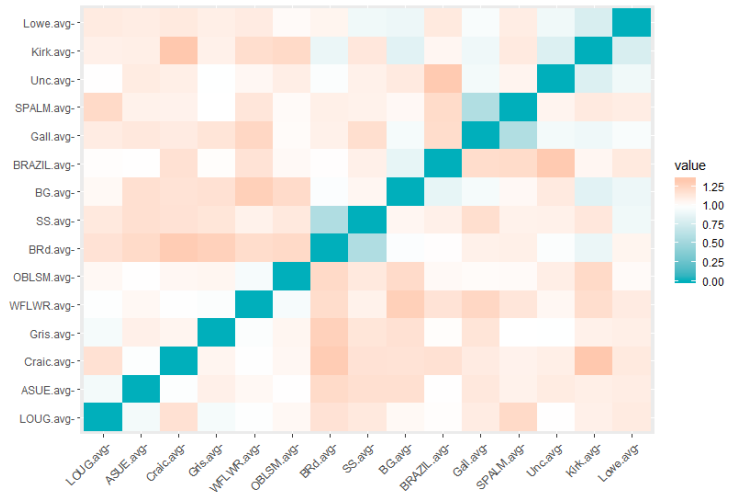
B) Distance Measure: Manhattan



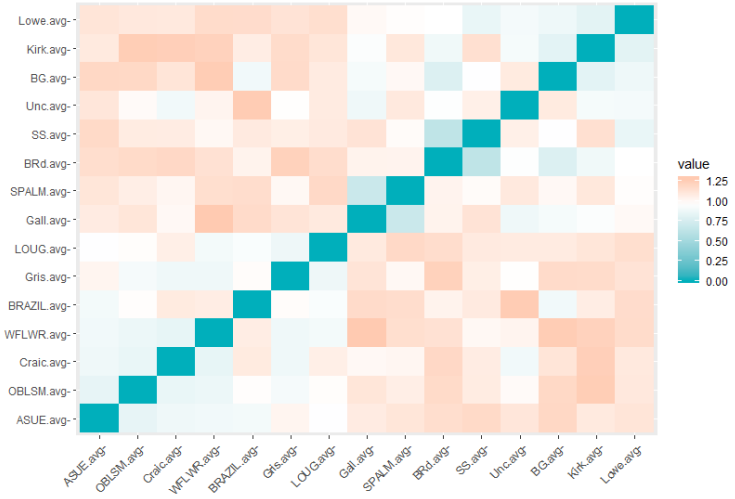
C) Distance Measure: Minowski



