Hindawi
Journal of Food Quality
Volume 2021, Article ID 6650990, 11 pages
https://doi.org/10.1155/2021/6650990

Research Article

Metabolomic Biomarkers Differentiate Soy Sauce Freshness under Conditions of Accelerated Storage

Thiruchelvi R. Reddy,[1] Katherine A. Overmyer,[1,2] Joshua J. Coon,[1,2,3] MaryAnne Drake,[4] Taro Horiba,[5] and Scott A. Rankin[6]

[1]*Department of Biomolecular Chemistry, University of Wisconsin-Madison, Madison, WI 53706, USA*
[2]*Morgridge Institute for Research, Madison, WI 53715, USA*
[3]*Department of Chemistry, University of Wisconsin-Madison, Madison, WI 53706, USA*
[4]*Department of Food,Bioprocessing,& Nutrition Sciences, North Carolina State University, Raleigh, NC 27695, USA*
[5]*Kikkoman USA, R & D Laboratory,Inc., Madison, WI 53719, USA*
[6]*Department of Food Science, University of Wisconsin-Madison, Madison, WI 53528, USA*

Correspondence should be addressed to Scott A. Rankin; sarankin@wisc.edu

Received 30 November 2020; Revised 26 March 2021; Accepted 8 April 2021; Published 5 May 2021

Academic Editor: Teresa Zotta

Naturally fermented soy sauce is one of the few globally valued food condiments. It is complex in its substrate, manufacturing processes, and chemical profile of salts and organic compounds, resulting from spontaneous, enzymatic and biochemical reactions. The overall chemical character of soy sauce has a few rivals relative to its chemical and bioactive complexity. Resulting from this complexity are unique sensory attributes contributing to the characteristic soy sauce flavor as well as potentiating other sensory sensations. Soy sauce is susceptible to deterioration after bottling during storage. This work examined soy sauces over an eight-month period using descriptive sensory methods and the discovery of metabolomic biomarkers with high resolution mass spectrometry, wherein samples were derivatized to enable volatility and identification of polar analytes. While several thousand metabolites were detected, only organic acids, amino acids, and various glycosylated metabolites were statistically defensible biomarkers of storage time. The relationships between sensory and metabolomic data were assessed using Kendall rank-based correlations to generate Kendall Tau correlation coefficients. A second approach filtered the data based on correlation significance and grouped molecules based on hierarchical clustering. Mass spectrometry analyses discovered several thousand unique analyte peaks with relevant changes denoted as significant relative to the fresh samples using volcano depictions of p values versus changes in compound abundances. We present a metabolomic approach for the analysis of complex food systems capable of differentiating a quantifiable extrinsic variable, which is, in this case, storage time with a correlation coefficient of 0.99. We further demonstrate that changes in soy sauce resulting from storage are characterized by sensory decreases in fruity/grape and nutty/sesame aroma and increases in methional/potato aroma and astringent attributes with concomitant changes in the concentrations of several key biomarkers.

1. Introduction

Analytical technologies that enable effective food quality characterization, differentiation, and management are based on accurate and reliable data generation and have been increasing in type and performance, including emerging methodologies, such as high-resolution mass spectrometry metabolomic-based approaches [1, 2]. Different types of analytical platforms are often applied in combination with nontargeted analyses to maximize and broaden detection capabilities for improved sensitivity and molecule identification. Such approaches generate discriminatory patterns of complex chemical components from different samples through the creation and comparison of metabolic fingerprints, thus providing the capability to differentiate chemically complex biological samples [3, 4]. In fact, several studies have reported attempts to relate these chemical fingerprints to the discriminatory sensory qualities of food

samples by examining these analytical data with various multivariate statistical designs [5–7].

In the case of exploratory studies for pursuing compounds in foods responsible for a particular sensory quality, sensory-guided techniques in conjunction with novel, advanced chromatographic and mass spectrometry analyses are applied. Such approaches can be useful for characterizing changes in sensory attributes with changes in chemical profiles resulting from independent experimental variables, such as product storage time [8]. However, such approaches include the risk of missing the more complicated collective or cumulative effects of biologically active substances that occur in authentic food systems. Yet, these exploratory approaches provide a critical foundation for further discoveries and can provide a more comprehensive understanding of the changing chemistries involved for the purposes of product differentiation. However, the elucidation of key, flavor-impact compound identities is not the universal goal of flavor assessments such as is the case with electronic nose technologies [9], where the compound-dependent signal can be used to differentiate treatment effects rather than elucidating the causative chemistries. Here, we present a similar approach using metabolomics to differentiate soy sauce (SS) samples exposed to an extended storage treatment.

Previous studies that characterize critical SS sensory attributes have been conducted, and some lexicons have been consequently developed by several research teams [10, 11]. These lexicons allow a common language for the evaluation of native SS sensory qualities, thus sharing new insights and applied discoveries among consumers, researchers, and manufacturers. The sensory attribute of freshness in SS is absent from existing SS lexicons, despite the fact that the loss of freshness has been anecdotally recognized as a key sensory attribute that determines SS quality as perceived by consumers, especially in markets that have high standards for SS sensory performance and value [12]. Furthermore, freshness as a key characteristic of foods has gained attention given its critical importance to consumers. However, it requires a complex and multisensory, cross-modal affective assessment [13].

As a condiment, SS is applied to a wide variety of foods ranging from common food to some of the most sophisticated meals crafted. After manufacturing and as a function of storage parameters, such as time, light exposure, and oxygen concentration, SS is anecdotally thought to lose its fresh character, manifested as a darkened color and altered sensory attributes, suggestive of spontaneous reactions that are especially detrimental, where SS is applied in fine culinary applications [12]. These adverse changes over storage time are typically expressed in the industry as a loss of freshness and are thought to be catalyzed when SS is exposed to oxygen, high temperatures, or excessive storage time. Chemical reactions generally associated with oxidative or Maillard reaction pathways have been suggested to contribute to these undesirable changes [12]; however, there is little information detailing discernable shifts in chemistry, and thus, there is no rationale to objectively mark its occurrence. In Japan, the trade association, called the Japan Soy Sauce Brewers Association (JSSB), proposed a standard "best before date" shelf life determination, in which freshness is a factor that must be sensorially determined. Yet, even in this JSSB method, there is no definitive sensory characterization for freshness aside from a general subjective loss of desirable character relative to a freshly prepared control anchor. This lack of definitive sensory and chemical change prevents the application of a means to prevent the loss of SS freshness. We further hypothesize that the native chemistries in flux over the course of a loss of freshness from storage can be elucidated through the comparison of differential metabolic fingerprints and multivariate analyses. Thus, the objectives for this work include the assessment of SS samples aged under conditions noted later as a means of elucidating sensory attribute changes with concomitant changes in chemical profiles. Study objectives were achieved based on changes in a limited set of metabolites, including organic acids, amino acids, and various glycosylated compounds, resulting from the aging process using a metabolomic approach.

