separations

MDPI

Article

Detection and Comparison of Volatile Organic Compounds in Four Varieties of Hawthorn Using HS-GC-IMS

Lijun Zhu [1,†], Feilin Ou [1,†], Yun Xiang [1], Bin Wang [1], Yingchao Mao [2], Lingfeng Zhu [2], Qun Zhang [2,*] and Chang Lei [1,3,*]

[1] State Key Laboratory of Chinese Medicine Powder and Medicine Innovation in Hunan (Incubation), Science and Technology Innovation Center, Hunan University of Chinese Medicine, Changsha 410208, China; 202104030124@stu.hnucm.edu.cn (L.Z.); oufeilin@stu.hnucm.edu.cn (F.O.); 003872@hnucm.edu.cn (Y.X.); 004146@hnucm.edu.cn (B.W.)
[2] Hunan Agricultural Product Processing Institute, Hunan Academy of Agricultural Sciences, Changsha 410125, China; mycjgs@hunaas.cn (Y.M.); zhulingfeng1988@163.com (L.Z.)
[3] Yunnan Key Laboratory of Screening and Research on Anti-Pathogenic Plant Resources from Western Yunnan, Dali 671000, China
* Correspondence: zqun208@163.com (Q.Z.); leichang@hnucm.edu.cn (C.L.)
† These authors contributed equally to this work.

Abstract: Hawthorn is a type of natural food with significant medicinal and nutritional properties; it has been listed in the "Both Food and Drug" list by the Chinese Ministry of Health Item List since 1997. However, hawthorn varieties have complex origins, and there are significant differences in the content, type, and medicinal efficacy of the chemically active ingredients in different varieties of hawthorn. This leads to the phenomenon of mixed varieties and substandard products being passed off as high-quality. In this work, by using headspace gas chromatography–ion mobility spectrometry (HS-GC-IMS), we identified and analyzed volatile organic compounds (VOCs) in four varieties of hawthorn, establishing their characteristic fingerprints. As a result, a total of 153 peaks were detected, and 139 VOCs were also identified. As shown by the fingerprint profiles, the different hawthorn samples contained different VOCs. Meanwhile, by using principal component analysis (PCA), Euclidean distance, and partial least-squares discriminant analysis (PLS-DA), the relationship between the VOCs found in the different varieties of hawthorn was revealed. This study developed a simple, fast, accurate, and sensitive method for identifying, tracking, and evaluating hawthorn varieties.

Keywords: hawthorn; HS-GC-IMS; volatile organic compounds; PCA; Euclidean distance; PLS-DA

Citation: Zhu, L.; Ou, F.; Xiang, Y.; Wang, B.; Mao, Y.; Zhu, L.; Zhang, Q.; Lei, C. Detection and Comparison of Volatile Organic Compounds in Four Varieties of Hawthorn Using HS-GC-IMS. *Separations* **2024**, *11*, 100. https://doi.org/10.3390/separations11040100

Academic Editors: Anna M. Kaczmarek and Constantinos G. Tsiafoulis

Received: 8 March 2024
Revised: 25 March 2024
Accepted: 26 March 2024
Published: 28 March 2024

1. Introduction

Hawthorn is the dried ripe form of *Crataegus pinnatifida* Bge. var *major* N.E. Br., or *Crataegus pinnatifida* Bge. [1], and it is known as a "nutritious fruit" due to its richness in bioactive substances [2]. Hawthorn is not only a food product but also a medicinal plant that has many active ingredients and health benefits. In 1997, the Chinese Ministry of Health included it on its "Both Food and Drug" list, which encompasses drugs, dietary supplements, and foods. Hawthorn mainly contains flavonoids, organic acids, proanthocyanidins, triterpenes, pectin, vitamins, minerals, and other chemical components [3,4]. It has anti-atherosclerosis effects, can lower blood lipids and lower blood pressure, has antioxidant effects, improves liver damage, and has other pharmacological effects [5–7]. Due to its unique flavor and health benefits, hawthorn has become widely used in health foods, snacks, additives, and teas [8–12], such as hawthorn preserves, hawthorn pectin, hawthorn seed oil, hawthorn functional drinks, hawthorn plum, etc.

The variety of sources for hawthorns are varied and complex. In addition to Shanlihong (*Crataegus pinnatifida* Bge. Var. *major* N. E. Br.) and Shanzha (*Crataegus pinnatifida* Bge.),

recorded in the 2020 edition of the "Chinese Pharmacopoeia", the most commonly used edible hawthorn varieties in the domestic market include "wild hawthorn" (*Crataegus cuneata* Sieb. et Zicc), *Malus doumeri* (Bois.) Chev. Shanlihong, and Shanzha, which are mostly grown in the north; therefore, they are commonly called "Northern Hawthorn", whereas wild hawthorn varieties mostly grow in the south; therefore, they are commonly known as "Nanshanzha". *Malus doumeri* (Bois.) Chev mostly grows in Guangdong, Guangxi, and other places, and it is commonly known as "Guangshanzha" [13]. These four varieties differ greatly in terms of source varieties, chemical composition, and pharmacological effects due to a series of factors, e.g., their origin, harvest period, variety, growth environment, soil, fertilizer, harvest, processing, and storage. Nanshanzha promotes qi, disperses blood stasis, astringes, and stops diarrhea; Guangshanzha regulates qi, strengthens the spleen, and digests stagnant food. The content and types of organic acid components in Northern Hawthorn, Nanshanzha, and Guangshanzha are very different. Using Nanshanzha and Guangshanzha instead of Northern Hawthorn will have a greater impact on research results and efficacy. The quality of the different varieties of Hawthorn is irregular, and so the market prices vary. This has led to common phenomena such as mixing production areas and passing off substandard products, resulting in greatly reduced efficacy and possibly even drug use and food safety accidents. However, in the 2020 edition, the Chinese Pharmacopoeia only uses acid–base titrations and indicators to evaluate the quality of hawthorn.

This method is not suitable for the qualitative and quantitative analysis of organic acid components, and it cannot reflect the overall quality of the medicinal materials and thus cannot be used for a reasonable evaluation; this makes it impossible to carry out a reasonable quality evaluation and control. Some researchers have also used empirical identification, microscopic identification, high-performance liquid chromatography (HPLC), molecular identification (such as DNA barcoding technology), and other techniques to identify hawthorn [14,15]. Although these methods provide more choices for identification, these approaches have their limitations. For example, an empirical identification is subjective, microscopic identification is not specific, HPLC identification is time-consuming and cumbersome, and molecular identification is relatively complex. Therefore, it is important to find a fast, easy, and green way to differentiate different varieties of hawthorn.

In addition to identifying food, volatile organic compounds (VOCs) also play an important role in determining a food's nutritional and sensory properties. At present, analyses of VOCs in food comprise two parts: sensory analysis and instrumental analysis. The sensory analysis of results is a kind of subjective sensory perception that analyzes the results. An instrumental analysis is a molecular analysis that is objective. However, the traditional odor and taste identification methods, which are highly subjective and empirical, can no longer match modern development. Moreover, most flavor substances and volatile odor substances are in the ppb range in food. Therefore, the detection of differentiation needs to be more sensitive and reliable. In the past, more and more researchers have used instrumental analysis techniques to detect VOCs in food, including gas chromatography–mass spectrometry (GC-MS), chromatography–olfactometry–mass spectrometry (GC-O-MS), headspace gas chromatography–ion mobility spectrometry (HS-GC-IMS), and electronic noses (E-noses) [16]. These methods can complement sensory evaluations by providing more objective information about the substance being tested at the molecular level. There are a number of gas chromatography and mass spectrometry methods available, but GC-MS has significant limitations in distinguishing isomeric molecules, especially cyclic isomers, due to its complex sample processing methods and high mass spectrometry resolution [17]. Aroma and flavor analysis can also be performed using GC-O-MS, but it shares the same disadvantages as GC-MS [18]. An E-nose is a new type of aroma detection technology that provides rapid detection, but it also suffers from low accuracy, sensor drift, and poor repeatability, as well as a high sensitivity to the surrounding environment [19].

A recent development in HS-GC-IMS technology incorporates the advantages of both GC and IMS, combining high separation capabilities with high resolution and high

sensitivity. It is equipped with a static headspace sampling device to detect trace organic components emitted from liquid or solid samples. As an emerging technology, this method is simple, fast, non-destructive, there is no need for sample pre-processing, and it has good reproducibility. It is used in the identification of foods, medicinal varieties [20,21], and food aroma analysis [22]. The impact of different drying methods and temperatures on food [23,24], component analyses of food during different harvest periods and storage periods [25,26], and other aspects have played an important role in this method. There have been very few studies conducted on HS-GC-IMS for the identification of hawthorn, and this difference is often one of the key factors contributing to the differences in quality between traditional Chinese medicines. The aim of this study was to identify and analyze the VOCs of different varieties of hawthorn using HS-GC-IMS, principal component analysis (PCA), Euclidean distance, and partial least-squares discriminant analysis (PLS-DA) methods. The characteristic VOCs in the different varieties of hawthorn were displayed in a visual form, providing technical support for the rapid analysis of VOCs and the identification of different varieties of hawthorn. Simultaneously, this study enriched the study of flavor compounds of hawthorn.

2. Materials and Methods

2.1. Medicinal Materials

Four kinds of hawthorn powder were purchased from the National Institute for Food and Drug Control, Beijing, China: Shanlihong (the dried ripe fruit of *Crataegus pinnatifida* Bge. Var. *major*, No. 121138-201206, named SZ-01), Shanzha (the dried ripe fruit of *Crataegus pinnatifida* Bge., No. 121626-201803, named SZ-02), Guangshanzha (the dried ripe fruit of *Cantonese Crataegus*, No. 120943-201903, named GSZ), and Nanshanzha (the dried ripe fruit of *South Crataegus*, No. 121055-201704, named NSZ).

2.2. Sample Preparation

For each kind of hawthorn powder, 1 g was accurately weighed and placed into a 20 mL headspace bottle and incubated at 80 °C for 20 min, and then the samples were injected. Each sample was measured in three parallel groups.

2.3. Headspace Sampling Conditions

The temperature of incubation was 80 °C. The incubation time was 20 min. The injection volume was 500 µL. The splitless injection method was used. The incubation speed was 500 rotations per min, and the injection needle temperature was 85 °C.

2.4. Chromatographic Conditions

In this study, we used a FlavourSpec® gas-phase ion mobility spectrometer (G.A.S, Dortmund, Germany), a CTC-PAL 3 static headspace automatic sampling device (CTC Analytics AG, Switzerland), and a 20 mL headspace bottle (Shandong Haineng Scientific Instrument Co., Ltd., Shandong, China); an MXT-WAX capillary chromatography column (15 m × 0.53 mm × 1 µm, Restek Company of the United States, Bellefonte, PA, USA) was also used. The temperature was 60 °C. The carrier gas was high-purity N_2 (purity ≥99.999%). The initial flow rate was 2.00 mL/min, which was maintained for 2 min. Then, it was linearly increased to 10.00 mL/min and then 100.00 mL/min within 10 min, and this was maintained for 40 min. The running time for the chromatography was 60 min, and the injection port temperature was 80 °C.