D) Distance Measure: Pearson



E) Distance Measure: Kendall



F) Distance Measure: Spearman

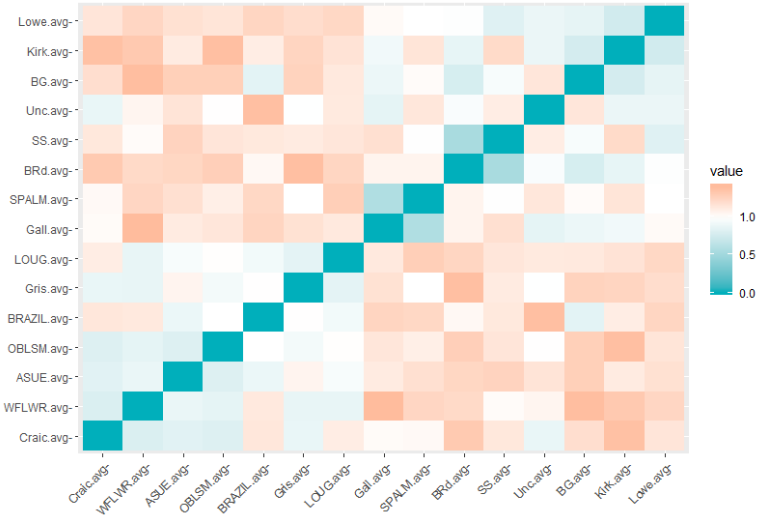