2. Materials and Methods

2.1. Soy Sauce Samples. Traditionally brewed SS was directly obtained from the manufacturer (Kikkoman, Chiba, Japan) bottled in sealed, 1-liter plastic containers for storage treatments. Samples were stored in a dark incubator held at 30°C, and triplicate samples were randomly removed at two-month intervals up to and including the eight-month storage event. Upon removal from incubation, samples were stored at −80°C until analyses were completed.

2.2. Sensory Analyses. A protocol for sensory analysis of SS was filed and approved (North Carolina State University, Institutional Review Board) prior to initiation of study. A trained panel ($n = 6$ panelists, each with at least 100 h of experience in the descriptive analysis of foods and beverages) assessed SS attributes from the literature [10, 14] and novel attributes from preliminary studies using a 15-point universal intensity scale consistent with the Spectrum™ method. Based on this documented complexity and other recently published literature, eighteen aroma and flavor attributes were selected for the study (Table 1), from which eleven attributes generated responses in the study (Table 2). Each SS sample was evaluated by each panelist in triplicate. Paper ballots were used for data collection and were manually transferred into statistical software. SS samples were diluted at 1 : 1 with deionized water prior to the sensory evaluation. Training sessions and preliminary evaluations indicated that dilution at this level minimized fatigue with no effect ($p > 0.05$) on perceived sensory attributes. Attribute intensities were scored on a 0- to 15-point universal intensity scale consistent with the Spectrum method; panelists were allowed to score beyond this range if warranted by a particular attribute intensity [19].

Table 1: Initially trained panel sensory attributes for soy sauce (terms adopted from previous studies are cited).

Term	Definition	Reference
Ashy [10]	Dry, dusty, smoky aromatics associated with the residual of burnt products	Ashes from burnt wood (from fireplace or outdoor fire pit)
Beefy/brothy [17]	Aromatics associated with beef stock or bouillon cubes	Bouillon cubes rehydrated in water, and then, 20 mL was placed in a 58 mL lidded soufflé cup
Brown fruity/ prune [14]	A sweet, floral aromatic associated with prunes	Mariani brand prunes
Brown spice clove [10]	Aromatics associated with nutmeg, clove, and cinnamon	Nutmeg, clove, and cinnamon
Caramel/sweet aromatic [10]	Sweet aromatics associated with vanilla, vanillin, burnt sugar, caramel, and molasses	Vanilla extract, vanillin, burnt sugar, caramel, and molasses
Chemical [10]	The aromatics associated with plastic and burnt plastics	Clear PET bottle
Fruity ester [16]	A sweet, floral aromatic blend of a variety of ripe fruits	Juicy juice Nestle all natural 100% kiwi strawberry ethyl hexanoate
Fruity fermented/ beer/malty [10]	A sour, sweet fermented aromatic associated with fermented fruit, beer, fruits, and malt	225 mL amber sniff jar with 10 ppm of 2-ethyl-3-methylbutanal, grape nuts soaked in water and beer
Fruity/grape	A sweet aromatic associated with grape juice	Welch's grape juice
Maple	Aromatic associated with maple syrup	Maple syrup
Methional/potato [17]	Aromatics associated with vegetable soup stock, bouillon cubes, or canned potatoes (methional)	Bouillon cubes rehydrated in water, and then, 20 mL was placed in a 58 mL lidded soufflé cup or a 225 mL amber sniff jar with 10 ppm of methional or potato flakes
Mushroom	The earthy aromatic associated with fresh mushrooms	1-octen-3-one shiitake mushroom
Nutty/sesame [10]	The nonspecific nutlike flavors that are characteristic of several different nuts, e.g., peanuts, hazelnuts, pecans, and almonds	Shamrock mixed crushed nuts or sesame seeds
Olive	Aromatic associated with green olives	Green olives
Phenolic [18]	Aromatic associated with phenol	225 mL amber sniff jar with 10 ppm of phenol or Band-Aids
Pyrazine green/ raw10	Aromatic associated with pyrazine or raw peanuts	225 mL amber sniff jar with 10 ppm of pyrazine
Smoky [10]	A sweet, pungent, dry, and woody aroma/aromatic associated with smoke	Liquid smoke with hickory barbecue flavor (Tone's brand)
Sour aromatic/ vinegar [10]	Sour aromatic associated with acetic acid	225 mL amber sniff jar with 10 ppm of acetic acid
Surimi (sweet aroma)	Sweet aromatic associated with imitation crab	Imitation crab (Fisherman's market)
Astringent mouthfeel [15]	Chemical feeling factor on the tongue or oral cavity described as puckering or dry	Alum (1% in water)
Mouth burn [15]	Trigeminal pain response to the activation of neural receptors on the tongue and soft palate	Soda water or ethanol
Sweet taste [15]	Fundamental taste sensation elicited by sugars	Sucrose (5% in water)
Sour taste [15]	Fundamental taste sensation elicited by citric acid	Citric acid (0.05% in water)
Salty taste [15]	Fundamental taste sensation elicited by salts	Sodium chloride (5% in water)
Bitter taste [15]	Fundamental taste sensation elicited by caffeine and quinine	Caffeine (0.05% in water)
Umami taste [15]	Fundamental taste sensation elicited by certain peptides and nucleotides	MSG (1% in water)

2.3. High-Resolution Gas Chromatography (HR-GC-MS) A preliminary study revealed several hundred polar compounds in SS by HR-GC-MS to determine if the instrumental analyses were able to detect relative changes in compound concentrations and refine sample preparation methodologies and instrument run parameters. We further note that traditional volatile chemistries associated with aroma character, such as esters or aldehydes, were not discoverable using the sample extraction and derivatization methods employed. Yet, this methodology allowed a relatively simple, rapid approach that discovered of a host of compounds as noted later, and thus, we deemed this approach suitable for the intention of this study, which is the ability to differentiate sample aging treatments. In general, classes of compounds shifted by the aging treatment included organic acids, amino acids, and various glycosylated metabolites.

2.4. Sample Preparation for HR-GC-MS Analyses. SS samples were kept on ice along with all other laboratory reagents used. Aliquots of SS samples were diluted 10-fold with DI water. Performed in triplicate, 20 μL of diluted sample was pipetted into a 1.5 mL Eppendorf tube and 225 μL MeOH was added. The mixture was vortexed for 10 s. Next, 750 μL methyl *tert*-butyl ether was added, and the mixture was vortexed for 10 s. The solution was mixed using an orbital

Table 2: Mean responses (0–15, where 0 = no response and 15 = extreme intensity) from trained panel analyses of SS samples aged for 8 months at 30°C for each of the attributes denoted. Scores for saltiness exceeded the 0–15 range, reflecting an extremely high intensity.