2.5. IMS Conditions

A tritium source (3H) was used; the flow rate was 75 mL/min; the power of the electric field was 500 V/cm; the voltage of the drift gas was 99.999%; the length of the migration tube was 53 mm; the electric field strength was 500 V/cm; the temperature of the migration tube was 45 °C; and the drift gas was high-purity nitrogen (purity = 99.999%).

2.6. Data Analysis

In order to characterize the VOCs in the samples, the GAS software (G.A.S, Dortmund, Germany, version 2.0.0), the built-in NIST retention index database, and the IMS drift time database were used. To compare VOCs between the samples, VOCal data processing software plug-ins, such as Reporter (Version 11.x), Gallery Plot (Version 1.1.0.2), and Dynamic PCA (version 0.0.3), were used to generate three-dimensional (3D) spectrums, two-dimensional (2D) spectrums, difference spectrums, fingerprints, and PCA maps, respectively. The PLS-DA VIPs were calculated using the SIMCA software (version 14.0) (Umea, Sweden).

3. Results and Discussion

3.1. GC-IMS Analysis of VOCs in Different Samples

We compared the spectral differences between the samples and the characteristic VOCs of the Shanlihong, Shanzha, Guangshanzha, and Nanshanzha samples using the VOCal software's Reporter plug-in. The X axis represents the ion migration time, the Y axis represents the retention time of the gas chromatograph, and the Z axis represents the peak intensity used for quantification. A comparison of the VOC contents of the Shanlihong, Shanzha, Guangshanzha, and Nanshanzha varieties is shown in Figure 1. In Figure 2, a top view is used for comparison due to the inconvenience of observation. After normalizing the data, the red vertical line at 1.0 represents the reactive ion peak (RIP peak). The gas chromatography retention time is represented by the ordinate. The dots on either side of the RIP represent the VOCs. The color represents the concentration of the substance; white represents a lower concentration, and red represents a higher concentration. In general, the darker the color, the greater the concentration. The samples from the Shanlihong, Shanzha, Guangshanzha, and Nanshanzha varieties differed in terms of their VOCs, as shown in Figure 2.

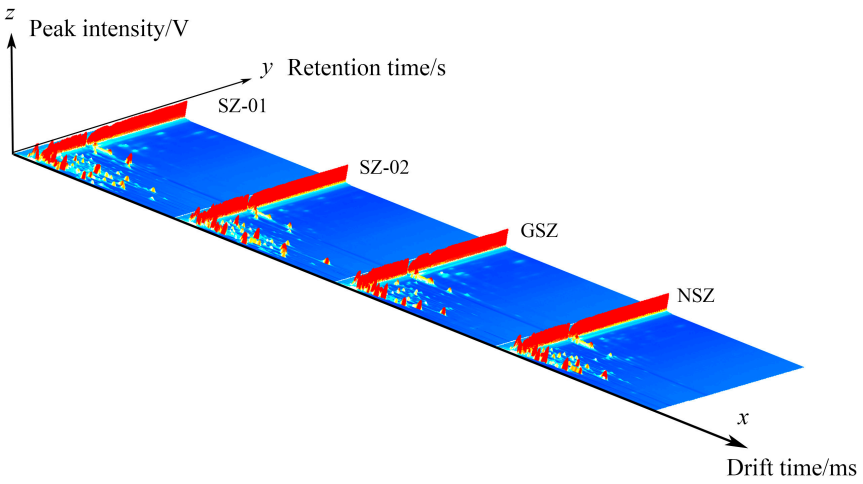

Figure 1. Three-dimensional spectra of hawthorn VOCs.

Using the spectrum of the Shanlihong (SZ-01) variety as a reference, we can visually compare the differences in the VOCs. From the spectra of the Shanzha, Guangshanzha, and Nanshanzha samples, we were able to determine their VOCs. The differences in organic matter are shown in Figure 3. In the figure, when the measured components of Shanlihong, Shanzha, Guangshanzha, and Nanshanzha were consistent, they are shown in white after deduction. Red and blue indicate that the concentrations of the measured components in the Shanzha, Guangshanzha, and Nanshanzha varieties were higher and lower than in the Shanlihong variety, respectively. With the difference comparison chart, it is easier to identify the differences between the four samples.

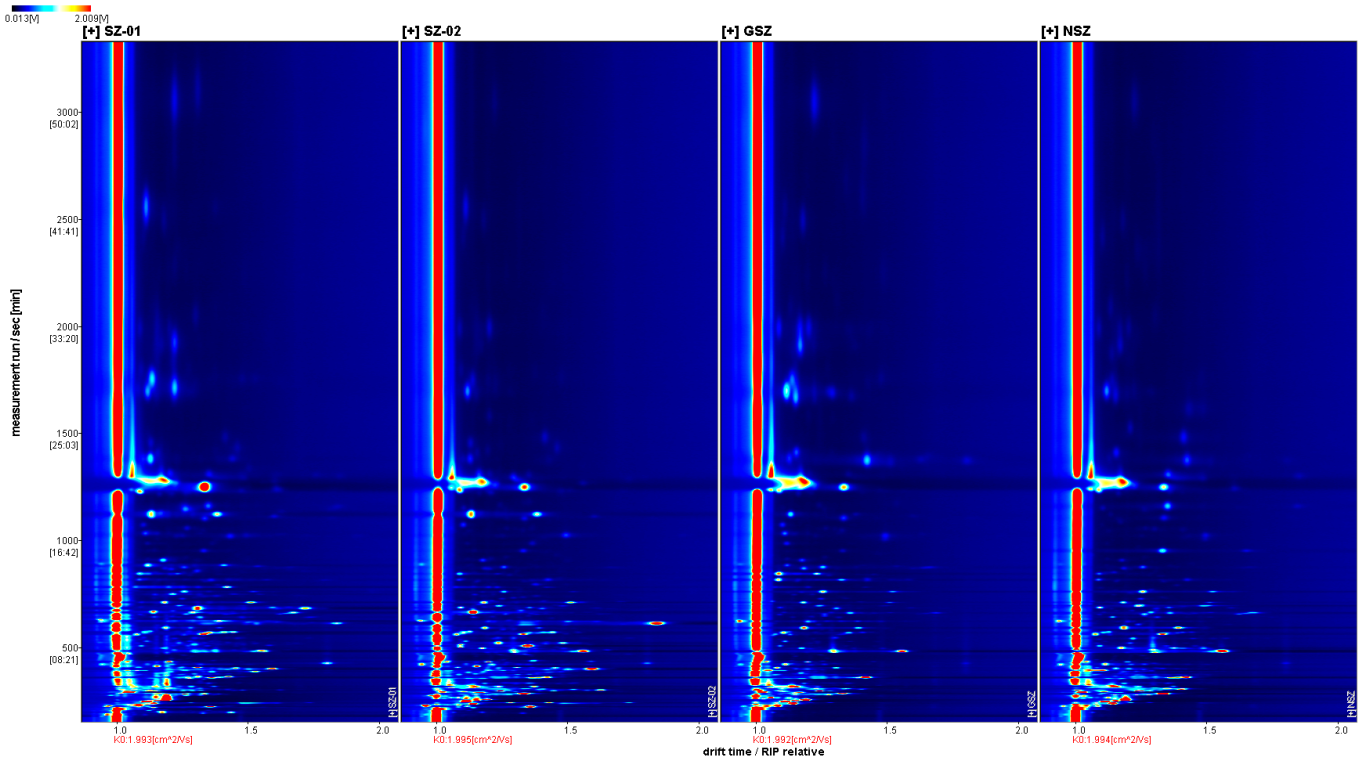

Figure 2. Two-dimensional spectra of hawthorn VOCs.

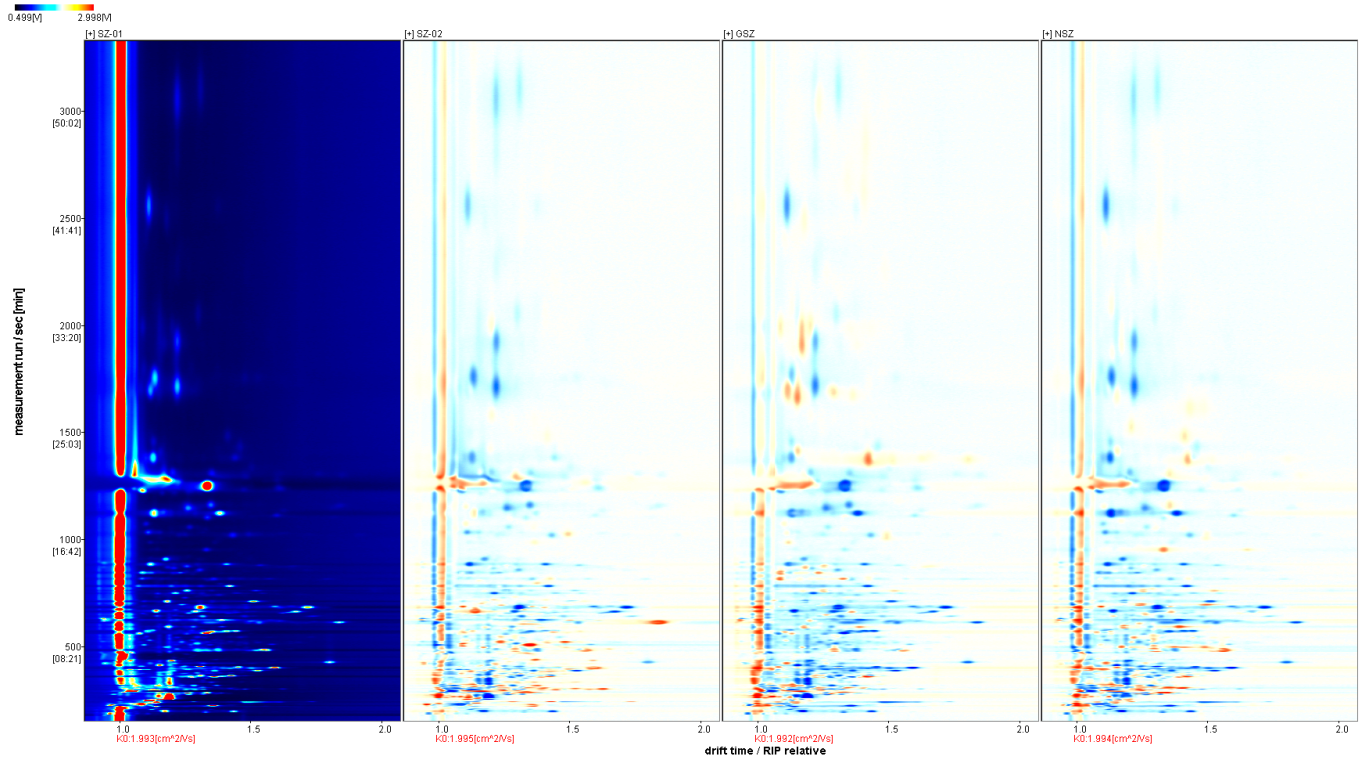

Figure 3. A comparison chart of the differences between varieties of hawthorn.