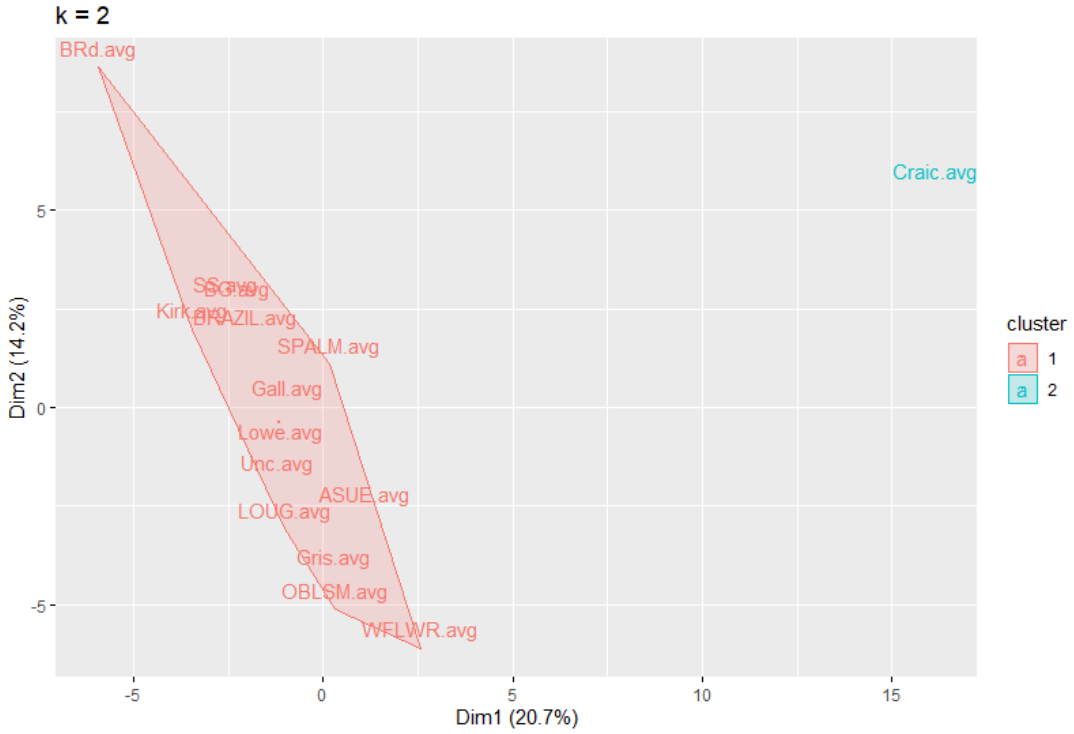
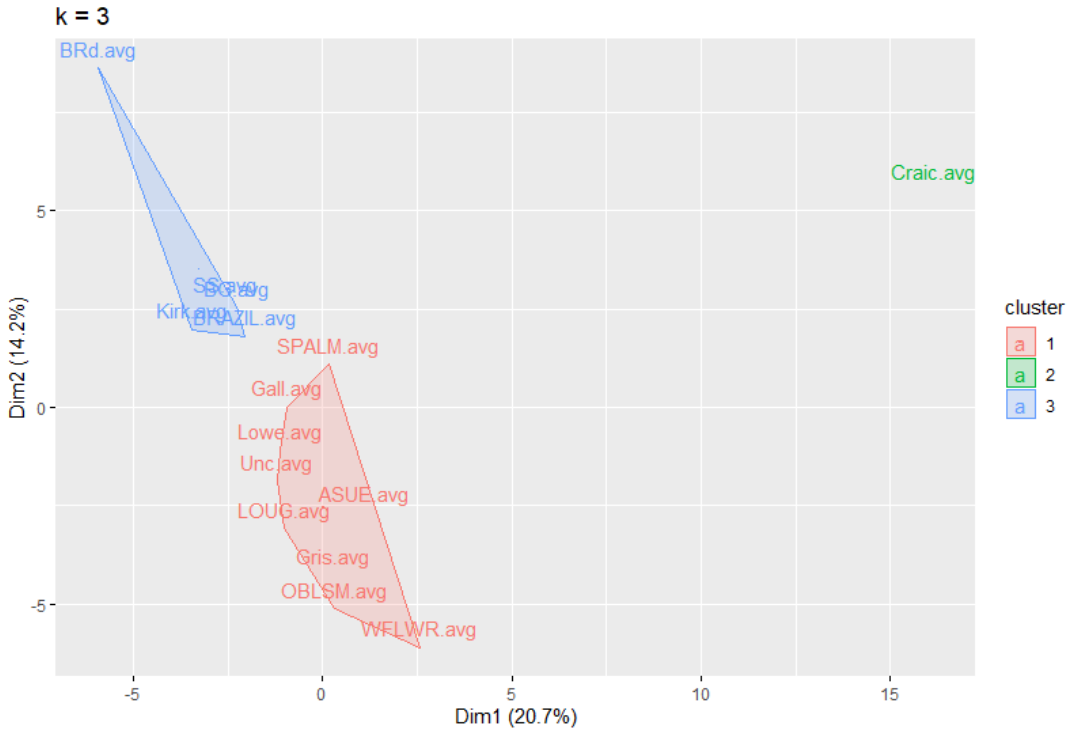
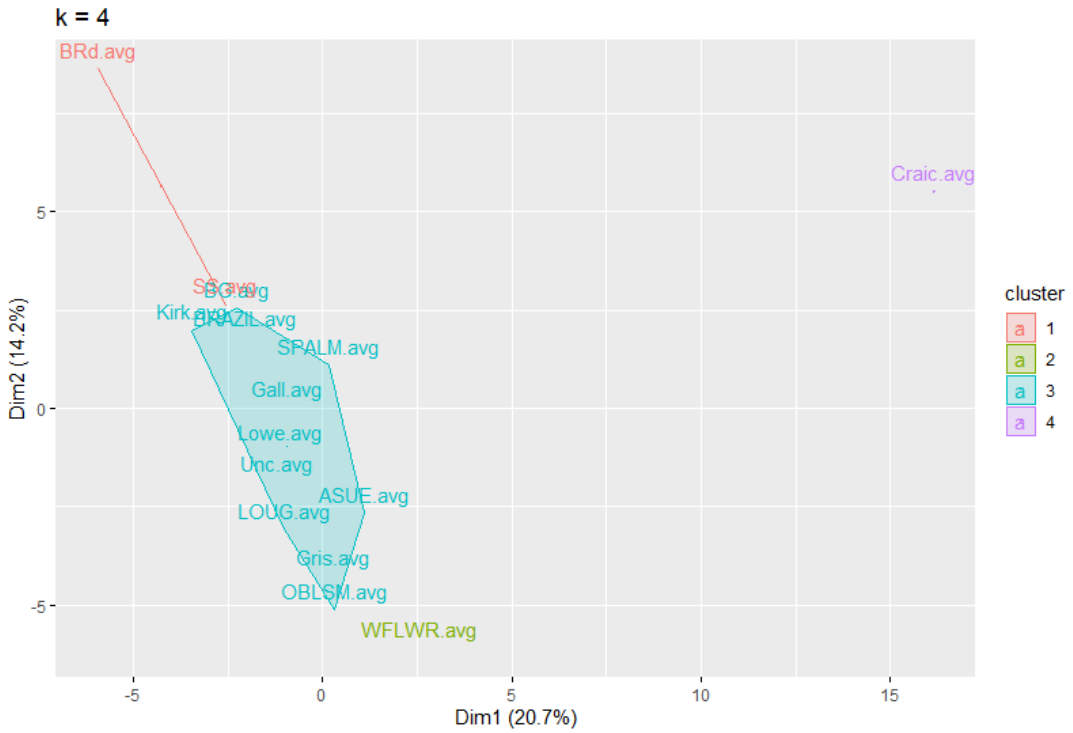
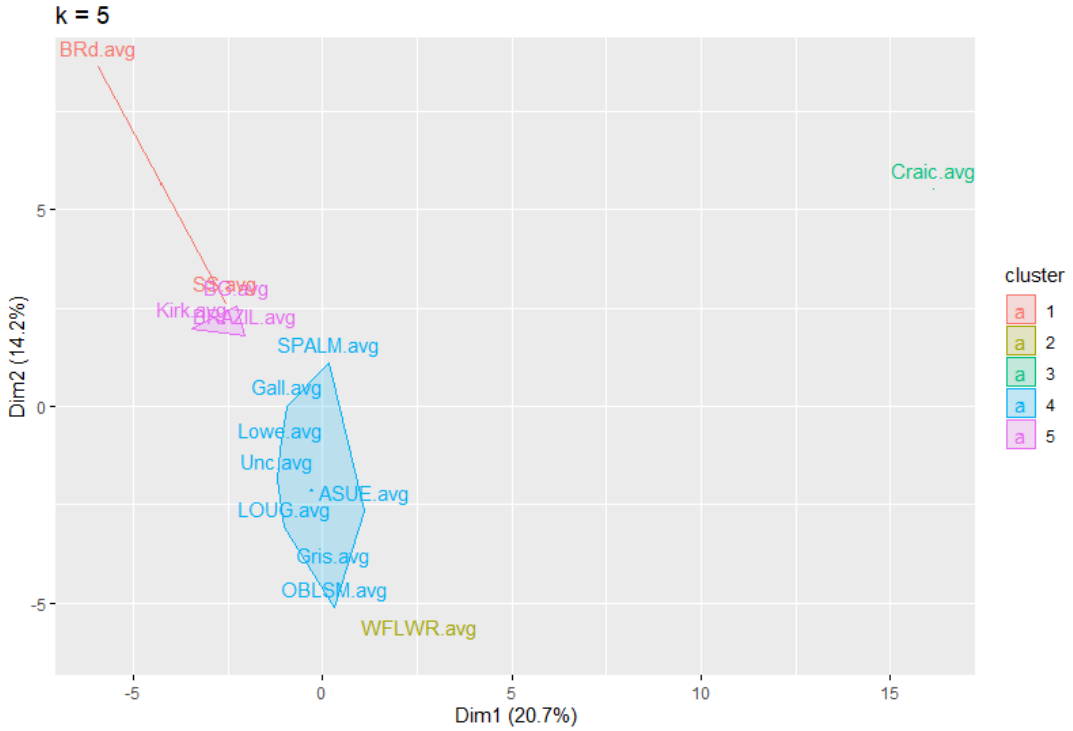