Attribute	$t = 0$ months	$t = 2$ months	$t = 4$ months	$t = 6$ months	$t = 8$ months
Astringent mouth feel	1.9	2.0	2.1	2.2	2.2
Beefy/brothy	4.0	3.9	3.9	3.9	4.0
Caramel/sweet aromatic	2.4a	2.2b	2.2b	2.1b	2.2b
Fruity/grape	2.7a*	2.5b	2.6ab	2.5ab	2.2c
Methional/potato	1.1c	1.5ab	1.5ab	1.4b	1.7a
Mouth burn	1.3	1.6	1.3	1.2	1.3
Nutty/sesame	1.7	0.9	0.6	0.7	ND
Salty taste	33.1	33.4	32.1	33.3	33.3
Sour taste	1.0	1.0	1.0	1.0	1.0
Sweet taste	2.7	2.6	2.5	2.5	2.6
Umami taste	6.7	6.7	6.7	6.7	6.7

*Mean values in a row followed by a different letter are significantly different ($p < 0.05$). Bolded terms in the attribute column denote a significant time effect on intensity score of sensory term.

shaker for 6 min. To induce phase separation, 187.5 μL water was added, and the solution was vortexed again for 20 s. The sample was then centrifuged for 2 min at 12,000×g and 4°C. The upper phase in the 1.5 mL Eppendorf tube was discarded, and 250 μL of the lower (aqueous) phase was removed and placed in a separate 1.5 mL Eppendorf tube. To this, 250 μL ACN was added to precipitate protein. The mixture was vortexed for 15 s and centrifuged at 13,000×g for 5 min at 4°C. Then, 300 μL of the supernatant was aliquoted into glass autosampler vials. The mixture was dried in a vacuum concentrator. Once dry, samples were resuspended in 50 μL methoxyamine hydrochloride solution (20 mg/mL, pyridine solution), vortexed for 15 s, and heated at 37°C for 90 min. Then, 100 μL of N-methyl-N- trimethylsilyltrifluoroacetamide (MSTFA) was added and vortexed for 15 s, and the mixture was heated at 60°C for 60 min.

2.5. HR-GC-MS Analysis. Samples were analyzed using a GC-MS instrument (Trace 1310 GC, Thermo Scientific, Waltham, MA) coupled to a mass spectrometer (Q Exactive Orbitrap, Thermo Scientific, Waltham, MA). A temperature gradient ranging from 100°C to 320°C was employed spanning a total runtime of 25 min. Analytes were injected onto a 30 m × 0.25 mm ID × 1 μm thickness column (TraceGOLD TG-5SILMS, Thermo Scientific, Waltham, MA) using a 1 : 10 split at a temperature of 275°C and ionized using electron ionization (EI). The mass spectrometer was operated in full scan mode using a resolution of 30,000 [3].

2.6. HR-GC-MS Data Processing. Data processing was done using a software suite developed in-house that is available at https://github.com/coongroup. Following data acquisition, raw EI-GC/MS spectral data were deconvolved into "features" and then grouped into individual spectra containing only product ions stemming from a singular parent molecule. Feature groups from samples and background were compared and those found in both were removed from further analyses. Compound identifications for the metabolites analyzed were assigned by comparing deconvolved high-resolution spectra against unit-resolution reference spectra present in the NIST 12 MS/EI library and authentic standards run in-house. To calculate spectral similarity between experimental and reference spectra, a weighted dot product calculation was used. Metabolites lacking a confident identification were classified as "unknown." Peak heights of specified quantified m/z ratios were used to represent metabolite abundance. The data set was further processed using a linear regression approach (non-log 2 transformed intensity values versus run order) to normalize for run order effects on signal.

2.7. Data Analysis. All MS measurements were integrated with sensory information and processed with data analysis and visualization software (see https://coonlabdatadev.com). p values displayed in figures and tables were calculated using Student's t-test comparing time points 2, 4, 6, and 8 months with 0 months. We performed covariant analysis in R using the "cor.test" function and method "Kendall"; with this method, we identified compounds from MS datasets correlating sensory descriptor scores ($p < 0.05$). From this subset of molecules which correlate with sensory terms, we performed hierarchical cluster analysis on the correlation matrix of molecule-sensory pairs using the R function "hclust" and k-means equal to 5. A heatmap of this matrix was generated using the R function "pheatmap" [20–22]. Both chemical and sensory data were further analyzed for the purpose of generating predictive models of their correlative value using multiple, stepwise linear regression (JMP vs. 14, SAS Institute Inc., Cary, NC, 1989-2020).

3. Results and Discussion

Overall, sensory and HR-GC-MS analyses provided novel analyte discoveries resulting from the aging or storage of SS as described by the aforementioned conditions. Correspondingly, there were limited, but significant age-based sensory changes in the SS including the loss of fruity/grape and nutty/sesame aroma characteristics and the increase of methional/potato aroma. Changes in biomarker profiles varied across the aging variable yet intensified in number and quantity of compounds captured as described in detail later. We note that the analyses conducted were designed to discern the most notable changes in polar biomarker profiles

discernable by the instrumental analyses as a function of the storage time variable. We further acknowledge that such an approach is not designed to or sufficient for establishing causal sensory changes; it is rather a means of discerning the influence of storage-based aging through specific biomarkers.

3.1. Sensory Profiles. Although eighteen attributes were initially considered by sensory panelists, eleven terms were selected for further consideration in that the other seven terms did not generate changes in sensory responses over the course of the study. Of the terms that generated responses, five were not affected by the treatment variable of storage time, leaving six that were affected. Two terms, namely, fruity/grape and nutty/sesame, displayed significant decreases in aroma intensity with increases in storage time. The term methional/potato increased with storage time, suggesting an increase in aroma resulting from sulfur-containing volatiles. Although the term caramel/sweet aromatic was influenced by storage time, the results were not consistent. The samples also displayed slight changes in the astringent and mouth burn sensations, yet the results were not consistent across storage time.

Changes in freshness in foods and beverages, such as spoilage in fluid milk, are complex and involve changes in various classes of compounds, which affect color, texture, aroma, and taste attributes created by spontaneous, enzymatic, and microbial activities. In isolation, none of these effects may be discernable by sensory assessment, yet collectively they can affect a notable sensory departure from the native state. A notable observation in this study is that newly manufactured or "fresh" SS is differentiated from aged SS by multiple distinguishable sensory attributes. Fresh SS was higher in caramel and fruity and nutty attributes and lower in methional notes. This latter aroma has been associated with deteriorative changes in other foods with complex compositional profiles [15, 21]. We further noted no changes in other SS attributes, such as umami or sweet tastes. The following is a regression model derived from sensory factors affected by age in months as a means of defining SS freshness loss:

$$\begin{aligned}\text{Freshness loss} = 17.0 - (5.61 * \text{caramel}) - (4.47 * \text{fruity}) \\ -(2.76 * \text{nutty}) + (11.5 * \text{methional}),\end{aligned} \quad (1)$$

and we propose that this model infers that SS freshness is characterized from a sensory standpoint from the collective changes in the sensory attributes of caramel and fruity and nutty attributes and it is lower in methional notes as weighted by the correlation coefficients derived from regression analysis. Although the sensory character is complex, these four attributes were able to predict 79% of the changes manifested across the storage time assessed based on correlation analysis. A complementary model was created using instrumental data wherein compounds were first selected based on their relative strength of correlation to the storage variable and then assessed using a stepwise regression model. The final model for the prediction of SS freshness loss using instrumental data is shown as

$$\begin{aligned}\text{Freshness loss} = 196 + (30.1 * \text{Sugar RT } 13.6) \\ +(-42.2 * \text{Sugar RT } 19.4) + (27.1 * \text{Sugar RT } 19.5) \\ +(-19.7 * \text{Arabinose}) + (0.737 * \text{L} - \text{Tryptophan}).\end{aligned} \quad (2)$$

Upon validation of the assessment using the aforementioned correlations, these five chemical variables were able to predict >99% of the variation in the storage variable.