3.2. Comparison of the Fingerprints of VOCs

To compare the specific VOCs from the hawthorn samples, the Gallery Plot plug-in of VOCal was used to select the peaks for fingerprint comparison (Figure 4). Each row in the figure represents a sample; the rightmost side is the name of the sample, each column

is a compound, and the bottom is the qualitative result of the samples. Some substances are followed by M and D, which indicate monomers and dimers of the same substance, respectively. The color indicates the concentration of the compounds, with darker colors indicating a higher concentration of the compounds, and black colors indicating a very low or undetectable concentration.

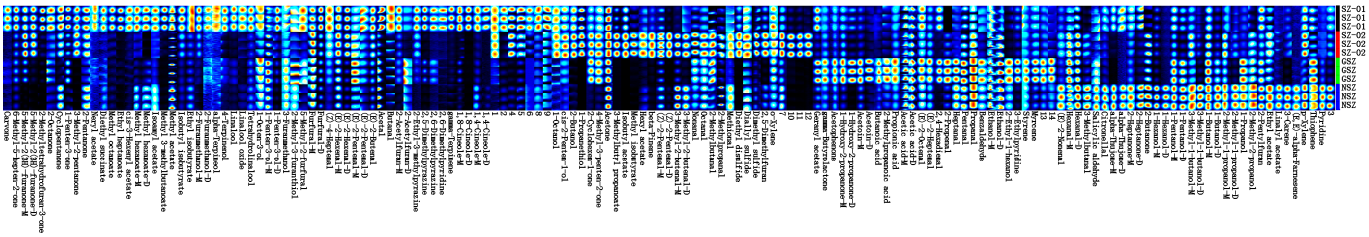

Figure 4. Fingerprints of VOCs of different varieties of hawthorn.

A comparison of the VOCs of the different hawthorn samples is shown in Figure 4. The contents of carvone, 6-methyl-5-hepten-2-one, 5-methyl-2(3H)-furanone, 2-methyltetrahydrofuran-3-one, 2-octanone, cyclopentanone, 1-penten-3-one, 3-methyl-2-pentanone, 2-pentanone, neryl acetate, diethyl succinate, methyl octanoate, ethyl enanthate, phytyl acetate, methyl caproate, isopentyl acetate, methyl 3-methylbutyrate, methyl acetate, isobutyl isobutyrate, ethyl isobutyrate, 2-furanmethanol, α-terpineol, 4-terpineol, linalool, linalool oxide, tetrahydrolinalool, 1-octen-3-ol, 1-Penten-3-ol, 3-furanmethanol, 2-methyl-3-furanthiol, 5-methyl-2-furfural, furfural, Z-4-heptenal, E-2-hexenal, E-2-pentenal, E-2-butenal, acetal, butyraldehyde, 2-acetylfuran, 2-ethyl-3-methylpyrazine, 2,5-dimethylpyrazine, 2,6-dimethylpyrazine, 2,6-dimethylpyridine, γ-terpinene, 1,8-cineole, 1,4-cineole, and 46 other substances were high in SZ-01.

The contents of 1-octanol, Z-2-penten-1-ol, 2-butanol, 1-propanethiol, 2-cyclohexen-1-one, 4-methyl-3-pentene-2-one, acetone, isoamyl propionate, isobutyl acetate, methyl isobutyrate, hexyl acetate, β-pinene, Z-2-pentenal, 3-methyl-2-butene aldehyde, nonanal, octanal, 2-methylbutyraldehyde, 2-methylpropionaldehyde, methylthiopropionaldehyde, diethyl disulfide, diallyl sulfide, dimethyl sulfide, 2,5-dimethylfuran, o-xylene, and 24 other substances were high in SZ-02.

The contents of geranyl acetate, γ-butyrolactone, acetophenone, 1-hydroxy-2-propanone, acetoin, butyric acid, 2-methyl propionic acid, propionic acid, acetic acid, E-2-octenal, E-2-heptenal, 2,4-hexadienal, 2-propenal, heptanal, valeraldehyde, propionaldehyde, benzaldehyde, ethanol, 2-ethyl-1-hexanol, 3-ethylpyridine, styrene, myrcene, and 22 other substances were high in GSZ.

The contents of E-2-nonenal, hexanal, 3-methyl butyraldehyde, salicylaldehyde, citronellal, α-thujone, 2-heptanone, 2-butanone, 1-hexanol, 1-pentanol, 3-methyl-1-butanol, 1-butanol alcohol, 2-methyl-1-propanol, 1-propanol, 2-methyl-2-propanol, 2-pentylfuran, ethyl acetate, linalyl acetate, 3-carene, (E, E), α-farnesene, p-xylene, thiophene, pyridine, and 23 other substances were high in NSZ.

3.3. Identifying VOCs in Different Samples

Our qualitative analysis was conducted using the NIST and IMS databases built into the VOCal software (Version 0.4.03, GAS Deutschland, Dortmund, Germany), as well as the Reporter, Gallery Plot, and Dynamic PCA plugins. Ultimately, we detected 153 peaks and identified 139 VOCs (Tables 1 and 2) from all the samples, including 31 aldehydes, 29 alcohols, 20 ketones, 20 esters, 16 terpenes, 5 acids, 4 furans, 3 pyridines, 3 pyrazines, 1 thiophene, and 7 other compounds.

Table 1. Results of the VOC analysis of different varieties of hawthorn.

No	Compounds	CAS	Molecular Formula	RI	Rt/s	Dt/ms	Comment
1	Carvone	C99490	$C_{10}H_{14}O$	1918.9	3096.478	1.31142	
2	Geranyl acetate	C105873	$C_{12}H_{20}O_2$	1911.1	3045.325	1.22205	
3	Neryl acetate	C141128	$C_{12}H_{20}O_2$	1866.5	2765.327	1.22205	
4	2-Furanmethanol	C98000	$C_5H_6O_2$	1829	2549.945	1.1134	Monomers
5	2-Furanmethanol	C98000	$C_5H_6O_2$	1829.5	2552.637	1.37275	Dimers
6	Acetophenone	C98862	C_8H_8O	1817.6	2488.023	1.17824	
7	alpha-Terpineol	C98555	$C_{10}H_{18}O$	1771.8	2253.795	1.2238	
8	Diethyl succinate	C123251	$C_8H_{14}O_4$	1727.1	2046.253	1.3028	
9	gamma-Butyrolactone	C96480	$C_4H_6O_2$	1713.6	1987.498	1.08941	
10	4-Terpinenol	C562743	$C_{10}H_{18}O$	1697.2	1917.95	1.2218	
11	5-Methyl-2-furfural	C620020	$C_6H_6O_2$	1656.9	1758.178	1.13623	
12	Propionic acid	C79094	$C_3H_6O_2$	1639.3	1692.39	1.11686	
13	Linalool	C78706	$C_{10}H_{18}O$	1646.4	1718.705	1.2218	
14	(E)-2-Nonenal	C18829566	$C_9H_{16}O$	1577.8	1481.866	1.41555	
15	2-Acetylfuran	C1192627	$C_6H_6O_2$	1546.6	1385.272	1.12833	Monomers
16	2-Acetylfuran	C1192627	$C_6H_6O_2$	1546.6	1385.29	1.44692	Dimers
17	1-Octanol	C111875	$C_8H_{18}O$	1562.1	1432.612	1.46191	
18	Acetic acid	C64197	$C_2H_4O_2$	1505.1	1266.629	1.05757	Monomers
19	Acetic acid	C64197	$C_2H_4O_2$	1505.6	1267.84	1.15537	Dimers
20	Furfural	C98011	$C_5H_4O_2$	1494.8	1238.763	1.08968	Monomers
21	Furfural	C98011	$C_5H_4O_2$	1494.8	1238.763	1.33199	Dimers
22	alpha-Thujone	C546805	$C_{10}H_{16}O$	1464.9	1161.223	1.35243	Monomers
23	alpha-Thujone	C546805	$C_{10}H_{16}O$	1464.4	1160.011	1.85457	Dimers
24	Linalool oxide	C60047178	$C_{10}H_{18}O_2$	1460.6	1150.319	1.26339	
25	5-Methyl-2(3H)-furanone	C591128	$C_5H_6O_2$	1450.2	1124.876	1.13055	Monomers
26	5-Methyl-2(3H)-furanone	C591128	$C_5H_6O_2$	1450.1	1124.55	1.38267	Dimers
27	(E)-2-Octenal	C2548870	$C_8H_{14}O$	1441.7	1104.36	1.33829	
28	2-Ethyl-3-methylpyrazine	C15707230	$C_7H_{10}N_2$	1445.7	1113.924	1.60293	
29	2-Cyclohexen-1-one	C930687	C_6H_8O	1428.1	1072.483	1.09996	
30	Tetrahydrolinalool	C78693	$C_{10}H_{22}O$	1422.1	1058.669	1.2709	
31	2,4-Hexadienal	C142836	C_6H_8O	1412.3	1036.355	1.10818	
32	Methyl octanoate	C111115	$C_9H_{18}O_2$	1405.1	1020.416	1.43527	
33	1-Hexanol	C111273	$C_6H_{14}O$	1373.7	953.473	1.33336	Monomers
34	6-Methyl-5-hepten-2-one	C110930	$C_8H_{14}O$	1352	909.719	1.17836	
35	2,5-Dimethylpyrazine	C123320	$C_6H_8N_2$	1340	886.579	1.12013	
36	Ethyl heptanoate	C106309	$C_9H_{18}O_2$	1340.4	887.259	1.4113	
37	cis-2-Penten-1-ol	C1576950	$C_5H_{10}O$	1344.3	894.746	0.94228	
38	(E)-2-Heptenal	C18829555	$C_7H_{12}O$	1335.8	878.412	1.25863	
39	1-Hydroxy-2-propanone	C116096	$C_3H_6O_2$	1319.7	848.466	1.07763	Monomers
40	1-Hydroxy-2-propanone	C116096	$C_3H_6O_2$	1319.3	847.786	1.23187	Dimers
41	cis-3-Hexenyl acetate	C3681718	$C_8H_{14}O_2$	1325.6	859.355	1.31057	
42	Acetoin	C513860	$C_4H_8O_2$	1304.2	820.562	1.08236	Monomers
43	2-Octanone	C111137	$C_8H_{16}O$	1302.3	817.159	1.36093	
44	2-Methyltetrahydrofuran-3-one	C3188009	$C_5H_8O_2$	1283.6	785.853	1.42389	
45	2,6-Dimethylpyrazine	C108509	$C_6H_8N_2$	1337.9	882.495	1.5372	
46	1-Pentanol	C71410	$C_5H_{12}O$	1270.7	765.435	1.2602	Monomers
47	1-Pentanol	C71410	$C_5H_{12}O$	1270.7	765.435	1.51989	Dimers
48	gamma-Terpinene	C99854	$C_{10}H_{16}$	1262.3	752.504	1.21456	
49	2,6-Dimethylpyridine	C108485	C_7H_9N	1261.8	751.823	1.44749	
50	2-Pentylfuran	C3777693	$C_9H_{14}O$	1246.9	729.364	1.25233	
51	(E)-2-Hexenal	C6728263	$C_6H_{10}O$	1235.8	713.03	1.18308	Monomers
52	(E)-2-Hexenal	C6728263	$C_6H_{10}O$	1235.8	713.03	1.51202	Dimers
53	3-Methyl-1-butanol	C123513	$C_5H_{12}O$	1225.9	698.738	1.2366	Monomers
54	1,8-Cineole	C470826	$C_{10}H_{18}O$	1217.1	686.342	1.30893	Monomers
55	1,8-Cineole	C470826	$C_{10}H_{18}O$	1217.6	687.087	1.71988	Dimers
56	Diethyl disulfide	C110816	$C_4H_{10}S_2$	1204.2	668.466	1.14134	
57	Methyl hexanoate	C106707	$C_7H_{14}O_2$	1203.6	667.721	1.2929	Monomers