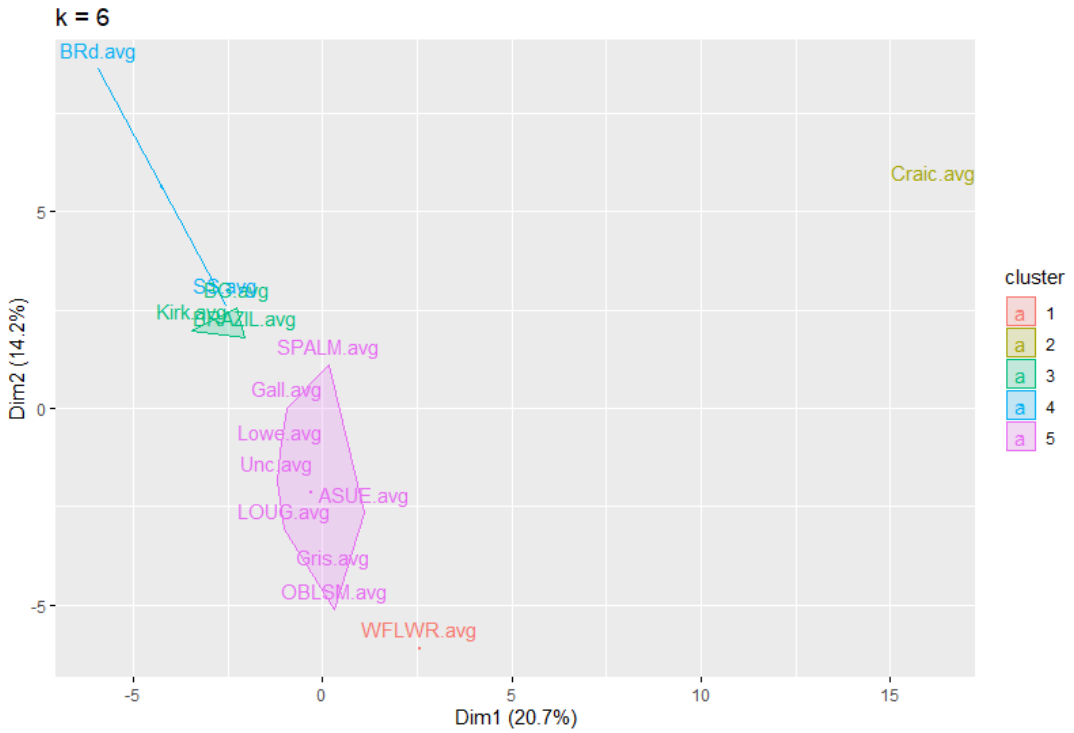
Figure 4: Clustering results for the Euclidean distance for different numbers of clusters (k), from 2 clusters to 7 clusters. Since no ideal cluster number was determined

