DOI:10.1021/jf903680d

JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY
ARTICLE

Correlation between Sensory Evaluation Scores of Japanese *Sake* and Metabolome Profiles

Masahiro Sugimoto,*[,†,§] Toshihiko Koseki,[#] Akiyoshi Hirayama,[†] Shinobu Abe,[†] Tomoyoshi Sano,[⊥] Masaru Tomita,[†,§] and Tomoyoshi Soga[†,§]

[†]Institute for Advanced Biosciences, Keio University, Tsuruoka, Yamagata 997-0052, Japan, [§]Systems Biology Program, Graduate School of Media and Governance, Keio University, Fujisawa, Kanagawa 252-8520, Japan, [#]Yamagata Research Institute of Technology, Matsuei, Yamagata 990-2473, Japan, and [⊥]Yamagata Integrated Agricultural Research Center, Tsuruoka, Yamagata 999-7601, Japan

The aim of this study was to explore the association between taste and metabolite profiles of Japanese refined *sake*. Nontarget metabolome analysis was conducted using capillary electrophoresis mass spectrometry. *Zatsumi*, an unpleasant not clear flavor, and sweetness, bitterness, and sourness were graded by four experienced panelists. Regression models based on support vector regression (SVR) were used to estimate the relationships among sensory evaluation scores and quantified metabolites and visualized as a nonlinear relationship between sensory scores and metabolite components. The SVR model was highly accurate and versatile: the correlation coefficients for whole training data, cross-validation, and separated validation data were 0.86, 0.73, and 0.73, respectively, for zatsumi. Other sensory scores were also analyzed and modeled by SVR. The methodology demonstrated here carries great potential for predicting the relevant parameters and quantitative relationships between charged metabolites and sensory evaluation in Japanese refined sake.

KEYWORDS: Japanese sake; sensory evaluation; support vector regression; metabolome; capillary electrophoresis mass spectrometry

INTRODUCTION

Elucidation of the relationship between the quality characteristics of foodstuffs and beverage products and their constitutive components is important for various research fields including food engineering, biochemistry, and medical sciences and may contribute to the improvement or investigation of food taste/flavor and to the study of the functional characteristics of foods. Although the food/beverage industry already carries out human assessment of taste/flavor, to fully understand these relationships, objective, reproducible, sensitive, and efficient tools are needed that mimic human sensory perceptions. To obtain sensitive, fair, and reproducible assessment criteria, the development of artificial tongues (*1*−*3*) and noses (*4*) is important for scoring taste and smell and for eliminating the bias of human evaluators.

Key components with sensory relevance have been investigated in amino acids (*5*, *6*), sugars, organic and inorganic acids (*7*, *8*), peptides (*9*), and larger nonvolatile metabolites (*10*) present in several beverages. Although attempts have been made to correlate one or more beverage components with sensory data obtained from human tasters, contradictory findings have been reported, and the analysis of multiple substances is both essential and intricate. To predict sensory evaluation scores or to classify the characteristics of a food or beverage by its constituents, pattern recognition methods have been applied routinely. Support vector machines (SVM) (*11*, *12*) have been used widely to classify the regional origin of wines using UV−visible spectrophotometry data (*13*). Artificial neural networks (ANN) and self-organizing maps (SOM) helped to predict the growth region of red wines on the basis of their constitutive metals (*14*); they were also used for the characteristic metabolome profiling of tomatoes (*15*). The ANN has been employed as an artificial nose in classifying the smell of wines (*4*) and teas (*5*). These studies concluded that pattern recognition yielded higher predictive performance than conventional statistical analyses such as principal component analysis (PCA) and partial least-squares discriminant analysis (PLS-DA) (*16*, *17*). The ability of these methods to solve nonlinear problems was confirmed; ANN can approximate any function with a single hidden layer (*18*), and SVM can separate given, nonlinearly distributed sample data by using hyperspace with a kernel (*19*). Currently such intelligent techniques, rather than linear methods, are used to obtain accurate discriminations and classifications.

Electric tongues and noses have been developed to estimate the taste or fragrance of water, green tea, coffee, orange juice, and miso soup (*1*,*20*−*22*). They usually predict the sensory evaluation scores assigned quantitatively by human tasters. Both linear predictive models such as multiple variable linear regression (MLR) and PLS (*22*) and nonlinear models, for example, ANN, have been used routinely to solve regression problems. Although the linear models facilitate simple interpretation, their

*Author to whom correspondence should be addressed (e-mail msugi@sfc.keio.ac.jp; telephone +81-235-29-0528; fax +81-235-29-0574).

 Published on Web 12/04/2009

regression ability is limited. On the other hand, although it is more difficult to interpret nonlinear models, they can potentially identify nonlinear relations. Once a mathematical model has been constructed for the prediction of sensory evaluation based on the constituents of a food or beverage, the contribution of constitutive variables to the sensory evaluation can be analyzed and the constituent profiles can be optimized by simulation to maximize or minimize sensory evaluations. Even with black box-type models such analyses can be conducted by systematically altering the variables and observing the models' response (*23*). Therefore, in the selection of applicable predictive models, their high predictive ability rather than their interpretability should be the primary focus.

Sake, a Japanese rice wine brewed with *Saccharomyces cerevisiae* yeast, is a traditional, favorite alcoholic beverage of Japanese consumers. It has played a central role in Japanese culture for at least 1700 years. Epidemiological studies indicated that moderate alcohol consumption may be beneficial; it reduces the risk for coronary heart disease and exerts antiaging effects on the skin (*24*−*28*). Sake is the product of a unique brewing process that starts by polishing rice kernels to remove their surface and separating the polished rice into two groups. One group is steamed, *Aspergillus oryzae* spores are added, and this is followed by a 2-day incubation to generate *koji* rice (malted rice). Then *shubo*, a yeast mash, is prepared by adding *S. cerevisiae* and water, and after 2 weeks, more water and the second group of preserved polished rice are added. After 1 month, shubo yields *moromi* (the final mash). Then the clear liquid is extracted by filtering; after pasteurization to inactivate microorganisms, the product is bottled as sake (*29*−*31*). In this process, koji produces enzymes that convert starch to sugar, proteolytic enzymes that break down proteins, and >50 other enzymes that are responsible for the flavor and taste of sake (*29*). Although there are several variations, sake is roughly categorized into two types depending on the filtration methods used in the final step: cloudy sake (*Nigorisake* or *Doburoku*) contains white particles due to the incomplete elimination of mash by rough filtration, and refined sake (*Seisyu*) is a clear and clean liquid.

To characterize perceivable attributes in Japanese refined sake, including sweetness, sourness, saltiness, bitterness, umami, and especially clearness or its inverse texture, *zatsumi*, is important (*32*−*34*). The taste parameter zatsumi, the inverse of a sense of clarity, is not limited to the physical turbidity of a liquid (*35*); it also indicates unsmoothed and unpleasant attributes, that is, negative sensory characteristics that produce a detracting flavor in clear beverages. This orosensory phenomenon is thought to be caused by low concentrations of bitter-tasting compounds (*33*). During the sake brewing process, white rice grains with a high protein content tend to produce low-quality sake because of the resultant zatsumi flavor. Although it has been thought that certain rice protein digestion products contribute to this unpleasant taste, the responsible compound has not yet been identified. Attempts have been made to correlate analytical data with distinct components in sake (*30*,*32*,*33*,*36*,*37*). Sake mainly contains alcoholic and esteric compounds that characterize its aroma, glucose, and especially amino and organic acids that determine its taste (*32*). Although individual amino acids are known for their contribution to enhancing a pleasant taste (*38*), an excess of total amino acids is expected to introduce unpleasant zatsumi (*30*, *33*, *39*). Thus, the relationship between amino acids and sensory scores is expected to be complex. In addition to amino acids, organic acids, especially lactate, which in some instances is added to shubo at the same time as yeast to prevent the growth of nonuseful microorganisms, are also major components that may affect the taste of sake (*29*, *40*). Moreover, peptides contribute importantly to the perceivable flavor of sake (*33*, *41*). Therefore, to identify the compounds other than amino acids that contribute to sensory evaluations, comprehensive *omics* analyses are necessary.

The aim of the present study was to elucidate the relationship between the metabolite profiles and the sensory evaluation of zatsumi in refined Japanese sake. We conducted nontarget metabolome analysis using capillary electrophoresis−time-of-flight mass spectrometry (CE-TOFMS) for the simultaneous quantification of charged and soluble metabolites such as amino acids, organic acids, and sugar phosphates (*42*,*43*) and performed sensory evaluations of zatsumi and related organoleptical criteria by four professional panelists. Besides clustering and correlation analyses that incorporate the obtained metabolic profiles and other available properties, we developed predictive models based on vector regression (SVR), an extended version of SVM (*13*, *19*), for estimating the assigned sensory scores. Although the methodology presented here is limited to mine the links among observable data and a sensory evaluation parameter, ours is a new example of an integrated study; it combines simultaneous metabolite measurements and a mathematical predictive technique to predict the quality of alcoholic beverages from their contents.