Table 1. *Cont.*

No	Compounds	CAS	Molecular Formula	RI	Rt/s	Dt/ms	Comment
58	2-Heptanone	C110430	$C_7H_{14}O$	1198.1	660.273	1.26375	Monomers
59	Methyl hexanoate	C106707	$C_7H_{14}O_2$	1203.1	666.976	1.6747	Dimers
60	1,4-Cineole	C470677	$C_{10}H_{18}O$	1189	645.376	1.32059	Monomers
61	1,4-Cineole	C470677	$C_{10}H_{18}O$	1189	645.376	1.73154	Dimers
62	1-Penten-3-ol	C616251	$C_5H_{10}O$	1181.1	628.245	0.93878	Monomers
63	1-Penten-3-ol	C616251	$C_5H_{10}O$	1181.1	628.245	1.37013	Dimers
64	3-Methylbutyl propanoate	C105680	$C_8H_{16}O_2$	1175.2	615.582	1.84229	
65	1-Butanol	C71363	$C_4H_{10}O$	1165.1	594.727	1.18798	Monomers
66	1-Butanol	C71363	$C_4H_{10}O$	1165.5	595.472	1.39054	Dimers
67	Diallyl sulfide	C592881	$C_6H_{10}S$	1157.3	579.085	1.13114	
68	(E)-2-Pentenal	C1576870	C_5H_8O	1154.4	573.307	1.10787	Monomers
69	p-Xylene	C106423	C_8H_{10}	1149.6	563.911	1.07888	
70	Cyclopentanone	C120923	C_5H_8O	1151.1	566.878	1.33287	
71	(E)-2-Pentenal	C1576870	C_5H_8O	1154.4	573.262	1.35804	Dimers
72	4-Methyl-3-penten-2-one	C141797	$C_6H_{10}O$	1133	532.755	1.11615	
73	beta-Pinene	C127913	$C_{10}H_{16}$	1128.1	523.853	1.239	
74	Isoamyl acetate	C123922	$C_7H_{14}O2$	1124.2	516.93	1.29836	
75	(Z)-2-Pentenal	C1576869	C_5H_8O	1120.6	510.501	1.09683	Monomers
76	(Z)-2-Pentenal	C1576869	C_5H_8O	1121.1	511.49	1.34943	Dimers
77	2-Methyl-1-propanol	C78831	$C_4H_{10}O$	1114.8	500.61	1.18379	Monomers
78	Hexanal	C66251	$C_6H_{12}O$	1105.2	484.291	1.29698	Monomers
79	Hexanal	C66251	$C_6H_{12}O$	1106.1	485.774	1.56477	Dimers
80	(E)-2-Butenal	C123739	C_4H_6O	1071.3	437.31	1.19897	
81	1-Propanol	C71238	C_3H_8O	1061.4	424.947	1.11753	
82	2-Butanol	C78922	$C_4H_{10}O$	1054.5	416.54	1.15342	
83	1-Penten-3-one	C1629589	C_5H_8O	1049.1	410.111	1.30802	
84	Isobutyl isobutyrate	C97858	$C_8H_{16}O_2$	1065	429.398	1.80496	
85	3-Methyl-2-pentanone	C565617	$C_6H_{12}O$	1046.5	407.144	1.45296	
86	Isobutyl acetate	C110190	$C_6H_{12}O_2$	1044	404.177	1.59514	
87	Methyl 3-methylbutanoate	C556241	$C_6H_{12}O_2$	1032.4	390.824	1.52612	
88	Pentanal	C110623	$C_5H_{10}O$	1010.2	366.592	1.42812	
89	2-Pentanone	C107879	$C_5H_{10}O$	1005.9	362.141	1.36462	
90	Acetal	C105577	$C_6H_{14}O_2$	914.3	294.39	1.02919	
91	2,5-Dimethylfuran	C625865	C_6H_8O	954.4	321.337	1.37323	
92	Ethyl acetate	C141786	$C_4H_8O_2$	903.7	287.695	1.33952	
93	Methyl acetate	C79209	$C_3H_6O_2$	862.7	263.024	1.19002	
94	Acetone	C67641	C_3H_6O	841.7	251.249	1.11087	
95	Propanal	C123386	C_3H_6O	824	241.717	1.14459	
96	Butanal	C123728	C_4H_8O	890.1	279.284	1.26624	
97	1-Propanethiol	C107039	C_3H_8S	794.4	226.578	1.35711	
98	Dimethyl sulfide	C75183	C_2H_6S	800	229.381	0.96577	
99	2-Propenal	C107028	C_3H_4O	872.3	268.631	1.05811	
100	2-Methylbutanal	C96173	$C_5H_{10}O$	912.6	293.302	1.16364	
101	2-Butanone	C78933	C_4H_8O	922.3	299.57	1.24652	
102	Ethanol	C64175	C_2H_6O	951.8	319.548	1.05719	Monomers
103	Ethanol	C64175	C_2H_6O	952.3	319.865	1.12819	Dimers
104	Methyl isobutyrate	C547637	$C_5H_{10}O_2$	925.6	301.79	1.44837	
105	Thiophene	C110021	C_4H_4S	1035.4	394.247	1.04484	
106	2-Methyl-2-propanol	C75650	$C_4H_{10}O$	939.4	311.032	1.34116	
107	Heptanal	C111717	$C_7H_{14}O$	1200.7	663.817	1.35886	
108	2-Methyl-1-propanol	C78831	$C_4H_{10}O$	1115.5	501.77	1.37666	Dimers
109	3-Carene	C13466789	$C_{10}H_{16}$	1139.7	545.163	1.22154	
110	2-Methylpropanal	C78842	C_4H_8O	833.6	246.832	1.28257	
111	3-Methylbutanal	C590863	$C_5H_{10}O$	934.7	307.854	1.40845	
112	Butanoic acid	C107926	$C_4H_8O_2$	1694.3	1906.285	1.16893	
113	2-Methylpropanoic acid	C79312	$C_4H_8O_2$	1631.8	1665.286	1.15156	
114	Nonanal	C124196	$C_9H_{18}O$	1406.2	1022.81	1.49511	

Table 1. *Cont.*

No	Compounds	CAS	Molecular Formula	RI	Rt/s	Dt/ms	Comment
115	3-Ethylpyridine	C536787	C_7H_9N	1390	987.657	1.11846	
116	Acetoin	C513860	$C_4H_8O_2$	1304.7	821.482	1.32632	Dimers
117	1-Hexanol	C111273	$C_6H_{14}O$	1374.4	954.885	1.65336	Dimers
118	2-Methyl-3-furanthiol	C28588741	C_5H_6OS	1323.2	854.833	1.14631	
119	Octanal	C124130	$C_8H_{16}O$	1303.5	819.364	1.428	
120	3-Methyl-1-butanol	C123513	$C_5H_{12}O$	1226.1	699.033	1.48637	Dimers
121	3-Methyl-2-butenal	C107868	C_5H_8O	1219.5	689.643	1.09452	Monomers
122	3-Methyl-2-butenal	C107868	C_5H_8O	1219.5	689.643	1.35947	Dimers
123	2-Heptanone	C110430	$C_7H_{14}O$	1198.2	660.366	1.62302	Dimers
124	Pyridine	C110861	C_5H_5N	1172.8	610.581	1.24515	
125	Styrene	C100425	C_8H_8	1240	719.162	1.05951	
126	Myrcene	C123353	$C_{10}H_{16}$	1178.9	623.482	1.20043	
127	o-Xylene	C95476	C_8H_{10}	1195.8	657.167	1.09339	
128	(Z)-4-Heptenal	C6728310	$C_7H_{12}O$	1260.5	749.747	1.14745	
129	Hexyl acetate	C142927	$C_8H_{16}O_2$	1290.9	797.599	1.37921	
130	3-Furanmethanol	C4412913	$C_5H_6O_2$	1802.7	2409.312	1.1025	
131	Salicylic aldehyde	C90028	$C_7H_6O_2$	1737	2090.266	1.13654	
132	(E,E)-alpha-Farnesene	C502614	$C_{15}H_{24}$	1618.9	1619.575	1.43366	
133	Linalyl acetate	C115957	$C_{12}H_{20}O_2$	1589.4	1519.491	1.21298	
134	Benzaldehyde	C100527	C_7H_6O	1556.6	1415.508	1.1583	
135	Citronellal	C106230	$C_{10}H_{18}O$	1520.3	1308.926	1.35554	
136	2-Ethyl-1-hexanol	C104767	$C_8H_{18}O$	1542.3	1372.615	1.4278	
137	1-Octen-3-ol	C3391864	$C_8H_{16}O$	1487.5	1219.24	1.16025	
138	Ethyl isobutyrate	C97621	$C_6H_{12}O_2$	974.9	336.085	1.18508	
139	Methional	C3268493	C_4H_8OS	1462.4	1154.981	1.10453	

Table 2. The average area of VOCs in different varieties of hawthorn.