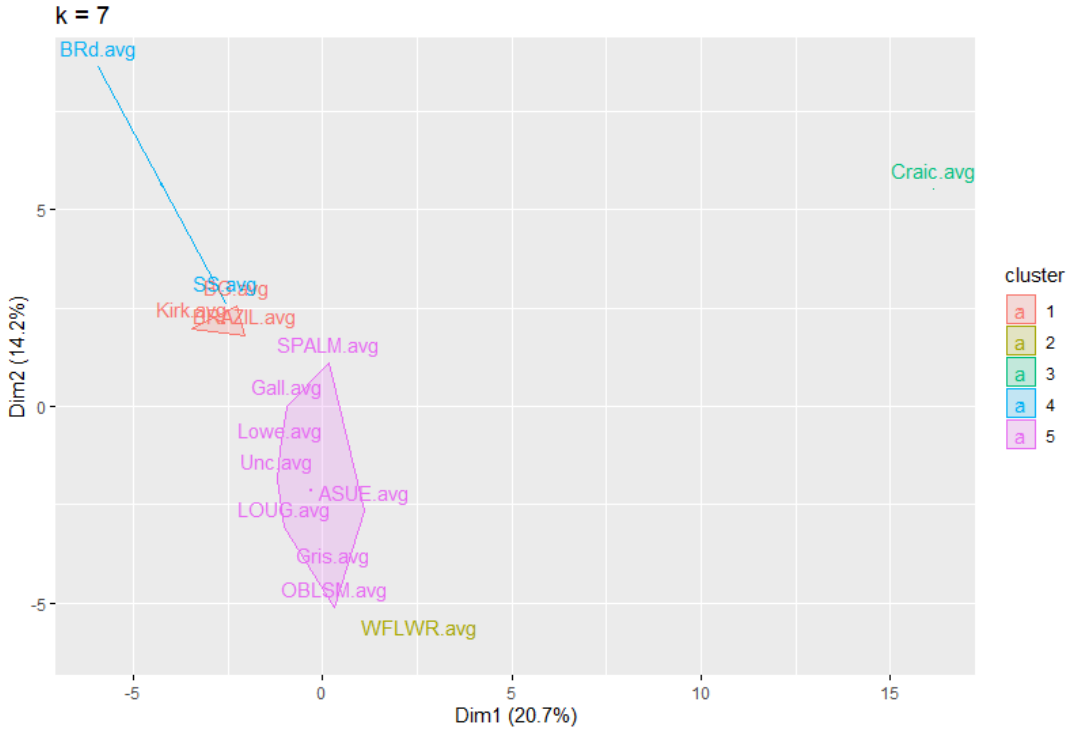





k = 4
BRd.avg
Craic.avg
SPALM.avg
Gall.avg
Lowe.avg
Unc.avg
ASUE.avg
LOUG.avg
Gris.avg
OBLSM.avg
WFLWR.avg
Dim2 (14.2%)
Dim1 (20.7%)
cluster
1
2
3
4


k = 5
BRd.avg
Craic.avg
SPALM.avg
Gall.avg
Lowe.avg
Unc.avg
ASUE.avg
LOUG.avg
Gris.avg
OBLSM.avg
WFLWR.avg
Dim2 (14.2%)
Dim1 (20.7%)
cluster
1
2
3
4
5


k = 6
BRd.avg
Craic.avg
SPALM.avg
Gall.avg
Lowe.avg
Unc.avg
ASUE.avg
LOUG.avg
Gris.avg
OBLSM.avg
WFLWR.avg
Dim2 (14.2%)
Dim1 (20.7%)
cluster
1
2
3
4
5


k = 7
BRd.avg
Craic.avg
SPALM.avg
Gall.avg
Lowe.avg
Unc.avg
ASUE.avg
LOUG.avg
Gris.avg
OBLSM.avg
WFLWR.avg
Dim2 (14.2%)
Dim1 (20.7%)
cluster
1
2
3
4
5