MATERIALS AND METHODS

Sample Preparation. We used 49 commercially available types of refined sake from Yamagata Prefecture in the Tohoku region of Japan; they included *junmai*, special junmai, and *junmai ginjo*, made without the artificial addition of alcohol. Junmai and special junmai are made from a combination of highly refined polished rice grains (polished down to 70% (junmai) or 60% (special junmai) of their original size), koji (malted rice (*29*)), and fresh spring water. Junmai ginjo is brewed for a relatively longer period at lower temperatures using highly polished rice (at least 60%) (*44*, *45*).

We took 100 μL from each sake sample and added 1000 μL of Milli-Q water (Millipore, Bedford, MA) containing internal standards with concentrations of up to 200 μM. Methionine sulfone and 3-aminopyrrodine were the positively charged (cationic) and 2-(*N*-morpholino)ethanesulfonic acid (MES) and Trimesate the negatively charged (anionic) internal standards. The samples were centrifuged at 12000 rpm for 15 min and poured through a 5 kDa cutoff membrane filter (Millipore) to remove suspended solids, and the filtrate was immediately measured by CE-MS.

Instrumentation. The instrumentation and measurement conditions of CE-TOFMS are described elsewhere (*46*, *47*). All CE-ESI-MS experiments were performed using an Agilent CE capillary electrophoresis system, an Agilent G6220A LC/MSD TOF system, an Agilent1100 series isocratic HPLC pump, a G1603A Agilent CE-MS adapter kit, and a G1607A Agilent CE-ESI-MS sprayer kit (Agilent Technologies, Waldbronn, Germany). The CE-MS adapter kit includes a capillary cassette that facilitates thermostating of the capillary. The CE-ESIMS sprayer kit simplifies coupling the CE system with the MS systems; it was equipped with an electrospray source. For system control and data acquisition, we used G2201AA Agilent ChemStation software for CE and Agilent MassHunter software for TOF-MS. With regard to the anion model, the original Agilent SST316Ti stainless steel ESI needle was replaced with passivated SST316Ti stainless steel (with 1% formic acid and 20% isopropanol aqueous solution at 80 °C for 30 min) and platinum.

CE-TOFMS Conditions for Cationic Metabolite Analysis. The CE-TOFMS conditions for cationic metabolite analysis were as described elsewhere (*46*). Sample separation was in fused-silica capillaries (50 μm i.d. $\times$ 100 cm total length) filled with 1 mol/L formic acid as the reference electrolyte. Sample solutions were injected at 50 mbar for 3 s, and a voltage of 30 kV was applied. The capillary temperature was maintained at 20 °C, and the temperature of the sample tray was kept below 5 °C. The sheath liquid, composed of methanol/water (50% v/v) and 0.1 μmol/L hexakis(2,2-difluoroethoxy)phosphazene (Hexakis), was delivered at 10 μL/min. ESI-TOF-MS was conducted in the positive ion mode.

The capillary voltage was set at 4 kV; the flow rate of nitrogen gas (heater temperature = 300 °C) was set at 10 psig. In TOF-MS, the fragmenter, skimmer, and OCT RF voltage were set at 75, 50, and 125 V, respectively. Automatic recalibration of each acquired spectrum was performed using reference masses of reference standards ([^{13}C isotopic ion of protonated methanol dimer (2MeOH + H)]$^+$, m/z 66.06371) and ([protonated Hexakis (M + H)]$^+$, m/z 622.02896). Mass spectra were acquired at the rate of 1.5 cycles/s over a m/z 50−1000 range.

CE-TOFMS Conditions for Anionic Metabolite Analysis. The CE-TOFMS conditions for anionic metabolite analysis were described elsewhere (*47*). A commercially available COSMO(+) capillary (50 μm i.d. × 110 cm) (Nacalai Tesque, Kyoto, Japan), chemically coated with a cationic polymer, was used as the separation capillary. A 50 mM ammonium acetate solution (pH 8.5) was the electrolyte for CE separation. Prior to the first use, a new capillary was flushed successively with the running electrolyte, 50 mmol/L acetic acid (pH 3.4), and then the electrolyte again for 10 min each. Before each injection, the capillary was equilibrated for 2 min by flushing with 50 mM acetic acid (pH 3.4) and then for 5 min by flushing with the running electrolyte. A sample solution (30 nL) was injected at 50 mbar for 30 s, and −30 kV of voltage was applied. The capillary temperature was thermostated to 20 °C, and the sample tray was cooled to below 5 °C. An Agilent 1100 series pump equipped with a 1:100 splitter was used to deliver 10 μL/min of 5 mM ammonium acetate in 50% (v/v) methanol/water containing 0.1 μM Hexakis to the CE interface, where it was used as a sheath liquid around the outside of the CE capillary to provide a stable electrical connection between the tip of the capillary and the grounded electrospray needle. ESI-TOF-MS was conducted in negative ionization mode; the capillary voltage was set at 3500 V. For TOF-MS, the fragmenter, skimmer, and Oct RF voltage were set at 100, 50, and 200 V, respectively. A flow rate of drying nitrogen gas (heater temperature = 300 °C) was maintained at 10 L/min. Automatic recalibration of each acquired spectrum was performed using reference masses of reference standards ([^{13}C isotopic ion of deprotonated acetic acid dimer ($2CH_3COOH - H$)]−, m/z 120.03841), and ([Hexakis + deprotonated acetic acid ($CH_3COOH - H$)]−, m/z 680.03554). Exact mass data were acquired at a rate of 1.5 spectra/s over a m/z 50−1000 range.

Data Processing. Raw data were analyzed with our proprietary software named MasterHands (*46*−*48*). Data analysis starting with noise-filtering, baseline correction, peak detection, and integration of the peak area from sliced electropherograms (m/z 0.02 width). Subsequently, the accurate m/z value for each peak detected within the time domain was calculated with Gaussian curve-fitting to the peak along the m/z axis. Dynamic programming and the simplex optimization method were used to explore the time normalization function for matching peaks in multiple measurements (*49*). The detected peaks with small differences in m/z (< 20 ppm) and normalized migration time (< 1 min) were treated as the peaks derived from a metabolite. Subsequently, neutral compounds, salt ions related to Na^+ and K^+, observed under our measurement conditions were removed. Redundant features such as fragments, adducts, isotopes, dimers, trimers, and their combination, for example, the adduct ion of dimers, were also eliminated on the basis of established m/z differences (*50*). Spike noise, CE-specific noise showing small and narrow peaks, and low-quality (not peak-like shape) results were also eliminated. For the remaining features, metabolite identities were assigned by matching their m/z values and migration times with those of standard compounds.

Sensory Evaluation. To enhance the variety in sensory evaluation scores simultaneously, we randomly selected sake samples with different sake meter value (SMV), acid degree, sake type, rice type, yeast, and sake brewing company. The sensory evaluation scores were assigned by a panel of four experts who applied typical sensory evaluation criteria for refined sake. Porcelain cups generally used for drinking sake were used. To adapt each sake sample to the cup, it was filled one-third to half with sake, and after several minutes, the content was disposed and the cup was immediately filled with the test sample. The temperature of all samples was controlled at 18 °C on water. The evaluators were blinded to the brand names to eliminate prejudicial bias. Although the need for visual masking to eliminate sensory evaluation bias attributable to the color of red wine has been reported (*51*), we did not mask our samples because they were almost achromatically clear. For sensory characterization, sweetness, sourness, bitterness, and zatsumi were graded from 1.0 to 5.0 with a score resolution of 0.5 point. Grade 5 indicated the highest grade of sweetness, sourness, bitterness, and zatsumi intensity and grade 1 the lowest grade of these parameters. To reduce panelist bias, a sake sample having SMV and acid degree closest to the average of 49 samples was selected as a reference, for all of which sensory scores were graded as 3, and all sensory scores for the other 48 samples were graded relative to this reference sample. The evaluators tasted the reference sample iteratively after tasting approximately 10 test samples; all test samples were evaluated at least twice, and the corrected final scores were used for subsequent analyses. Although others assessed total quality and several other criteria such as umami, aftertaste, bulge, and smoothness in addition to the four criteria used in this study as perceivable sake characteristics (*32*), we confined our study to sweetness, sourness, bitterness, and zatsumi because the definition of the other qualities is more abstract and thus less subject to consensus.

Heatmap Visualization and Statistical Analysis. The measured metabolite concentrations were transferred to Z scores, clustered on the basis of Euclidean distance, and visualized as a heatmap representation using Mev TM4 software (Dana-Farber Cancer Institute, Boston, MA) (*52*). A relevance network representing networks among metabolites with a high correlation coefficient was also visualized with the same software. The Pearson method with two-tailed p values calculated by GraphPad Prism ver. 5.02 (GraphPad Software, Inc., San Diego, CA) was used to evaluate the correlation between two parameters. To develop the multiple linear regression (MLR) model, we first removed the absolute values of correlation coefficients below a specified threshold between the concentration of each metabolite and the sensory evaluation score because for subsequent analysis the number of features had to be lower than the number of samples. Then we performed stepwise feature selection (forward and backward) to select the minimum feature sets to eliminate their multicolinearity; a new feature was added at $p < 0.25$, and a feature was removed at $p > 0.25$. JMP version 7.0.1 (SAS Institute, Cary, NC) was used for feature selection and to develop the multiple linear regression model (MLR).