No	Compounds	CAS	Molecular Formula	SZ-01	SZ-02	GSZ	NSZ	Comment
1	Carvone	C99490	$C_{10}H_{14}O$	2193.56	541.96	823.61	642.46	
2	Geranyl acetate	C105873	$C_{12}H_{20}O_2$	5019.10	1463.21	5307.01	2248.35	
3	Neryl acetate	C141128	$C_{12}H_{20}O_2$	917.78	389.55	428.69	378.18	
4	2-Furanmethanol	C98000	$C_5H_6O_2$	5669.53	2512.74	972.12	632.27	Monomers
5	2-Furanmethanol	C98000	$C_5H_6O_2$	475.41	258.86	327.52	218.86	Dimers
6	Acetophenone	C98862	C_8H_8O	1023.00	669.31	1408.41	707.06	
7	alpha-Terpineol	C98555	$C_{10}H_{18}O$	240.43	84.07	97.26	108.22	
8	Diethyl succinate	C123251	$C_8H_{14}O_4$	1107.30	366.69	505.19	618.82	
9	gamma-Butyrolactone	C96480	$C_4H_6O_2$	1109.29	993.19	1921.69	975.16	
10	4-Terpinenol	C562743	$C_{10}H_{18}O$	2903.52	525.24	1004.35	445.81	
11	5-Methyl-2-furfural	C620020	$C_6H_6O_2$	2629.86	739.76	2082.71	597.91	
12	Propionic acid	C79094	$C_3H_6O_2$	1765.67	1282.93	2963.09	1424.15	
13	Linalool	C78706	$C_{10}H_{18}O$	4410.20	487.86	981.68	510.19	
14	(E)-2-Nonenal	C18829566	$C_9H_{16}O$	355.43	531.43	358.10	902.93	
15	2-Acetylfuran	C1192627	$C_6H_6O_2$	2575.65	922.08	1010.97	320.80	Monomers
16	2-Acetylfuran	C1192627	$C_6H_6O_2$	249.79	60.10	329.26	220.72	Dimers
17	1-Octanol	C111875	$C_8H_{18}O$	364.64	342.25	144.64	118.95	
18	Acetic acid	C64197	$C_2H_4O_2$	13428.62	13566.71	16484.00	16149.05	Monomers
19	Acetic acid	C64197	$C_2H_4O_2$	17099.25	15889.10	26134.83	20763.35	Dimers
20	Furfural	C98011	$C_5H_4O_2$	3003.64	3394.42	3123.56	2878.95	Monomers
21	Furfural	C98011	$C_5H_4O_2$	19115.34	8236.68	6371.44	5202.22	Dimers
22	alpha-Thujone	C546805	$C_{10}H_{16}O$	1330.87	140.79	82.44	2013.87	Monomers
23	alpha-Thujone	C546805	$C_{10}H_{16}O$	274.58	74.63	79.60	278.58	Dimers
24	Linalool oxide	C60047178	$C_{10}H_{18}O_2$	984.97	195.45	114.28	254.53	
25	5-Methyl-2(3H)-furanone	C591128	$C_5H_6O_2$	4697.91	5499.15	1945.43	247.30	Monomers

Table 2. *Cont.*

No	Compounds	CAS	Molecular Formula	SZ-01	SZ-02	GSZ	NSZ	Comment
26	5-Methyl-2(3H)-furanone	C591128	$C_5H_6O_2$	3054.88	2789.99	300.81	122.59	Dimers
27	(E)-2-Octenal	C2548870	$C_8H_{14}O$	620.58	399.09	618.27	481.19	
28	2-Ethyl-3-methylpyrazine	C15707230	$C_7H_{10}N_2$	355.92	216.98	107.11	49.75	
29	2-Cyclohexen-1-one	C930687	C_6H_8O	321.72	486.49	480.64	88.71	
30	Tetrahydrolinalool	C78693	$C_{10}H_{22}O$	328.98	104.38	140.67	159.87	
31	2,4-Hexadienal	C142836	C_6H_8O	685.98	101.89	769.25	89.31	
32	Methyl octanoate	C111115	$C_9H_{18}O_2$	739.69	205.33	149.65	134.87	
33	1-Hexanol	C111273	$C_6H_{14}O$	601.63	563.04	340.31	2029.10	Monomers
34	6-Methyl-5-hepten-2-one	C110930	$C_8H_{14}O$	1315.11	513.84	540.21	291.59	
35	2,5-Dimethylpyrazine	C123320	$C_6H_8N_2$	1888.63	162.56	612.51	133.67	
36	Ethyl heptanoate	C106309	$C_9H_{18}O_2$	948.68	69.21	135.69	71.20	
37	cis-2-Penten-1-ol	C1576950	$C_5H_{10}O$	222.05	232.78	109.55	73.88	
38	(E)-2-Heptenal	C18829555	$C_7H_{12}O$	445.84	487.73	883.89	518.61	
39	1-Hydroxy-2-propanone	C116096	$C_3H_6O_2$	1123.98	920.47	2165.55	492.28	Monomers
40	1-Hydroxy-2-propanone	C116096	$C_3H_6O_2$	351.30	142.87	972.10	38.89	Dimers
41	cis-3-Hexenyl acetate	C3681718	$C_8H_{14}O_2$	319.19	53.84	51.66	43.60	
42	Acetoin	C513860	$C_4H_8O_2$	520.34	365.75	1561.63	496.44	Monomers
43	2-Octanone	C111137	$C_8H_{16}O$	385.46	163.09	186.42	101.27	
44	2-Methyltetrahydrofuran-3-one	C3188009	$C_5H_8O_2$	1370.05	187.85	97.48	58.10	
45	2,6-Dimethylpyrazine	C108509	$C_6H_8N_2$	187.98	30.34	94.02	32.06	
46	1-Pentanol	C71410	$C_5H_{12}O$	1102.19	706.17	1512.61	1594.35	Monomers
47	1-Pentanol	C71410	$C_5H_{12}O$	284.89	109.55	502.58	583.88	Dimers
48	gamma-Terpinene	C99854	$C_{10}H_{16}$	880.50	83.20	133.79	123.04	
49	2,6-Dimethylpyridine	C108485	C_7H_9N	116.96	16.15	20.26	18.56	
50	2-Pentylfuran	C3777693	$C_9H_{14}O$	212.42	140.66	215.24	250.38	
51	(E)-2-Hexenal	C6728263	$C_6H_{10}O$	1503.23	1528.99	997.49	967.21	Monomers
52	(E)-2-Hexenal	C6728263	$C_6H_{10}O$	1610.16	1391.31	344.95	320.48	Dimers
53	3-Methyl-1-butanol	C123513	$C_5H_{12}O$	484.07	735.00	563.46	1280.72	Monomers
54	1,8-Cineole	C470826	$C_{10}H_{18}O$	4195.97	319.15	698.21	455.41	Monomers
55	1,8-Cineole	C470826	$C_{10}H_{18}O$	2361.24	75.23	49.00	78.11	Dimers
56	Diethyl disulfide	C110816	$C_4H_{10}S_2$	1966.57	5444.81	99.44	127.56	
57	Methyl hexanoate	C106707	$C_7H_{14}O_2$	1443.58	338.08	222.98	680.66	Monomers
58	2-Heptanone	C110430	$C_7H_{14}O$	356.57	215.85	180.50	454.44	Monomers
59	Methyl hexanoate	C106707	$C_7H_{14}O_2$	1080.43	38.42	241.15	679.24	Dimers
60	1,4-Cineole	C470677	$C_{10}H_{18}O$	663.29	78.52	129.59	298.24	Monomers
61	1,4-Cineole	C470677	$C_{10}H_{18}O$	221.18	21.13	22.03	44.06	Dimers
62	1-Penten-3-ol	C616251	$C_5H_{10}O$	1793.51	1585.35	1798.64	1259.45	Monomers
63	1-Penten-3-ol	C616251	$C_5H_{10}O$	2045.22	613.22	421.33	168.38	Dimers
64	3-Methylbutyl propanoate	C105680	$C_8H_{16}O_2$	1705.70	6833.11	103.03	70.18	
65	1-Butanol	C71363	$C_4H_{10}O$	1472.10	1392.35	2185.12	2529.75	Monomers
66	1-Butanol	C71363	$C_4H_{10}O$	273.52	337.82	838.25	1687.61	Dimers
67	Diallyl sulfide	C592881	$C_6H_{10}S$	367.91	553.83	204.87	334.69	
68	(E)-2-Pentenal	C1576870	C_5H_8O	655.72	757.33	988.04	515.80	Monomers
69	p-Xylene	C106423	C_8H_{10}	1189.37	1521.00	649.73	1639.21	
70	Cyclopentanone	C120923	C_5H_8O	3230.16	1918.17	411.52	2212.65	
71	(E)-2-Pentenal	C1576870	C_5H_8O	1004.82	443.33	467.68	274.07	Dimers
72	4-Methyl-3-penten-2-one	C141797	$C_6H_{10}O$	230.97	321.60	312.89	45.95	
73	beta-Pinene	C127913	$C_{10}H_{16}$	402.86	1641.86	85.24	97.13	
74	Isoamyl acetate	C123922	$C_7H_{14}O2$	593.09	164.07	214.31	497.19	
75	(Z)-2-Pentenal	C1576869	C_5H_8O	433.34	1556.07	357.24	603.99	Monomers
76	(Z)-2-Pentenal	C1576869	C_5H_8O	309.11	5188.38	93.61	349.68	Dimers
77	2-Methyl-1-propanol	C78831	$C_4H_{10}O$	583.87	852.56	628.62	1177.18	Monomers
78	Hexanal	C66251	$C_6H_{12}O$	1944.26	2803.51	3203.67	3114.45	Monomers
79	Hexanal	C66251	$C_6H_{12}O$	2331.59	3735.09	3204.66	5285.60	Dimers
80	(E)-2-Butenal	C123739	C_4H_6O	1630.63	910.14	567.29	438.94	
81	1-Propanol	C71238	C_3H_8O	328.31	322.34	1224.25	1262.25	
82	2-Butanol	C78922	$C_4H_{10}O$	1298.13	1530.66	105.56	203.92	