Discussion

The analysis showed 116 unique compounds present in at least half of the honey sampled. When looking at the selected compounds (1-methyl-butanal, furfural, benzaldehyde, and octanal) found in 100% of the samples have specific aroma profiles. Some of these are consistent with monofloral and polyfloral honeys that have heather and buckwheat profiles. (5) Furfural corresponds to a sweet profile while benzaldehyde and octanal correspond to a fruity profile. (4) Since these were all found in all of the samples, this further shows the polyfloral quality of the samples by showing that they all have a mixture of different aroma profiles. The other compounds that were found in the samples were specific to each sample and made it so each one could be differentiated. Comparison between local honeys and non-local honeys are recorded here to show differences in chemical composition in different regions based on observed concentration. This supports the idea that statistical measures of a list of compounds should allow for regionality.

When the results of the statistical analysis of the 116 compounds are examined, it is difficult to form conclusions of any meaningful clustering among these seventeen samples of honey. The study was still able to provide information nonetheless. The PCA analysis showed that the it is possible to separate compounds based on these 116 compounds present. Unfortunately, it appears that the relatively small number of honey sample, both in total and outside the local region, may be limiting the usefulness of the analysis. Within the cluster analysis, it is clear that there is a difference between the more traditional distances (Figure 3a-c) and the correlation based measures (Figure 3d-f). The first three traditional measures should be used when it is appropriate to group observations with high values of features together and low values of features together. Correlation based measures help identify clusters with similar

profiles regardless of magnitudes (ex. on/off both occur or don't). In addition, form this analysis it appears that Craic is very different from the other types of honey. Mild differences exist with BRd, WFLWR, and BRAZIL. Lowe appears to have some similarities with a group of honeys. But again, this appears to be limited by the low sample number.

This work does prove the use and the potential of using HS-SPME coupled with GC-MS to find regional markers that establish the geographic location of honeys. In the future, NMR information will also be used to further aide in cluster analysis but will need to be put on hold until the appropriate equipment is available for use. Also, additional honey samples regionally and nationally will be used. The addition of honeys will provide more data and more precise statistical analysis and could show potential for geographical clustering.

Conclusions:

This initial study was performed to investigate the ability to assign regional differences in wildflower honey though the HS-SPME coupled with GC-MS analysis of chemicals present in the aroma. 116 compounds were identified as being present in at least half the honey samples, and the quantification of 1-methyl-butanal, furfural, benzaldehyde, and octanal showed different concentration in all honey samples. The incorporation of either PCA or cluster analysis failed to produce this geographical regionality. However, with an increased number of honey sample, both locally and outside the region, and with the potential addition of information from complementary analysis using NMR it is likely that this regionality can be achieved.

References:

(1) Oddo, Livia Persano, et al. "Invertase Activity in Honey." *Apidologie*, vol. 30, no. 1, 23 Sept. 1999, pp. 57–65., doi:10.1051/apido:19990107.

(2) Karabagias, I.K., Nikolaou, C. & Karabagias, V.K. Volatile fingerprints of common and rare honeys produced in Greece: in search of PHVMs with implementation of the honey code. *Eur Food Res Technol* **245,** 23–39 (2019). https://doi.org/10.1007/s00217-018-3137-x

(3) Pino, Jorge A. "Analysis of Odour-Active Compounds of Black Mangrove (Avicennia Germinans L.) Honey by Solid-Phase Microextraction Combined with Gas Chromatography–Mass Spectrometry and Gas Chromatography–Olfactometry." *International Journal of Food Science and Technology*, 25 Feb. 2012, pp. 1688–1694.

(4) Huaixiang Tian, Yongbo Shen, Haiyan Yu & Chen Chen (2018) Aroma features of honey measured by sensory evaluation, gas chromatography-mass spectrometry, and electronic nose, *International Journal of Food Properties*, 21:1, 1755-1768, DOI: 10.1080/10942912.2016.1213744

(5) Seisonen, Sirli, et al. "Characterisation of the Aroma Profiles of Different Honeys and Corresponding Flowers Using Solid-Phase Microextraction and Gas Chromatography–Mass Spectrometry/Olfactometry." *Food Chemistry*, vol. 169, 2015, pp. 34–40., doi:10.1016/j.foodchem.2014.07.125.