Support Vector Regression (SVR). As the theories underlying SVM and SVR have been detailed elsewhere (*13*, *23*, *53*), we provide only a brief description. SVM maps place data sets into a high-dimensional space and then classify them using a hyperplane of computed maximal margins between them. The optimally identified hyperplane in the feature space corresponds to a nonlinear decision boundary in the input space. SVR uses the same principles as SVM for classification with only a few minor modifications; it applies a margin of tolerance set in approximation to the SVM that would have arisen from the problem. SVR is solved by a quadratic optimization problem.

The SVM is a binary discriminant classification tool that maximizes the margin using a hyperplane between sample data sets that exist in high-order dimensional space. SVM for the regression problem employs the principle

$$\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{M}(\xi_i + \xi_i^*) \quad (1)$$

$$y_i - \langle w^T \times \phi(x_k)\rangle - b \le \varepsilon + \xi_i \quad (2)$$

$$\langle w^T \times \phi(x_k)\rangle + b - y_i \le \varepsilon + \xi_i^* \quad (3)$$

$$i = 1, 2, \ldots, M, \xi_i, \xi_i^* \ge 0 \quad (4)$$

where $\phi(x)$ is the kernel function, w is the separation hyperplane, the pairing (x_i, y_i) is the sample and sensory evaluation score, M is the number of variables in the training data set, C is a trade-off between training error and margin, and ξ and ξ^* are the slack variables representing upper and lower constraints on system output. SVR generates the predictions using the formula

$$f(x) \equiv \sum_{i=1}^{n} \theta_i \phi(x, x_i) + b \quad (5)$$

The radial bias function (RBF) [6], generally used for regression problems, was used as the kernel function:

$$\phi(x, y) = \exp(-\gamma\|x - y\|^2) \quad (6)$$

We trained SVR models in the ranges of 0.1, 0.2, ..., 1.0 for γ, 0.00, 0.02, ..., 0.18 for ε, and 0.001, 0.005, 0.01, 0.05, ..., 100.0 for C to search the

parameter set. First, we randomly separated all data sets into two groups, a training and a validation data set. Cross-validation (CV) was performed with only training data sets to optimize the training parameters yielding the best correlation coefficients for observed and predicted data sets. In this process, the training data set was randomly separated into two groups including 90% (A) and 10% (B) of the data, and a model was developed using data sets (A); performance was validated by using data sets (B). This process was repeated 10 times for selecting all data sets as (B) (10-fold CV). Subsequently, a predictive model was developed using the optimized training parameters,and the data sets preserved as validated tests were predicted. The correlation coefficients obtained using all training data sets, CV, and validation data are described as R_t, R_{cv}, and R_v in the following sections. For analysis using SVR, SVM light (v6.02, Cornel University, Ithaca, NY) (*54*) was employed with default options except for γ, C, and ε.

Feature Selection for the SVR Model. Although predictive models such as ANN and SVR can be applied to multiple-feature problems, feature selection, that is, the reduction of features, is usually performed prior to the development of the model. Feature selection introduces a reduction in model complexity; this prevents the overfitting of the model to a specific problem and also eliminates unnecessary parameters, thereby contributing to improved interpretability of the final model (*55*). Here we compared the predictive accuracy of SVR using data sets without feature selection and three feature selection methods: (i) stepwise feature selection; (ii) correlation-based feature subset selection (CFS) that evaluates the worth of a subset of attributes by considering the individual predictive ability of each feature along with the degree of redundancy among features; and (iii) a relief algorithm (RA) that evaluates the worth of a parameter by repeatedly sampling an instance and considering the value of the given parameter for the nearest instance of the same and different class (*53*, *56*, *57*). We used JMP for procedure i and weka data-mining software (*56*) with default parameters for procedures ii and iii.

Simulation of the SVR Model with Feature Alterations. Variations in the evaluation scores predicted by the trained SVR were computationally simulated by systematically altering the concentration of each metabolite. For each of the 49 samples, one metabolite concentration was systematically altered within a range from its minimum to its maximum; the concentration of the other metabolites corresponded to the actual measured values. This procedure was repeated until all metabolites had been selected for variation.

RESULTS AND DISCUSSION

Sensory Evaluation Test and Resultant Metabolite Profiles. CE-TOFMS analyses identified 536 ± 17 (SD) and 201 (±9) non-redundant features in positively and negatively charged mode, respectively. The features that metabolite identities presented in our standard library were assigned, and >25 samples were visualized as a heatmap with relevant information including the sensory evaluation scores and their correlation coefficients (**Figure 1**). Overall, amino acids and organic acids (pink and green in the bar graph) showed relatively higher concentrations than other metabolites. Consistent with previous results (*7*, *36*) among amino acids we observed relatively higher concentrations of alanine (2093 ± 504.5 μmol/L), glycine (1166 ± 227.6 μmol/L), and proline (1034 ± 151.4 μmol/L) and lower concentrations of tryptophan (20.10 ± 11.72 μmol/L) and methionine (36.22 ± 37.47 μmol/L). Among organic acids, especially lactate (2590 ± 817.1 μmol/L), succinate (1855 ± 252.7 μmol/L), and malate (783.4 ± 258.0 μmol/L) manifested relatively higher concentrations. The large standard deviation (SD) of arginine compared to glutamate (816.8 ± 227.7 μmol/L) and aspartate (394.3 ± 171.5 μmol/L) was also consistent with previous reports (*36*) as was the high level of agmatine (957.5 ± 414.4 μmol/L), which is produced from arginine by arginine decarboxylase (EC 4.1.1.19) during the sake brewing process and is also a precursor of polyamine that is prevalent in wine but not detected in sake (*58*). Arginine and other amino acids in sake originated from koji, where the amino acids were produced by rice protein degradation and not used by yeast, and the product of yeast was not used.

The average sensory evaluation scores were 3.2 ± 0.7 (sweetness), 3.1 ± 0.7 (sourness), 3.2 ± 0.7 (bitterness), and 3.3 ± 0.6 (zatsumi), implying that difference in the evaluation scores assigned by the different panelists was <2 points in most cases because there was a resolution of 0.5 between neighboring scores. Among the sensory evaluation scores, only zatumi and sourness showed a relatively weak but positive correlation (R = 0.477, p = 0.0005); the other pairs of sensory scores showed independence. The average value of these scores was close to 3.0. These score values were the same as those assigned to the reference sake having SMV, acid degree, and alcoholicity of an almost average value, implying their correlation with the sensory evaluation scores. A small SMV value is indicative of sweetness due to a higher sugar content, whereas a larger value indicates bitterness due to a low sugar content. A large degree of acidity and alcoholicity yields a bitter and dense taste, whereas low degrees produce a sweet or mild flavor. We found that the degree of acidity was positively correlated with sensory evaluation scores for sourness (R = 0.413, p = 0.0059) and zatsumi (R = 0.447, p = 0.0026), and SMV was negatively correlated with sourness (R = −0.423, p = 0.0028); the correlation was not strong, and other parameters also manifested few correlations ($|R| < 0.4$). Thus, these three parameters alone might not be sufficient to characterize the flavor of sake.

On the heatmap (**Figure 1**), most amino acids are shown as relatively clear clusters compared to other metabolites (B in **Figure 1**). Another metabolite cluster (A) that includes seven organic acids and a sweet amino acid, proline, was also observed at the center. Corresponding to metabolite clusters A and B, sake samples are clustered in D or E. Among amino acids, only 4-aminobutyrate (GABA), a common nitrogen in sake, and 20 types of protein-containing amino acids and ornithine (*9*) were not included in these clusters, although there was a positive correlation with zatsumi (R = 0.319, p = 0.026). As demonstrated on the correlation heatmap (right of **Figure 1**), the metabolites in both clusters frequently correlated positively with sourness, zatsumi, or the degree of acidity. The metabolites in cluster C also exhibited unique profiles; their concentration in samples F and G was strikingly high. The correlation between total concentrations or these clusters and each sensory score is summarized in Table S1 of the Supporting Information. Sweetness and bitterness manifested no significant correlation. In contrast, sourness was significantly correlated with clusters A (R = 0.671, $p < 0.0001$) and B (R = 0.453, p = 0.0011) and total concentration (R = 0.559, $p < 0.0001$). Zatsumi was also significantly correlated with clusters A (R = 0.483, p = 0.0004), B (R = 0.594, $p < 0.0001$), and C (R = 0.344, p = 0.0156) and total concentration (R = 0.596, $p < 0.0001$). Thus, it is reasonable to suggest that metabolite cluster A and total concentration contribute to sweetness and zatsumi, respectively, and sourness and bitterness were not or less strongly related with these clusters.