Table 2. *Cont.*

No	Compounds	CAS	Molecular Formula	SZ-01	SZ-02	GSZ	NSZ	Comment
83	1-Penten-3-one	C1629589	C_5H_8O	1272.76	578.92	725.65	220.10	
84	Isobutyl isobutyrate	C97858	$C_8H_{16}O_2$	1220.46	568.65	206.85	224.66	
85	3-Methyl-2-pentanone	C565617	$C_6H_{12}O$	932.60	826.90	24.55	29.98	
86	Isobutyl acetate	C110190	$C_6H_{12}O_2$	2312.15	4457.47	34.51	42.00	
87	Methyl 3-methylbutanoate	C556241	$C_6H_{12}O_2$	1182.99	127.62	38.62	82.05	
88	Pentanal	C110623	$C_5H_{10}O$	1427.20	1488.45	2802.92	2309.67	
89	2-Pentanone	C107879	$C_5H_{10}O$	1316.21	1179.53	570.79	1223.85	
90	Acetal	C105577	$C_6H_{14}O_2$	6995.16	1799.58	2950.66	2889.02	
91	2,5-Dimethylfuran	C625865	C_6H_8O	310.92	960.26	45.48	296.73	
92	Ethyl acetate	C141786	$C_4H_8O_2$	1302.33	2567.23	1565.33	3776.09	
93	Methyl acetate	C79209	$C_3H_6O_2$	15754.19	4078.68	5572.34	10602.05	
94	Acetone	C67641	C_3H_6O	5488.83	6531.17	4739.95	4322.90	
95	Propanal	C123386	C_3H_6O	1740.43	2326.29	3105.40	3006.52	
96	Butanal	C123728	C_4H_8O	228.42	73.54	76.38	60.76	
97	1-Propanethiol	C107039	C_3H_8S	1862.28	2854.62	30.31	18.37	
98	Dimethyl sulfide	C75183	C_2H_6S	1394.34	3434.90	791.13	1232.80	
99	2-Propenal	C107028	C_3H_4O	35.38	438.45	1562.47	364.22	
100	2-Methylbutanal	C96173	$C_5H_{10}O$	525.16	1865.45	1875.52	740.19	
101	2-Butanone	C78933	C_4H_8O	476.24	1406.98	1107.86	2272.08	
102	Ethanol	C64175	C_2H_6O	7729.16	6583.94	7764.54	7470.68	Monomers
103	Ethanol	C64175	C_2H_6O	2829.49	4731.35	5563.51	5461.13	Dimers
104	Methyl isobutyrate	C547637	$C_5H_{10}O_2$	145.10	618.81	19.14	21.81	
105	Thiophene	C110021	C_4H_4S	1051.26	688.43	2039.39	2418.48	
106	2-Methyl-2-propanol	C75650	$C_4H_{10}O$	209.37	524.83	655.14	925.26	
107	Heptanal	C111717	$C_7H_{14}O$	318.10	203.81	729.21	638.26	
108	2-Methyl-1-propanol	C78831	$C_4H_{10}O$	150.32	243.20	75.85	608.16	Dimers
109	3-Carene	C13466789	$C_{10}H_{16}$	191.59	253.67	43.51	646.50	
110	2-Methylpropanal	C78842	C_4H_8O	273.40	831.43	203.10	442.13	
111	3-Methylbutanal	C590863	$C_5H_{10}O$	188.66	766.00	365.90	1686.55	
112	Butanoic acid	C107926	$C_4H_8O_2$	1409.48	1213.95	5082.70	1589.55	
113	2-Methylpropanoic acid	C79312	$C_4H_8O_2$	1035.89	664.33	3651.45	1295.05	
114	Nonanal	C124196	$C_9H_{18}O$	358.80	871.42	779.24	553.25	
115	3-Ethylpyridine	C536787	C_7H_9N	166.39	102.28	412.40	62.95	
116	Acetoin	C513860	$C_4H_8O_2$	168.77	93.75	697.27	115.47	Dimers
117	1-Hexanol	C111273	$C_6H_{14}O$	66.88	66.76	63.03	314.99	Dimers
118	2-Methyl-3-furanthiol	C28588741	C_5H_6OS	100.71	54.57	77.06	91.73	
119	Octanal	C124130	$C_8H_{16}O$	164.88	328.96	289.08	251.28	
120	3-Methyl-1-butanol	C123513	$C_5H_{12}O$	196.78	210.35	104.43	983.60	Dimers
121	3-Methyl-2-butenal	C107868	C_5H_8O	100.87	789.99	224.29	885.33	Monomers
122	3-Methyl-2-butenal	C107868	C_5H_8O	128.27	332.57	34.78	299.32	Dimers
123	2-Heptanone	C110430	$C_7H_{14}O$	254.58	50.93	50.04	303.64	Dimers
124	Pyridine	C110861	C_5H_5N	71.09	95.75	86.10	258.29	
125	Styrene	C100425	C_8H_8	171.75	144.71	289.07	225.05	
126	Myrcene	C123353	$C_{10}H_{16}$	343.73	330.98	619.30	231.94	
127	o-Xylene	C95476	C_8H_{10}	100.68	170.75	119.71	96.15	
128	(Z)-4-Heptenal	C6728310	$C_7H_{12}O$	120.71	107.95	66.20	39.55	
129	Hexyl acetate	C142927	$C_8H_{16}O_2$	79.82	212.47	22.47	58.66	
130	3-Furanmethanol	C4412913	$C_5H_6O_2$	544.54	437.53	526.63	507.57	
131	Salicylic aldehyde	C90028	$C_7H_6O_2$	454.43	371.03	557.35	686.03	
132	(E,E)-alpha-Farnesene	C502614	$C_{15}H_{24}$	236.64	211.38	247.35	895.47	
133	Linalyl acetate	C115957	$C_{12}H_{20}O_2$	192.30	223.63	194.29	672.84	
134	Benzaldehyde	C100527	C_7H_6O	415.99	282.18	1085.63	761.03	
135	Citronellal	C106230	$C_{10}H_{18}O$	966.16	443.24	391.84	1637.24	
136	2-Ethyl-1-hexanol	C104767	$C_8H_{18}O$	379.36	212.65	2173.18	1290.30	
137	1-Octen-3-ol	C3391864	$C_8H_{16}O$	190.36	107.23	179.83	153.22	
138	Ethyl isobutyrate	C97621	$C_6H_{12}O_2$	1233.05	298.40	265.62	396.27	
139	Methional	C3268493	C_4H_8OS	80.68	185.66	73.75	155.20	

3.4. PCA of the Four Hawthorn Samples

The PCA method is a multivariate statistical analysis method that reduces the dimensions of a dataset by linearly combining variables and retaining the variability of the original data [27], in which the relationship between samples is more clearly and intuitively compared [28]. An analysis of PCA plots shows that the greater the similarity between samples, the greater the clustering degree, while the greater the difference between samples, the greater the distance [29]. As shown in Figure 5, the PCA of the VOCs in the four varieties of hawthorn was performed. Each color represents a different hawthorn sample, and the distance between each point indicates how similar they were.

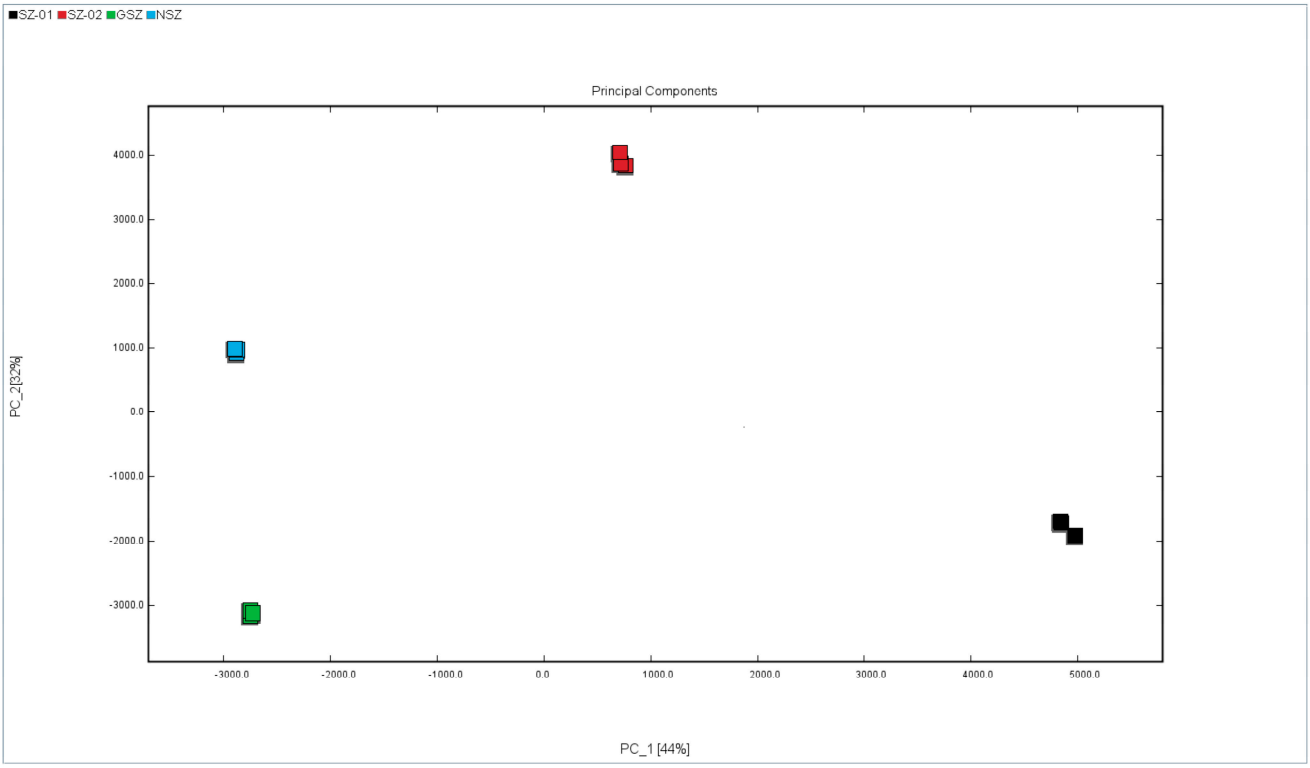

Figure 5. PCA analysis of different hawthorn varieties.

We can obviously see that the differentiation of the four samples and the distance between the Nanshanzha and Guangshanzha varieties was relatively close, which was because there were similarities in the VOC parts of the two samples, which was basically consistent with the signal peak information of the samples in the fingerprint spectrum. These results indicate that PCA combined with HS-GC-IMS could clearly distinguish differences in the VOCs among the four varieties of hawthorn.

3.5. Fingerprint Similarity Analysis Using Euclidean Distance

The distance coefficient is used to determine similarity in terms of the Euclidean distance, as is carried out in a PCA. This means that the difference is small if the distance between the samples is close, and this means that the difference is obvious if the distance is far [30]. Based on the Euclidean distance analysis, we created a fingerprint map with more intuitive characteristics than the PCA analysis. Based on the Euclidean distances between the four hawthorn varieties, Figure 6 depicts the fingerprint similarity. Thus, the Shanlihong and Nanshanzha varieties were close to each other, as were the Nanshanzha and Guangshanzha varieties, but the difference in VOCs between the Shanlihong and Nanshanzha varieties was of most significance. The results were in good agreement with the PCA analysis results. This may be because the Shanlihong and Shanzha varieties (the two are commonly known as "Northern Hawthorn") mostly grow in north China, while

the Nanshanzha and Guangshanzha varieties mostly grow in south China, such as in Guangdong and Guangxi. The north and south of China have very different climates, soils, altitudes, and so on, which causes the VOC levels to differ between varieties of hawthorn.

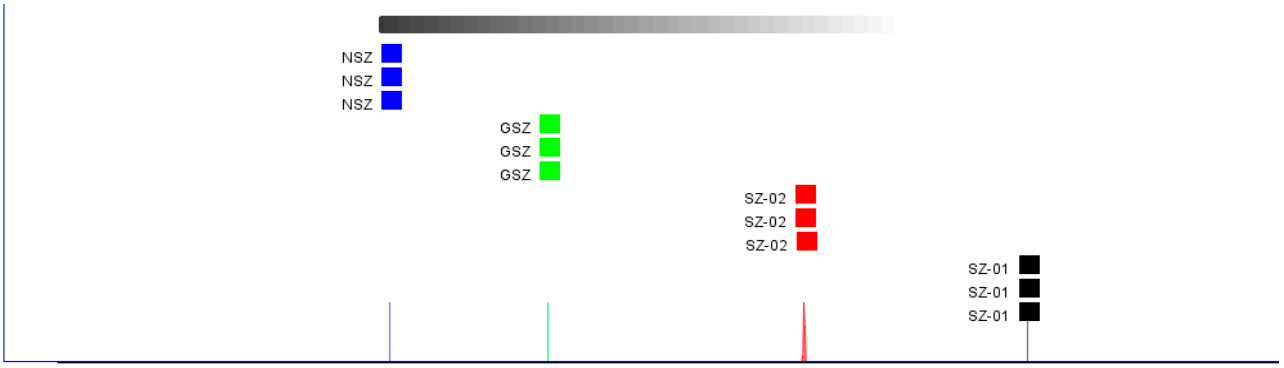

Figure 6. Fingerprint similarity based on the Euclidean distance of different varieties of hawthorn.