3.2. HR-GC-MS Metabolomics. As anticipated, chromatograms were complex and yielded several thousand resolved analytes. While many analytes were tentatively identified, a significant number were not, namely, those with sugar moieties where the type and degree of ligand substitution were not differentiable through molecular weight or mass spectra database comparison (Table 3). A graphical depiction of the type and complexity of variation from the $t = 0$ control SS sample is presented in Figures 1(a)–1(d). To better understand which metabolites are associated with the critical sensory parameters, we performed a correlation analysis between metabolite abundance and sensory measurements for the 15 samples (0–8 mo). The Kendall Tau correlation coefficient was used to assess strength of association and was chosen to better account for the nonparametric nature of the sensory data. The resulting significant correlations are visualized using a heatmap (Figure 2(a)), where the strength of the correlation is indicated by heat color ranging from blue (indicating a strong negative correlation) to red color (strong positive correlation).

After applying hierarchical clustering and k-means ($k = 5$) clustering techniques, we were able to identify distinct patterns of associations between metabolites (Figure 2(a)) and sensory attributes. For example, in the yellow cluster, there are strong positive correlations between cluster member metabolites, like L-tryptophan, and sensory attributes, nutty/sesame and fruity grape. In contrast, the orange cluster metabolites, such as glyceric acid, show more negative correlations with nutty/sesame and positive correlations with astringent mouthfeel scores. Next, we explored how these cluster members (Figure 2(b)) changed over time, and we plotted the identified metabolite features abundance relative to the time zero months (Figures 2(c)–2(g)). The orange cluster members positively associated with astringent mouthfeel scores increased between 0 and 8 months, while the yellow cluster members associated with nutty/sesame decreased between 0 and 8 months. The green cluster members also decreased with time, and cluster members of blue and gray clusters showed less change with time. Lastly, we wanted to integrate these results to the model using metabolite features, sugar RT 13.6, sugar RT 19.4, sugar RT 19.5, arabinose, and L-tryptophan, to quantify "freshness." Notably, the first three components of the model are members of the orange cluster, and the last two components of the model, arabinose and L-tryptophan, are members of the yellow cluster.

This work provides a depiction of modern metabolomic technologies applied to the subject of assessing storage-based aging of SS with time and sensory-based variables. Metabolite profiling such as the one applied in this study offers a promising technology to unravel complex, yet

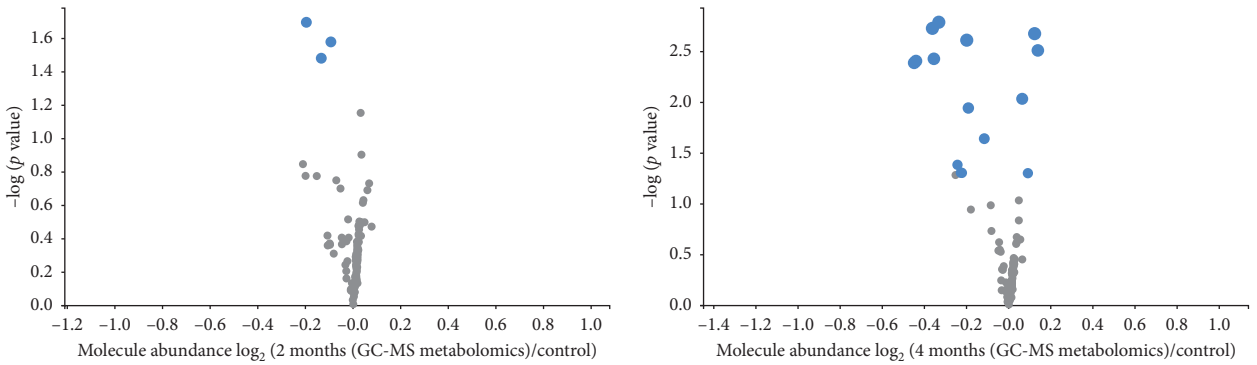
Table 3: Average peak heights (log 2 converted) of analytes resolved from aged soy sauce samples selected on the basis that their relative concentrations changed as a result of aging.