Figure 2 depicts the relevance networks of detected metabolites; highly correlated metabolites are visualized in close proximity with links showing the intensity of the square of the correlation coefficient. Notably, all metabolites in this network, except for agmatine and glutamine, correlated positively with zatsumi, and amino acids with known sensory feature mapped in relatively close nodes. Regarding sourness, glutamine and asparagine were neighbors of hypoxanthine (*59*), a metabolite indicating umami. Guanine, also relevant to umami (*60*), mapped between the bitter metabolites tyrosine and phenylalanine. With respect to sweetness, although two dipeptides, alanine−alanine

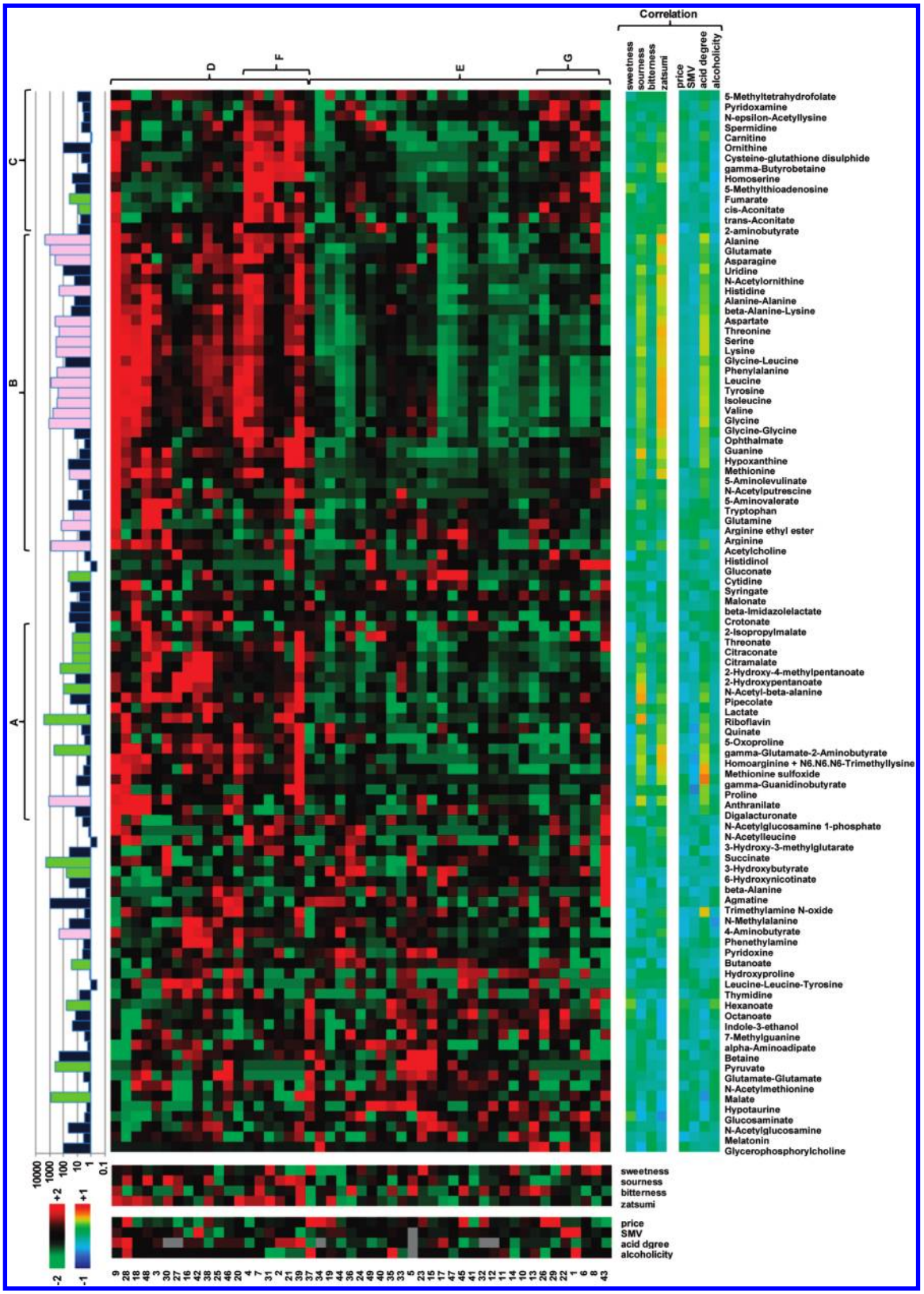

Figure 1. Heatmap showing the quantified metabolic profiles of 49 refined sake samples, sensory evaluation scores for zatumi, bitterness, sourness, and sweetness, and other relevant information (price (tax exclusive), SMV, acid degree, alcoholicity (in black−green−red scheme)) obtained from the bottle label. The bar graph shows the absolute concentration (μmol/L) of individual metabolites on the log axis. Amino acids and organic acids are pink and green, respectively. Numbers from 1 to 49 indicate sample number. Sake sample 1 was used as the reference in sensory evaluation tests. All concentration and sensory score values in the heatmap were transformed into *Z*-scores and clustered with Euclidean distances. Gray areas in the heatmap of relevant information indicate missing data. The correlation between each metabolite and sensory scores or relevant information is shown separately at the bottom of the heatmap (in rainbow scheme). See main document for alphabetical labels.

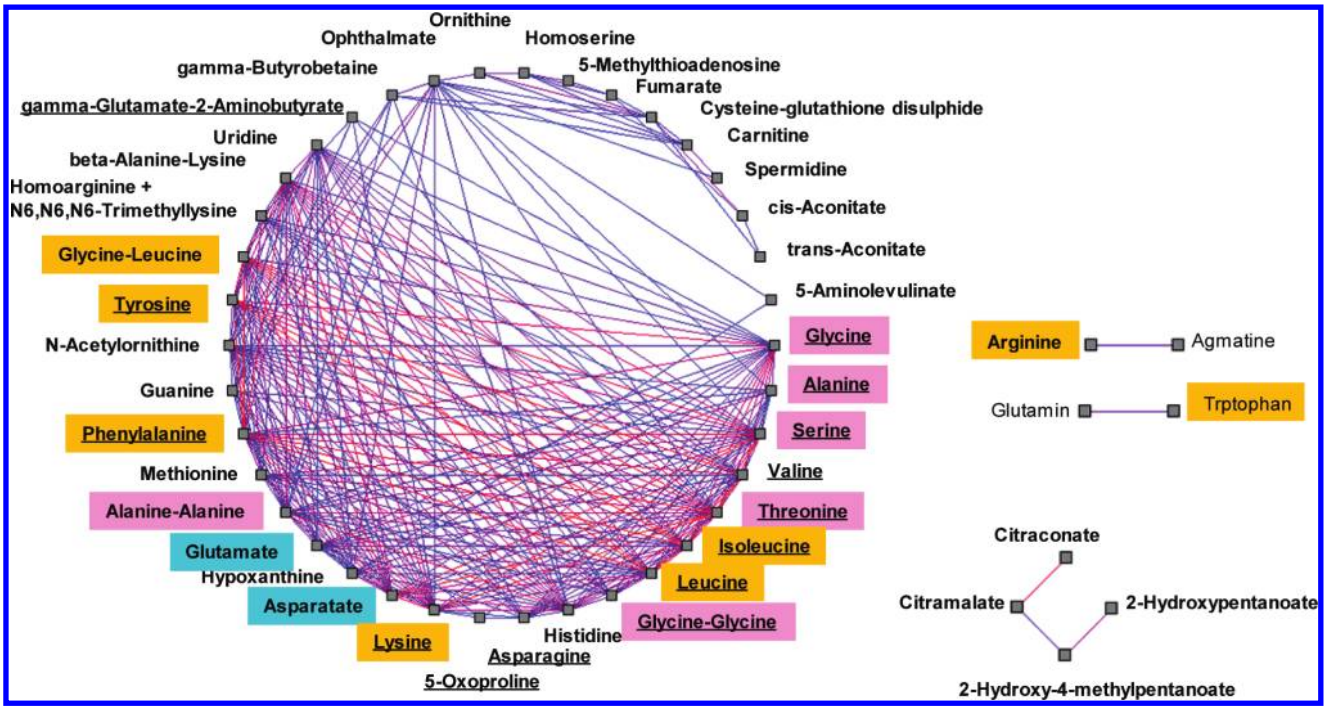

Figure 2. Relevance network of quantified metabolites. Metabolites with their links are shown when the square correlation coefficient (R^2) between two metabolites was >0.51 by the permutation test. The link color indicates the intensity of R^2 from 0.51 (blue) to 1 (red). The background color of metabolite names indicates taste, that is, sweet (pink), bitter (orange), and sour (bright blue). Metabolites in bold and regular font indicate positive and negative correlation with zatsumi, respectively, and underlined metabolites exhibited a strong correlation with zatsumi at $R > 0.55$.