3.6. PLS-DA Analysis of the Four Hawthorn Samples

PLS-DA is a discriminant-analysis-supervised model that identifies hidden feature variables that damage model robustness and highlight the differences between groups [31]. Similarities and differences between samples are directly reflected in the PLS-DA score plot. In the score plot, a farther distance between two locations indicates a greater difference between two samples. The PLS-DA score plots of the four hawthorn samples were used to explore the differences in the VOC levels between the Shanlihong, Shanzha, Guangshanzha, and Nanshanzha varieties. The results are shown in Figure 7. It is evident that the samples of the different varieties of hawthorn were clearly distinguished based on the results of the principal component analysis. The SIMCA software (version 14.0) revealed that $R^2X = 0.993$ and $Q^2 = 0.999$, indicating that the model had a relatively accurate generalization and prediction ability and could both distinguish between the different varieties and identify their differences. Figure 8 shows the validation of the model carried out by the PLS-DA analysis after the permutation test ($n = 200$ times), with $R^2 = 0.0309$ and $Q^2 = -0.605$. It is evident from the slopes of the two regression lines that the model has a good prediction ability and does not show overfitting.

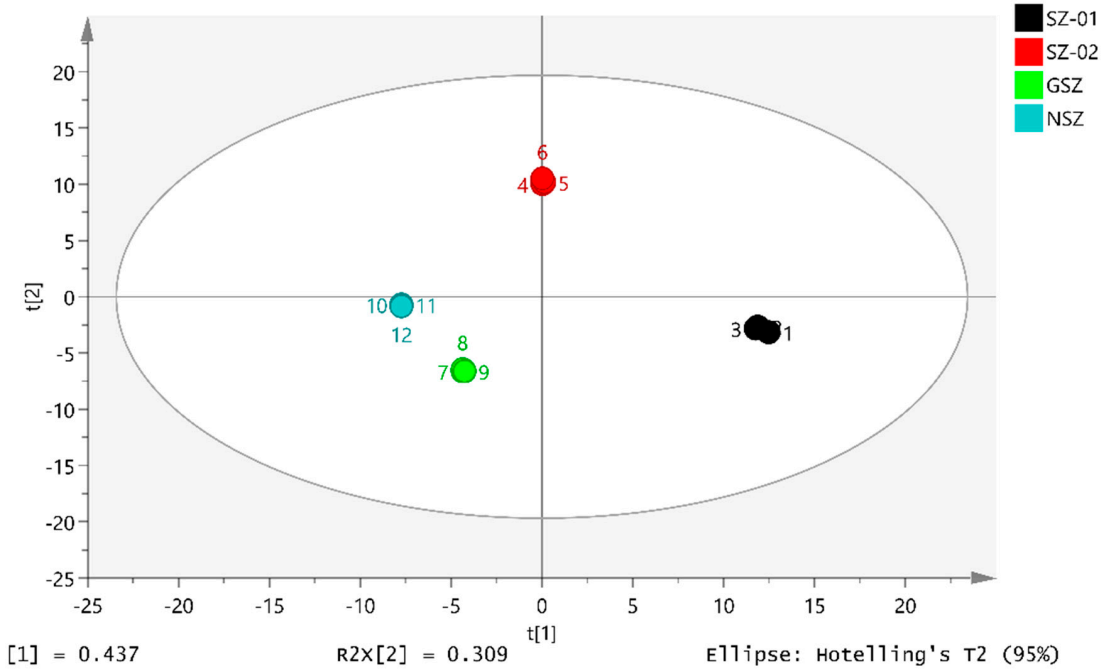

Figure 7. PLS-DA score chart.

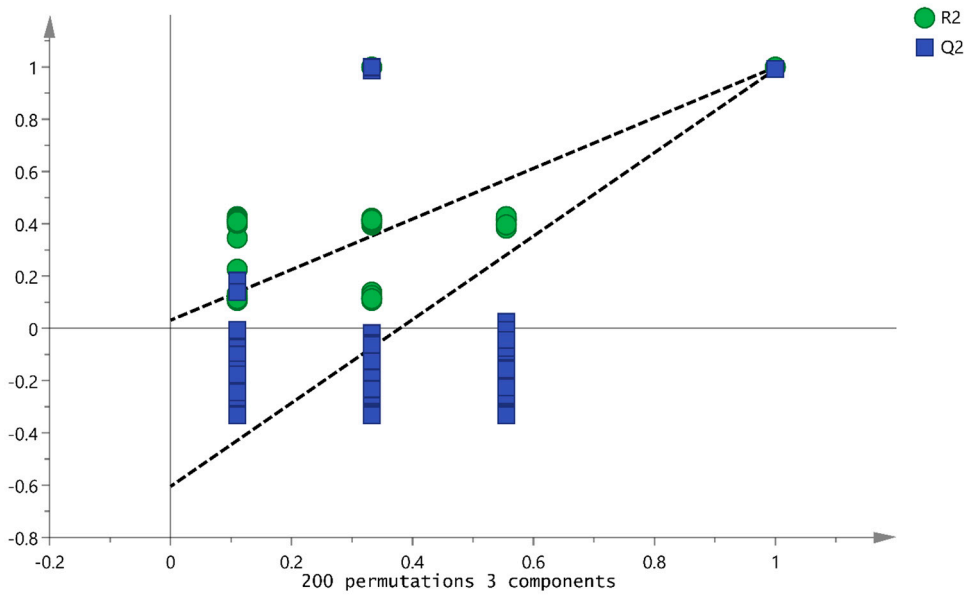

Figure 8. PLS-DA fitting curve.

It can be seen from the PLS-DA score chart that the supervised analysis method could better distinguish the four varieties of hawthorn and could obtain the variable weight value (variable important in projection: VIP). The results are shown in Figure 9. The VIP value is usually used to reflect the importance of PLS-DA model variables. The higher the column height and contribution value to the model in the figure, the more significant the difference [32]. The results show that there were 70 types of VOC with VIP >1 (marked in red boxes). Among them, the VOCs with the highest VIP values were 2-heptanone-M, alpha-thujone-M, (E)-2-pentenal-M, citronellal, 2-heptanone-D, and 2-cyclohexen-1-one. The abovementioned VOCs can be used as important indicators for the classification and identification of hawthorn, and they can provide a reference for the rapid identification of the four different types of hawthorn.

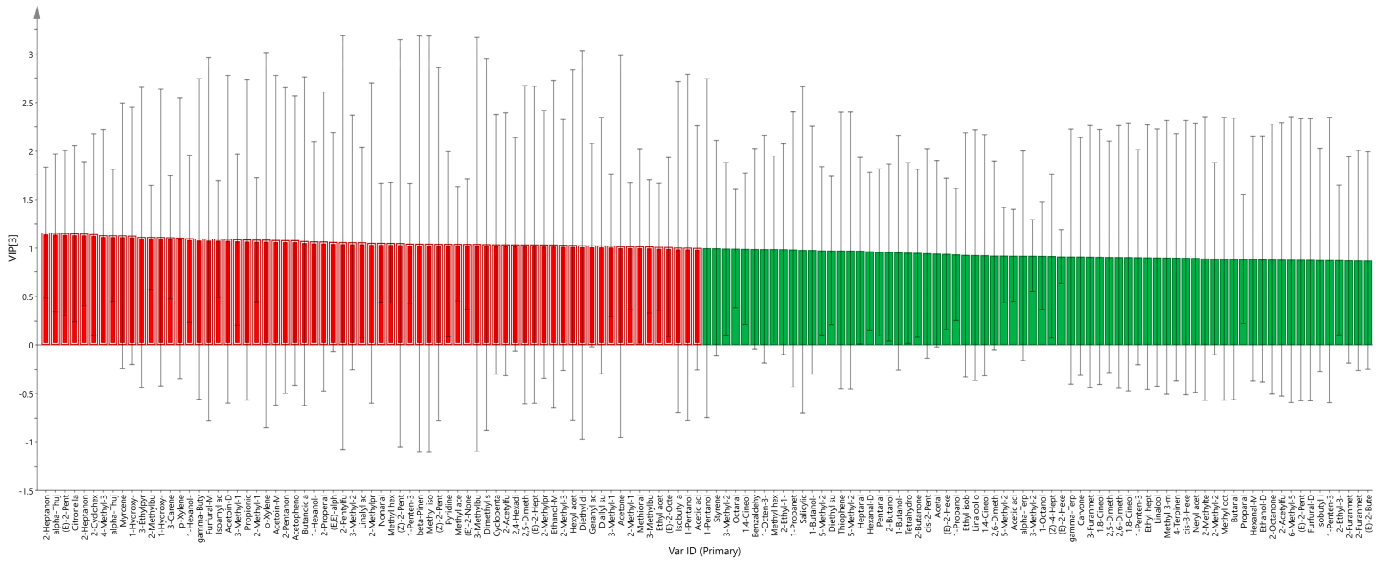

Figure 9. VIP diagram of the PLS-DA model.

4. Conclusions

This study used HS-GC-IMS for detecting and analyzing VOCs in hawthorn varieties. A total of 153 peaks were detected, and 139 VOCs were identified, including aldehydes, esters, alcohols, ketones, terpenes, acids, furans, pyridines, pyrazines, and thiophenes. By

using the fingerprints, PCA, Euclidean distance, and PLS-DA, we were able to quickly and effectively distinguish the four hawthorns. According to the results, the four hawthorns could be distinguished accurately and objectively without using appearance features. This study used HS-GC-IMS to analyze the VOCs in four varieties of hawthorn, which is a simple, rapid, and accurate method. In addition, this approach requires fewer sample preparation steps and less analysis time than other analytical methods. As a result, the detection and comparison methods applied in this study provide a valuable reference tool for identifying hawthorns from food VOCs. Simultaneously, this research enriched the study of flavor compounds in hawthorn.

However, there were still some shortcomings in this study, such as the limited sample size of the test, the fact that established model may not be the optimal model, the fact that the ion migration spectrum database is not perfect, and the fact that the database of various commercial instruments is not universal, resulting in the identification of some ions. Next, we will carry out a study with a larger sample size and establish a comprehensive and detailed database of GC-IMS, which can also be combined with mass spectrometry to further improve the qualitative ability.