Feature ID/sample	Average peak height (log2)					p value relative to 0 month treatment			
	0 months	2 months	4 months	6 months	8 months	2 months	4 months	6 months	8 months
Aconitic acid 3TMS	22.682	22.688	22.659	22.645	22.603	0.758	0.415	0.196	**0.018**
Arabinose 4TMS 1MOX	29.813	29.743	29.622	29.573	29.456	0.178	**0.011**	**0.007**	**0.001**
Glyceric acid 3TMS	23.978	24.013	24.026	24.071	24.031	0.125	0.091	**0.008**	**0.042**
Glycolic acid 2TMS	23.467	23.500	23.532	23.576	23.537	0.071	**0.009**	**0.001**	**0.005**
Glycylglycine 4TMS	26.950	26.949	26.953	26.934	26.900	0.941	0.887	0.377	**0.019**
Hydroxyproline 3TMS	24.003	24.028	24.042	24.076	24.038	0.343	0.205	**0.045**	0.209
L-Tryptophan 3TMS	26.401	26.205	26.039	25.897	25.761	**0.020**	**0.002**	**0.001**	**≤≤0.001**
Sugar RT 10.763	27.066	26.914	26.711	26.553	26.433	0.167	**0.004**	**0.002**	**≤0.001**
Sugar RT 10.884	27.991	27.858	27.660	27.591	27.465	**0.033**	**0.002**	**0.001**	**≤0.001**
Sugar RT 11.176	25.703	25.610	25.504	25.460	25.362	**0.027**	**0.002**	**0.002**	**≤0.001**
Sugar RT 11.635	24.734	24.536	24.286	24.097	24.003	0.167	**0.004**	**0.002**	**≤0.001**
Sugar RT 11.94	28.916	28.928	28.923	28.960	28.911	0.493	0.634	**0.042**	0.740
Sugar RT 12.287	25.094	24.884	24.655	24.474	24.382	0.142	**0.004**	**0.002**	**≤0.001**
Sugar RT 13.44	28.475	28.368	28.232	28.104	28.003	0.380	**0.041**	**0.040**	**0.001**
Sugar RT 13.594	25.911	25.955	26.003	26.058	26.048	0.232	0.050	**0.017**	**0.004**
Sugar RT 13.824	30.200	30.191	30.152	30.147	30.058	0.796	0.285	0.267	**0.010**
Sugar RT 13.847	30.172	30.075	29.950	29.868	29.758	0.432	**0.049**	0.079	**0.003**
Sugar RT 13.966	29.289	29.183	29.037	28.905	28.798	0.437	0.051	**0.049**	**0.001**
Sugar RT 14.867	30.611	30.513	30.384	30.294	30.186	0.425	**0.049**	0.084	**0.003**
Sugar RT 15.414	26.738	26.728	26.692	26.712	26.639	0.806	0.236	0.510	**0.026**
Sugar RT 15.75	23.527	23.474	23.411	23.401	23.300	0.199	**0.023**	**0.014**	**0.002**
Sugar RT 19.354	23.797	23.857	23.936	24.007	24.001	0.206	**0.003**	**0.003**	**0.001**
Sugar RT 19.525	22.587	22.655	22.711	22.799	22.790	0.185	**0.002**	**0.002**	**0.001**
Sugar RT 20.37	24.714	24.633	24.535	24.420	24.338	0.487	0.113	0.073	**0.003**
Sugar RT 21.555	25.564	25.532	25.482	25.456	25.380	0.567	0.183	0.181	**0.015**
Sugar RT 21.798	22.691	22.663	22.659	22.616	22.543	0.624	0.440	0.095	**0.017**
Unknown RI1426 m/z 189.112	20.615	20.685	20.714	20.785	20.780	**0.007**	**0.004**	**≤0.001**	**≤0.001**
Unknown RI 1428 m/z 186.131	22.611	22.585	22.578	22.594	22.565	**0.043**	**0.015**	0.102	**0.020**
Unknown RI 1448 m/z 218.115	20.486	20.248	20.083	19.797	19.772	0.321	**0.049**	**0.026**	**0.003**
Unknown RI 1465 m/z 232.118	23.360	23.308	23.274	23.242	23.188	**0.026**	**≤0.001**	**0.019**	**≤0.001**
Unknown RI 1475 m/z 156.084	18.805	18.732	18.718	18.761	18.686	**0.013**	**0.001**	0.063	**0.016**
Unknown RI 1490 m/z 144.084	23.012	22.949	22.859	22.804	22.693	0.090	**0.009**	**0.003**	**≤0.001**
Unknown RI 1545 m/z 123.044.	21.027	21.116	21.216	21.340	21.379	**0.002**	**≤0.001**	**≤0.001**	**≤0.001**
Unknown RI 1554 m/z 292.134	21.464	21.508	21.526	21.569	21.529	0.132	0.059	**0.021**	**0.036**
Unknown RI 1563 m/z 345.137	21.223	20.905	20.873	20.896	20.836	**≤0.001**	**≤0.001**	**≤0.001**	**≤0.001**
Unknown RI 1576 m/z 142.105	23.665	23.635	23.569	23.565	23.455	0.536	**0.027**	0.071	**0.002**
Unknown RI 1596 m/z 231.087	20.565	20.585	20.613	20.670	20.634	0.434	0.115	**0.009**	**0.034**
Unknown RI 1643 m/z 245.102	26.367	26.436	26.483	26.566	26.554	**0.033**	**0.006**	**0.001**	**0.002**
Unknown RI 1664 m/z 204.1	23.365	23.303	23.229	23.170	23.129	0.466	0.081	0.071	**0.003**
Unknown RI 1671 m/z 173.087	22.758	22.728	22.658	22.637	22.540	0.419	0.051	**0.031**	**0.002**
Unknown RI 1674 m/z 214.126	21.175	21.159	21.125	21.163	21.104	0.571	**0.042**	0.639	**0.011**
Unknown RI 1691 m/z 129.037	22.474	22.477	22.427	22.410	22.357	0.894	0.141	0.081	**0.008**
Unknown RI 1693 m/z 275.161	21.823	21.872	21.897	21.954	21.929	0.140	0.060	**0.033**	**0.011**
Unknown RI 1769 m/z 129.073	22.545	22.592	22.654	22.731	22.758	**0.029**	**≤0.001**	**≤0.001**	**≤0.001**
Unknown RI 1769 m/z 286.122	21.364	21.369	21.407	21.447	21.491	0.901	0.163	0.054	**0.014**
Unknown RI 1773 m/z 225.105	21.137	21.148	21.169	21.206	21.198	0.677	0.207	0.059	**0.042**
Unknown RI 1775 m/z 129.037	21.828	21.875	21.894	21.940	21.910	0.069	0.051	**0.006**	**0.014**
Unknown RI 1831 m/z 362.163	20.409	20.407	20.380	20.379	20.307	0.942	0.339	0.330	**0.017**
Unknown RI 1838 m/z 217.107	24.666	24.592	24.463	24.388	24.276	0.688	0.156	0.190	**0.011**
Unknown RI 1846 m/z 245.102	25.195	25.276	25.354	25.446	25.462	**0.015**	**0.002**	**≤0.001**	**≤0.001**
Unknown RI 1855 m/z 245.102	27.646	27.732	27.810	27.912	27.908	**0.020**	**0.003**	**≤0.001**	**≤0.001**
Unknown RI 1873 m/z 226.113	23.724	23.760	23.797	23.855	23.839	0.238	**0.029**	**0.004**	**0.006**
Unknown RI 1876 m/z 307.157	22.276	22.338	22.375	22.428	22.406	0.072	**0.024**	**0.020**	**0.003**
Unknown RI 1881 m/z 157.068	24.998	25.050	25.091	25.162	25.148	0.142	**0.043**	**0.009**	**0.003**
Unknown RI 1937 m/z 142.105	22.673	22.663	22.658	22.656	22.622	0.492	0.156	0.501	**0.008**
Unknown RI 1941 m/z 70.065	22.556	22.587	22.606	22.632	22.614	0.249	0.062	**0.013**	**0.022**
Unknown RI 1955 m/z 143.052	21.127	21.104	21.082	21.090	21.042	0.383	0.086	0.202	**0.014**
Unknown RI 1961 m/z 333.137	21.955	22.006	22.053	22.123	22.097	0.089	**0.019**	**0.002**	**0.004**
Unknown RI 1961 m/z 387.158	21.043	21.091	21.142	21.218	21.216	0.097	**0.002**	**0.002**	**0.001**

Table 3: Continued.