Table 1. Multiple Regression Model for Predicting Sensory Evaluation Values of Zatsumi from Selected Metabolite Concentrations

metabolite	coefficient	error	*p* value	water solubility (g L^{-1} at 25 °C)
(intercept)	0.273	0.112	0.0201	
alanine	0.863	0.225	0.0006	167
homoarginine + *N*6,*N*6,*N*6-trimethyllysine	−0.606	0.174	0.0015	50, >50[a]
glucosaminate	−0.291	0.135	0.0385	100
γ-butyrobetaine	0.215	0.178	0.2346	>50[a]
riboflavin	0.377	0.132	0.0074	68×10^{-3}
carnitine	−0.431	0.130	0.0022	100[b]
5-oxoproline	0.522	0.142	0.0008	129[a]
N-acetylglucosamine-1-phosphate	0.430	0.103	0.0002	>50[a]

[a] Water solubility was obtained from Sigma-Aldrich Inc. (St. Louis, MO). Data for metabolites without this remark were obtained from the CAS database (http://www.cas-japan.jp/expertise/cascontent/registry/index.html). [b] The value was obtained at 20°C.

and glycine−glycine were slightly isolated, the amino acids glycine, alanine, serine, and tyrosine were strongly correlated.

Supplementary Figure S1 of the Supporting Information shows the plots of PCA. In the classification of sake (Figure S1A), only junmai ginjo (red) was not clearly separated; rather, it slightly converged compared to the other types. This observation indicates that the test samples had relatively similar metabolic profiles. However, the accumulative contribution of this PCA was low (47.7% even in the third PC), and neither the yeast type (Figure S1B, Supporting Information) nor the rice type (Figure S1C, Supporting Information) exhibited type-specific features.

Prediction Results Using the Multiple Regression Model. The multiple regression model (MLR) with the metabolites selected by stepwise feature selection is presented in **Table 1** for zatsumi and in Supporting Information Table S2 for other sensory scores. The correlation coefficient matrices between two metabolites are shown in Supporting Information Table S3. The selected metabolites were independent of one another because the highest correlation between two parameters was 0.717 between γ-butyrobetaine and carnitine. The correlation coefficient between predicted and observed scores was $R_t = 0.849$ ($p < 0.0001$) using complete training data sets, $R_{cv} = 0.724$ ($p < 0.0001$) in CV, and $R_t = 0.654$ ($p = 0.0787$) using test data sets.

In the MLR model (**Table 1**), alanine, γ-butyrobetaine, riboflavin, 5-oxoproline, and *N*-acetylglucosamine 1-phosphate exhibited positive coefficients, possibly resulting in increased zatsumi. On the other hand, homoarginine + *N*6,*N*6,*N*6-trimethyllysine, which was not separated and observed as a single peak by CE-MS, glucosaminate, and carnitine manifested negative coefficients, possibly leading to decreased zatsumi. Except for alanine and 5-oxoproline (471.7 ± 246.9 μmol/L), the concentration of other metabolites was relatively low (< 5.164 μmol/L). In addition, the significance of alanine ($p = 0.0006$) and 5-oxoproline ($p = 0.0008$) was relatively large, suggesting that these two metabolites contributed mainly to zatsumi. Among metabolites of the established MLR, only alanine produces sweetness. Although 5-oxoproline is reacted to glutamate, known by its umami flavor, 5-oxoprolinase (EC 3.5.2.9) was observed in wine samples (*61*), its sensory property is not known. The water solubility of the metabolites in the MLR model was large

(> 50 g/L) except for riboflavin (68 mg/L), known as vitamin B2; this suggests that the zatsumi flavor is not attributable to insoluble particles. Iwano et al. (*32*) used liquid chromatography−mass spectrometry (LC-MS) to profile chemical components; consistent with our findings, they reported that most of the amino acid-related components were positively correlated with zatsumi in sake samples (**Figure 1**). In the MLR model, only alanine was selected among amino acids; the other metabolites were eliminated in the feature selection because these amino acids were highly correlated (**Figure 2**). Iwano et al. also reported that only a few metabolites including carnosine were negatively correlated (*32*). Carnosine was detected by CE-TOFMS, but its peak was noted in only a few samples.

Predictions Using SVR. In the training SVR models, we also used randomly separated data sets including training and validation data sets; CV was conducted with only training data sets. Comparative results using feature selection methods are summarized in **Table 2**. The SVR models with 108 metabolites (without feature selection) showed the best correlation coefficients between observed and predicted sensory scores (R_t = 0.860). However, the correlation at CV dramatically deteriorated (R_{cv} = 0.472), which implies overtraining of the SVR model specific for the given data sets and a decrease in versatility. Results obtained during both the training phase and CV with CFS and RA methods showed lower R_t and R_{cv} values than did stepwise methods. Thus, the stepwise method exhibited the best correlation at CV for the data sets used in this study (R_{cv} = 0.729).

Supplementary Figure S2 of the Supporting Information presents the correlation coefficients (R_{cv}) obtained with the CV procedure between the predicted and actual sensory evaluation scores using the SVR parameters (trade-off between training error and margin C, coefficient of RBF kernel γ, and ε width of tube for regression ε). The optimized parameters with C = 10.0, γ = 0.1, and ε = 0.1 yielded the highest correlation coefficient in CV. The R_{cv} values were relatively constant except when C was close to 0 (Figure S2A, Supporting Information). In Figure S2B (Supporting Information), although a small γ value tended to produce a better R_{cv} value, the changes in R_{cv} (Figure S2B, Supporting Information) were smaller than in Figure S2A (Supporting Information). Thus, the optimized parameters were insensitive to changes in the value of the parameters, suggesting the stability of the model against tuning parameters.

Table 2. Prediction Performance of Each Variable Selection Method Using Support Vector Regression Models

		correlation		training options		
variable selection	no. of features	R_t	R_{cv}	γ	ε	C
all variables	108	0.830	0.472	0.1	0.12	0.1
stepwise	8	0.860	0.729	0.1	0.1	10
CFS	12	0.779	0.646	0.2	0	0.1
RA	13	0.756	0.589	0.1	0	0.5

Figure 3 depicts the predicted and observed sensory evaluation scores for zatsumi obtained by SVR and the MLR model, respectively. Supporting Information Table S4 summarizes the correlation coefficients and errors between predicted and observed sensory evaluation scores using whole samples or CV in training and validation data sets. The SVR model yielded a better prediction in CV for zatsumi (R_{cv} = 0.729, mean error (ME) = 0.138), sweetness (R_{cv} = 0.853, ME = 0.120), and sourness (R_{cv} = 0.835, ME = 0.078) than MLR (R_{cv} = 0.724, 0.770, and 0.802, ME = 0.138, 0.112, 0.079). In contrast, MLR was more accurate (R_{cv} = 0.719, ME = 0.167) than SVR (R_{cv} = 0.654, ME = 0.183) for bitterness. In the validation test, the SVR for zatsumi (R_v = 0.728, ME = 0.161) and sourness (R_{cv} = 0.735, ME = 0.192) was better than MLR (R_{cv} = 0.654 and 0.705, ME = 0.188 and 0.173). With regard to bitterness, the correlation coefficients in the validation test were considerably worse in both MLR (R_v = 0.054, ME = 0.324) and SVR (R_v = 0.115, ME = 0.291); the values were derived from a few samples in the validation data set; nonetheless, SVR was better considering the mean errors.

To overcome this problem in the evaluation of prediction accuracy and versatility of the developed model using a relatively small sample number, we artificially added random noise to the measured data. We artificially generated different intensities (5, 10, and 15% relative to the average of the metabolite concentration or sensory evaluation socres) of white noise to the metabolites, to the sensory evaluation score, or to both. For each case, 200 different trials were conducted with different random numbers; averaged results are summarized in Supporting Information Table S5. For zatsumi and sweetness, the mean errors yielded by SVR models were lower than those of MLR models in all instances subjected to validation tests. In contrast, the mean errors produced by SVR models were almost the same with respect to sourness and greater with respect to bitterness compared to MLR models. Thus, with the data we used, the prediction performance and versatility of SVR was not always superior to that of MLR. On the other hand, the SVR can express more complex associations among features. The highly correlated metabolites can be exchanged with the features used in the currently developed models of both SVR and MLR. As shown in **Figures 1** and **2**, a number of metabolites correlated with other metabolites, and we obtained a sensory evaluation score for

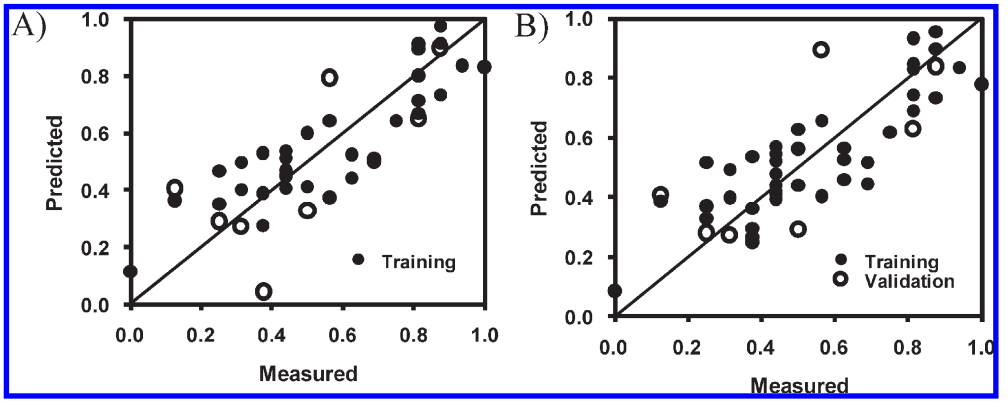

Figure 3. Predicted and actual normalized sensory evaluation scores of SVR (**A**) and MLR (**B**). Solid circles were predicted using complete data sets and the training options yielding the best correlation coefficient through the CV procedure. Open circles show the results of validation data sets predicted by the trained predictive model. In panel **B**, a plot in validation data sets ((*x*, *y*) = (0.375, −0.017)) was omitted.