Author Contributions: Conceptualization, L.Z. (Lijun Zhu); methodology, L.Z. (Lijun Zhu) and F.O.; software, Y.M. and L.Z. (Lingfeng Zhu); validation, Y.X. and B.W.; formal analysis, L.Z. (Lijun Zhu) and F.O.; investigation, Y.X.; resources, B.W.; data curation, C.L.; writing—original draft preparation, L.Z. (Lijun Zhu); writing—review and editing, C.L.; visualization, Y.M. and Q.Z.; supervision, Q.Z.; project administration, Y.M. and C.L.; funding acquisition, C.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Yunnan Key Laboratory of Screening and Research on Anti-pathogenic Plant Resources from Western Yunnan (No. APR202302), the Hunan Agricultural Science and Technology Innovation Fund Project (No. 2023CX30), the Hunan Provincial Natural Science Foundation of China (No. 2021JJ40406), the Undergraduate Research and Innovation Fund Project of Hunan University of Chinese Medicine (No. 2023), and the domestic First-Class Construction Discipline of Chinese Medicine in Hunan University of Chinese Medicine (No. 2024).

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. National Pharmacopoeia Committee. *Pharmacopoeia of the People's Republic of China*; China Medical Science and Technology Press: Beijing, China, 2020.
2. Lou, X.; Yuan, B.; Wang, L.; Xu, H.; Hanna, M.; Yuan, L. Evaluation of Physicochemical Characteristics, Nutritional Composition and Antioxidant Capacity of Chinese Organic Hawthorn Berry (*Crataegus pinnatifida*). *Int. J. Food Sci. Technol.* **2020**, *55*, 1679–1688. [CrossRef]
3. Duan, L.; Xiong, H.; Du, Y.; Wang, Z.; Li, Y.; Zhao, S.; Chen, J.; Si, D.; Pan, H. High-throughput LC-MS method for the rapid characterisation and comparative analysis of multiple ingredients of four hawthorn leaf extracts. *Phytochem. Anal.* **2022**, *33*, 635–643. [CrossRef] [PubMed]
4. Zhang, Q.; Zhao, P.; Sun, J.; Zhao, C.; Song, Y.; Wu, J. Research progress on chemical constituents and pharmacological action of hawthorn. *J. Northwest Pharm.* **2021**, *3*, 521–523.
5. Wang, S.Z.; Wu, M.; Chen, K.J.; Liu, Y.; Sun, J.; Sun, Z.; Ma, H.; Liu, L.T. Hawthorn Extract Alleviates Atherosclerosis through Regulating Inflammation and Apoptosis Related Factors: An Experimental Study. *Chin. J. Integr. Med.* **2019**, *25*, 108–115. [CrossRef] [PubMed]
6. Feng, Y.; Gao, S.; Zhu, T.; Sun, G.; Zhang, P.; Huang, Y.; Qu, S.; Du, X.; Mou, D. Hawthorn fruit acid consumption attenuates hyperlipidemia-associated oxidative damage in rats. *Front. Nutr.* **2022**, *9*, 936229. [CrossRef] [PubMed]
7. Kim, E.; Jang, E.; Lee, J.H. Potential Roles and Key Mechanisms of Hawthorn Extract against Various Liver Diseases. *Nutrients* **2022**, *14*, 867. [CrossRef] [PubMed]
8. Li, T.; Fu, S.; Huang, X.; Zhang, X.; Cui, Y.; Zhang, Z.; Ma, Y.; Zhang, X.; Yu, Q.; Yang, S. Biological Properties and Potential Application of Hawthorn and Its Major Functional Components: A Review. *J. Funct. Foods* **2022**, *90*, 104988. [CrossRef]
9. Nazhand, A.; Lucarini, M.; Durazzo, A.; Zaccardelli, M.; Cristarella, S.; Souto, S.B.; Silva, A.M.; Severino, P.; Souto, E.B.; Santini, A. Hawthorn (*Crataegus* spp.): An Updated Overview on Its Beneficial Properties. *Forests* **2020**, *11*, 564. [CrossRef]
10. Zhang, J.; Chai, X.; Zhao, F.; Hou, G.; Meng, Q. Food Applications and Potential Health Benefits of Hawthorn. *Foods* **2022**, *11*, 2861. [CrossRef]

11. Li, L.; Gao, X.; Liu, J.; Chitrakar, B.; Wang, B.; Wang, Y. Hawthorn pectin: Extraction, function and utilization. *Curr. Res. Food Sci.* **2021**, *4*, 429–435. [CrossRef]
12. Kwek, E.; Yan, C.; Ding, H.; Hao, W.; He, Z.; Liu, J.; Ma, K.Y.; Zhu, H.; Chen, Z.Y. Effects of hawthorn seed oil on plasma cholesterol and gut microbiota. *Nutr. Metab. (Lond.)* **2022**, *19*, 55. [CrossRef] [PubMed]
13. Jiang, H. Research Progress on Character Identification and Organic Acid Compositions of Shanzha (*Crataegus pinnatifida*), Nanshanzha (South *Crataegus*) and Guangshanzha (Cantonese *Crataegus*). *J. Liaoning Univ. Chin. Med.* **2023**, *25*, 132–137. [CrossRef]
14. Fei, C.; Dai, H.; Wu, X.; Li, L.; Lu, T.; Li, W.; Cai, B.; Yin, W.; Yin, F. Quality evaluation of raw and processed Crataegi Fructus by color measurement and fingerprint analysis. *J. Sep. Sci.* **2018**, *2*, 582–589. [CrossRef] [PubMed]
15. Gao, L.; Zhong, L.; Huang, R.; Yue, J.; Li, L.; Nie, L.; Wu, A.; Huang, S.; Yang, C.; Cao, G.; et al. Identification and determination of different processed products and their extracts of Crataegi Fructus by infrared spectroscopy combined with two-dimensional correlation analysis. *Spectrochim. Acta A Mol. Biomol. Spectrosc.* **2024**, *310*, 123922. [CrossRef]
16. Yin, J.; Wu, M.; Lin, R.; Li, X.; Ding, H.; Han, L.; Yang, W.; Song, X.; Li, W.; Qu, H. Application and development trends of gas chromatography-ion mobility spectrometry for traditional Chinese medicine, clinical, food and environmental analysis. *Microchem. J.* **2021**, *168*, 106527. [CrossRef]
17. Guo, Y.; Chen, D.; Dong, Y.; Ju, H.; Wu, C.; Lin, S. Characteristic volatiles fingerprints and changes of volatile compounds in fresh and dried Tricholoma matsutake Singer by HS-GC-IMS and HS-SPME-GC-MS. *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **2018**, *1099*, 46–55. [CrossRef]
18. Song, H.; Liu, J. GC-O-MS technique and its applications in food flavor analysis. *Food Res. Int.* **2018**, *114*, 187–198. [CrossRef] [PubMed]
19. Huang, X.H.; Zheng, X.; Chen, Z.H.; Zhang, Y.Y.; Du, M.; Dong, X.P.; Qin, L.; Zhu, B.W. Fresh and grilled eel volatile fingerprinting by e-Nose, GC-O, GC-MS and GC×GC-QTOF combined with purge and trap and solvent-assisted flavor evaporation. *Food Res. Int.* **2019**, *115*, 32–43. [CrossRef]
20. Zhou, S.; Feng, D.; Zhou, Y.; Duan, H.; He, Y.; Jiang, Y.; Yan, W. Characteristic Volatile Organic Compound Analysis of Different Cistanches Based on HS-GC-IMS. *Molecules* **2022**, *27*, 6789. [CrossRef]
21. Dai, Y.; Liu, S.; Yang, L.; He, Y.; Guo, X.; Ma, Y.; Li, S.; Huang, D. Explorative study for the rapid detection of Fritillaria using gas chromatography-ion mobility spectrometry. *Front. Nutr.* **2024**, *11*, 1361668. [CrossRef]
22. Feng, X.; Wang, H.; Wang, Z.; Huang, P.; Kan, J. Discrimination and characterization of the volatile organic compounds in eight kinds of huajiao with geographical indication of China using electronic nose, HS-GC-IMS and HS-SPME-GC-MS. *Food Chem.* **2022**, *375*, 131671. [CrossRef]
23. Feng, D.; Wang, J.; He, Y.; Ji, X.J.; Tang, H.; Dong, Y.M.; Yan, W.J. HS-GC-IMS detection of volatile organic compounds in Acacia honey powders under vacuum belt drying at different temperatures. *Food Sci. Nutr.* **2021**, *9*, 4085–4093. [CrossRef]
24. Yang, B.; Huang, J.; Jin, W.; Sun, S.; Hu, K.; Li, J. Effects of Drying Methods on the Physicochemical Aspects and Volatile Compounds of *Lyophyllum decastes*. *Foods* **2022**, *11*, 3249. [CrossRef]
25. Duan, H.; Zhou, S.; Guo, J.; Yan, W. HS-GC-IMS Analysis of Volatile Organic Compounds in Different Varieties and Harvesting Times of *Rhizoma gastrodiae* (Tian Ma) in Yunnan Province. *Molecules* **2023**, *28*, 6705. [CrossRef]
26. Zhang, Y.; Tong, X.; Chen, B.; Wu, S.; Wang, X.; Zheng, Q.; Jiang, F.; Qiao, Y. Novel application of HS-GC-IMS for characteristic fingerprints and flavor compound variations in citrus reticulatae pericarpium during storage with different *Aspergillus niger* fermentation. *Food Chem. X* **2023**, *18*, 100653. [CrossRef] [PubMed]
27. Dong, W.; Tan, L.; Zhao, J.; Hu, R.; Lu, M. Characterization of Fatty Acid, Amino Acid and Volatile Compound Compositions and Bioactive Components of Seven Coffee (*Coffea robusta*) Cultivars Grown in Hainan Province, China. *Molecules* **2015**, *20*, 16687–16708. [CrossRef] [PubMed]
28. Christmann, J.; Rohn, S.; Weller, P. gc-ims-tools—A new Python package for chemometric analysis of GC–IMS data. *Food Chem.* **2022**, *394*, 133476. [CrossRef]
29. Zhou, Y.; Wang, D.; Duan, H.; Zhou, S.; Guo, J.H.; Yan, W.J. Detection and analysis of volatile flavor compounds in different varieties and origins of goji berries using HS-GC-IMS. *LWT* **2023**, *187*, 115322. [CrossRef]
30. Zhou, S.Q.; Feng, D.; Zhou, Y.X.; Zhao, J.; Zhao, J.Y.; Guo, Y.; Yan, W.J. HS-GC-IMS detection of volatile organic compounds in cistanche powders under different treatment methods. *LWT—Food Sci. Technol.* **2022**, *165*, 113730. [CrossRef]
31. Jiménez-Carvelo, A.M.; Martín-Torres, S.; Ortega-Gavilán, F.; Camacho, J. PLS-DA vs sparse PLS-DA in food traceability. A case study: Authentication of avocado samples. *Talanta* **2021**, *224*, 121904. [CrossRef]
32. Li, H.; Mao, Y.; Liu, Y.; Li, D.; Wang, M.; Ren, X.; Dou, Z. Comparative investigation of raw and processed products of Gardeniae Fructus and *Gardenia jasminoides* var. *radicans* using HPLC coupled with chemometric methods. *Biomed. Chromatogr.* **2021**, *35*, e5051. [CrossRef]