Feature ID/sample	Average peak height (log2)					p value relative to 0 month treatment			
	0 months	2 months	4 months	6 months	8 months	2 months	4 months	6 months	8 months
Unknown RI 2013 m/z 174.113	22.964	23.008	23.033	23.092	23.061	0.263	0.074	**0.043**	**0.015**
Unknown RI 2026 m/z 188.128	23.381	23.243	23.179	23.193	23.140	0.060	**0.003**	**0.010**	**0.001**
Unknown RI 2036 m/z 311.124	21.072	21.114	21.190	21.252	21.261	0.279	**0.003**	**≤0.001**	**0.002**
Unknown RI 2052 m/z 174.113	25.791	25.787	25.753	25.751	25.675	0.924	0.387	0.360	**0.049**
Unknown RI 2069 m/z 224.074	20.424	20.302	20.240	20.180	20.037	0.241	0.128	**0.040**	**0.021**
Unknown RI 2082 m/z 238.113	21.850	21.880	21.920	21.987	21.956	0.619	0.233	**0.012**	0.123
Unknown RI 2127 m/z 142.068	21.357	21.255	21.226	21.245	21.142	0.087	0.052	**0.015**	**0.021**
Unknown RI 2144 m/z 326.19	23.523	23.191	23.138	23.153	23.068	**0.030**	**0.016**	**0.023**	**0.008**
Unknown RI 2147 m/z 475.229	21.205	21.008	20.898	20.890	20.803	**0.026**	**0.003**	**0.007**	**0.001**
Unknown RI 2158 m/z 232.118	22.864	22.773	22.709	22.643	22.487	0.217	0.093	**0.026**	**0.015**
Unknown RI 2219 m/z 142.068	21.682	21.553	21.558	21.545	21.491	0.104	0.107	**0.022**	**0.024**
Unknown RI 2252 m/z 217.107	24.622	24.662	24.703	24.737	24.701	0.246	**0.042**	**0.017**	0.060
Unknown RI 2256 m/z 160.079	21.564	21.646	21.716	21.819	21.777	0.230	**0.002**	**0.002**	**0.001**
Unknown RI 2262 m/z 255.134	21.612	21.708	21.787	21.913	21.899	0.148	**0.001**	**0.004**	**≤0.001**
Unknown RI 2280 m/z 232.118	24.412	24.469	24.522	24.551	24.501	0.196	**0.049**	**0.039**	0.063
Unknown RI 2308 m/z 269.15	21.981	22.075	22.163	22.301	22.298	0.108	**0.001**	**0.003**	**≤0.001**
Unknown RI 2310 m/z 387.213	22.158	22.272	22.340	22.409	22.390	**0.042**	**0.002**	**0.004**	**0.001**
Unknown RI 2323 m/z 116.089	19.736	19.845	19.925	20.020	19.999	0.109	**0.004**	**0.002**	**0.001**
Unknown RI 2326 m/z 343.15	21.586	21.724	21.817	21.868	21.868	0.143	**0.002**	**0.001**	**0.001**
Unknown RI 2356 m/z 142.068	20.013	20.081	20.130	20.188	20.145	0.216	**0.002**	**0.016**	**0.006**
Unknown RI 2381 m/z 84.081	22.358	22.383	22.419	22.521	22.495	0.785	0.190	0.085	**0.006**
Unknown RI 2385 m/z 174.113	24.701	24.704	24.700	24.674	24.656	0.687	0.941	0.202	**0.011**
Unknown RI 2446 m/z 98.096	20.979	20.987	21.036	21.155	21.141	0.929	0.077	0.069	**0.014**
Unknown RI 2454 ma 327.131	22.215	22.255	22.351	22.390	22.422	0.496	0.121	0.165	**0.027**
Unknown RI 2470 m/z 217.107	23.859	23.950	24.071	24.153	24.157	0.341	**0.014**	**0.007**	**0.001**
Unknown RI 2477 m/z 204.1	23.443	23.529	23.630	23.725	23.773	0.269	**0.033**	**0.007**	**0.002**
Unknown RI 2489 m/z 217.107	23.387	23.530	23.687	23.804	23.841	0.197	**0.004**	**0.002**	**≤0.001**
Unknown RI 2493 m/z 156.084	25.002	25.151	25.302	25.411	25.439	0.298	**0.005**	**0.004**	**≤0.001**
Unknown RI 2525 m/z 417.182	20.587	20.621	20.711	20.820	20.791	0.577	0.101	**0.034**	**0.018**
Unknown RI 2635 m/z 132.081	21.575	21.645	21.743	21.811	21.764	0.411	**0.013**	**0.006**	**0.006**
Unknown RI 2652 m/z 132.081	20.145	20.274	20.357	20.448	20.365	0.200	**0.035**	**0.009**	**0.027**
Unknown RI 2758 m/z 191.092	23.313	23.268	23.232	23.190	23.118	0.315	0.140	0.065	**0.010**
Unknown RI 2940 m/z 174.113	20.482	20.582	20.663	20.782	20.738	0.212	**≤0.001**	**0.007**	**0.001**
Unknown RI1602 m/z 200.11	21.879	21.870	21.854	21.892	21.813	0.677	0.213	0.477	**0.010**

p value columns depict the result of a statistical comparison with unaged soy sauce sample. Bold values denote p values less than 0.05. For each "unknown" peak, RI represents a relative retention index. TMS: trimethylsiloxane; MOX: methoxamine.



(a) (b)

Figure 1: Continued.

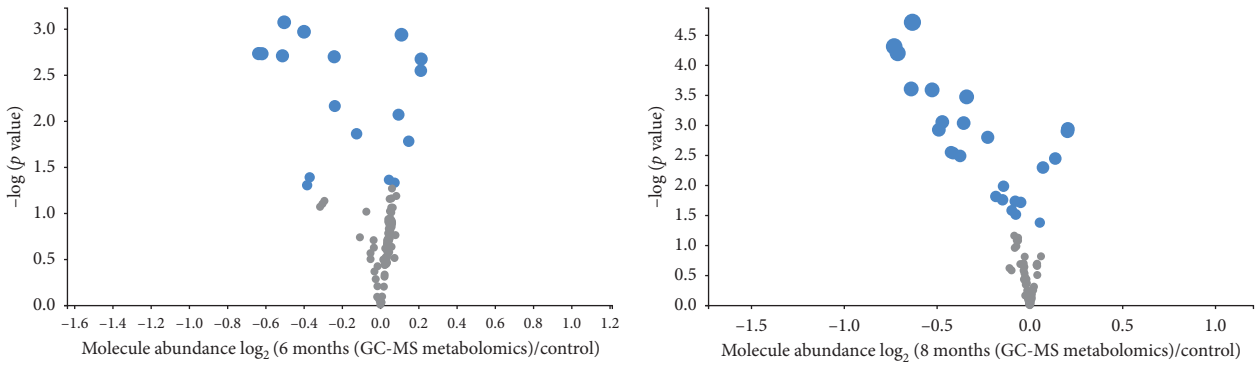


Figure 1: (a–d) Volcano plots of analytes in SS samples stored for eight months at 30°C. Each dot within the graph represents a unique analyte. The x-axis units are the log2-fold changes of the molecular abundances of the aged samples relative to the control $t = 0$ sample. The y-axis units are the negative log of p values comparing the concentrations of the denoted analyte to that of the control. When there were significant differences between the abundances of the analyte compared with the control, meaning a p value less than 0.05, the analyte was denoted with a blue dot; gray dots denote analytes not significantly different in abundance from the control. Thus, as storage time increased from two to eight months, the number of blue analytes increased.