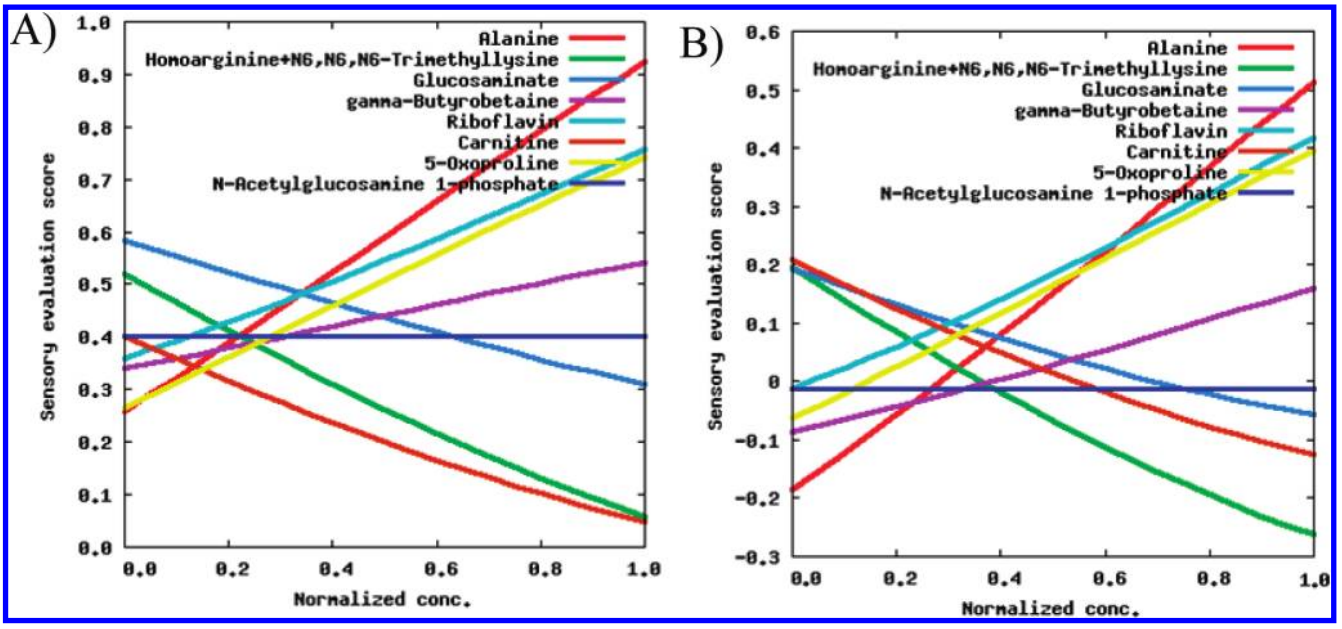

Figure 4. Relationships between sensory evaluation scores of zatsumi and systematically altered concentrations of each metabolite. These values were calculated by the trained SVR with metabolites selected by the stepwise feature selection method and yielded the best correlation coefficient by the CV procedure. Panels **A** and **B** show the results from sake samples 35 and 37 as examples. The concentration on the *X* axis and the sensory evaluation score on the *Y* axis are linearly normalized from 0 to 1.

sourness and zatsumi, implying that other models that incorporate alternative metabolites can be developed. In contrast, there were few metabolites that correlated with sweetness or bitterness (**Figure 1**). Therefore, it is possible that charged metabolites do not contribute significantly to these sensory flavors.

Simulation with Sensory Evaluation Values and Metabolite Concentrations Using Trained SVR. Figure 4 presents examples of the quantitative relationship between the predicted sensory evaluation scores of zatsumi and altered concentrations of each metabolite selected by stepwise feature selection using the trained SVR model yielding the best R_{cv} value. The simulated trajectories were almost straight lines, although the degree of the gradient decreased slightly in a nonlinear manner when the normalized concentration of glucosaminate, carnitine, and homoarginine + *N*6,*N*6,*N*6-trimethyllysine was high. Overall positive/negative correlations between individual metabolites and their sensory evaluation scores in the MLR and SVR models showed a similar trend (**Table 1**). For example, an increase in alanine raised the sensory evaluation score drastically, whereas an increase in homoarginine + *N*6,*N*6,*N*6-trimethyllysine and carnitine decreased the score, although the change was relatively smaller than for alanine. The coefficients of the MLR model for these three metabolites were 0.863 (p = 0.0006), −0.606 (p = 0.015), and −0.431 (p = 0.022), indicating that the absolute value of the coefficients and the degree of the gradient of the simulated curve exhibited consistency. However, the sensory evaluation score remained almost constant even though the normalized concentration of *N*-acetylglucosamine-1-phosphate was changed. This conflicts with the coefficients in the MLR model (0.430, p = 0.0002), possibly due to the relatively low concentration of this metabolite. Thus, compared to the MLR, the SVR modeling methodology provides two advantages: (1) it captures the nonlinear relationship between metabolites and sensory evaluation scores and (2) it reduces the risk of overestimating the effect of metabolites such as *N*-acetylglucosamine-1-phosphate on zatsumi that contribute to the sensory evaluation. For further understanding of the relations among sensory evaluation and selected features, multiple features of SVR models should be altered simultaneously and visualized as their sensitivities.

Limitation of This Study. To explore the relationship between sensory evaluation scores and the constitutive components in diverse beverages, artificial perturbation, that is, omission or excessive adjunction of a single component, and changes in the sensory evaluation criteria and taste recognition threshold concentrations of constituent components are traditional strategies (*10*,*36*). Although this bottom-up approach can catalogue the sensory properties of individual molecules, combinations of components hinder investigation of the individual sensory contribution made by multiple components. In contrast, the method demonstrated in our study represents a top-down approach that captures the overview or general patterns in a feature space, here the metabolite profiles of sake samples, and their association with sensory scores. Needless to say, this approach cannot substantiate the contribution to sensory perceptions attributable to individual metabolites.

Representing a limitation in our profiling technique, under our CE-MS measurement conditions, D- and L-isomers of amino acids could not be separated. On the other hand, humans have receptors that yield a distinct taste of these isomers (*62*). D- and L-lactate in sake, the most abundant charged metabolites that are expected to show different organoleptic properties, were separately quantified. A high correlation between L-lactate and total lactate and a relatively constant D-lactate level have been reported during the sake brewing process (*29*). Integrated analyses with profiling that include individual quantification of these isomers, low-polarity molecules, or volatile compounds (*63*) and involve other hyphenated MS technology are necessary for a deeper understanding of these phenomena. Although we conducted nontarget analysis, a number of unannotated peaks were observed, and only 14% of the observed peaks were identified and used in our subsequent analyses. Because the carrying capacity of the capillaries we used was low, we used sample aliquots for transfer to other measurement instrumentation. In addition, systematic MS/MS identification is difficult due to the wide diversity of the chemical structures and the limited availability of reference databases.

Here we first demonstrate an integrated analysis with large-scale charged metabolome profiling, statistical analysis, and SVR modeling of refined sake samples. For sensory evaluation, zatsumi and sourness were used, and relatively positively correlated metabolites including not only amino acids but also organic acids were found; these metabolites were correlated with each other. The modeling based on the SVR developed here successfully captured the nonlinear relationship between sensory evaluation scores, taking into account both the absolute quantity and relative changes in the concentration of the metabolites. In addition, numerical simulation using the established model clearly visualized these relations. The methodology demonstrated in this study may facilitate prediction of quantitative relations reflected in the quality of foodstuffs and beverages other than sake.

ABBREVIATIONS USED

ANN, artificial neural networks; CE-MS, capillary electrophoresis-mass spectrometry; CV, cross-validation; ESI, electrospray ionization; MLR, multiple logistic regression; SOM, self-organizing maps; PCA, principal component analysis; PLS-DA, partial least-squares-discriminant analysis; SMV, sake meter value; SVM, support vector machine; SVR, support vector regression.

ACKNOWLEDGMENT

We thank the professional panelists including Drs. Ishigaki, Muraoka, and Kudo of the Yamagata Research Institute of Technology for conducting the sensory evaluation tests and Drs. Sugawara, Ikeda, Suzuki, and Ueda. We also thank Dr. Kazuki Saito at Riken Plant Science Center and Dr. Goto at the Institute of Advanced Biosciences for fruitful discussions.

Supporting Information Available: Figures S1 and S2 and Tables S1−S5. This material is available free of charge via the Internet at http://pubs.acs.org.

LITERATURE CITED

(1) Ishihara, S.; Ikeda, A.; Citterio, D.; Maruyama, K.; Hagiwara, M.; Suzuki, K. Smart chemical taste sensor for determination and prediction of taste qualities based on a two-phase optimized radial basis function network. *Anal. Chem.* **2005**, *77*, 7908–7915.

(2) de Sousa, H. C.; Riul, A., Jr. In *Using MLP Networks to Classify Red Wines and Water Readings of an Electronic Tongue*, VII Brazilian Symposium on Neural Networks (SBRN'02), **2002**.

(3) Ferreira, E. J.; Pereira, R. C. T.; Delbem, A. C. B.; O.N., O. J.; Mattoso, L. H. C. Random subspace method for analysing coffee with electronic tongue. *Electron. Lett.* **2007**, *43* (21), 1138–1139.

(4) Yamagaki, A.; Leudermir, T. B.; de Souto, M. C. P. Classification of vintages of wine by artificial nose using time delay neural networks. *IEE Electron. Lett.* **2001**, *37* (24), 1466–1467.

(5) Alcazar, A.; Ballesteros, O.; Jurado, J. M.; Pablos, F.; Martin, M. J.; Vilches, J. L.; Navalon, A. Differentiation of green, white, black, oolong, and pu-erh teas according to their free amino acids content. *J. Agric. Food Chem.* **2007**, *55*, 5960–5965.

(6) Syu, K. Y.; Lin, C. L.; Huang, H. C.; Lin, J. K. Determination of theanine, GABA, and other amino acids in green, oolong, black, and pu-erh teas with dabsylation and high-performance liquid chromatography. *J. Agric. Food Chem.* **2008**, *56*, 7637–7643.

(7) Yu, H.; Ding, Y. S.; Mou, S. F. Direct and simultaneous determination of amino acids and sugars in rice wine by high-performance anion-exchange chromatography with integrated pulsed amperometric detection. *Chromatographia* **2003**, *57*, 721–728.

(8) Nose, A.; Myojin, M.; Hojo, M.; Ueda, T.; Okuda, T. Proton nuclear magnetic resonance and Raman spectroscopic studies of Japanese sake, an alcoholic beverage. *J. Biosci. Bioeng.* **2005**, *99* (5), 493–501.

(9) Yamada, T.; Furukawa, K.; Hara, S.; Mizoguchi, H. Effect of amino acids on peptide transport. *J. Biosci. Bioeng.* **2005**, *99* (4), 383–389.

(10) Hufnagel, J. C.; Hofmann, T. Quantitative reconstruction of the nonvolatile sensometabolome of a red wine. *J. Agric. Food Chem.* **2008**, *56*, 9190–9199.

(11) Mahadevan, S.; Shah, S. L.; Marrie, T. J.; Slupsky, C. M. Analysis of metabolomic data using support vector machines. *Anal. Chem.* **2008**, *80*, 7562–7570.

(12) Chen, Q.; Guo, Z.; Zhao, J. Identification of green tea's (*Camellia sinensis* (L.)) quality level according to measurement of main catechins and caffeine contents by HPLC and support vector classification pattern recognition. *J. Pharm. Biomed. Anal.* **2008**, *48* (5), 1321–1325.

(13) Acevedo, F. J.; Jimenez, J.; Maldonado, S.; Dominguez, E.; Narvaez, A. Classification of wines produced in specific regions by UV-visible spectroscopy combined with support vector machines. *J. Agric. Food Chem.* **2007**, *55*, 6842–6849.

(14) Diaz, C.; Conde, J. E.; Estevez, D.; Perez Olivero, S. J.; Perez Trujillo, J. P. Application of multivariate analysis and artificial neural networks for the differentiation of red wines from the Canary Islands according to the island of origin. *J. Agric. Food Chem.* **2003**, *51*, 4303–4307.

(15) Mounet, F.; Lemaire-Chamley, M.; Maucourt, M.; Cabasson, C.; Giraudel, J.-L.; Deborde, C.; Lessire, R.; Gallusci, P.; Bertrand, A.; Gaudillere, M.; Rothan, C.; Rolin, D.; Moinga, A. Quantitative metabolic profiles of tomato flesh and seeds during fruit development: complementary analysis with ANN and PCA. *Metabolomics* **2007**, *3* (3), 273–288.

(16) Son, H. S.; Hwang, G. S.; Kim, K. M.; Ahn, H. J.; Park, W. M.; Van Den Berg, F.; Hong, Y. S.; Lee, C. H. Metabolomic studies on geographical grapes and their wines using ^{1}H NMR analysis coupled with multivariate statistics. *J. Agric. Food Chem.* **2009**, *57*, 1481–1490.

(17) Son, H. S.; Kim, K. M.; van den Berg, F.; Hwang, G. S.; Park, W. M.; Lee, C. H.; Hong, Y. S. 1H nuclear magnetic resonance-based metabolomic characterization of wines by grape varieties and production areas. *J. Agric. Food Chem.* **2008**, *56*, 8007–8016.

(18) Funahashi, K. On the approximate realization of continuous mappings by neural networks. *Neural Networks* **1989**, *2* (3), 183–192.

(19) Vapnik, V. N. *The Nature of Statistical Learning Theory*; Springer: Berlin, Germany, 1995.

(20) Ferreira, E. J.; Pereira, R. C. T.; Delbem, A. C. B.; O. N., O. J.; Mattoso, L. H. C. Random subspace method for analysing coffee with electronic tongue. *Electron. Lett.* **2007**, *43* (21), 1138–1139.

(21) Hayashi, N.; Chen, R.; Ikezaki, H.; Ujihara, T. Evaluation of the umami taste intensity of green tea by a taste sensor. *J. Agric. Food Chem.* **2008**, *56*, 7384–7387.

(22) Cozzolino, D.; Cowey, G.; Lattey, K. A.; Godden, P.; Cynkar, W. U.; Dambergs, R. G.; Janik, L.; Gishen, M. Relationship between wine scores and visible-near-infrared spectra of Australian red wines. *Anal. Bioanal. Chem.* **2008**, *391*, 975–981.

(23) Ustun, B.; Melssen, W. J.; Buydens, L. M. Visualisation and interpretation of support vector regression models. *Anal. Chim. Acta* **2007**, *595* (1−2), 299–309.

(24) Seo, M. Y.; Chung, S. Y.; Choi, W. K.; Seo, Y. K.; Jung, S. H.; Park, J. M.; Seo, M. J.; Park, J. K.; Kim, J. W.; Park, C. S. Anti-aging effect of rice wine in cultured human fibroblasts and keratinocytes. *J. Biosci. Bioeng.* **2009**, *107* (3), 266–271.

(25) Hirotsune, M.; Haratake, A.; Komiya, A.; Sugita, J.; Tachihara, T.; Komai, T.; Hizume, K.; Ozeki, K.; Ikemoto, T. Effect of ingested concentrate and components of sake on epidermal permeability barrier disruption by UVB irradiation. *J. Agric. Food Chem.* **2005**, *53*, 948–952.

(26) Gronbaek, M.; Deis, A.; Sorensen, T. I.; Becker, U.; Schnohr, P.; Jensen, G. Mortality associated with moderate intakes of wine, beer, or spirits. *Br. Med. J.* **1995**, *310* (6988), 1165–1169.

(27) Fuchs, C. S.; Stampfer, M. J.; Colditz, G. A.; Giovannucci, E. L.; Manson, J. E.; Kawachi, I.; Hunter, D. J.; Hankinson, S. E.; Hennekens, C. H.; Rosner, B. Alcohol consumption and mortality among women. *N. Engl. J. Med.* **1995**, *332* (19), 1245–1250.

(28) Thun, M. J.; Peto, R.; Lopez, A. D.; Monaco, J. H.; Henley, S. J.; Heath, C. W., Jr.; Doll, R. Alcohol consumption and mortality among middle-aged and elderly U.S. adults. *N. Engl. J. Med.* **1997**, *337* (24), 1705–1714.

(29) Kodama, S.; Yamamoto, A.; Matsunaga, A.; Matsui, K.; Nakagomi, K.; Hayakawa, K. Behaviors of D- and L-lactic acids during the brewing process of sake (Japanese rice wine). *J. Agric. Food Chem.* **2002**, *50*, 767–770.

(30) Tanimoto, S.; Matsumoto, H.; Jujii, K.; Ohdoi, R.; Sakamoto, K.; Yamane, Y.; Miyake, M.; Shimoda, M.; Osajima, Y. Enzyme inactivation and quality preservation of sake by high-pressure carbonation at a moderate temperature. *Biosci., Biotechnol., Biochem.* **2008**, *72* (1), 22–28.

(31) Izawa, S.; Kita, T.; Ikeda, K.; Miki, T.; Inoue, Y. Formation of cytoplasmic P-bodies in sake yeast during Japanese sake brewing and wine making. *Biosci., Biotechnol., Biochem.* **2007**, *71* (11), 2800–2807.