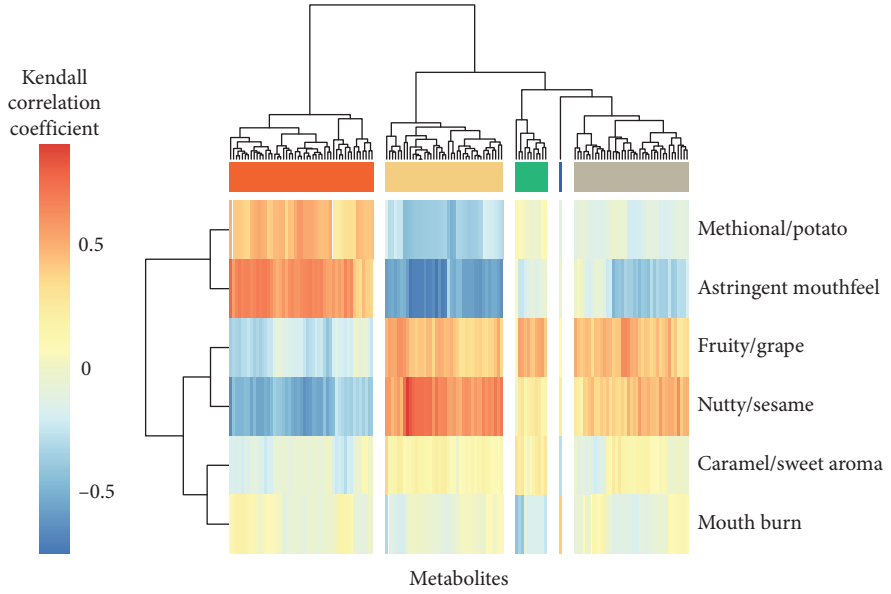


Figure 2: Continued.

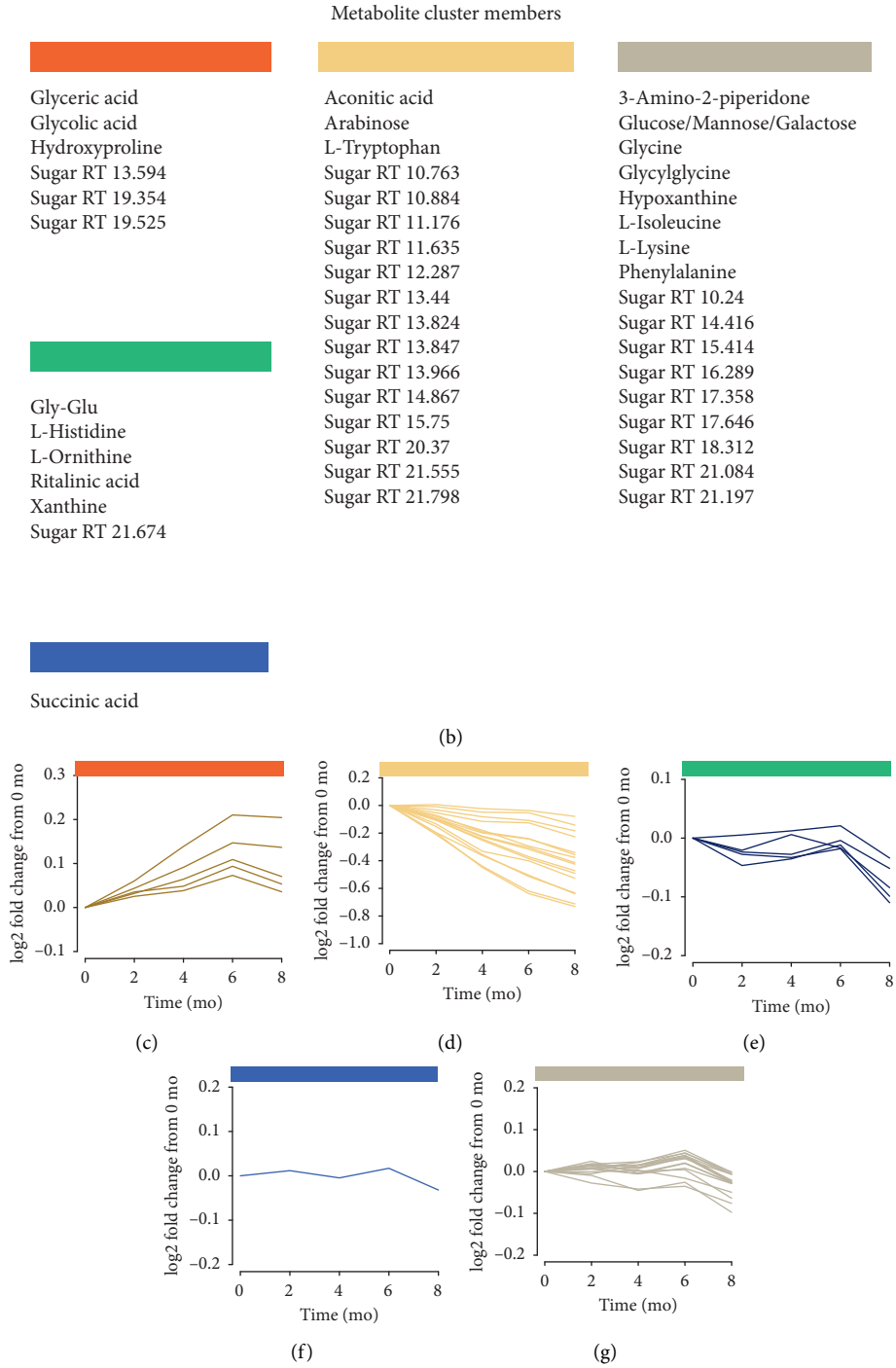


Figure 2: Heat map of correlations between sensory terms affected by storage time and metabolites. Heat coloring denotes the strength and direction of the correlation based on the Kendall correlation coefficient scale. Metabolites with significant correlations ($p < 0.05$) to one or more sensory terms were split into five groups based on hierarchical clustering denoted by the color bars orange, yellow, green, blue, and gray. Sensory variables were similarly grouped by hierarchal clustering based on correlations with metabolites (a). Identified metabolites associated within each hierarchical cluster (b). Unidentified features are not listed and sugar molecules are only confidently assigned at a molecular class level and are differentiated by chromatographic retention time. For each cluster, the changes in abundance for cluster members (molecules) are plotted over time as log2 of the fold change from time 0 month (c–g). The members of the yellow cluster have the greatest change over time.

critical aroma-based attributes in foods such as SS. We further suggest that SS freshness, as a sensory attribute, is a complex term defined by the presence of specific aromatic attributes of fruity/grape and nutty/sesame and absence of methional/potato aroma. More quantitatively, it can be expressed as the models derived from the regression analyses mentioned earlier. Furthermore, we recognize that consumer perceptions of freshness and value of finished food products such as SS may vary based on prior experiences, familiarization, and exposure factors [23]. However, we suggest that a sensory basis for quality assessment provides a key means of discriminating important quality parameters such as those influenced by extended storage or product spoilage. We acknowledge that the changes in sensory attributes were not rationally associated with the changes noted in the chemical profiles, whereas changes in the fruity attribute did not correlate with compounds noted for fruity character, such as ethyl esters. We propose that these finer discoveries, at least on the volatile fraction, are the subject of future investigations such as those done in other works [23], wherein more comprehensive analytical examinations are conducted complemented by sensory validation using the addition of authentic standards to induce specific sensory attributes.

Yet, three novel contributions are served in this work. First, we have established that changes in sensory character are discernable in SS as a function of age or storage time. Second, the sensory character of SS freshness is indeed a multidimensional or meta-term attribute affected by changes in several key aromatic attributes as noted in other foods [24]. Third, a rapid, analytical assessment of SS compounds was developed that can predict or discern aging in SS across storage time with high precision using changes in polar metabolite profiles.

Data Availability

Data can be accessed through contacting author's institution at https://www.minds.wisconsin.edu.

Conflicts of Interest

Author Coon is a consultant for Thermo Fisher Scientific. The remaining authors declare no conflicts of interest related to this work.

Acknowledgments

The authors acknowledge the funding and technical assistance for this work provided by the Kikkoman USA R&D Laboratory (Madison, WI). This work was supported in part by the National Institutes of Health (NIH) grant P41GM108538.

References

[1] D. I. Ellis, H. Muhamadali, D. P. Allen, C. T. Elliott, and R. Goodacre, "A flavour of omics approaches for the detection of food fraud," *Current Opinion in Food Science*, vol. 10, pp. 7–15, 2016.

[2] G. P. Danezis, A. S. Tsagkaris, V. Brusic, and C. A. Georgiou, "Food authentication: state of the art and prospects," *Current Opinion in Food Science*, vol. 10, pp. 22–31, 2016.

[3] J. A. Stefely, N. W. Kwiecien, E. C. Freiberger et al., "Mitochondrial protein functions elucidated by multi-omic mass spectrometry profiling," *Nature biotechnology*, vol. 34, no. 11, pp. 1191–1197, 2016.

[4] S. K. Manier, A. Keller, J. Schäper, and M. R. Meyer, "Untargeted metabolomics by high resolution mass spectrometry coupled to normal and reversed phase liquid chromatography as a tool to study the in vitro biotransformation of new psychoactive substances," *Scientific Reports*, vol. 9, p. 2741, 2019.

[5] J. Ledauphin, C. Le Milbeau, D. Barillier, and D. Hennequin, "Differences in the volatile compositions of French labeled brandies (Armagnac, Calvados, Cognac, and Mirabelle) using GC-MS and PLS-DA," *Journal of Agricultural and Food Chemistry*, vol. 58, pp. 782–7793, 2010.

[6] N. Zhao, C. Zhang, Q. Yang et al., "Selection of taste markers related to lactic acid bacteria microflora metabolism for Chinese traditional Paocai: a gas chromatography-mass spectrometry-based metabolomics approach," *Journal of Agricultural and Food Chemistry*, vol. 64, no. 11, pp. 2415–2422, 2016.

[7] S. Yamamoto, K. Shiga, Y. Kodama et al., "Analysis of the correlation between dipeptides and taste differences among soy sauces by using metabolomics-based component profiling," *Journal of Bioscience and Bioengineering*, vol. 118, no. 1, pp. 56–63, 2014.

[8] E. Chambers and K. Koppel, "Associations of volatile compounds with sensory aroma and flavor: the complex nature of flavor," *Molecules*, vol. 18, no. 5, pp. 4887–4905, 2013.

[9] A. Loutfi, S. Coradeschi, G. K. Mani, P. Shankar, J. B. B. Rayappan, and B. Rayappan, "Electronic noses for food quality: a review," *Journal of Food Engineering*, vol. 144, pp. 103–111, 2015.

[10] P. Cherdchu, E. Chambers, and T. Suwonsichon, "Sensory Lexicon Development Using Trained Panelists in Thailand and the U.S.A.: Soy Sauce," *Journal of Sensory Studies*, vol. 28, no. 3, pp. 248–255, 2013.

[11] M. Imamura, "Descriptive terminology for the sensory evaluation of soy sauce," *Journal of Sensory Studies*, vol. 31, no. 5, pp. 393–407, 2016.

[12] H. Motai, "Soy sauce colors and its significance for quality control," *Journal of the Society of Brewing Japan*, vol. 72, no. 1, pp. 18–20, 1977.

[13] J. Roque, M. Auvray, and J. Lafraire, "Understanding freshness perception from the cognitive mechanisms of flavor: the case of beverages," *Frontiers in Psychology*, vol. 8, p. 2360, 2018.

[14] T. Pujchakarn, S. Suwonsichon, and T. Suwonsichon, "Development of a sensory lexicon for a specific subcategory of soy sauce: Seasoning soy sauce," *Journal of Sensory Studies*, vol. 31, no. 6, pp. 443–452, 2016.

[15] M. C. Meilgaard, G. V. Civille, and B. T. Carr, *Selection and Training of Panel Members*, CRC Press/Taylor & Francis, Boca Raton, FL, USA, 5th edition, 2016.

[16] E. Chambers, K. Sanchez, U. X. T. Phan, R. Miller, G. V. Civille, and B. Di Donfrancesco, "Development of a "living" lexicon for descriptive sensory analysis of brewed coffee," *Journal of Sensory Studies*, vol. 31, no. 6, pp. 465–480, 2016.

[17] S. L. Drake, M. D. Yates, and M. A. Drake, "Development of a flavor lexicon for processed and imitation cheeses," *Journal of Sensory Studies*, vol. 25, no. 5, pp. 720–739, 2010.

[18] M. A. Drake, S. C. McIngvale, P. D. Gerard, K. R. Cadwallader, and G. V. Civille, "Development of a descriptive language for cheddar cheese," *Journal of Food Science*, vol. 66, no. 9, pp. 1422–1427, 2001.

[19] M. M. Meilgaard, G. V. Civille, and T. Carr, "The spectrum™ descriptive analysis method," in *Sensory Evaluation Techniques*, pp. 189–253, CRC Press, New York, NY, USA, 4th edition, 2007.

[20] R Foundation for Statistical Computing, *A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2017.

[21] N. Kiyoko, K. Kamizake, L. Catarini, P. Silva, and H. Prudencio, "Impact of soybean aging conditions on tofu sensory characteristics and acceptance," *Journal of the Science of Food and Agriculture*, vol. 98, pp. 1132–1139, 2017.

[22] P. Dazeley and C. Houston-Price, "Exposure to foods' non-taste sensory properties. A nursery intervention to increase children's willingness to try fruit and vegetables," *Appetite*, vol. 84, pp. 1–6, 2015.

[23] P. Schieberle and W. Grosch, "Evaluation of the flavour of wheat and rye bread crusts by aroma extract dilution analysis," *Zeitschrift für Lebensmittel-Untersuchung und -Forschung*, vol. 185, no. 2, pp. 111–113, 1987.

[24] P. Hongsoongnern and E. Chambers, "A Lexicon For Texture And Flavor Characteristics of Fresh And Processed Tomatoes," *Journal of Sensory Studies*, vol. 23, no. 5, pp. 583–599, 2008.