(32) Iwano, K.; Ito, T.; Nakazawa, N. Correlation analysis of a sensory evaliation and the chemical components of Ginjo-shu. *J. Brew. Soc. Jpn.* **2005**, *100* (9), 639−649 (in Japanese).

(33) Hashizume, K.; Okuda, M.; Numata, M.; Iwashita, K. Bitter-tasting sake peptides derived from the N-terminus of the rice glutelin acidic subunit. *Food Sci. Technol. Res.* **2007**, *13* (3), 270–274.

(34) Okamoto, M.; Yamauchi, T.; Yano, S.; Kurose, N.; Kawakita, S.; Takahashi, K.; Nakamura, T. Preservation of sake quality by decreasing the dissolved oxygen concentration. *J. Brew. Soc. Jpn.* **1999**, *94* (10), 827−832 (in Japanese).

(35) Moreno, N. J.; Marco, A. G.; Azpilicueta, C. A. Influence of wine turbidity on the accumulation of volatile compounds from the oak barrel. *J. Agric. Food Chem.* **2007**, *55*, 6244–6251.

(36) Iwano, K.; Takahashi, K.; Ito, T.; Nakazawa, N. Search for amino acids affecting the taste of Japanese sake. *J. Brew. Soc. Jpn.* **2004**, *99* (9), 659−664 (in Japanese).

(37) Matsuura, K.; Hirotsune, M.; Hamachi, M.; Nunokawa, Y. Modeling of the sensory evaluation of sake by Dempster−Shafer's measure and genetic algorithm. *J. Ferment. Bioeng.* **1995**, *79* (1), 45–53.

(38) Kawai, M.; Okiyama, A.; Ueda, Y. Taste enhancements between various amino acids and IMP. *Chem. Senses* **2002**, *27*, 739–745.

(39) Otsuka, H. In *Analysis for Evaluated Expressions of Tasting Japanese Sake*, The 18th Annual Conference of the Japanese Society for Artificial Intelligence, **2004**; pp 2D2−11 (in Japanese).

(40) Arikawa, Y.; Kobayashi, M.; Kodaira, R.; Shimosaka, M.; Muratsubaki, H.; Enomoto, K.; Okazaki, M. Isolation of sake yeast strains possessing various levels of succinate- and/or malate-producing abilities by gene disruption or mutation. *J. Biosci. Bioeng.* **1999**, *87* (3), 333–339.

(41) Kirimura, J.; Shimizu, A.; Kimizuka, A.; Ninomiya, T.; Katsuya, N. Contribution of peptides and amino acids to the taste of foods. *J. Agric. Food Chem.* **1969**, *17*, 689–695.

(42) Soga, T.; Ohashi, Y.; Ueno, Y.; Naraoka, H.; Tomita, M.; Nishioka, T. Quantitative metabolome analysis using capillary electrophoresis mass spectrometry. *J. Proteome Res.* **2003**, *2*, 488–494.

(43) Soga, T.; Baran, R.; Suematsu, M.; Ueno, Y.; Ikeda, S.; Sakurakawa, T.; Kakazu, Y.; Ishikawa, T.; Robert, M.; Nishioka, T.; Tomita, M. Differential metabolomics reveals ophthalmic acid as an oxidative stress biomarker indicating hepatic glutathione consumption. *J. Biol. Chem.* **2006**, *281* (24), 16768–16776.

(44) Toko, K. Taste sensor with global selectivity. *Mater. Sci. Eng. C* **1996**, *4*, 69–82.

(45) Manabe, Y.; Utsunomiya, H.; Gotoh, K.; Kurosu, T.; Fushiki, T. Relationship between the preference for sake (Japanese rice wine) and the movements of metabolic parameters coinciding with sake intake. *Biosci., Biotechnol., Biochem.* **2004**, *68* (4), 796–802.

(46) Hirayama, A.; Kami, K.; Sugimoto, M.; Sugawara, M.; Toki, N.; Onozuka, H.; Kinoshita, T.; Saito, N.; Ochiai, A.; Tomita, M.; Esumi, H.; Soga, T. Quantitative metabolome profiling of colon and stomach cancer microenvironment by capillary electrophoresis time-of-flight mass spectrometry. *Cancer Res.* **2009**, *69* (11), 4918–4925.

(47) Soga, T.; Igarashi, K.; Ito, C.; Mizobuchi, K.; Zimmermann, H. P.; Tomita, M. Metabolomic profiling of anionic metabolites by capillary electrophoresis mass spectrometry. *Anal. Chem.* **2009**, *81*, 6165−6174

(48) Sugimoto, M.; Wong, T. D.; Hirayama, A.; Soga, T.; Tomita, M. Capillary electrophoresis mass spectrometry-based saliva metabolomics identified oral, breast and pancreatic cancer-specific profiles. *Metabolomics* **2009**, DOI: 10.1007/s11306-009-0178-y.

(49) Sugimoto, M.; Hirayama, A.; Takamasa, I.; Robert, M.; Baran, R.; Uehara, K.; Kawai, K.; Soga, T.; Tomita, M. Differential metabolomics software for capillary electrophoresis-mass spectrometry data analysis. . *Metabolomics* **2009**, DOI: 10.1007/s11306-009-0175-1.

(50) Brown, M.; Dunn, W. B.; Dobson, P.; Patel, Y.; Winder, C. L.; Francis-McIntyre, S.; Begley, P.; Carroll, K.; Broadhurst, D.; Tseng, A.; Swainston, N.; Spasic, I.; Goodacre, R.; Kell, D. B. Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics. *Analyst* **2009**, *134* (7), 1322–1332.

(51) Kitamoto, K.; Oda, K.; Gomi, K.; Takahashi, K. Genetic engineering of a sake yeast producing no urea by successive disruption of arginase gene. *Appl. Environ. Microbiol.* **1991**, *57* (1), 301–306.

(52) Saeed, A. I.; Sharov, V.; White, J.; Li, J.; Liang, W.; Bhagabati, N.; Braisted, J.; Klapa, M.; Currier, T.; Thiagarajan, M.; Sturn, A.; Snuffin, M.; Rezantsev, A.; Popov, D.; Ryltsov, A.; Kostukovich, E.; Borisovsky, I.; Liu, Z.; Vinsavich, A.; Trush, V.; Quackenbush, J. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques* **2003**, *34* (2), 374–378.

(53) Chiu, S. H.; Chen, C. C.; Lin, T. H. Using support vector regression to model the correlation between the clinical metastases time and gene expression profile for breast cancer. *Artif. Intell. Med.* **2008**, *44* (3), 221–231.

(54) Joachims, T. *Making Large-Scale SVM Learning Practical. Advances in Kernel Methods − Support Vector Learning*; MIT Press: Cambridge, MA, 1999.

(55) Hermes, L.; Buhmann, J. M. Feature selection for support vector machine. *Pattern Recognition, 2000*, Proceedings of the 15th International Conference; **2001**; Vol. *2*, pp 712−715.

(56) Witten, I. H.; Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed.; Morgan Kaufman: Amsterdam, The Netherlands, 2005.

(57) Kononenko, I. In *Estimating Attributes: Analysis and Extensions of RELIEF*, European Conference on Machine Learning, **1994**; pp 171−182.

(58) Okamoto, A.; Sugi, E.; Koizumi, Y.; Yanagida, F.; Udaka, S. Polyamine content of ordinary foodstuffs and various fermented foods. *Biosci., Biotechnol., Biochem.* **1997**, *61* (9), 1582–1584.

(59) Tikk, M.; Tikk, K.; Torngren, M. A.; Meinert, L.; Aaslyng, M. D.; Karlsson, A. H.; Andersen, H. J. Development of inosine monophosphate and its degradation products during aging of pork of different qualities in relation to basic taste and retronasal flavor perception of the meat. *J. Agric. Food Chem.* **2006**, *54*, 7769–7777.

(60) Cairoli, P.; Pieraccini, S.; Sironi, M.; Morelli, C. F.; Speranza, G.; Manitto, P. Studies on umami taste. Synthesis of new guanosine 5′-phosphate derivatives and their synergistic effect with monosodium glutamate. *J. Agric. Food Chem.* **2008**, *56*, 1043–1050.

(61) Pfeiffer, P.; Konig, H. *Pyroglutamic Acid: A Novel Compound in Wines*; Springer: Berlin, Germany, 2009; pp 233−240.

(62) Nelson, G.; Chandrashekar, J.; Hoon, M. A.; Feng, L.; Zhao, G.; Ryba, N. J.; Zuker, C. S. An amino-acid taste receptor. *Nature* **2002**, *416* (6877), 199–202.

(63) Isogai, A.; Utsunomiya, H.; Kanda, R.; Iwata, H. Changes in the aroma compounds of sake during aging. *J. Agric. Food Chem.* **2005**, *53*, 4118–4123.

Received for review May 19, 2009. Revised manuscript received November 13, 2009. Accepted November 16, 2009. This work was supported by research funds from the Yamagata Prefectural Government and Tsuruoka City.